

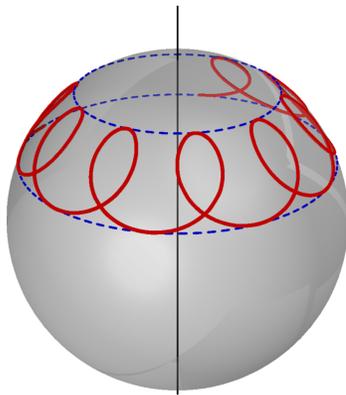
MÉCANIQUE II

PHQ414

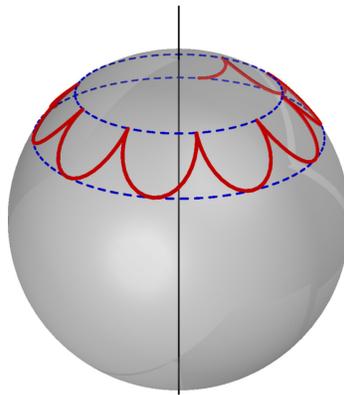
par

David SÉNÉCHAL

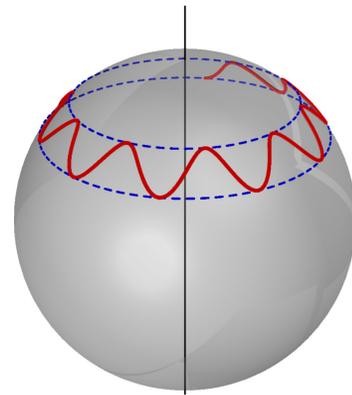
Ph.D., Professeur Titulaire



(A)



(B)



(C)



UNIVERSITÉ DE SHERBROOKE

Faculté des sciences

Département de physique

26 juin 2018

Table des matières

1 Mécanique de Lagrange	9
A Équations de Lagrange	9
1.A.1 Coordonnées généralisées	9
1.A.2 Forces de contrainte et déplacements virtuels	11
1.A.3 Équations de Lagrange	12
B Applications élémentaires	14
C Principe de la moindre action	20
1.C.1 Énoncé et démonstration	20
1.C.2 La brachistochrone	22
1.C.3 Méthode des multiplicateurs de Lagrange	26
D Application aux petites oscillations	28
1.D.1 Système masses-ressorts	28
1.D.2 Modes propres	30
2 Mécanique de Hamilton	39
A Équations de Hamilton	39
2.A.1 Transformation de Legendre	39
2.A.2 Moments conjugués et hamiltonien	41
2.A.3 Exemples élémentaires	43
2.A.4 Principe de la moindre action	46
B Formalisme canonique	47
2.B.1 Crochets de Poisson	47
2.B.2 Théorèmes de Liouville et de Poincaré	49
2.B.3 Lois de conservation et symétries	51
C Transformations canoniques	53
2.C.1 Définitions et propriétés générales	53
2.C.2 Fonctions génératrices	54
2.C.3 Transformations canoniques infinitésimales	57
D Théorie de Hamilton-Jacobi	59
2.D.1 Équation de Hamilton-Jacobi	59
2.D.2 Fonction caractéristique de Hamilton	62
2.D.3 La mécanique classique comme limite de la mécanique ondulatoire	63
3 Mouvement des corps rigides	71
A Cinématique des corps rigides	71
3.A.1 Matrices de Rotation	71
3.A.2 Rotations passives et actives	74
B Vitesse angulaire	75
3.B.1 Rotations infinitésimales	75
3.B.2 Caractère vectoriel de la vitesse angulaire	76

3.B.3	Composition des vitesses	76
3.B.4	Angles d'Euler	78
C	Théorème du moment cinétique	80
D	Tenseur d'inertie	82
3.D.1	Relation entre vitesse angulaire et moment cinétique	82
3.D.2	Tenseurs et axes principaux	83
E	Mouvement libre d'un objet rigide	86
3.E.1	Équations d'Euler	86
3.E.2	Rotation libre d'un objet symétrique	86
3.E.3	Rotation libre d'un objet asymétrique	88
F	Mouvement d'une toupie symétrique	90
3.F.1	Lagrangien, hamiltonien et problème effectif à une variable	90
3.F.2	Précession uniforme	92
3.F.3	Solution générale	94
3.F.4	Toupie dormante	96
G	Précession des équinoxes	97
3.G.1	Énergie potentielle d'un objet plongé dans un champ gravitationnel	97
3.G.2	Fréquence de précession	100
4	Forces centrales	107
A	Solution générale	107
4.A.1	Réduction du problème à deux corps	107
4.A.2	Problème radial effectif	109
4.A.3	Intégrale de la trajectoire	111
4.A.4	Théorème du viriel	112
B	Le problème de Kepler	114
4.B.1	Sections coniques	114
4.B.2	Équation de Kepler	117
4.B.3	Repérage des orbites dans l'espace	120
C	Théorie classique de la diffusion par un potentiel central	122
4.C.1	Section efficace	122
4.C.2	Formule de Rutherford	124
D	Forces centrales et équation de Hamilton-Jacobi	126
4.D.1	Équation de Hamilton-Jacobi dans le cas d'un potentiel azimutal	126
4.D.2	Cas d'un potentiel central	128
4.D.3	Variables d'action-angle	128
4.D.4	Variables d'action-angle dans le problème de Kepler	130
E	Introduction à la théorie des perturbations	134
4.E.1	Méthode de variation des constantes	134
4.E.2	Précession des orbites	136
5	Relativité et électromagnétisme	143
A	Théorie de la relativité restreinte	143
5.A.1	Espace-temps	143
5.A.2	Exemples d'invariants et de quadrivecteurs	145
B	Dynamique relativiste	148
5.B.1	Action d'une particule libre	148

5.B.2	Particule chargée dans un champ électromagnétique	149
5.B.3	Forme covariante de l'équation du mouvement	151
5.B.4	Tenseur de Faraday et transformation des champs	154
C	Théorie des circuits électriques	156
5.C.1	Description générale des circuits	156
5.C.2	Lagrangien et hamiltonien d'un circuit	159
6	Introduction aux systèmes chaotiques	167
A	Problème de Hénon-Heiles	167
6.A.1	Potentiel de Hénon-Heiles	167
6.A.2	Sections de Poincaré	168
6.A.3	Stabilité et exposant de Liapounov	171
B	Dynamique discrète et approche du chaos	172
7	Annexes	177
A	Rappels d'algèbre linéaire	177
7.A.1	Notations indicielle et matricielle	177
7.A.2	Changements de base	180
7.A.3	Vecteurs et valeurs propres d'une matrice	180
B	Vecteurs et tenseurs	184

Table des problèmes

1.1	Machine d'Atwood généralisée	34
1.2	Mouvement le long d'une cycloïde	34
1.3	Pendule à pivot tournant	35
1.4	Masse pesante se déplaçant sur un cône	36
1.5	Masse pesante sur une sphère	36
1.6	Objet glissant le long d'un cylindre	36
1.7	Roulement d'un objet cylindrique sur un autre	36
1.8	Bille sur un anneau	37
1.9	Géodésiques	37
1.10	Systèmes masses-ressorts	38
1.11	Pendules couplés	38
2.1	Film de savon	66
2.2	Pendule avec tige élastique	66
2.3	Pendule à pivot tournant : hamiltonien	67
2.4	Masse glissant le long d'une hélice libre de tourner	67
2.5	Masse glissant le long d'une hélice	68
2.6	Coordonnées elliptiques	68
2.7	Crochets de Poisson et moment cinétique	68
2.8	Transformations canoniques	69
2.9	Équation de Hamilton-Jacobi pour une force constante	69
2.10	Équation de Hamilton-Jacobi en coordonnées elliptiques	69
3.1	Matrice de rotation générale	102
3.2	Rotations et matrices de Pauli	102
3.3	Matrices de rotations	103
3.4	Vitesse angulaire dans le repère fixe	103
3.5	Tenseur d'inertie de la molécule de méthane	103
3.6	Moment d'inertie d'un disque mince	104
3.7	Moments d'inertie d'un ellipsoïde	104
3.8	Angles d'Euler et rotation libre	104
3.9	Toupie dormante et potentiel effectif	104
3.10	Toupie reposant sur une plaque sans frottement	104
3.11	Précession rapide	105
3.12	Rotation libre d'un objet asymétrique	105
3.13	Mouvement de billes magnétiques	105
4.1	Diffusion de Rutherford : cas attractif	139
4.2	Diffusion par un potentiel en $1/r^2$	139
4.3	Diffusion par une sphère molle	139
4.4	Théorie des perturbations pour l'oscillateur anharmonique	139
4.5	Perturbation d'une orbite par une force constante	140
4.6	Précession de l'orbite d'un satellite	140

Table des matières

5.1	Forme générale de la transformation de Lorentz	163
5.2	Tenseur métrique et élément de volume	163
5.3	Effet Compton	163
5.4	Force de Lorentz	163
5.5	Mouvement dans des champs électrique et magnétique croisés	164
5.6	Hamiltonien d'une particule dans un champ magnétique uniforme	164
5.7	Problème de Kepler relativiste	164
5.8	Ligne de transmission	165

CHAPITRE 1

Mécanique de Lagrange

A Équations de Lagrange

1.A.1 Coordonnées généralisées

Considérons un système mécanique comportant N particules évoluant en trois dimensions. Les positions de ces particules seront notées \mathbf{r}_i ($i = 1, 2, \dots, N$). Chaque vecteur position comportant trois composantes, il faut $3N$ coordonnées pour spécifier la configuration du système dans son entier. Supposons en outre que ces $3N$ coordonnées ne sont pas indépendantes, c'est-à-dire qu'elles ne sont pas libres d'évoluer indépendamment, mais qu'elles sont reliées par un certain nombre K de contraintes, qui pourraient s'exprimer comme un ensemble de relations mathématiques explicites :

$$C_a(\mathbf{r}_1, \dots, \mathbf{r}_N, t) = 0 \quad (a = 1, 2, \dots, K) \quad (1.1)$$

Par exemple, deux particules de positions \mathbf{r}_1 et \mathbf{r}_2 pourraient être contraintes de manière à ce que leur distance relative soit constante. Dans ce cas, il n'y a qu'une seule contrainte ($K = 1$) qui s'exprime comme suit :

$$|\mathbf{r}_1 - \mathbf{r}_2| - d = 0 \quad (1.2)$$

Un autre exemple est celui du pendule simple contraint d'osciller dans le plan xy , et suspendu à l'origine par une tige infiniment mince et rigide de longueur l . Il y a alors deux contraintes ($K = 2$) imposées sur la coordonnée \mathbf{r} de la masse du pendule :

$$|\mathbf{r}| - l = 0 \quad \text{et} \quad \mathbf{r} \cdot \mathbf{z} = 0 \quad (1.3)$$

En général, on suppose que les contraintes peuvent aussi dépendre explicitement du temps. Ce serait le cas, par exemple, si le point de suspension d'un pendule oscillait de manière bien définie par rapport à une position moyenne, en raison d'un mécanisme externe au système étudié. Les contraintes qui peuvent s'exprimer comme une relation fonctionnelle explicite (1.1) sont qualifiées d'*holonomes*.

L'imposition de contraintes sur les positions des particules rend l'utilisation des coordonnées \mathbf{r}_i beaucoup moins pratique. Il est alors éminemment souhaitable de procéder à un changement de variables et d'introduire un autre ensemble de $3N$ variables, qu'on décomposera en deux groupes : un groupe de $n = 3N - K$ variables qu'on notera q_α ($\alpha = 1, 2, \dots, n$) et un autre groupe de K variables qu'on

notera précisément C_a ($a = 1, \dots, K$). Ces $3N$ variables sont en principe des fonctions connues des coordonnées des particules et, accessoirement, du temps :

$$\begin{aligned} q_\alpha &= q_\alpha(\mathbf{r}_1, \dots, \mathbf{r}_N, t) & (\alpha = 1, 2, \dots, n) \\ C_a &= C_a(\mathbf{r}_1, \dots, \mathbf{r}_N, t) & (a = 1, 2, \dots, K) \end{aligned} \quad (1.4)$$

Elles sont indépendantes, ce qui signifie que les coordonnées peuvent en retour s'exprimer en fonction de ces variables :

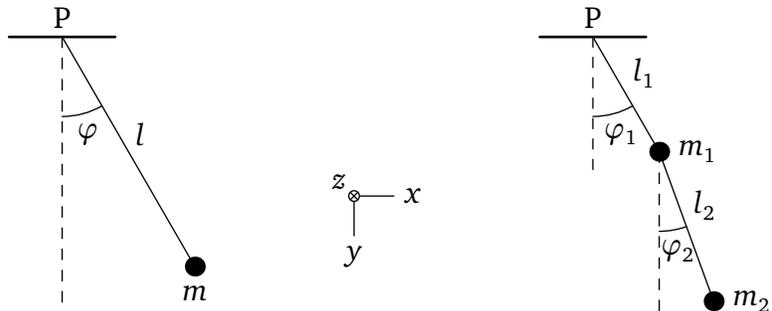
$$\mathbf{r}_i = \mathbf{r}_i(q_\alpha, C_a, t) \quad (1.5)$$

Ce qui rend ces variables utiles est que les K contraintes sont exprimées simplement par la relation $C_a = 0$, alors que les n autres variables q_α ne sont pas contraintes et donc libres de varier indépendamment les unes des autres. On appelle $n = 3N - K$ le nombre de *degrés de liberté* du système mécanique étudié. Les n variables q_α sont appelées *coordonnées généralisées*. L'espace défini par ces n variables contient toutes les configurations possibles du système respectant les contraintes et est appelé *espace des configurations*. Comme les K variables C_a sont toujours nulles, on peut simplement considérer les $3N$ coordonnées comme des fonctions seulement des coordonnées généralisées et du temps (la dépendance explicite en t n'est pas obligatoire et n'intervient que dans certaines situations seulement) :

$$\mathbf{r}_i = \mathbf{r}_i(q_\alpha, t) \quad (1.6)$$

La mécanique de Lagrange est une prescription générale qui permet d'obtenir directement des équations différentielles pour les coordonnées généralisées q_α , à partir d'une expression des énergies cinétique et potentielle en fonction des coordonnées généralisées.

FIGURE 1.1
Le pendule simple (à gauche) et le pendule double (à droite).



Exemple 1.1 Pendule simple

Considérons une masse m suspendue à un point P , qu'on prend comme origine, par une tige rigide de longueur l et de masse négligeable. Supposons en outre que cette tige est suspendue de manière à empêcher tout mouvement hors du plan xy . Ces deux contraintes réduisent le nombre de degrés de liberté à un seul. On peut choisir l'angle φ des coordonnées cylindriques comme coordonnée généralisée. Les trois variables q_α et C_a sont alors

$$q_1 = \varphi = \arctan(x/y) \quad (1.7)$$

$$C_1 = \sqrt{x^2 + y^2} - l \quad (1.8)$$

$$C_2 = z \quad (1.9)$$

et les 2 contraintes imposées au système sont $C_1 = 0$ et $C_2 = 0$.

Exemple 1.2 Le pendule double dans un plan

Comme extension de l'exemple précédent, ajoutons une deuxième masse (m_2), suspendue à la masse m_1 du pendule simple par une autre tige rigide de masse négligeable et de longueur l_2 , elle aussi contrainte de pivoter dans le plan xy seulement. Désignons par φ_2 l'angle que fait la deuxième tige par rapport à la verticale (contre φ_1 pour la première tige, de longueur l_1). Désignons par (x_1, y_1, z_1) et (x_2, y_2, z_2) les coordonnées cartésiennes des deux masses. Les six variables q_α et C_α sont alors

$$q_1 = \varphi_1 = \arctan\left(\frac{x_1}{y_1}\right) \quad (1.10)$$

$$q_2 = \varphi_2 = \arctan\left(\frac{x_2 - x_1}{y_2 - y_1}\right) \quad (1.11)$$

$$C_1 = \sqrt{x_1^2 + y_1^2} - l_1 \quad (1.12)$$

$$C_2 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} - l_2 \quad (1.13)$$

$$C_3 = z_1 \quad (1.14)$$

$$C_4 = z_2 \quad (1.15)$$

et les 4 contraintes sont $C_1 = C_2 = C_3 = C_4 = 0$.

1.A.2 Forces de contrainte et déplacements virtuels

Les contraintes $C_a(\mathbf{r}_i, t) = 0$ ($a = 1, \dots, K$) s'exerçant sur les $3N$ particules du système sont imposées par des forces, dont la valeur précise dépend généralement du mouvement suivi par les particules. Il est utile d'imaginer ces contraintes comme provenant d'une force de rappel extrêmement forte, qui ramène toujours les particules sur la surface $C_a(\mathbf{r}_i, t) = 0$. Supposons que cette force de rappel provienne d'un potentiel extrêmement serré, qu'on notera $V^c(\mathbf{r}_i)$. Ce potentiel a la propriété essentielle qu'on pourrait l'exprimer seulement en fonction des variables C_a (c'est-à-dire qu'il serait indépendant des q_α) et que son minimum serait situé au point $C_a = 0$. La force de contrainte agissant sur la particule i est alors

$$\mathbf{f}_i^c = -\frac{\partial V^c}{\partial \mathbf{r}_i} = -\sum_a \frac{\partial V^c}{\partial C_a} \frac{\partial C_a}{\partial \mathbf{r}_i} = \sum_a \lambda_a \frac{\partial C_a}{\partial \mathbf{r}_i} \quad (1.16)$$

où les constantes λ_a sont a priori inconnues, mais peuvent être obtenues une fois connue la solution des équations du mouvement $\mathbf{r}_i(t)$. Ceci est vrai même si l'équation de la contrainte dépend explicitement du temps.

Un déplacement infinitésimal $d\mathbf{r}_i$ de la particule i qui respecte les contraintes, pendant un temps dt , peut donc s'écrire

$$d\mathbf{r}_i = \sum_\alpha \frac{\partial \mathbf{r}_i}{\partial q_\alpha} dq_\alpha + \frac{\partial \mathbf{r}_i}{\partial t} dt \quad (1.17)$$

Le premier groupe de termes correspond à un *déplacement virtuel*, c'est-à-dire une variation de la position, compatible avec les contraintes à un moment donné, mais qui ne correspond pas à une véritable évolution dans le temps. Le deuxième terme ajoute à ce déplacement virtuel ce qu'il faut pour obtenir un déplacement réel, qui tient compte de l'évolution explicite des contraintes en fonction

du temps. Dans le cas où toutes les contraintes sont indépendantes du temps, les deux types de déplacements (réel et virtuel) sont équivalents. Comme les variables C_a sont indépendantes des coordonnées généralisées q_α , on peut écrire la relation suivante, par une application de la règle d'enchaînement :

$$0 = \frac{\partial C_a}{\partial q_\alpha} = \sum_i \frac{\partial C_a}{\partial \mathbf{r}_i} \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \quad (1.18)$$

En combinant les relations (1.16) et (1.18), on obtient

$$0 = \sum_a \sum_i \lambda_a \frac{\partial C_a}{\partial \mathbf{r}_i} \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} = \sum_i \mathbf{f}_i^{(c)} \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \quad (1.19)$$

Cette relation signifie que les forces de contrainte ne font aucun travail sous un déplacement virtuel (ou aucun 'travail virtuel') ; elle nous sera utile plus bas.

1.A.3 Équations de Lagrange

La force totale s'exerçant sur la particule i est la somme d'une force appliquée $\mathbf{f}_i^{(a)}$, associée par exemple à une certaine énergie potentielle V , et d'une force de contrainte. On peut donc écrire la deuxième loi de Newton comme

$$m_i \ddot{\mathbf{r}}_i - \mathbf{f}_i^{(a)} - \mathbf{f}_i^{(c)} = 0 \quad (1.20)$$

Appliquons maintenant la relation (1.19), en sommant sur les particules :

$$\sum_i (m_i \ddot{\mathbf{r}}_i - \mathbf{f}_i^{(a)}) \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} = 0 \quad (1.21)$$

Concentrons-nous sur le premier terme de cette équation :

$$\sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} = \sum_i \left(\frac{d}{dt} \left(m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \right) - m_i \mathbf{v}_i \cdot \frac{d}{dt} \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \right) \quad (1.22)$$

Les vitesses \mathbf{v}_i peuvent être exprimées en fonction des coordonnées généralisées q_α , de leurs dérivées par rapport au temps et, accessoirement, du temps :

$$\mathbf{v}_i = \sum_\alpha \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial \mathbf{r}_i}{\partial t} \quad (1.23)$$

Il ressort de cette équation que

$$\frac{\partial \mathbf{v}_i}{\partial \dot{q}_\alpha} = \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \quad (1.24)$$

et ceci sera utile dans le traitement du premier terme de (1.22). L'expression figurant dans le dernier terme de (1.22) peut s'exprimer comme suit :

$$\begin{aligned} \frac{d}{dt} \frac{\partial \mathbf{r}_i}{\partial q_\alpha} &= \sum_\beta \frac{\partial^2 \mathbf{r}_i}{\partial q_\alpha \partial q_\beta} \dot{q}_\beta + \frac{\partial^2 \mathbf{r}_i}{\partial q_\alpha \partial t} \\ &= \frac{\partial \mathbf{v}_i}{\partial q_\alpha} \end{aligned} \quad (1.25)$$

où nous avons utilisé l'éq. (1.23) dans la dernière égalité. Donc l'expression (1.22) peut être réécrite comme

$$\begin{aligned} \sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} &= \sum_i \left(\frac{d}{dt} \left(m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial \dot{q}_\alpha} \right) - m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial q_\alpha} \right) \\ &= \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} \end{aligned} \quad (1.26)$$

où T désigne l'énergie cinétique :

$$T = \sum_i \frac{1}{2} m_i \mathbf{v}_i^2 \quad (1.27)$$

L'énergie cinétique T peut être exprimée en fonction des coordonnées généralisées q_α , de leurs dérivées \dot{q}_α et du temps : $T(q_\alpha, \dot{q}_\alpha, t)$.

L'équation du mouvement de Newton peut donc être réécrite ainsi :

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} - \sum_i \mathbf{f}_i^{(a)} \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} = 0 \quad (1.28)$$

Supposons en outre que les forces appliquées $\mathbf{f}_i^{(a)}$ peuvent être dérivées d'une énergie potentielle :

$$\mathbf{f}_i^{(a)} = - \frac{\partial V}{\partial \mathbf{r}_i} \quad (1.29)$$

L'équation du mouvement devient alors

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} + \sum_i \frac{\partial V}{\partial \mathbf{r}_i} \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} = \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} + \frac{\partial V}{\partial q_\alpha} = 0 \quad (1.30)$$

Définissons alors la *fonction de Lagrange* ou *lagrangien*

$$L(q_\alpha, \dot{q}_\alpha, t) = T(q_\alpha, \dot{q}_\alpha, t) - V(q_\alpha, t) \quad (1.31)$$

Les équations du mouvement de Newton sont alors équivalentes à l'ensemble d'équations suivantes, appelées *équations de Lagrange* :

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} - \frac{\partial L}{\partial q_\alpha} = 0 \quad (1.32)$$

Forces non conservatives

Dans certains cas, la force appliquée n'est pas totalement conservative et contient une partie $\mathbf{f}_i^{(nc)}$ qui ne dérive pas d'un potentiel :

$$\mathbf{f}_i^{(a)} = - \frac{\partial V}{\partial \mathbf{r}_i} + \mathbf{f}_i^{(nc)} \quad (1.33)$$

Dans ce cas, l'équation du mouvement prend la forme

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} + \frac{\partial V}{\partial q_\alpha} - \sum_i \mathbf{f}_i^{(nc)} \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} = 0 \quad (1.34)$$

On définit alors la *force généralisée* comme

$$Q_\alpha = \sum_i \mathbf{f}_i^{(nc)} \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \quad (1.35)$$

de sorte qu'on écrit une version plus générale des équations de Lagrange comme

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} - \frac{\partial L}{\partial q_\alpha} = Q_\alpha \quad (1.36)$$

La méthode de Lagrange, mise de l'avant en 1788 dans son célèbre traité *La mécanique analytique*, a pour but d'établir de manière systématique les équations différentielles déterminant le mouvement du système mécanique étudié, en fonction des coordonnées généralisées, à partir simplement de l'expression de l'énergie cinétique et de l'énergie potentielle. En un sens, elle schématise au maximum l'étude des problèmes mécaniques en offrant le chemin le plus court et le plus sûr vers les équations du mouvement. Lagrange se vantait que son traité ne contenait aucune illustration ou schéma et que la méthode qu'il proposait était purement *analytique*, par opposition à *géométrique* (voir figure 1.2).

B Applications élémentaires

Dans cette section nous donnerons quelques exemples d'applications de la méthode de Lagrange pour obtenir les équations du mouvement. Il est bien sûr beaucoup plus simple d'obtenir ces équations différentielles que de les résoudre !

Exemple 1.3 Pendule simple

Considérons le pendule simple de l'exemple 1.1 ci-dessus. Les coordonnées (x, y, z) du pendule s'expriment ainsi en fonction de la coordonnée généralisée φ :

$$x = l \sin \varphi \quad y = l \cos \varphi \quad z = 0 \quad (1.37)$$

Notez que y est défini positif vers le bas. Donc les composantes de la vitesse dépendent de φ et $\dot{\varphi}$ de la manière suivante :

$$\dot{x} = l \dot{\varphi} \cos \varphi \quad \dot{y} = -l \dot{\varphi} \sin \varphi \quad \dot{z} = 0 \quad (1.38)$$

et l'énergie cinétique prend la forme

$$T = \frac{1}{2} m \mathbf{v}^2 = \frac{1}{2} m l^2 \dot{\varphi}^2 (\cos^2 \varphi + \sin^2 \varphi) = \frac{1}{2} m l^2 \dot{\varphi}^2 \quad (1.39)$$

L'énergie potentielle gravitationnelle est, quant à elle,

$$V = -mgy = -mgl \cos \varphi \quad (1.40)$$

Le lagrangien est donc

$$L = \frac{1}{2} m l^2 \dot{\varphi}^2 + mgl \cos \varphi \quad (1.41)$$

MÉCHANIQUE ANALITIQUE;

Par M. DE LA GRANGE, de l'Académie des Sciences de Paris,
de celles de Berlin, de Pétersbourg, de Turin, &c.



A PARIS,

Chez LA VEUVE DESAINT, Libraire,
rue du Foin S. Jacques.

M. DCC. LXXXVIII.

AVEC APPROBATION ET PRIVILEGE DU ROI.

AVERTISSEMENT.

On a déjà plusieurs Traités de Méchanique, mais le plan de celui-ci est entièrement neuf. Je me suis proposé de réduire la théorie de cette Science, & l'art de résoudre les problèmes qui s'y rapportent, à des formules générales, dont le simple développement donne toutes les équations nécessaires pour la solution de chaque problème. J'espère que la manière dont j'ai tâché de remplir cet objet, ne laissera rien à désirer.

Cet Ouvrage aura d'ailleurs une autre utilité; il réunira & présentera sous un même point de vue, les différens Principes trouvés jusqu'ici pour faciliter la solution des questions de Méchanique, en montrera la liaison & la dépendance mutuelle, & mettra à portée de juger de leur justesse & de leur étendue.

Je le divise en deux Parties; la Statique ou la Théorie de l'Équilibre, & la Dynamique ou la Théorie

EXTRAIT DES REGISTRES DE L'ACADÉMIE ROYALE DES SCIENCES.

Du vingt-sept Février mil sept cent quatre-vingt-huit.

MESSEURS DE LA PLACE, COUSIN, LE GENDRE & moi, ayant rendu compte d'un Ouvrage intitulé: *Méchanique analytique*, par M. DE LA GRANGE, l'Académie a jugé cet Ouvrage digne de son Approbation, & d'être imprimé sous son Privilège.

Je certifie cet Extrait conforme aux registres de l'Académie. A Paris, ce 27 Février 1788.

LE MARQUIS DE CONDORCET.

PRIVILEGE DU ROI.

LOUIS, PAR LA GRACE DE DIEU, ROI DE FRANCE ET DE NAVARRE; A nos amés & féaux Conseillers, les Gens tenant nos Cours de Parlement, Maîtres des Requêtes ordinaires de notre Hôtel, Grand-Conseil, Prévôt de Paris, Baillifs, Sénéchaux, leurs Lieutenans Civils, & autres nos Jufficiers qu'il appartiendra, SALUT. Nos bien-amés LES MEMBRES DE L'ACADÉMIE ROYALE DES SCIENCES de notre bonne Ville de Paris, Nous ont fait exposer qu'ils auroient besoin de nos Lettres de Privilège pour l'impression de leurs Ouvrages: A ces causes, voulant favorablement traiter les Exposans, Nous leur avons permis & permettons par ces Présentes, de faire imprimer, par tel Imprimeur qu'ils voudront choisir, toutes les Recherches ou Observations journalières, ou Relations annuelles de tout ce qui aura été fait dans les Affemblées de ladite Académie Royale des Sciences, les Ouvrages, Mémoires ou Traités de chacun des Particuliers qui la composent, & généralement tout ce que ladite Académie voudra faire paroître, après avoir fait examiner lesdits Ouvrages, & jugé qu'ils seront dignes de l'impression, en tels volumes, forme, marge, caractères, conjointement, ou séparément, & autant de fois que bon leur semblera, & de les faire vendre & débiter par tout notre Royaume, pendant le tems de vingt années consécutives, à compter du jour de la date des Présentes; sans toutefois qu'à l'occasion des Ouvrages ci-dessus spécifiés, il en puisse être imprimé d'autres qui ne soient pas de ladite Académie: Faisons défenses à toutes sortes de personnes, de quelque qualité & condition qu'elles soient, d'en introduire d'im-

vj AVERTISSEMENT.

du Mouvement; & chacune de ces Parties traitera séparément des Corps solides & des fluides.

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



FIGURE 1.2

Frontispice et préface de la *Mécanique analytique* de Lagrange (1788).

et l'équation de Lagrange prend alors la forme suivante :

$$\frac{\partial L}{\partial \varphi} = -mgl \sin \varphi \quad , \quad \frac{\partial L}{\partial \dot{\varphi}} = ml^2 \dot{\varphi} \implies \frac{d}{dt} \frac{\partial L}{\partial \dot{\varphi}} - \frac{\partial L}{\partial \varphi} = ml^2 \ddot{\varphi} + mgl \sin \varphi = 0 \quad (1.42)$$

ou encore

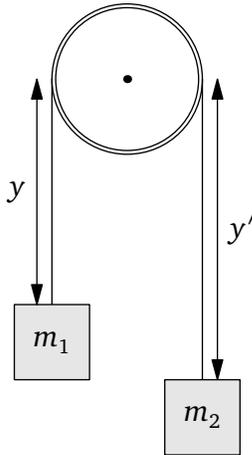
$$\ddot{\varphi} + \omega^2 \sin \varphi = 0 \quad \text{où} \quad \omega \equiv \sqrt{\frac{g}{l}} \quad (1.43)$$

C'est l'équation bien connue qui gouverne l'oscillation d'un pendule. Sa solution ne peut pas s'exprimer par fonctions élémentaires, sauf dans l'approximation des petites oscillations ($\varphi \ll 1$), ou l'équation devient

$$\ddot{\varphi} + \omega^2 \varphi = 0 \implies \varphi(t) = A \cos(\omega t + \xi) \quad (1.44)$$

les constantes A et ξ étant fixées par les valeurs initiales $\varphi(0)$ et $\dot{\varphi}(0)$.

Exemple 1.4 Machine d'Atwood



Considérons deux masses m_1 et m_2 reliées par une corde inextensible de masse négligeable, qui passe autour d'une poulie (voir figure). Les deux coordonnées verticales y et y' , positives vers le bas, ne sont pas indépendantes, mais reliées par la contrainte $y + y' = l$, où l est une constante. Les vitesses sont donc contraintes par la relation

$$\dot{y} + \dot{y}' = 0 \quad (1.45)$$

L'énergie potentielle peut s'exprimer uniquement en fonction de y :

$$V = -m_1 g y - m_2 g y' = -(m_1 - m_2) g y - m_2 g l \quad (1.46)$$

et l'énergie cinétique s'exprime comme suit :

$$T = \frac{1}{2} m_1 \dot{y}^2 + \frac{1}{2} m_2 \dot{y}'^2 = \frac{1}{2} (m_1 + m_2) \dot{y}^2 \quad (1.47)$$

Le lagrangien est alors

$$L = T - V = \frac{1}{2} (m_1 + m_2) \dot{y}^2 + (m_1 - m_2) g y + m_2 g l \quad (1.48)$$

et l'équation de Lagrange est

$$\frac{\partial L}{\partial y} = (m_1 - m_2) g \quad , \quad \frac{\partial L}{\partial \dot{y}} = (m_1 + m_2) \dot{y} \implies \frac{d}{dt} \frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = (m_1 + m_2) \ddot{y} - (m_1 - m_2) g = 0 \quad (1.49)$$

Donc l'accélération (constante) de la masse m_1 vers le bas est

$$\ddot{y} = \frac{m_1 - m_2}{m_1 + m_2} g \quad (1.50)$$

Exemple 1.5 Pendule forcé

Revenons au pendule simple, et supposons maintenant que le point d'attache P est en mouvement et que sa hauteur y_p soit une fonction $y_p(t)$ connue à l'avance. Trouvons les équations de Lagrange dans ce cas. Les coordonnées de la masse sont

$$x = l \sin \varphi \quad y = l \cos \varphi + y_p(t) \quad (1.51)$$

et les vitesses sont

$$\dot{x} = l\dot{\varphi} \cos \varphi \quad \dot{y} = -l\dot{\varphi} \sin \varphi + \dot{y}_p(t) \quad (1.52)$$

et l'énergie cinétique prend la forme

$$\begin{aligned} T &= \frac{1}{2} m \mathbf{v}^2 = \frac{1}{2} m l^2 \dot{\varphi}^2 (\cos^2 \varphi + \sin^2 \varphi) + \frac{1}{2} m \dot{y}_p^2 - m l \dot{y}_p \dot{\varphi} \sin \varphi \\ &= \frac{1}{2} m l^2 \dot{\varphi}^2 + \frac{1}{2} m \dot{y}_p^2 - m l \dot{y}_p \dot{\varphi} \sin \varphi \end{aligned} \quad (1.53)$$

L'énergie potentielle gravitationnelle est, quant à elle,

$$V = -mgy = -mgl \cos \varphi - mgy_p \quad (1.54)$$

Le lagrangien est donc

$$L = \frac{1}{2} m l^2 \dot{\varphi}^2 + \frac{1}{2} m \dot{y}_p^2 - m l \dot{y}_p \dot{\varphi} \sin \varphi + mgl \cos \varphi + mgy_p \quad (1.55)$$

Les termes de l'équation de Lagrange sont

$$\begin{aligned} \frac{\partial L}{\partial \varphi} &= -mgl \sin \varphi - m l \dot{y}_p \dot{\varphi} \cos \varphi \\ \frac{\partial L}{\partial \dot{\varphi}} &= m l^2 \dot{\varphi} - m l \dot{y}_p \sin \varphi \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{\varphi}} &= m l^2 \ddot{\varphi} - m l \ddot{y}_p \sin \varphi - m l \dot{y}_p \dot{\varphi} \cos \varphi \end{aligned} \quad (1.56)$$

Finalement, l'équation de Lagrange est

$$m l^2 \ddot{\varphi} + mgl \sin \varphi - m l \ddot{y}_p \sin \varphi = 0 \quad (1.57)$$

La seule différence d'avec l'équation (1.44) du pendule simple est le dernier terme, qu'on peut facilement interpréter : \ddot{y}_p étant l'accélération (vers le bas) du point d'attache et $\ddot{y}_p \sin \varphi$ la projection de cette accélération dans la direction du mouvement du pendule (la direction azimutale), ce terme n'est que la projection sur cette direction de la force transmise à la masse du pendule par le point d'attache en accélération. En outre, ce terme supplémentaire peut être obtenu de l'équation du pendule à pivot fixe en procédant à la substitution $g \rightarrow g - \ddot{y}_p$, c'est-à-dire en remplaçant le champ gravitationnel constant par un champ variable dans le temps, ce qui est entièrement compréhensible si on se place dans le référentiel accéléré du pivot et qu'on applique le principe d'équivalence.

Exemple 1.6 Pendule double

Le lagrangien du pendule double (exemple 1.2 ci-haut) dépend des deux angles φ_1 et φ_2 et de leurs dérivées. Les coordonnées des deux masses sont :

$$\begin{aligned} x_1 &= l_1 \sin \varphi_1 & y_1 &= l_1 \cos \varphi_1 \\ x_2 &= l_2 \sin \varphi_2 + x_1 & y_2 &= l_2 \cos \varphi_2 + y_1 \end{aligned} \quad (1.58)$$

et leurs vitesses sont

$$\begin{aligned} \dot{x}_1 &= l_1 \dot{\varphi}_1 \cos \varphi_1 & \dot{y}_1 &= -l_1 \dot{\varphi}_1 \sin \varphi_1 \\ \dot{x}_2 &= l_2 \dot{\varphi}_2 \cos \varphi_2 + \dot{x}_1 & \dot{y}_2 &= -l_2 \dot{\varphi}_2 \sin \varphi_2 + \dot{y}_1 \end{aligned} \quad (1.59)$$

De sorte que l'énergie cinétique prend la forme

$$T = \frac{1}{2} m_1 \mathbf{v}_1^2 + \frac{1}{2} m_2 \mathbf{v}_2^2 = \frac{1}{2} (m_1 + m_2) l_1^2 \dot{\varphi}_1^2 + \frac{1}{2} m_2 l_2^2 \dot{\varphi}_2^2 + m_2 l_1 l_2 \dot{\varphi}_1 \dot{\varphi}_2 \cos(\varphi_1 - \varphi_2) \quad (1.60)$$

et l'énergie potentielle la forme

$$V = -g(m_1 y_1 + m_2 y_2) = -(m_1 + m_2) g l_1 \cos \varphi_1 - m_2 g l_2 \cos \varphi_2 \quad (1.61)$$

Donc

$$L = \frac{1}{2} (m_1 + m_2) l_1^2 \dot{\varphi}_1^2 + \frac{1}{2} m_2 l_2^2 \dot{\varphi}_2^2 + m_2 l_1 l_2 \dot{\varphi}_1 \dot{\varphi}_2 \cos(\varphi_1 - \varphi_2) + (m_1 + m_2) g l_1 \cos \varphi_1 + m_2 g l_2 \cos \varphi_2$$

Les termes de l'équation de Lagrange sont

$$\frac{\partial L}{\partial \varphi_1} = -(m_1 + m_2) g l_1 \sin \varphi_1 - m_2 l_1 l_2 \dot{\varphi}_1 \dot{\varphi}_2 \sin(\varphi_1 - \varphi_2) \quad (1.62)$$

$$\frac{\partial L}{\partial \varphi_2} = -m_2 g l_2 \sin \varphi_2 + m_2 l_1 l_2 \dot{\varphi}_1 \dot{\varphi}_2 \sin(\varphi_1 - \varphi_2) \quad (1.63)$$

$$\frac{\partial L}{\partial \dot{\varphi}_1} = (m_1 + m_2) l_1^2 \dot{\varphi}_1 + m_2 l_1 l_2 \dot{\varphi}_2 \cos(\varphi_1 - \varphi_2) \quad (1.64)$$

$$\frac{\partial L}{\partial \dot{\varphi}_2} = m_2 l_2^2 \dot{\varphi}_2 + m_2 l_1 l_2 \dot{\varphi}_1 \cos(\varphi_1 - \varphi_2) \quad (1.65)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\varphi}_1} = (m_1 + m_2) l_1^2 \ddot{\varphi}_1 + m_2 l_1 l_2 \ddot{\varphi}_2 \cos(\varphi_1 - \varphi_2) - m_2 l_1 l_2 \dot{\varphi}_2 \sin(\varphi_1 - \varphi_2) (\dot{\varphi}_1 - \dot{\varphi}_2) \quad (1.66)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\varphi}_2} = m_2 l_2^2 \ddot{\varphi}_2 + m_2 l_1 l_2 \ddot{\varphi}_1 \cos(\varphi_1 - \varphi_2) - m_2 l_1 l_2 \dot{\varphi}_1 \sin(\varphi_1 - \varphi_2) (\dot{\varphi}_1 - \dot{\varphi}_2) \quad (1.67)$$

Les équations de Lagrange pour les deux coordonnées généralisées sont donc

$$\begin{aligned} (m_1 + m_2) l_1^2 \ddot{\varphi}_1 + m_2 l_1 l_2 \ddot{\varphi}_2 \cos(\varphi_1 - \varphi_2) + m_2 l_1 l_2 \dot{\varphi}_2^2 \sin(\varphi_1 - \varphi_2) + (m_1 + m_2) g l_1 \sin \varphi_1 &= 0 \\ m_2 l_2^2 \ddot{\varphi}_2 + m_2 l_1 l_2 \ddot{\varphi}_1 \cos(\varphi_1 - \varphi_2) - m_2 l_1 l_2 \dot{\varphi}_1^2 \sin(\varphi_1 - \varphi_2) + m_2 g l_2 \sin \varphi_2 &= 0 \end{aligned} \quad (1.68)$$

Il s'agit d'un système de deux équations différentielles du deuxième ordre couplées, dont la solution est bien évidemment impossible. Une simplification importante intervient si on suppose que les angles φ_1 et φ_2 sont petits. Dans ce cas, on peut se limiter aux termes du premier ordre en

$\varphi_{1,2}$ et en leurs dérivées. On peut également remplacer $\sin \varphi_{1,2}$ par $\varphi_{1,2}$ et $\cos \varphi_{1,2}$ par 1. Dans cette approximation, on obtient donc l'équivalent de deux oscillateurs harmoniques couplés et les équations du mouvement deviennent alors

$$\begin{aligned} (m_1 + m_2)l_1^2\ddot{\varphi}_1 + m_2l_1l_2\ddot{\varphi}_2 + (m_1 + m_2)gl_1\varphi_1 &= 0 \\ m_2l_2^2\ddot{\varphi}_2 + m_2l_1l_2\ddot{\varphi}_1 + m_2gl_2\varphi_2 &= 0 \end{aligned} \quad (1.69)$$

On peut récrire ce système sous forme matricielle :

$$\begin{pmatrix} (m_1 + m_2)l_1^2 & m_2l_1l_2 \\ m_2l_1l_2 & m_2l_2^2 \end{pmatrix} \begin{pmatrix} \ddot{\varphi}_1 \\ \ddot{\varphi}_2 \end{pmatrix} = - \begin{pmatrix} (m_1 + m_2)gl_1 & 0 \\ 0 & m_2gl_2 \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} \quad (1.70)$$

On peut résoudre ce système de la manière habituelle : on part de l'hypothèse que le mouvement est harmonique de fréquence ω , de sorte que $\ddot{\varphi}_1$ puisse être remplacé par $-\omega^2\varphi_1$ et de même pour φ_2 . En substituant, on trouve

$$\begin{pmatrix} \omega^2(m_1 + m_2)l_1^2 - (m_1 + m_2)gl_1 & \omega^2m_2l_1l_2 \\ \omega^2m_2l_1l_2 & \omega^2m_2l_2^2 - m_2gl_2 \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = 0 \quad (1.71)$$

Pour que cette équation matricielle ait une solution non nulle, il faut absolument que la matrice soit singulière, c'est-à-dire que son déterminant soit nul. Nous ne donnerons ici que les solutions particulières associées aux valeurs $m_1 = m_2 = m$ et $l_1 = l_2 = l$:

$$\omega = \sqrt{(2 \pm \sqrt{2}) \frac{g}{l}} \quad (1.72)$$

On peut montrer que le signe + donne une fréquence plus élevée correspondant à un mouvement alterné des deux pendules (mode antisymétrique), alors que le signe – donne une fréquence plus basse correspondant à un mouvement solidaire des deux pendules (mode symétrique).

Exemple 1.7

Coordonnées sphériques Écrivons le lagrangien d'une particule en coordonnées sphériques. Les coordonnées cartésiennes sont reliées aux coordonnées sphériques par les relations suivantes :

$$\begin{aligned} x &= r \sin \theta \cos \varphi \\ y &= r \sin \theta \sin \varphi \\ z &= r \cos \theta \end{aligned} \quad (1.73)$$

Donc les composantes de la vitesse sont

$$\begin{aligned} \dot{x} &= \dot{r} \sin \theta \cos \varphi + r \dot{\theta} \cos \theta \cos \varphi - r \dot{\varphi} \sin \theta \sin \varphi \\ \dot{y} &= \dot{r} \sin \theta \sin \varphi + r \dot{\theta} \cos \theta \sin \varphi + r \dot{\varphi} \sin \theta \cos \varphi \\ \dot{z} &= \dot{r} \cos \theta - r \dot{\theta} \sin \theta \end{aligned} \quad (1.74)$$

Le lagrangien prend donc la forme suivante (après quelques simplifications élémentaires) :

$$L = \frac{1}{2}m \{ \dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2 \theta \dot{\varphi}^2 \} - V(\mathbf{r}) \quad (1.75)$$

Notez que si le potentiel V ne dépend que de la coordonnée radiale r , alors le lagrangien ne dépend pas de φ et donc la quantité suivante est conservée dans le temps :

$$\frac{\partial L}{\partial \dot{\varphi}} = mr^2 \sin^2 \theta \dot{\varphi} \quad (1.76)$$

On montrera en page 45 comment cette quantité est reliée au moment cinétique.

C Principe de la moindre action

Nous allons démontrer dans cette section que les équations de Lagrange peuvent aussi être considérées comme la conséquence d'un principe fondamental, appelé le plus souvent *principe de la moindre action*. Ce principe est logiquement équivalent aux lois de Newton.

1.C.1 Énoncé et démonstration

Considérons un système mécanique décrit par des coordonnées généralisées q_α . Un *mouvement* défini entre les temps $t = t_i$ et $t = t_f$ correspond alors à un ensemble de n fonctions $q_\alpha(t)$. Nous utiliserons souvent la notation \mathbf{q} pour désigner l'ensemble des coordonnées généralisées q_α . Cette notation est l'analogue dans l'espace des configurations de la notation vectorielle (nous réserverons les caractères gras habituels, comme \mathbf{q} , aux vecteurs dans l'espace à trois dimensions). Considérons l'ensemble des mouvements ayant des valeurs données de la configuration initiale $\mathbf{q}(t_i)$ et de la configuration finale $\mathbf{q}(t_f)$. Le principe de la moindre action stipule que, de tous ces mouvements, celui qui sera effectivement suivi par le système et qui donc respecte les équations de Lagrange, est celui qui minimise l'intégrale suivante, appelée *action* :

$$S = \int_{t_i}^{t_f} dt L(\mathbf{q}, \dot{\mathbf{q}}, t) \quad (1.77)$$

L'action est une *fonctionnelle*, c'est-à-dire une fonction dont le domaine est lui-même un ensemble de fonctions. Pour distinguer les fonctionnelles des fonctions courantes, on utilise des crochets ($[\dots]$) pour encadrer leur argument. Par exemple, l'action associée à un mouvement $\mathbf{q}(t)$ sera notée $S[\mathbf{q}(t)]$.

Procédons maintenant à la démonstration des équations de Lagrange à partir du principe de la moindre action. À cette fin, supposons que le mouvement $\mathbf{q}(t)$ minimise l'action S . Cela signifie qu'un mouvement qui diffère de $\mathbf{q}(t)$ par une fonction infinitésimale du temps $\delta\mathbf{q}(t)$ aurait une valeur égale de l'action, au premier ordre en $\delta\mathbf{q}$:

$$S[\mathbf{q} + \delta\mathbf{q}] = S[\mathbf{q}] + \mathcal{O}(\delta\mathbf{q}^2) \quad (1.78)$$

Il s'agit simplement d'une généralisation du critère qui définit le minimum d'une fonction ordinaire : si la fonction $f(x)$ admet un minimum à $x = x_0$, alors, pour une variation infinitésimale δx , on trouve

$$f(x_0 + \delta x) = f(x_0) + \mathcal{O}(\delta x^2) \quad (1.79)$$

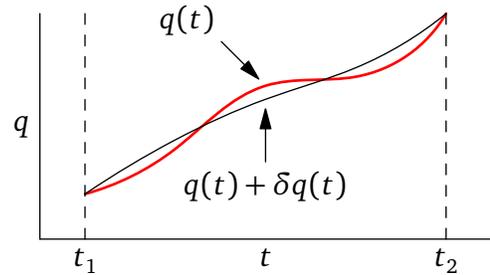
En effet, en supposant que la fonction est différentiable et en appliquant un développement de Taylor limité, on trouve

$$f(x_0 + \delta x) = f(x_0) + f'(x_0)\delta x + \frac{1}{2}f''(x_0)\delta x^2 + \dots \quad (1.80)$$

La valeur x_0 correspond effectivement à un minimum si le terme linéaire en δx s'annule, ce qui implique l'annulation de la première dérivée : $f'(x_0) = 0$.

FIGURE 1.3

Mouvement réel $q(t)$ et mouvement varié $q(t) + \delta q(t)$.



Le développement limité peut aussi être appliqué à l'action, car cette dernière s'exprime par une fonction connue des coordonnées q et de leurs dérivées \dot{q} :

$$\begin{aligned} S[q + \delta q] &= \int dt L(q + \delta q, \dot{q} + \delta \dot{q}, t) \\ &= \int dt \left\{ L(q, \dot{q}, t) + \frac{\partial L}{\partial q} \cdot \delta q + \frac{\partial L}{\partial \dot{q}} \cdot \delta \dot{q} \right\} \\ &= S[q] + \int dt \frac{\partial L}{\partial q} \cdot \delta q + \int dt \frac{\partial L}{\partial \dot{q}} \cdot \frac{d}{dt} \delta q \\ &= S[q] + \int dt \frac{\partial L}{\partial q} \cdot \delta q + \int dt \left\{ \frac{d}{dt} \left[\frac{\partial L}{\partial \dot{q}} \cdot \delta q \right] - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \cdot \delta q \right\} \\ &= S[q] + \int dt \left\{ \frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right\} \cdot \delta q + \left[\frac{\partial L}{\partial \dot{q}} \cdot \delta q \right]_{t_i}^{t_f} \end{aligned} \quad (1.81)$$

Nous avons utilisé le fait que

$$\delta \dot{q} = \frac{d}{dt} \delta q \quad (1.82)$$

et nous avons intégré par parties à la quatrième ligne. La notation du produit scalaire est utilisée pour désigner une somme sur l'indice α ; ainsi,

$$\frac{\partial L}{\partial q} \cdot \delta q = \sum_{\alpha} \frac{\partial L}{\partial q_{\alpha}} \delta q_{\alpha} \quad (1.83)$$

Comme le principe de la moindre action stipule que les mouvements considérés ont tous la même valeur de q à $t = t_i$ et $t = t_f$, la variation δq doit s'annuler à ces points : $\delta q(t_i) = \delta q(t_f) = 0$ (voir

fig. 1.3). Le dernier terme de la dernière équation est donc nul et le principe de la moindre action stipule donc que

$$\int dt \left\{ \frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right\} \cdot \delta q(t) = 0 \quad (1.84)$$

où le lagrangien L et ses dérivées sont évalués au mouvement $q(t)$. Cette expression doit s'annuler, quelle que soit la variation $\delta q(t)$. En particulier, on pourrait choisir une variation $\delta q(t)$ qui est nulle partout sauf dans un intervalle temporel Δt aussi petit que l'on veut autour d'un temps t quelconque, tout en étant infiniment différentiable. La seule façon d'annuler l'intégrale ci-haut dans ces conditions est que l'expression entre accolades soit identiquement nulle :

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = 0 \quad (1.85)$$

Il s'agit bien sûr des équations de Lagrange. Nous avons donc démontré que le principe de la moindre action est équivalent aux équations de Lagrange.¹

Remarques:

- ◆ En fait, nous n'avons pas démontré que l'action est véritablement un minimum à la trajectoire physique $q(t)$; uniquement qu'elle est un *point stationnaire*. Mais c'est tout ce qui compte au fond, et le principe devrait plus correctement s'appeler *principe de l'action stationnaire*.
- ◆ La démonstration suppose que tous les mouvements considérés, incluant les variations arbitraires δq , sont des fonctions différentiables de t . Nous avons en plus affirmé qu'on pouvait choisir la variation comme étant nulle partout sauf dans un intervalle Δt aussi petit que l'on veut autour d'un temps t quelconque. Ce point mérite d'être étayé. Considérons à cet effet la fonction

$$f(x) = \begin{cases} \exp((1/a^2) - (1/a^2 - x^2)) & -a < x < a \\ 0 & \text{autrement} \end{cases} \quad (1.86)$$

Cette fonction est nulle en dehors de l'intervalle $[-a, a]$ et est infiniment différentiable et positive partout dans cet intervalle, avec un maximum de 1 à $x = 0$. En translatant cette fonction comme on veut et prenant a aussi petit qu'on veut, et en l'appliquant à la variable q_α qu'on désire, on obtient une variation qui obéit aux critères que nous avons posés ci-haut.

- ◆ La démonstration repose sur l'hypothèse que les coordonnées généralisées sont vraiment indépendantes les unes des autres, c'est-à-dire qu'elles ne sont pas reliées par des contraintes supplémentaires. Sinon, les variations δq_α ne sont pas indépendantes et on ne peut conclure que l'intégrand de (1.84) doit s'annuler pour chaque valeur de α séparée.
- ◆ Nous allons dorénavant considérer que le principe de la moindre action est fondamental, et qu'il constitue un point de départ aussi valable que la deuxième loi de Newton à l'étude de la mécanique.

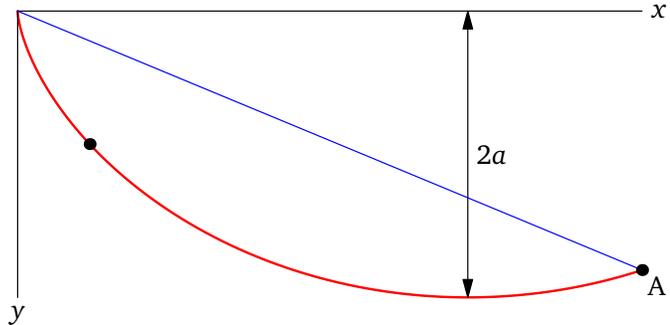
1.C.2 La brachistochrone

Le principe de la moindre action est une application particulière d'une méthode plus générale appelée *calcul des variations*, qui vise à trouver les points stationnaires d'une fonctionnelle. Historiquement,

1. Il est évident que si les équations de Lagrange sont satisfaites, alors le principe de moindre action s'applique.

ce type de problème tire son origine d'un défi lancé à la communauté des mathématiciens en 1696 par Jean Bernoulli, et auquel répondirent les plus grands mathématiciens de cette époque (Jean et Jacques Bernoulli, Newton, L'Hôpital, Leibniz et Tschirnhaus). Ce problème est celui de la *brachistochrone*, c'est-à-dire de la courbe de moindre temps.

FIGURE 1.4
La courbe brachistochrone



Expliquons : supposons qu'un point matériel, relâché de l'origine, soit contraint de se mouvoir sur une courbe descendante fixe, tout en étant soumis à un champ gravitationnel constant. Le problème est de déterminer la courbe qui minimise le temps de parcours entre l'origine et un point A donné (en rouge sur la fig. 1.4).

Cette courbe peut être décrite à l'aide de deux fonctions $x(\phi)$ et $y(\phi)$ d'un paramètre ϕ . Le temps pris pour parcourir un arc infinitésimal ds de cette courbe est

$$dt = \frac{ds}{v} \quad (1.87)$$

où v est la vitesse du point matériel à cet endroit de la courbe. En fonction des dérivées paramétriques

$$\dot{x} = \frac{dx}{d\phi} \quad \text{et} \quad \dot{y} = \frac{dy}{d\phi} \quad (1.88)$$

la longueur d'arc est

$$ds = \sqrt{\dot{x}^2 + \dot{y}^2} d\phi \quad (1.89)$$

et le temps total de descente est l'intégrale

$$T = \int \frac{ds}{v} = \int \frac{\sqrt{\dot{x}^2 + \dot{y}^2}}{v} d\phi \quad (1.90)$$

La vitesse, par conservation de l'énergie, s'obtient de la hauteur du point par rapport au point de départ :

$$\frac{1}{2}v^2 = gy \implies v = \sqrt{2gy} \quad (1.91)$$

où y est positif vers le bas. L'intégrale à minimiser est donc

$$T = \int \frac{ds}{v} = \int \frac{\sqrt{\dot{x}^2 + \dot{y}^2}}{\sqrt{2gy}} d\phi = \int F(y, \dot{x}, \dot{y}) d\phi \quad (1.92)$$

La variation de cette intégrale mène aux équations de Lagrange, cette fois exprimées en fonction de F :

$$\frac{d}{d\phi} \frac{\partial F}{\partial \dot{x}} - \frac{\partial F}{\partial x} = 0 \quad \frac{d}{d\phi} \frac{\partial F}{\partial \dot{y}} - \frac{\partial F}{\partial y} = 0 \quad (1.93)$$

Comme F ne dépend pas explicitement de x , l'expression

$$p_x = \frac{\partial F}{\partial \dot{x}} = \frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}} \frac{1}{\sqrt{2gy}} \quad (1.94)$$

est indépendante de ϕ . La deuxième des équations (1.93) peut donc s'écrire

$$\frac{d}{d\phi} \left(\frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}} \frac{1}{\sqrt{2gy}} \right) + \frac{\sqrt{\dot{x}^2 + \dot{y}^2}}{2\sqrt{2gy^3}} = 0 \quad (1.95)$$

ou encore, en utilisant la conservation de p_x ,

$$\frac{d}{d\phi} \left(\frac{\dot{y}}{\dot{x}} \right) + \frac{\dot{x}}{4p_x^2 g y^2} = 0 \quad (1.96)$$

Avant d'aller plus loin, nous devons considérer les conditions aux limites, ainsi que l'allure générale de la courbe. Considérons encore la fig. 1.4. Si la première réaction est de penser que la courbe recherchée est une droite allant de l'origine au point A, on se détrompe rapidement en réalisant que ce n'est pas la distance qu'il faut minimiser, mais le temps de parcours, et que pour ce faire on a avantage à infléchir la courbe vers le bas au début de la descente de manière à faire gagner rapidement de la vitesse à la particule, afin qu'elle puisse parcourir le reste de la courbe plus vite, même si la longueur est aussi plus grande.

Nous allons maintenant introduire une constante a telle que

$$p_x = \frac{1}{\sqrt{4ga}} \quad (1.97)$$

Il ressort de l'expression de p_x que $2a$ est la valeur de y quand $\dot{y} = 0$, ce qui correspond à la hauteur de la courbe à son minimum (voir fig. 1.4), si elle a un minimum.² L'équation différentielle devient donc, dans cette notation,

$$\frac{d}{d\phi} \left(\frac{\dot{y}}{\dot{x}} \right) + \frac{a\dot{x}}{y^2} = 0 \quad (1.98)$$

Une autre propriété de cette équation différentielle est qu'elle est indépendante de la manière dont on paramètre la courbe : si on procède à un changement de variable $\phi \rightarrow \phi'$ et que la courbe est maintenant décrite par une fonction $y(\phi')$, l'équation (1.98) demeure inchangée, mais cette fois exprimée en fonction de la variable ϕ' . On appelle cela *l'invariance par reparamétrisation*. Ceci nous donne une certaine liberté qui nous permet en retour d'imposer certaines conditions (qu'on pourrait appeler *conditions de jauge* par analogie avec l'électromagnétisme).

En particulier, nous imposerons la condition suivante :

$$\dot{x} = y \quad (1.99)$$

Il est toujours possible d'imposer cette condition en choisissant une paramétrisation particulière. De même, on pourrait décider, par exemple, d'utiliser x comme paramètre décrivant la courbe, ce qui

2. Si la courbe ne comporte pas de minimum (du moins avant le point A), cette définition est quand même valable : ce n'est qu'un changement de paramètre. Nous verrons plus loin que la courbe aura un minimum de toute manière, qui peut être situé soit avant ou après le point A.

mènerait à $\dot{x} \equiv 1$. Mais la condition ci-haut est plus utile, comme on le verra. L'équation (1.98) s'écrit alors

$$\frac{d}{d\phi} \left(\frac{\dot{y}}{y} \right) + \frac{a}{y} = 0 \quad \text{ou encore} \quad \frac{d^2}{d\phi^2} \ln y + \frac{a}{y} = 0 \quad (1.100)$$

Posons alors $u = \ln y$. L'équation devient

$$\ddot{u} + a e^{-u} = 0 \quad (1.101)$$

On multiplie ensuite par le facteur intégrant \dot{u} :

$$\dot{u}\ddot{u} + a\dot{u}e^{-u} = 0 \implies \frac{1}{2} \frac{d}{d\phi} (\dot{u}^2) - a \frac{d}{d\phi} e^{-u} = 0 \quad (1.102)$$

dont la solution est

$$\dot{u}^2 = 2a e^{-u} + C \quad \text{ou} \quad \left(\frac{\dot{y}}{y} \right)^2 = \frac{2a}{y} + C \quad (1.103)$$

On peut déterminer la constante C en posant que $\dot{y} = 0$ en même temps que $y = 2a$, ce qui est nécessairement le cas d'après la définition de la constante a . Ceci fixe $C = -1$. Il reste donc

$$\dot{u} = \sqrt{2a e^{-u} - 1} \quad \text{ou} \quad \frac{du}{\sqrt{2a e^{-u} - 1}} = d\phi \quad (1.104)$$

En intégrant et en choisissant une constante d'intégration nulle (ce qui revient à choisir l'origine des ϕ), on trouve

$$\begin{aligned} \phi &= \int \frac{du}{\sqrt{2a e^{-u} - 1}} \\ &= \int \frac{dy}{\sqrt{2ay - y^2}} \\ &= \int \frac{dy}{\sqrt{a^2 - (y-a)^2}} \\ &= \arccos \frac{|y-a|}{a} \end{aligned} \quad (1.105)$$

Donc

$$y = a + a \cos \phi \quad \text{ou} \quad y = a - a \cos \phi \quad (1.106)$$

Le choix entre ces deux possibilités est arbitraire. En choisissant la solution de droite, l'origine de la courbe correspond à $\phi = 0$. Comme $\dot{x} = y = a - a \cos \phi$, il s'ensuit que $x = a\phi - a \sin \phi + \text{cte}$, la constante d'intégration étant nulle si la courbe débute à l'origine. La solution est donc

$$\begin{aligned} x &= a\phi - a \sin \phi \\ y &= a - a \cos \phi \end{aligned}$$

(1.107)

Il s'agit de l'équation paramétrique bien connue de la cycloïde de rayon a , qui est précisément illustrée à la fig. 1.4. Selon la position du point A, c'est-à-dire le rapport entre x_A et y_A , la bille parviendra au minimum de la courbe avant le point A (si $x_A > \pi y_A/2$) ou après le point A (si $x_A < \pi y_A/2$).

1.C.3 Méthode des multiplicateurs de Lagrange

Supposons que les coordonnées généralisées q_a soient sujettes à M contraintes supplémentaires. Ces contraintes peuvent être holonomes, dans lequel cas il serait possible de réduire le nombre de coordonnées généralisées, ou elles peuvent être *semi-holonomes*, c'est-à-dire qu'elles peuvent être exprimées par une relation entre les coordonnées généralisées et les vitesses généralisées :

$$f_a(q, \dot{q}) = 0 \quad (a = 1, 2, \dots, M) \quad (1.108)$$

Un exemple de contrainte semi-holonome est la contrainte de roulement, qui implique une relation entre la vitesse de translation et la vitesse de rotation d'un objet. Dans un cas simple de roulement, on pourrait *intégrer* la contrainte et la transformer en contrainte holonome entre les coordonnées seulement, sans faire intervenir les vitesses, mais pas dans tous les cas.

Une manière élégante de tenir compte de ces contraintes semi-holonomes est la méthode des *multiplicateurs de Lagrange*. On ajoute au lagrangien une combinaison linéaire des contraintes :

$$L(q, \dot{q}, t) \rightarrow L(q, \dot{q}, t) + \sum_{a=1}^M \lambda_a f_a(q, \dot{q}, t) \quad (1.109)$$

où on considère les M paramètres λ_a comme des paramètres variationnels, au même titre que le mouvement $q(t)$. En appliquant le principe de moindre action

$$\delta \int dt \left(L + \sum_{a=1}^M \lambda_a f_a \right) = 0 \quad (1.110)$$

où à la fois q et λ_a sont sujets à une variation, on obtient les équations suivantes :

$$\begin{cases} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = \sum_{a=1}^M \left\{ \lambda_a \left(\frac{\partial f_a}{\partial q} - \frac{d}{dt} \frac{\partial f_a}{\partial \dot{q}} \right) - \dot{\lambda}_a \frac{\partial f_a}{\partial \dot{q}} \right\} \\ f_a(q, \dot{q}) = 0 \end{cases} \quad (1.111)$$

Cet ensemble d'équations comporte M équations supplémentaires (les équations de contrainte), qui permettent donc de résoudre le mouvement complètement. Noter ici qu'au sens des variations, les coordonnées généralisées sont considérées comme indépendantes, ce qui est rendu possible par l'introduction des paramètres λ_a .

Exemple 1.8 Minimisation d'une fonction simple avec contraintes

Pour comprendre un peu mieux comment fonctionne la méthode des multiplicateurs de Lagrange, considérons un problème de minimisation ordinaire. Par exemple, soit la fonction à deux variables $f(x, y) = x^2 + y^2$, où les variables x et y sont sujettes à la contrainte $x + y = 1$. La stratégie la plus directe pour trouver le minimum de f sujet à cette contrainte est d'éliminer la variable y , ce qui mène à l'expression suivante de f en fonction de x seul : $f(x) = x^2 + (1-x)^2 = 2x^2 - 2x + 1$. En appliquant la condition de dérivée nulle, on trouve

$$4x - 2 = 0 \implies x = \frac{1}{2} \quad \text{et donc} \quad y = 1 - \frac{1}{2} = \frac{1}{2} \quad (1.112)$$

Par contre, la méthode des multiplicateurs de Lagrange, dans ce cas, nous amène à considérer la fonction à trois variables x, y, λ

$$f(x, y) + \lambda(x + y - 1) = x^2 + y^2 + \lambda(x + y - 1) \quad (1.113)$$

La condition de dérivée nulle sur chacune des trois variables nous donne

$$\begin{aligned} 2x + \lambda &= 0 \\ 2y + \lambda &= 0 \\ x + y - 1 &= 0 \end{aligned} \quad (1.114)$$

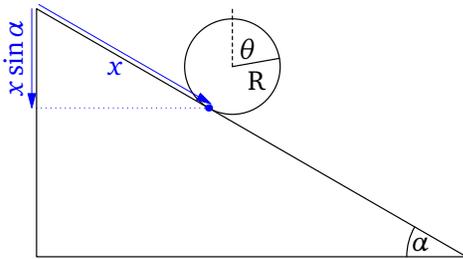
et qui possède la solution suivante :

$$x = \frac{1}{2} \quad y = \frac{1}{2} \quad \lambda = -1 \quad (1.115)$$

La méthode semble superflue ici, mais illustre que la symétrie entre x et y n'est pas brisée par la méthode. Par contre, la méthode des multiplicateurs de Lagrange peut être très utile dans des exemples plus complexes, où l'élimination d'une variable n'est pas possible.

Exemple 1.9

Objet roulant sur un plan incliné



Considérons un objet de rayon R en roulement sur un plan incliné d'un angle α . La distance entre le point de contact du cylindre sur le plan et le haut du plan définit une première coordonnée généralisée x . L'angle de rotation de l'objet par rapport à la verticale est θ . La contrainte de roulement est par nature semi-holonome, et relie la vitesse \dot{x} à la vitesse angulaire $\dot{\theta}$:

$$R\dot{\theta} = \dot{x} \implies f(x, \dot{x}, \theta, \dot{\theta}) = \dot{x} - R\dot{\theta} \quad (1.116)$$

L'énergie cinétique est ici la somme de l'énergie cinétique de translation $\frac{1}{2}m\dot{x}^2$ et de l'énergie cinétique de rotation $\frac{1}{2}I\dot{\theta}^2$, où I est le moment d'inertie de l'objet. L'énergie potentielle est $V = -mgx \sin \alpha$. Donc le lagrangien, plus la fonction de contrainte, devient

$$L + \lambda f = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}I\dot{\theta}^2 + mgx \sin \alpha + \lambda(\dot{x} - R\dot{\theta}) \quad (1.117)$$

Les équations du mouvement (1.111) sont alors

$$\begin{aligned} m\ddot{x} - mg \sin \alpha &= -\dot{\lambda} \\ I\ddot{\theta} &= \dot{\lambda}R \\ \dot{x} &= R\dot{\theta} \end{aligned} \quad (1.118)$$

La dernière équation peut être utilisée pour éliminer θ , notamment en la dérivant par rapport

au temps : $\ddot{x} = R\ddot{\theta}$. On obtient alors

$$\begin{aligned} m\ddot{x} - mg \sin \alpha &= -\dot{\lambda} \\ \frac{I}{R}\ddot{x} &= R\dot{\lambda} \end{aligned} \quad (1.119)$$

et donc, en combinant ces deux équations,

$$m\ddot{x} - mg \sin \alpha + \frac{I}{R^2}\ddot{x} = 0 \implies \left(m + \frac{I}{R^2}\right)\ddot{x} = mg \sin \alpha \quad (1.120)$$

Si k désigne le rayon de giration de l'objet, alors $I = mk^2$ et l'équation ci-haut se réduit à

$$\ddot{x} = \frac{g \sin \alpha}{1 + k^2/R^2} \quad (1.121)$$

On en déduit aussi que

$$\dot{\lambda} = \frac{mgk^2 \sin \alpha}{R^2 + k^2} \quad (1.122)$$

On voit d'après la première des éq. (1.119) que $\dot{\lambda}$ est la force de contrainte appliquée vers le haut du plan. L'un des avantages de la méthode des multiplicateurs de Lagrange est précisément qu'elle donne accès à la valeur des forces de contraintes.

D Application aux petites oscillations

1.D.1 Système masses-ressorts

Considérons un système formé de n masses différentes reliées par des ressorts, comme illustré à la figure 1.5. Les ressorts ont des constantes et des longueurs à l'équilibre différentes. Désignons par \mathbf{r}_i la position de la particule no i par rapport à sa position d'équilibre. Obtenons une expression pour le lagrangien de ce système.

L'énergie cinétique de ce système sera bien évidemment donnée par

$$T = \frac{1}{2} \sum_i m_i \dot{\mathbf{r}}_i^2 \quad (1.123)$$

où m_i est la masse de la particule no i , alors que l'énergie potentielle élastique est

$$V = \frac{1}{2} \sum_{i < j} k_{ij} (|\mathbf{r}_i - \mathbf{r}_j| - \ell_{ij})^2 \quad (1.124)$$

où k_{ij} est la constante du ressort qui relie les masses i et j et ℓ_{ij} est sa longueur à l'équilibre (la somme est effectuée sur toutes les paires distinctes de particules). Autrement dit, l'énergie potentielle élastique résidant dans chaque ressort est proportionnelle au carré de la déviation de sa longueur

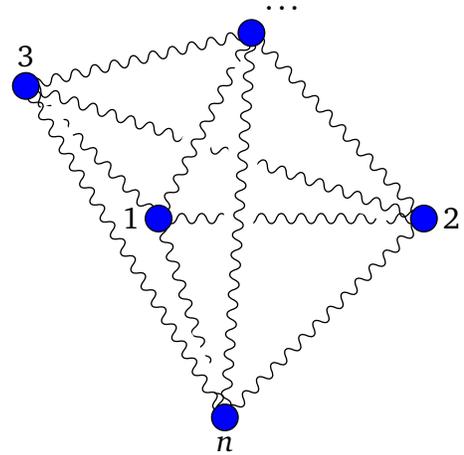


FIGURE 1.5

par rapport à sa longueur d'équilibre ℓ_{ij} . Nous pourrions en plus supposer que chaque masse est reliée par un autre ressort, de constante k_i et de longueur à l'équilibre ℓ_i , à un point d'attache fixe de position \mathbf{R}_i , de sorte qu'un terme supplémentaire de la forme suivante serait ajouté à l'énergie potentielle :

$$\frac{1}{2} \sum_i k_i (|\mathbf{r}_i - \mathbf{R}_i| - \ell_i)^2 \quad (1.125)$$

Au total, nous avons une expression relativement complexe de l'énergie potentielle $V(\mathbf{r}_1, \dots, \mathbf{r}_n)$ en fonction des n coordonnées \mathbf{r}_i .

L'utilisation des ressorts ne vise qu'à véhiculer une certaine intuition du système et n'est absolument pas nécessaire. De manière plus réaliste, on pourrait considérer que ces n masses sont en fait n atomes qui forment une molécule polyatomique avec un état fondamental dont l'énergie dépend de la position des n noyaux atomiques.

Nous allons maintenant supposer qu'il existe une configuration à l'équilibre $\bar{\mathbf{r}}_i$ de ce système, c'est-à-dire une configuration dans laquelle toutes les forces sont nulles. Le gradient de l'énergie potentielle est donc nul en ce point :

$$\left. \frac{\partial V}{\partial \mathbf{r}_i} \right|_{\bar{\mathbf{r}}_i} = 0 \quad (1.126)$$

Désignons par \mathbf{u}_i la déviation de chaque particule par rapport à cette position d'équilibre : $\mathbf{r}_i = \bar{\mathbf{r}}_i + \mathbf{u}_i$. En fait, introduisons plutôt un ensemble de $3N$ coordonnées généralisées q_α , qui sont les composantes cartésiennes des vecteurs \mathbf{u}_i , de sorte que $q_1 = u_{1x}$, $q_2 = u_{1y}$, $q_3 = u_{1z}$, $q_4 = u_{2x}$, etc. . L'énergie potentielle est maintenant une fonction $V(\mathbf{q})$ dont le gradient s'annule à l'origine ($q_\alpha = 0$). En effectuant un développement de Taylor de l'énergie potentielle par rapport à la configuration d'équilibre, on obtient, au deuxième ordre,

$$V(\mathbf{q}) = V(0) + \frac{1}{2} \sum_{\alpha, \beta} \left. \frac{\partial^2 V}{\partial q_\alpha \partial q_\beta} \right|_{\mathbf{q}=0} q_\alpha q_\beta + \dots \quad (1.127)$$

Le terme proportionnel aux dérivées premières qui devraient apparaître dans cette expression est nul en raison de l'annulation du gradient de V à $\mathbf{q} = 0$.

L'énergie cinétique de ce système aurait la forme

$$T = \sum_{\alpha} \frac{1}{2} m_{\alpha} \dot{q}_{\alpha}^2 \quad (1.128)$$

où m_{α} est la masse associée à la coordonnée q_{α} . Nous allons poser, plus généralement, que l'énergie cinétique est de la forme

$$T = \frac{1}{2} \sum_{\alpha\beta} M_{\alpha\beta} \dot{q}_{\alpha} \dot{q}_{\beta} \quad (1.129)$$

où M est une matrice symétrique, qui serait diagonale dans le cas d'un système masses-ressorts. En introduisant la matrice symétrique K définie par

$$K_{\alpha\beta} := \left. \frac{\partial^2 V}{\partial q_{\alpha} \partial q_{\beta}} \right|_{q=0} \quad (1.130)$$

on peut écrire le lagrangien de la manière suivante :

$$L = T - V = \frac{1}{2} \sum_{\alpha,\beta} M_{\alpha\beta} \dot{q}_{\alpha} \dot{q}_{\beta} - \frac{1}{2} \sum_{\alpha,\beta} K_{\alpha\beta} q_{\alpha} q_{\beta} = \frac{1}{2} \dot{\tilde{q}} M \dot{q} - \frac{1}{2} \tilde{q} K q \quad (1.131)$$

Notons que cette expression est une approximation, obtenue en développant l'énergie potentielle au deuxième ordre par rapport à la configuration d'équilibre. Nous devons donc supposer que les déviations par rapport à l'équilibre sont petites afin que ce lagrangien approximatif représente adéquatement le système physique étudié.

1.D.2 Modes propres

Obtenons l'équation de Lagrange provenant du lagrangien (1.131). Pour commencer,

$$\begin{aligned} \frac{\partial L}{\partial \dot{q}_{\gamma}} &= \frac{1}{2} \sum_{\alpha,\beta} M_{\alpha\beta} \delta_{\alpha\gamma} \dot{q}_{\beta} + \frac{1}{2} \sum_{\alpha,\beta} M_{\alpha\beta} \dot{q}_{\alpha} \delta_{\beta\gamma} \\ &= \frac{1}{2} \sum_{\beta} M_{\gamma\beta} \dot{q}_{\beta} + \frac{1}{2} \sum_{\alpha} M_{\alpha\gamma} \dot{q}_{\alpha} \\ &= \sum_{\beta} M_{\gamma\beta} \dot{q}_{\beta} \end{aligned} \quad (1.132)$$

où on a utilisé la symétrie de la matrice M : $M_{\alpha\beta} = M_{\beta\alpha}$. De la même manière, on voit que

$$\frac{\partial L}{\partial q_{\gamma}} = - \sum_{\beta} K_{\gamma\beta} q_{\beta} \quad (1.133)$$

L'équation de Lagrange est donc

$$\frac{d}{dt} \sum_{\beta} M_{\gamma\beta} \dot{q}_{\beta} + \sum_{\beta} K_{\gamma\beta} q_{\beta} = \sum_{\beta} (M_{\gamma\beta} \ddot{q}_{\beta} + K_{\gamma\beta} q_{\beta}) = 0 \quad (1.134)$$

En notation matricielle, cette équation s'écrit comme suit :

$$M \ddot{q} + K q = 0 \quad (1.135)$$

Cette équation matricielle constitue un système d'équations différentielles linéaires à coefficients constants. La façon classique de trouver les solutions à ce type d'équation est de supposer une solution oscillante :

$$\mathbf{q}(t) = \mathbf{u} \cos(\omega t + \xi) \quad (1.136)$$

où ω est la fréquence d'oscillation, \mathbf{u} est un vecteur constant et ξ une phase déterminée par les conditions initiales. En substituant dans l'équation différentielle, sachant que $\ddot{\mathbf{q}} = -\omega^2 \mathbf{q}$, on trouve

$$-\omega^2 \mathbf{M} \mathbf{u} \cos(\omega t + \xi) + \mathbf{K} \mathbf{u} \cos(\omega t + \xi) = 0 \quad (1.137)$$

Comme cette équation est valable en tout temps, on peut simplifier le cosinus et on se retrouve avec une *équation aux valeurs propres généralisées* :

$$\mathbf{K} \mathbf{u} = \lambda \mathbf{M} \mathbf{u} \quad \lambda := \omega^2 \quad (1.138)$$

Il existe plusieurs solutions distinctes à l'équation (1.138); elles seront notées ω_r et \mathbf{u}_r ($r = 1, 2, \dots, 3N$). Notons aussi que la matrice \mathbf{M} est définie positive, c'est-à-dire que l'énergie cinétique $\frac{1}{2} \dot{\mathbf{q}}^T \mathbf{M} \dot{\mathbf{q}}$ est toujours positive. Comme démontré dans l'annexe A, il s'ensuit que les différentes valeurs propres $\lambda_r = \omega_r^2$ sont nécessairement réelles (et donc ω_r est soit purement réel ou purement imaginaire) et que les différents vecteurs propres \mathbf{u}_r sont orthogonaux via la matrice \mathbf{M} :

$$\lambda_r \in \mathbb{R} \quad \tilde{\mathbf{u}}_r^T \mathbf{M} \mathbf{u}_s = 0 \quad \text{si } \lambda_r \neq \lambda_s \quad (1.139)$$

En définissant une matrice \mathbf{U} dont les colonnes sont les différents vecteurs propres \mathbf{u}_r , la propriété d'orthogonalité via \mathbf{M} s'exprime comme suit :

$$\sum_{\alpha\beta} U_{\alpha r} M_{\alpha\beta} U_{\beta s} = \delta_{rs} \quad \text{ou encore} \quad \tilde{\mathbf{U}} \mathbf{M} \mathbf{U} = \mathbf{1} \quad (1.140)$$

Donc la matrice inverse $\mathbf{U}^{-1} = \tilde{\mathbf{U}} \mathbf{M}$ existe et on peut procéder à un changement de base de sorte que tout vecteur \mathbf{q} puisse être exprimé comme une combinaison linéaire des vecteurs propres \mathbf{u}_r :

$$\mathbf{q} = \sum_r Q_r \mathbf{u}_r \quad \text{ou encore} \quad \mathbf{q} = \mathbf{U} \mathbf{Q} \quad (1.141)$$

où les nouvelles coordonnées Q_r ($r = 1, \dots, 3N$) sont les coefficients de chaque vecteur propre dans l'expression de \mathbf{q} . En fonction de \mathbf{Q} , l'équation (1.135) s'écrit

$$\mathbf{M} \mathbf{U} \ddot{\mathbf{Q}} + \mathbf{K} \mathbf{U} \mathbf{Q} = 0 \quad (1.142)$$

et en même temps l'équation (1.138) s'écrit

$$\mathbf{K} \mathbf{U} = \mathbf{M} \mathbf{U} \mathbf{D} \quad (1.143)$$

où \mathbf{D} est la matrice diagonale dont les éléments sont les valeurs propres ω_r^2 . En combinant ces deux dernières équations, on trouve

$$\mathbf{M} \mathbf{U} \ddot{\mathbf{Q}} + \mathbf{M} \mathbf{U} \mathbf{D} \mathbf{Q} = 0 \quad \text{ou encore} \quad \ddot{\mathbf{Q}} + \mathbf{D} \mathbf{Q} = 0 \quad (1.144)$$

Composante par composante, cela s'exprime comme

$$\ddot{Q}_r + \omega_r^2 Q_r = 0 \quad (1.145)$$

dont la solution est

$$Q_r(t) = \begin{cases} A_r \cos(\omega_r t + \xi_r) & \omega_r^2 > 0 \\ A_r e^{|\omega_r|t} + B_r e^{-|\omega_r|t} & \omega_r^2 < 0 \\ A_r t + B_r & \omega_r = 0 \end{cases} \quad (1.146)$$

Si $\omega_r^2 > 0$, la coordonnée Q_r oscille autour du point d'équilibre. Au contraire, si $\omega_r^2 < 0$, la coordonnée s'éloigne exponentiellement rapidement du point d'équilibre. L'équilibre est donc instable et le Lagrangien approximatif (1.131) ne peut fidèlement représenter le système réel, car la petitesse des écarts par rapport au point d'équilibre n'est pas une hypothèse valable. Enfin, si $\omega_r = 0$, alors la coordonnée Q_r augmente linéairement dans le temps. Dans ce cas, nous devons distinguer deux possibilités :

1. L'énergie potentielle reçoit des corrections d'ordre supérieur dans un développement en puissances de Q_r ; le système est alors à un point de transition entre un équilibre stable et un équilibre instable.
2. L'énergie potentielle est indépendante de Q_r même dans sa forme exacte. Dans ce cas Q_r est une coordonnée cyclique (voir la section 2.B.3) et la solution $Q_r(t) = A_r t + B_r$ est exacte. C'est par exemple le cas si Q_r est l'une des coordonnées du centre de masse du système illustré à la figure 1.5.

Notons qu'il suffit qu'une seule des N valeurs propres soit négative pour que le système dans son ensemble soit instable. Ce cas correspond généralement à un point d'inflexion (ou *col*) du potentiel V situé au point $q = 0$. Si V admet un vrai minimum au point $q = 0$, alors la matrice hessienne K est définie positive, c'est-à-dire que toutes ses valeurs propres sont positives, et l'équilibre est stable.

Supposons donc, pour continuer, que toutes les valeurs propres ω_r soient réelles, avec l'exception possible de quelques valeurs propres nulles correspondant à des variables Q_r qui n'apparaissent pas du tout dans l'énergie potentielle. En général, ces valeurs propres sont toutes différentes, et les différents mouvements d'oscillation qui se superposent ont des fréquences différentes également. Chaque fréquence correspond à un *mode propre*, c'est-à-dire à une certaine configuration oscillante du système, caractérisée par une seule coordonnée Q_r . En fonction de ce mode propre, les coordonnées q_α du système sont exprimées comme suit :

$$q_\alpha(t) = \sum_r A_r U_{\alpha r} \cos(\omega_r t + \xi_r) \quad (1.147)$$

où les constantes A_r et ξ_r sont déterminées par les conditions initiales du problème.

En conclusion :

Les solutions à l'équation aux valeurs propres généralisées $Ku = \omega^2 Mu$ sont les fréquences et les modes propres d'oscillation du système quand on l'éloigne de son équilibre.

C'est là le résultat essentiel de cette section.

Exemple 1.10 Pendule double

Retournons à l'exemple 1.2 (page 11). Le lagrangien de ce système est

$$L = \frac{1}{2}(m_1 + m_2)l_1^2\dot{\varphi}_1^2 + \frac{1}{2}m_2l_2^2\dot{\varphi}_2^2 + m_2l_1l_2\dot{\varphi}_1\dot{\varphi}_2 \cos(\varphi_1 - \varphi_2) + (m_1 + m_2)gl_1 \cos \varphi_1 + m_2gl_2 \cos \varphi_2 \quad (1.148)$$

Les deux angles φ_1 et φ_2 sont nuls dans la configuration d'équilibre, et donc le vecteur des déviations par rapport à l'équilibre est $q = (\varphi_1, \varphi_2)$. Le lagrangien dans l'approximation des petites oscillations s'obtient en développant l'expression ci-dessus en puissances de $\varphi_{1,2}$ et $\dot{\varphi}_{1,2}$ et en ne conservant que les termes quadratiques. Les deux premiers termes sont déjà sous la forme voulue. Dans le troisième terme, ce développement revient à remplacer le cosinus par l'unité. Enfin, dans les deux derniers termes, qui constituent l'énergie potentielle, il suffit de remplacer $\cos x$ par $1 - \frac{1}{2}x^2$. On obtient alors la forme (1.131), avec les matrices suivantes :

$$M = \begin{pmatrix} (m_1 + m_2)l_1^2 & m_2l_1l_2 \\ m_2l_1l_2 & m_2l_2^2 \end{pmatrix} \quad K = \begin{pmatrix} gl_1(m_1 + m_2) & 0 \\ 0 & gl_2m_2 \end{pmatrix} \quad (1.149)$$

Pour simplifier le problème, supposons que $l_1 = l_2 = \ell$ et $m_1 = m_2 = m$. Ces matrices deviennent alors

$$M = m\ell^2 \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad K = mg\ell \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \quad (1.150)$$

Nous devons maintenant trouver les vecteurs propres u et fréquences propres ω telles que $\omega^2 Mu = Ku$. On doit alors résoudre l'équation aux valeurs propres

$$K^{-1}Mu = \frac{1}{\omega^2}u \quad \text{où} \quad K^{-1} = \frac{1}{mg\ell} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix} \quad \text{et donc} \quad K^{-1}M = \frac{\ell}{g} \begin{pmatrix} 1 & \frac{1}{2} \\ 1 & 1 \end{pmatrix} \quad (1.151)$$

Les valeurs propres λ de cette dernière matrice sont les racines du polynôme caractéristique

$$0 = \begin{vmatrix} 1 - \lambda & \frac{1}{2} \\ 1 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 - \frac{1}{2} \implies \lambda = 1 \pm \frac{1}{\sqrt{2}} \quad (1.152)$$

Les vecteurs propres correspondants sont $(1, \pm\sqrt{2})$. Les fréquences propres sont

$$\omega^2 = \frac{g}{\ell} \frac{1}{1 \pm 1/\sqrt{2}} = \frac{g}{\ell} \frac{1 \mp 1/\sqrt{2}}{1 - \frac{1}{2}} = \frac{g}{\ell} (2 \mp \sqrt{2}) \implies \omega = \sqrt{\frac{g}{\ell}} \sqrt{2 \mp \sqrt{2}} \quad (1.153)$$

(le signe supérieur de l'expression précédent correspond au signe supérieur dans l'expression des vecteurs d'ondes). Donc les deux fréquences propres et les vecteurs propres associés sont :

$$\begin{aligned} \omega_1 &= \sqrt{\frac{g}{\ell}} \sqrt{2 - \sqrt{2}} & \omega_2 &= \sqrt{\frac{g}{\ell}} \sqrt{2 + \sqrt{2}} \\ u_1 &= \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix} & u_2 &= \begin{pmatrix} 1 \\ -\sqrt{2} \end{pmatrix} \end{aligned} \quad (1.154)$$

La première solution, de plus basse fréquence, correspond à une oscillation en phase des deux pendules. Dans la deuxième solution, de plus haute fréquence, les deux pendules sont en anti-phase. Notons aussi que dans les deux cas l'amplitude du pendule inférieur est $\sqrt{2}$ fois plus grande que celle du pendule supérieur; ceci n'est vrai bien sûr que si $m_1 = m_2$ et entre $l_1 = l_2$.

Problèmes

Problème 1.1 Machine d'Atwood généralisée

Considérez le système illustré, dans lequel une masse m_1 est reliée, via la poulie A, à une deuxième poulie B, elle-même reliant deux autres masses m_2 et m_3 . On néglige les masses des fils et des poulies et la gravité g agit vers le bas. Les coordonnées verticales des trois masses sont y_1 , y_2 et y_3 , positives vers le haut ; on néglige tout mouvement suivant tout autre axe. La poulie A est fixe, et les fils sont de longueurs constantes.

- A** Montrez que ce problème compte deux degrés de liberté.
- B** Écrivez le lagrangien de ce système, en utilisant y_1 et y_2 comme coordonnées généralisées.
- C** Trouvez une expression explicite pour les accélérations \ddot{y}_1 et \ddot{y}_2 en fonction de g et des trois masses. Sous quelle condition \ddot{y}_1 est-il nul ?

Problème 1.2 Mouvement le long d'une cycloïde

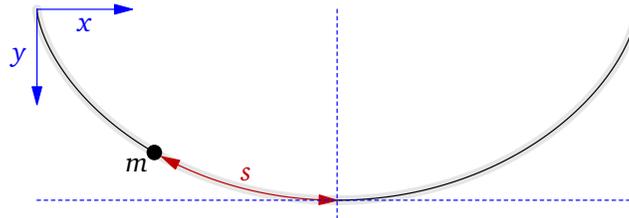
Une particule de masse m est contrainte de se déplacer le long d'une cycloïde représentée par les équations paramétriques suivantes :

$$x = a(\phi - \sin \phi) \quad y = a(1 - \cos \phi) \quad (1.155)$$

La particule est aussi sous l'influence d'un champ gravitationnel gy (y est positif vers le bas). Nous utiliserons comme coordonnée généralisée la longueur d'arc s mesurée depuis le bas de la cycloïde (voir figure). Notez que le bas de la cycloïde correspond à $\phi = \pi$ et est situé à $(x, y) = (a\pi, 2a)$

- A** Exprimez s en fonction de ϕ .
- B** Écrivez le lagrangien du problème en fonction de s et \dot{s} .
- C** Montrez que la particule oscille par rapport au bas de la cycloïde et que la période d'oscillation est indépendante de l'amplitude (aucune approximation ne doit être faite). Cela signifie que la cycloïde est une courbe isochrone, c'est-à-dire que le temps mis par la particule pour atteindre le bas de la cycloïde si on la relâche du repos est le même, quelle que soit la distance à laquelle on

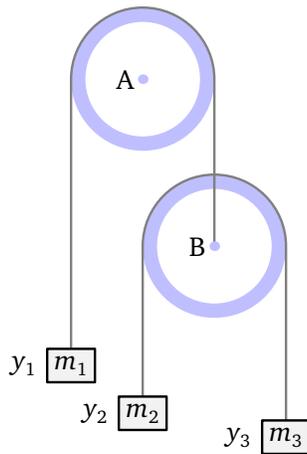
la relâche.



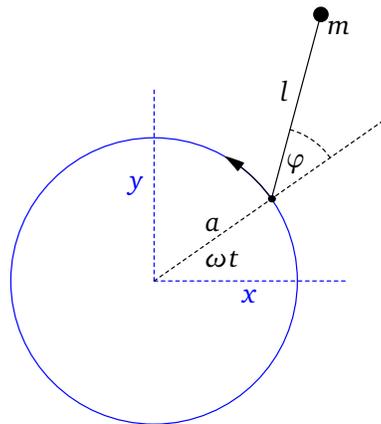
Problème 1.3 Pendule à pivot tournant

Un pendule rigide simple de longueur l et de masse m est attaché à un pivot, lui-même en mouvement circulaire uniforme de fréquence ω et de rayon a (voir figure). Le système n'est pas sous l'influence de la gravité. On utilise l'angle φ de déviation par rapport à la direction radiale comme coordonnée généralisée.

- A** Écrivez le lagrangien de ce système.
- B** Écrivez l'équation de Lagrange. Comment se compare-t-elle à celle qui régit le mouvement d'un pendule simple suspendu à un pivot fixe, mais dans un champ gravitationnel uniforme? Expliquez physiquement la ressemblance.
- C** Ajoutons maintenant un véritable champ gravitationnel $-gy$ au problème. Comment changent le lagrangien et l'équation de Lagrange? En vous basant sur votre intuition et non sur l'équation de Lagrange, identifiez deux régimes opposés des paramètres ω et g et décrivez qualitativement le mouvement dans chaque régime.



Prob. 1.1



Prob. 1.3

Problème 1.4 Masse pesante se déplaçant sur un cône

Une particule de masse m est contrainte de se déplacer sur un cône inversé d'angle α (l'angle entre la paroi du cône et la verticale). On peut placer le sommet du cône à l'origine, et le cône s'ouvre vers le haut. La particule est soumise à un champ gravitationnel uniforme g dirigé vers le bas. Adoptons comme coordonnées généralisées la distance r entre la particule et l'origine et l'angle azimutal φ que fait la projection de la position de la particule sur le plan xy avec l'axe des x .

- A** Écrivez le lagrangien du problème et obtenez les équations de Lagrange.
- B** Identifiez la ou les quantités conservées dans ce problème.
- C** Trouvez une relation différentielle entre r et φ qui pourrait en principe nous permettre de déterminer la trajectoire de la particule. Ne pas intégrer cette relation (c'est d'ailleurs impossible).
- D** Posez $g = 0$ (particule libre) et intégrez cette relation. À quelle courbe cela correspond-il ?

Problème 1.5 Masse pesante sur une sphère

Une particule de masse m est contrainte de se déplacer sur la surface d'une sphère de rayon a . La particule est soumise à un champ gravitationnel uniforme g dirigé vers le bas. Adoptons comme coordonnées généralisées les coordonnées sphériques, avec l'origine au centre de la sphère et $r = a$.

- A** Écrivez le lagrangien du problème et obtenez les équations de Lagrange.
- B** Identifiez la ou les quantités conservées dans ce problème.
- C** Trouvez une relation différentielle entre θ et φ qui pourrait en principe nous permettre de déterminer la trajectoire de la particule. Ne pas intégrer cette relation (c'est d'ailleurs impossible).
- D** Posez $g = 0$ (particule libre) et intégrez cette relation. À quelle courbe cela correspond-il ?

Problème 1.6 Objet glissant le long d'un cylindre

Un objet de masse m glisse à la surface d'un demi-cylindre, sous l'influence de la gravité (voir figure).

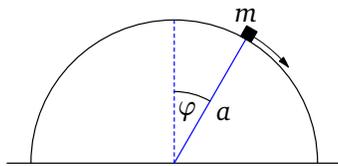
- A** En utilisant les coordonnées r et φ , écrivez le lagrangien associé à ce problème à l'aide de la méthode des multiplicateurs de Lagrange pour imposer la contrainte $r = a$.
- B** Trouvez l'angle φ auquel l'objet quitte la surface du cylindre en demandant que la force de contrainte, reliée au multiplicateur de Lagrange, s'annule.

Problème 1.7 Roulement d'un objet cylindrique sur un autre

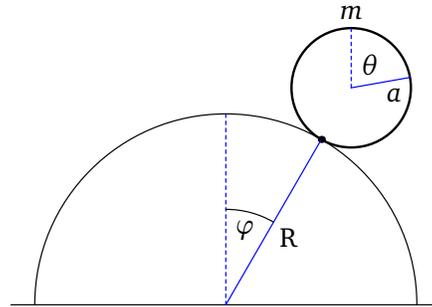
Un objet cylindrique de rayon a et de masse m roule à la surface d'un demi-cylindre de rayon R , sous l'influence de la gravité (voir figure). L'objet roule sans glisser.

A En utilisant les coordonnées r et φ du centre de l'objet, ainsi que la coordonnée angulaire θ décrivant l'orientation de l'objet, écrivez le lagrangien associé à ce problème à l'aide de la méthode des multiplicateurs de Lagrange pour imposer la contrainte $r = a + R$ ainsi que la contrainte de roulement (deux multiplicateurs sont requis).

B Trouvez l'angle φ auquel l'objet quitte la surface du cylindre en demandant que la force de contrainte, reliée au multiplicateur de Lagrange, s'annule.



Prob. 1.6



Prob. 1.7

Problème 1.8 Bille sur un anneau

Une bille de masse m est libre de glisser le long d'un anneau de rayon a et de masse μ . Cet anneau repose sur un plan et est libre de glisser sur ce plan, sauf qu'il est fixé à l'origine par un pivot sans frottement (voir figure). On peut donc utiliser les coordonnées généralisées φ (l'angle d'un diamètre de l'anneau par rapport à une direction de référence) et θ (l'angle de la bille le long de l'anneau, mesuré à partir du pivot).

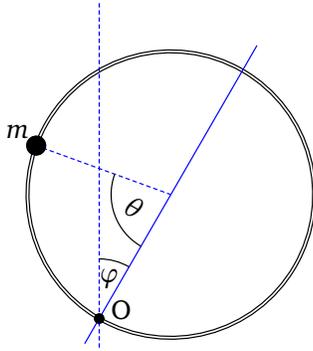
A Écrivez le lagrangien de ce système.

B Ce problème comporte une coordonnée cyclique. Écrivez une expression pour le moment conjugué à cette coordonnée, qui est donc une quantité conservée.

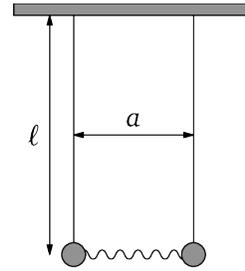
C Supposons que la bille et l'anneau sont initialement au repos, et que la bille soit munie d'un petit moteur qui la fasse avancer le long de l'anneau, jusqu'à ce qu'elle accomplisse un tour complet. Ce moteur microscopique n'injecte pas d'énergie nette dans le système, car l'énergie qu'il donne au démarrage sera récupérée au freinage. Il ne peut pas non plus modifier le moment cinétique du système, car il en fait partie, en quelque sorte. Trouvez une expression pour le changement d'orientation $\Delta\varphi$ de l'anneau suite au périple de la bille.

Problème 1.9 Géodésiques

Montrez que, lorsqu'une particule est contrainte de se mouvoir sur une surface et en l'absence de force externe, les courbes décrites par la particule sont les géodésiques de la surface. Utilisez pour ce faire le principe de moindre action.



Prob. 1.8



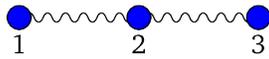
Prob. 1.11

Problème 1.10 Systèmes masses-ressorts

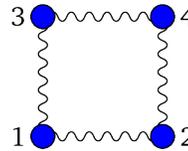
Les systèmes masses-ressorts illustrés sont constitués de particules ayant toutes la même masse m et de ressorts ayant tous la même constante k et la même longueur à l'équilibre ℓ .

A Déterminez les modes propres et les fréquences associées pour le système de trois masses illustré en (A). Supposez que les trois masses sont contraintes de se déplacer en une dimension seulement, le long des ressorts. N'hésitez pas à vous aider d'un logiciel de calcul symbolique (ex. Mathematica) au besoin, pour cette partie ou la suivante.

B Faites de même pour le système de quatre masses illustré en (B). Supposez que les masses ne peuvent se déplacer qu'en deux dimensions, sur le plan défini par le carré illustré.



(A)



(B)

Problème 1.11 Pendules couplés

Deux pendules simples, de longueur ℓ et de masse m , sont séparés latéralement d'une distance a et reliés par un ressort de constante k dont la longueur à l'équilibre est également a . La figure représente la configuration à l'équilibre du système. Soit φ_1 et φ_2 les angles que font chacun des deux pendules par rapport à la verticale.

A Dans l'approximation des petites oscillations ($\varphi_{1,2} \ll 1$), obtenez une expression pour le lagrangien de ce système, en fonction notamment de $\Omega = \sqrt{g/\ell}$ et de $\omega = \sqrt{k/\ell}$.

B Trouvez les fréquences propres et les modes propres d'oscillation.

CHAPITRE 2

Mécanique de Hamilton

A Équations de Hamilton

2.A.1 Transformation de Legendre

Considérons une fonction $f(x)$ d'une variable x . La *transformée de Legendre* de f est une autre fonction $f^*(p)$, définie par la relation suivante :

$$f^*(p) = px - f(x) \quad \text{où} \quad p = \frac{\partial f}{\partial x} \quad (2.1)$$

Ceci signifie que la relation de droite doit être utilisée pour exprimer x en fonction de p et ensuite exprimer f^* en fonction de p seulement. La preuve que f^* ne dépend pas explicitement de x lorsque cette substitution est accomplie est immédiate si on calcule la dérivée df^*/dx en tenant compte de la règle d'enchaînement :

$$\frac{\partial f^*}{\partial x} = p - \frac{\partial f}{\partial x} = 0 \quad (2.2)$$

Un exemple particulièrement simple est celui d'une fonction quadratique :

$$f(x) = \frac{1}{2}ax^2 + bx + c \implies p = ax + b \quad \text{et} \quad f^*(p) = \frac{1}{2}ax^2 - c = \frac{(p-b)^2}{2a} - c \quad (2.3)$$

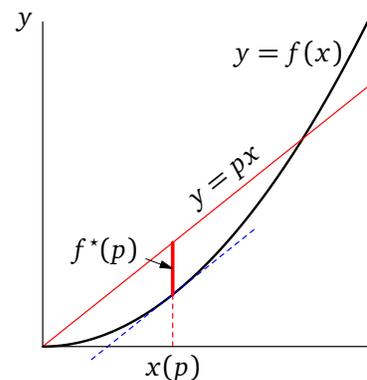


FIGURE 2.1
Interprétation géométrique de la transformation de Legendre.

La transformation de Legendre se prête à une interprétation géométrique. Considérons à cet effet la figure 2.1 : Étant donné une fonction $f(x)$ et un nombre p , on trace la droite $y = xp$ et on trouve le point sur la courbe $y = f(x)$ dont la dérivée est la même que la pente p de la droite. Ce point, $x(p)$, dépend évidemment de p . La valeur de la transformée de Legendre $f^*(p)$ est alors la longueur du segment indiqué, soit la différence entre $px(p)$ et $f(x(p))$: $f^*(p) = px(p) - f(x(p))$. Pour qu'une seule valeur $x(p)$ existe pour une valeur donnée de p , il faut que la fonction $f(x)$ soit convexe, c'est-à-dire que sa dérivée première soit strictement croissante, ou que sa dérivée seconde soit toujours positive. Pour s'en convaincre, il suffit de considérer l'équivalent de la figure 2.1 pour une courbe non convexe et de constater que plusieurs solutions $x(p)$ peuvent exister pour une valeur donnée de p .

La transformation de Legendre est *involutive*, c'est-à-dire qu'en l'appliquant deux fois de suite, on retrouve la fonction initiale. Autrement dit, $f^{**} = f$. Ceci se démontre aisément : soit $q = df^*/dp$ et la deuxième transformée

$$f^{**}(q) = qp - f^*(p) \quad q = \frac{df^*}{dp} \quad (2.4)$$

D'après l'expression de $f^*(p)$ et la définition de p , on constate que

$$q = \frac{df^*}{dp} = x + p \frac{dx}{dp} - \frac{df}{dx} \frac{dx}{dp} = x + p \frac{dx}{dp} - p \frac{dx}{dp} = x \quad (2.5)$$

D'autre part,

$$f^{**}(q) = qp - px + f(x) = xp - px + f(x) = f(x) \quad (2.6)$$

ce qui démontre la propriété d'involution.

Remarques:

- ◆ Si on définit une famille de droites par l'équation $y = px - f^*(p)$, p étant ici considéré comme un paramètre, alors l'équation de l'enveloppe de cette famille de droites est précisément $y = f(x)$, $f^*(p)$ étant la transformée de Legendre de $f(x)$ (ou vice-versa). L'un des membres de cette famille est précisément la droite tangente indiquée en bleu sur la figure 2.1.
- ◆ Considérons l'expression $px - f(x)$ pour des valeurs quelconques de p et de x , c'est-à-dire qui ne sont pas nécessairement reliées par la condition $p = df/dx$. D'après l'interprétation géométrique de la figure 2.1, $f^*(p)$ correspond à la distance maximale qu'il peut y avoir entre la droite $y = px$ et la courbe $y = f(x)$. Donc $px - f(x) \leq f^*(p)$, ce qui constitue l'inégalité de Young :

$$px \leq f(x) + f^*(p) \quad p \text{ et } x \text{ quelconques} \quad (2.7)$$

Par exemple, en l'appliquant à la fonction $f(x) = \frac{1}{2}x^2$, on trouve $f^*(p) = \frac{1}{2}p^2$ et donc

$$px \leq \frac{1}{2}x^2 + \frac{1}{2}p^2 \quad \text{ou encore} \quad x^2 + p^2 - 2px = (x - p)^2 \geq 0 \quad (2.8)$$

- ◆ La définition de la transformée de Legendre s'étend immédiatement à des fonctions de N variables $f(x)$, où $x = (x_1, \dots, x_N)$. On définit alors

$$p_i = \frac{df}{dx_i} \quad \text{et} \quad f^*(p) = \sum_i p_i x_i - f(x) = p \cdot x - f(x) \quad (2.9)$$

- ◆ La transformation de Legendre est utile lorsque la fonction $f(x)$ est strictement convexe et permet alors d'exprimer l'information contenue dans f en fonction de la dérivée $p = df/dx$. Elle est très utile en thermodynamique lorsqu'on veut définir des quantités exprimées, par exemple, en fonction de la pression au lieu du volume, ou en fonction de la température au lieu de l'entropie.

2.A.2 Moments conjugués et hamiltonien

Considérons un système mécanique à N degrés de liberté, décrit par un lagrangien $L(q, \dot{q}, t)$. On définit le *moment conjugué* p_α comme

$$p_\alpha := \frac{\partial L}{\partial \dot{q}_\alpha} \quad (2.10)$$

et la fonction de Hamilton, ou *hamiltonien*, comme suit :

$$H(q, p, t) := \sum_{\alpha} p_{\alpha} \dot{q}_{\alpha} - L(q, \dot{q}, t) = p \cdot \dot{q} - L(q, \dot{q}, t) \quad (2.11)$$

Le hamiltonien est la transformée de Legendre du lagrangien, considéré comme une fonction de la vitesse généralisée \dot{q} . La coordonnée généralisée q et t ne jouent aucun rôle dans cette transformation (ce sont en quelque sorte des spectateurs). Le hamiltonien, en tant que transformée de Legendre, doit s'exprimer en fonction de (q, p, t) et non de (q, \dot{q}, t) .

L'équation de Lagrange peut alors s'écrire

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = \frac{dp}{dt} = \frac{\partial L}{\partial q} \quad (2.12)$$

D'après la définition (2.11), $\partial L / \partial q = -\partial H / \partial q$. D'autre part, comme la transformation de Legendre est involutive, on a la relation $\dot{q} = \partial H / \partial p$. Ces deux relations forment ce qu'on appelle les *équations de Hamilton* :

$$\dot{q} = \frac{\partial H}{\partial p} \quad \dot{p} = -\frac{\partial H}{\partial q} \quad (2.13)$$

Elles sont équivalentes aux équations de Lagrange.

La différence essentielle entre les équations de Hamilton et celles de Lagrange est la suivante : pour un système à N degrés de liberté, il y a N équations de Lagrange du deuxième ordre, mais $2N$ équations de Hamilton du premier ordre. Au total, $2N$ conditions initiales sont nécessaires pour déterminer la solution de manière unique, quel que soit le jeu d'équations utilisé. On peut voir la formulation de Hamilton comme une ré expression des N équations de Lagrange sous la forme de $2N$ équations du premier ordre. Cette formulation fut proposée par William Rowan Hamilton en 1834.¹

Espace des phases

L'avantage conceptuel de la formulation de Hamilton est la notion d'*espace des phases*, défini comme l'espace à $2N$ dimensions décrit par les coordonnées généralisées q_α et les moments conjugués p_α . Un point $\eta = (q, p)$ dans l'espace des phases définit l'*état dynamique* d'un système mécanique, c'est-à-dire une donnée qui détermine à elle seule l'état futur du système. Autrement dit, les équations de Hamilton possèdent une solution unique $\eta(t) = (q(t), p(t))$ qui passe par le point $\eta_0 = (q_0, p_0)$ au temps $t = 0$, parce que les équations sont du premier ordre en dérivées.

1. Philosophical Transactions of the Royal Society, part II for 1834, pp. 247–308.

L'évolution dans le temps du système définit un *flot* (ou *écoulement*) dans l'espace des phases : par chaque point $\eta = (q, p)$ de l'espace des phases passe une courbe unique $\eta(t) = (q(t), p(t))$ qui décrit l'évolution temporelle du système mécanique étudié. Ceci est analogue à l'écoulement stationnaire d'un fluide. L'analogue de la vitesse du fluide à un point donné est la dérivée $\dot{\eta} = (\dot{q}, \dot{p})$ à ce point.

Hamiltonien et énergie

La fonction de Hamilton est égale à l'énergie totale (cinétique et potentielle) du système mécanique. Cela se voit de la manière suivante : l'énergie cinétique T est généralement une fonction quadratique des vitesses généralisées, avec la forme générale suivante :

$$T = \frac{1}{2} \sum_{\alpha, \beta} M_{\alpha\beta}(q) \dot{q}_\alpha \dot{q}_\beta \quad (2.14)$$

alors que l'énergie potentielle $V(q)$ ne dépend que des coordonnées. Dans ce cas, comme $L = T - V$, on trouve

$$p_\alpha = \sum_{\beta} M_{\alpha\beta}(q) \dot{q}_\beta \quad (2.15)$$

et donc

$$\begin{aligned} H &= \sum_{\alpha} p_\alpha \dot{q}_\alpha - \frac{1}{2} \sum_{\alpha, \beta} M_{\alpha\beta}(q) \dot{q}_\alpha \dot{q}_\beta + V(q) \\ &= \frac{1}{2} \sum_{\alpha, \beta} M_{\alpha\beta}(q) \dot{q}_\alpha \dot{q}_\beta + V(q) \\ &= T + V \end{aligned} \quad (2.16)$$

Cependant, le hamiltonien doit être exprimé en fonction des moments conjugués, et non des vitesses. En notation matricielle, on a $p = M\dot{q}$ et donc $\dot{q} = M^{-1}p$, ce qui nous permet d'écrire

$$\begin{aligned} H &= \frac{1}{2} \tilde{q} M \dot{q} + V(q) \\ &= \frac{1}{2} \overbrace{(M^{-1}p)}^{\tilde{p}} M M^{-1} p + V(q) \\ &= \frac{1}{2} \tilde{p} M^{-1} M M^{-1} p + V(q) \\ &= \frac{1}{2} \tilde{p} M^{-1} p + V(q) \end{aligned} \quad (2.17)$$

où on a utilisé le fait que la matrice M est symétrique ($\tilde{M} = M$ et $\tilde{M}^{-1} = M^{-1}$). Au total, en restaurant les indices, on obtient

$$H = \frac{1}{2} \sum_{\alpha, \beta} M^{-1}(q)_{\alpha\beta} p_\alpha p_\beta + V(q) \quad (2.18)$$

On appelle M la *matrice de masse*, car elle correspond vraiment à la masse dans le cas d'une seule particule en coordonnées cartésiennes. Mais en général cette matrice dépend des coordonnées et n'a pas nécessairement les dimensions de la masse.

Forme symplectique des équations de Hamilton

On peut écrire les équations de Hamilton de manière plus concise en introduisant le vecteur à $2N$ composantes

$$\eta = (q_1, q_2, \dots, q_N, p_1, p_2, \dots, p_N) \quad (2.19)$$

et la matrice antisymétrique $2N \times 2N$

$$J = \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix} \quad (2.20)$$

où \mathbb{I} représente ici la matrice unité $N \times N$. À l'aide ces symboles, les équations de Hamilton s'écrivent ainsi :

$$\dot{\eta} = J \frac{\partial H}{\partial \eta} \quad (2.21)$$

2.A.3 Exemples élémentaires

Exemple 2.1 Particule ponctuelle en coordonnées cartésiennes

Considérons une particule de masse m dans un champ de forces dérivant d'une énergie potentielle $V(x, y, z)$, décrite en coordonnées cartésiennes. Le lagrangien est

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V(x, y, z) \quad (2.22)$$

Les moments conjugués sont

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} \quad p_y = \frac{\partial L}{\partial \dot{y}} = m\dot{y} \quad p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z} \quad (2.23)$$

On voit donc que ces moments sont les composantes de la quantité de mouvement. Le hamiltonien est alors

$$\begin{aligned} H &= p_x \dot{x} + p_y \dot{y} + p_z \dot{z} - L \\ &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + V(x, y, z) \\ &= \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + V(x, y, z) \end{aligned} \quad (2.24)$$

et les équations de Hamilton se réduisent aux équations de Newton habituelles, ainsi qu'à la définition de la quantité de mouvement :

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial p_x} = \frac{p_x}{m} & \dot{p}_x &= -\frac{\partial H}{\partial x} = -\frac{\partial V}{\partial x} = F_x \\ \dot{y} &= \frac{\partial H}{\partial p_y} = \frac{p_y}{m} & \dot{p}_y &= -\frac{\partial H}{\partial y} = -\frac{\partial V}{\partial y} = F_y \\ \dot{z} &= \frac{\partial H}{\partial p_z} = \frac{p_z}{m} & \dot{p}_z &= -\frac{\partial H}{\partial z} = -\frac{\partial V}{\partial z} = F_z \end{aligned} \quad (2.25)$$

Exemple 2.2 Pendule simple

Considérons le pendule simple de l'exemple 1.1. La fonction de Lagrange est

$$L = \frac{1}{2}ml^2\dot{\varphi}^2 + mgl \cos \varphi \quad (2.26)$$

en fonction de la coordonnée φ et de sa dérivée $\dot{\varphi}$. Le moment conjugué est alors

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = ml^2 \dot{\varphi} \quad (2.27)$$

et le Hamiltonien est

$$H = p_\varphi \dot{\varphi} - L = \frac{1}{2} ml^2 \dot{\varphi}^2 - mgl \cos \varphi = \frac{p_\varphi^2}{2ml^2} - mgl \cos \varphi \quad (2.28)$$

Les équations de Hamilton sont alors

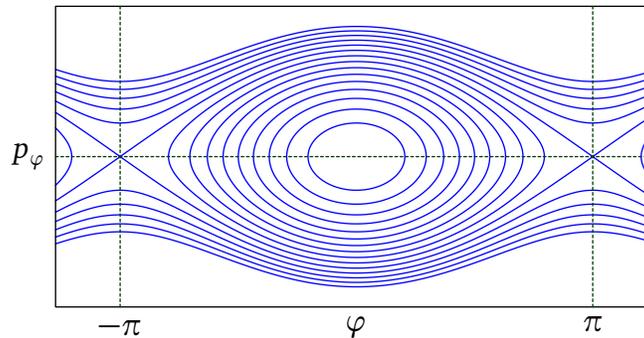
$$\dot{\varphi} = \frac{\partial H}{\partial p_\varphi} = \frac{p_\varphi}{ml^2} \quad \dot{p}_\varphi = -\frac{\partial H}{\partial \varphi} = -mgl \sin \varphi \quad (2.29)$$

La solution de ces équations ne s'obtient pas différemment que pour les équations de Lagrange : en fait, il faut différentier la première et y substituer la deuxième, ce qui nous ramène à l'équation originale de Lagrange :

$$ml^2 \ddot{\varphi} + mgl \sin \varphi = 0 \quad (2.30)$$

Il n'y a donc pas d'avantage apparent à utiliser le formalisme de Hamilton. Cependant, l'avantage de ce formalisme est ailleurs, en particulier dans la démonstration de théorèmes généraux et dans l'emploi de changements de variables particuliers, comme nous allons le voir plus loin.

La figure ci-dessous illustre les courbes d'énergie constante dans l'espace des phases pour le pendule. Comme la coordonnée angulaire φ est périodique, le diagramme l'est aussi et devrait être idéalement représenté sur un cylindre enroulé dans la direction φ , alors qu'il est « déroulé » sur la figure. Pour une énergie inférieure à mgl , les trajectoires suivies dans l'espace des phases sont fermées, alors qu'elles sont ouvertes dans le cas contraire, la courbe séparatrice correspondant à $H = mgl$. Les courbes illustrées sont parcourues dans le sens horaire, c'est-à-dire vers la gauche quand $p_\varphi < 0$ et vice-versa.



Exemple 2.3

Coordonnées sphériques Considérons une particule de masse m en dimension trois, soumise à une force centrale dérivant d'un potentiel $V(\mathbf{r})$. Utilisons les coordonnées sphériques (r, φ, θ) pour décrire la position de la particule. L'expression du lagrangien a déjà été obtenue à l'exemple

1.7 (page 19) : Les coordonnées cartésiennes sont reliées aux coordonnées sphériques par les relations suivantes :

$$L = \frac{1}{2}m \{ \dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\varphi}^2 \} - V(\mathbf{r}) \quad (2.31)$$

Les moments conjugués sont :

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r} \quad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2 \dot{\theta} \quad p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = mr^2 \sin^2 \theta \dot{\varphi} \quad (2.32)$$

On constate que p_r est la composante radiale de la quantité de mouvement. Les deux autres moments conjugués sont reliés au moment cinétique. Rappelons que le moment cinétique orbital est donné par le vecteur $\mathbf{L} = m\mathbf{r} \wedge \mathbf{v}$, dont les composantes cartésiennes sont :

$$\begin{aligned} L_x &= m(y\dot{z} - z\dot{y}) \\ L_y &= m(z\dot{x} - x\dot{z}) \\ L_z &= m(x\dot{y} - y\dot{x}) \end{aligned} \quad (2.33)$$

On vérifie aisément en substituant directement les expressions (1.73) et (1.74) que

$$L_z = mr^2 \sin^2 \theta \dot{\varphi} = p_\varphi \quad (2.34)$$

et que la grandeur au carré du moment cinétique est

$$L^2 = L_x^2 + L_y^2 + L_z^2 = m^2 r^4 (\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2) = p_\theta^2 + \frac{L_z^2}{\sin^2 \theta} \quad (2.35)$$

Le hamiltonien est

$$H = p_r \dot{r} + p_\theta \dot{\theta} + p_\varphi \dot{\varphi} - L = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\varphi^2}{2mr^2 \sin^2 \theta} + V(\mathbf{r}) \quad (2.36)$$

Les équations de Hamilton sont les suivantes :

$$\begin{aligned} \dot{r} &= \frac{\partial H}{\partial p_r} = \frac{p_r}{m} & \dot{p}_r &= -\frac{\partial H}{\partial r} = -\frac{\partial V}{\partial r} + \frac{1}{mr^3} \left(p_\theta^2 + \frac{p_\varphi^2}{\sin^2 \theta} \right) \\ \dot{\theta} &= \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{mr^2} & \dot{p}_\theta &= -\frac{\partial H}{\partial \theta} = -\frac{\partial V}{\partial \theta} + \frac{p_\varphi^2 \cos \theta}{mr^2 \sin^3 \theta} \\ \dot{\varphi} &= \frac{\partial H}{\partial p_\varphi} = \frac{p_\varphi}{mr^2 \sin^2 \theta} & \dot{p}_\varphi &= -\frac{\partial H}{\partial \varphi} = -\frac{\partial V}{\partial \varphi} \end{aligned} \quad (2.37)$$

Si V ne dépend pas de φ (symétrie azimutale), ces équations sont simplifiées du fait que p_φ est une constante du mouvement. De plus, on vérifie aisément que la parenthèse de la première équation, qui n'est autre que L^2 , est aussi une constante du mouvement, si le potentiel est central, c'est-à-dire s'il ne dépend que de r :

$$\frac{d}{dt} \left(p_\theta^2 + \frac{p_\varphi^2}{\sin^2 \theta} \right) = 2p_\theta \dot{p}_\theta - \frac{2p_\varphi^2 \cos \theta \dot{\theta}}{\sin^3 \theta} = 0 \quad (2.38)$$

(l'annulation se vérifie en substituant les équations ci-haut pour $\dot{\theta}$ et \dot{p}_θ). Le hamiltonien peut donc s'écrire

$$H = \frac{p_r^2}{2m} + \frac{L^2}{2mr^2} + V(r) \quad (2.39)$$

La constance de L^2 dans le temps fait que la première paire des équations de Hamilton est autonome des deux autres :

$$\dot{r} = \frac{p_r}{m} \quad \dot{p}_r = -\frac{\partial V}{\partial r} + \frac{L^2}{mr^3} = -\frac{\partial}{\partial r} \left(V + \frac{L^2}{2mr^2} \right) \quad (2.40)$$

Donc le problème devient effectivement unidimensionnel, la fonction

$$V_{\text{eff.}} = V + \frac{L^2}{2mr^2} \quad (2.41)$$

faisant office de potentiel effectif dans un problème unidimensionnel avec coordonnée r .

2.A.4 Principe de la moindre action

Le principe de la moindre action (voir Sec. C) peut s'exprimer également dans le langage hamiltonien. En fait, en exprimant le lagrangien en fonction du hamiltonien, on peut affirmer que l'action est donnée par l'expression

$$S = \int dt (p \cdot \dot{q} - H(q, p, t)) = \int (p \cdot dq - H(q, p, t) dt) \quad (2.42)$$

Le principe de la moindre action stipule que si on effectue une variation infinitésimale de la trajectoire réelle, caractérisée par des variations indépendantes $\delta q(t)$ et $\delta p(t)$, tout en maintenant des points initial et final fixes dans l'espace des phases, alors la variation correspondante de l'action s'annule au premier ordre. Les équations de Hamilton sont une conséquence de ce principe, comme les équations de Lagrange l'étaient dans le formalisme lagrangien. Ceci ne devrait pas nous surprendre, mais donnons tout de même la démonstration :

En appliquant les variations $q \rightarrow q + \delta q$ et $p \rightarrow p + \delta p$ à l'expression de l'action, on trouve la variation

$$\delta S = \int dt \left(\delta p \cdot \dot{q} + p \cdot \delta \dot{q} - \frac{\partial H}{\partial q} \cdot \delta q - \frac{\partial H}{\partial p} \cdot \delta p \right) \quad (2.43)$$

$$= [p \cdot \delta q]_{t_1}^{t_2} + \int dt \left(\dot{q} \cdot \delta p - \dot{p} \cdot \delta q - \frac{\partial H}{\partial q} \cdot \delta q - \frac{\partial H}{\partial p} \cdot \delta p \right) \quad (2.44)$$

$$= [p \cdot \delta q]_{t_1}^{t_2} + \int dt \left[\left(\dot{q} - \frac{\partial H}{\partial p} \right) \cdot \delta p - \left(\dot{p} + \frac{\partial H}{\partial q} \right) \cdot \delta q \right] \quad (2.45)$$

Comme les variations δq et δp sont indépendantes, les coefficients de chacun doivent s'annuler identiquement, ce qui nous mène aux équations de Hamilton :

$$\dot{q} = \frac{\partial H}{\partial p} \quad \text{et} \quad \dot{p} = -\frac{\partial H}{\partial q} \quad (2.46)$$

B Formalisme canonique

2.B.1 Crochets de Poisson

Considérons une fonction générale $G(q, p, t)$ dans l'espace des phases, qui peut accessoirement avoir une dépendance explicite dans le temps. Cette fonction possède en outre une dépendance implicite dans le temps parce que les variables q et p dépendent elles-mêmes du temps. La dérivée temporelle de G s'exprime alors de la manière suivante :

$$\begin{aligned} \frac{dG}{dt} &= \frac{\partial G}{\partial q} \cdot \dot{q} + \frac{\partial G}{\partial p} \cdot \dot{p} + \frac{\partial G}{\partial t} \\ &= \frac{\partial G}{\partial q} \cdot \frac{\partial H}{\partial p} - \frac{\partial G}{\partial p} \cdot \frac{\partial H}{\partial q} + \frac{\partial G}{\partial t} \\ &= [G, H] + \frac{\partial G}{\partial t} \end{aligned} \quad (2.47)$$

où on a défini le *crochet de Poisson* $[A, B]$ entre deux fonctions A et B :

$$[A, B] = \frac{\partial A}{\partial q} \cdot \frac{\partial B}{\partial p} - \frac{\partial B}{\partial q} \cdot \frac{\partial A}{\partial p} \quad (2.48)$$

En fonction de la coordonnée $\eta = (q, p)$, le crochet de Poisson prend la forme

$$[A, B] = \widetilde{\frac{\partial A}{\partial \eta}} \cdot \frac{\partial B}{\partial \eta} \quad (2.49)$$

où le tilde (\sim) désigne la transposée d'une matrice ou d'un vecteur colonne. Le crochet de Poisson possède les propriétés suivantes :

1. L'antisymétrie : $[A, B] = -[B, A]$.
2. L'identité de Jacobi : $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$
3. La dérivée d'un produit : $[A, BC] = B[A, C] + [A, B]C$.

Les équations de Hamilton sont un cas particulier de la relation (2.47) :

$$\begin{aligned} \dot{p}_\alpha &= [p_\alpha, H] = \sum_\beta \left\{ \frac{\partial p_\alpha}{\partial q_\beta} \frac{\partial H}{\partial p_\beta} - \frac{\partial H}{\partial q_\beta} \frac{\partial p_\alpha}{\partial p_\beta} \right\} = -\frac{\partial H}{\partial q_\alpha} \\ \dot{q}_\alpha &= [q_\alpha, H] = \sum_\beta \left\{ \frac{\partial q_\alpha}{\partial q_\beta} \frac{\partial H}{\partial p_\beta} - \frac{\partial H}{\partial q_\beta} \frac{\partial q_\alpha}{\partial p_\beta} \right\} = \frac{\partial H}{\partial p_\alpha} \end{aligned} \quad (2.50)$$

Une quantité F sera donc conservée si (1) elle ne dépend pas explicitement du temps et (2) son crochet de Poisson avec le hamiltonien s'annule : $[F, H] = 0$.

Les variables p_α et q_α ont un crochet de Poisson mutuel très simple :

$$[q_\alpha, p_\beta] = \delta_{\alpha\beta} \quad [q_\alpha, q_\beta] = 0 \quad [p_\alpha, p_\beta] = 0 \quad (2.51)$$

Tout ensemble de $2N$ variables décrivant l'espace des phases et dont les crochets de Poisson ont la forme ci-haut est qualifié de *canonique*.

Influence sur le formalisme de la mécanique quantique

La formulation la plus courante de la mécanique quantique est héritée de la *mécanique des matrices* de Heisenberg, et est calquée sur la mécanique de Hamilton. Une fonction F sur l'espace des phases devient, en mécanique quantique, une observable \hat{F} , un opérateur agissant sur l'espace de Hilbert des états possibles du système physique considéré. Le crochet de Poisson $[F, G]$ de deux quantités physiques devient, en mécanique quantique, le commutateur $[\hat{F}, \hat{G}]$ des deux observables correspondants, à un facteur $i\hbar$ près :

$$[F, G] \longrightarrow \frac{1}{i\hbar}[\hat{F}, \hat{G}] \quad (2.52)$$

En particulier, le crochet de Poisson des coordonnées cartésiennes x_α avec les impulsions correspondantes p_β , qui est $\delta_{\alpha\beta}$, devient le commutateur suivant :

$$[x_\alpha, p_\beta] \rightarrow \frac{1}{i\hbar}[\hat{x}_\alpha, \hat{p}_\beta] = \delta_{\alpha\beta} \quad \text{ou} \quad [\hat{x}_\alpha, \hat{p}_\beta] = i\hbar\delta_{\alpha\beta} \quad (2.53)$$

Enfin, l'équation (2.47) pour la dérivée par rapport au temps d'une quantité physique F , devient l'équation de Heisenberg en mécanique quantique :

$$\frac{d\hat{F}}{dt} = \frac{1}{i\hbar}[\hat{F}, \hat{H}] + \frac{\partial \hat{F}}{\partial t} \quad (2.54)$$

Exemple 2.4 Oscillateur harmonique

L'oscillateur harmonique peut être traité de manière particulièrement élégante en utilisant les crochets de Poisson. Cette démonstration trouve son parallèle exact dans la théorie quantique de l'oscillateur harmonique. Le hamiltonien d'un oscillateur harmonique en une dimension est

$$H = T + V = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \quad \text{où} \quad [x, p] = 1 \quad (2.55)$$

Plutôt que de travailler avec les variables canoniques x et p , définissons la variable complexe

$$a = \sqrt{\frac{m\omega}{2}} \left(x + i \frac{p}{m\omega} \right) \quad a^* = \sqrt{\frac{m\omega}{2}} \left(x - i \frac{p}{m\omega} \right) \quad (2.56)$$

Les variables x et p peuvent bien sûr s'exprimer en fonction de a en inversant le système ci-dessus :

$$x = \frac{1}{\sqrt{2m\omega}}(a + a^*) \quad p = \frac{1}{i} \sqrt{\frac{m\omega}{2}}(a - a^*) \quad (2.57)$$

En fonction de a et a^* , le hamiltonien prend la forme $H = \omega a^* a$ et le crochet de Poisson de cette variable avec son conjugué est particulièrement simple :

$$[a, a^*] = \frac{1}{2}[x, -ip] + \frac{1}{2}[ip, x] = -i \quad (2.58)$$

L'évolution temporelle de cette variable est alors gouvernée par l'équation

$$\dot{a} = [a, H] = \omega [a, a^* a] = \omega a [a, a^*] = -i\omega a \quad (2.59)$$

dont la solution est $a(t) = a(0)e^{-i\omega t}$. De cette solution, on retrouve les valeurs de x et p :

$$\begin{aligned} x(t) &= \frac{1}{\sqrt{2m\omega}} (a(t) + a^*(t)) = \sqrt{\frac{2}{m\omega}} a(0) \cos \omega t \\ p(t) &= \frac{1}{i} \sqrt{\frac{m\omega}{2}} (a(t) - a^*(t)) = \sqrt{2m\omega} a(0) \sin \omega t \end{aligned} \quad (2.60)$$

où nous avons supposé, dans la deuxième colonne d'équations, que $a(0)$ est réel. Manifestement l'amplitude et la phase de la variable a correspondent à l'amplitude et la phase de l'oscillation. D'autre part, la variable complexe a est autonome : l'équation différentielle de son mouvement peut être résolue d'elle-même, sans la combiner à une autre équation, ce qui serait le cas si on avait utilisé les variables canoniques x et p directement.

2.B.2 Théorèmes de Liouville et de Poincaré

Théorème de Liouville

Le vecteur $v = \dot{\eta}$ définit un champ vectoriel sur l'espace des phases, l'analogie du champ de vitesse dans la description du mouvement d'un fluide. Nous allons montrer que ce flot est celui d'un liquide incompressible. Ceci revient à dire que le champ vectoriel v possède une divergence nulle, ce qui découle immédiatement des équations de Hamilton :

$$\nabla \cdot v = \frac{\partial}{\partial q} \cdot \dot{q} + \frac{\partial}{\partial p} \cdot \dot{p} = \frac{\partial}{\partial q} \cdot \frac{\partial H}{\partial p} - \frac{\partial}{\partial p} \cdot \frac{\partial H}{\partial q} = 0 \quad (2.61)$$

en raison de l'égalité des dérivées secondes croisées :

$$\frac{\partial^2 H}{\partial q_\alpha \partial p_\alpha} = \frac{\partial^2 H}{\partial p_\alpha \partial q_\alpha} \quad (2.62)$$

Démontrons maintenant que l'annulation de la divergence du flot entraîne que le fluide est incompressible, c'est-à-dire que l'évolution dans le temps préserve le volume dans l'espace des phases. Considérons à cet effet une surface fermée dans l'espace des phases, qu'on désigne par ∂V , alors que le volume intérieur est désigné par V . Comment ce volume est-il affecté par l'évolution temporelle du système sur un temps infinitésimal δt ? Chaque point de la surface se trouve à effectuer un déplacement $\delta \eta = \dot{\eta} \delta t$; si n est la normale à la surface ∂V au point η , alors le changement de volume causé par le déplacement de la surface est

$$\delta V = \int_{\partial V} da n \cdot \delta \eta = \delta t \int_{\partial V} da n \cdot v \quad (2.63)$$

Cette intégrale de surface n'est rien d'autre que le flux du champ vectoriel v à travers la surface, et se réduit donc à une intégrale de volume par le théorème de la divergence (ou théorème de Gauss) :

$$\delta V = \delta t \int_V d^N q d^N p \nabla \cdot v \quad (2.64)$$

ce qui s'annule car $\nabla \cdot v = 0$. Notons que le mot « volume » est mal choisi, car la dimension de l'espace des phases est toujours paire, et de toute manière les unités de ce « volume » sont les unités de l'action

à la puissance N , le nombre de degrés de liberté. L'expression *extension de phase* devrait être utilisée en lieu et place. Ceci constitue l'important *théorème de Liouville*

Théorème 2.1 Théorème de Liouville

L'évolution temporelle en mécanique préserve l'extension de phase d'une région donnée de l'espace des phases.

Le théorème de Liouville permet de démontrer le théorème suivant :

Théorème 2.2 Théorème du retour (Poincaré)

Si un système mécanique évolue dans un sous-ensemble fini D de l'espace des phases, alors il va un jour revenir aussi proche que l'on veut de son état de départ.

Démonstration : Désignons par g l'opérateur qui propage le système un temps τ dans le futur, et U un voisinage de l'état initial η_0 du système considéré. Dans ce langage, gU représente l'évolution de ce voisinage un temps τ plus tard, g^2U son évolution un temps 2τ plus tard, etc. Considérons donc la suite des sous-ensembles $\{U, gU, g^2U, g^3U, \dots\}$. Chacun des membres de cette suite possède le même volume dans l'espace des phases, en vertu du théorème de Liouville. Comme le volume D est fini et préservé par l'évolution temporelle ($gD = D$), différents membres de cette suite ont certainement une intersection non nulle ; c'est-à-dire qu'il existe nécessairement deux entiers m et n tels que $g^nU \cap g^mU \neq \emptyset$, ce qui est équivalent à $U \cap g^{m-n}U \neq \emptyset$ (en supposant que $m > n$). Ceci signifie que certains points du voisinage U vont revenir dans U au bout d'un temps fini $(m - n)\tau$. Comme U peut être pris arbitrairement petit autour de η_0 , cela démontre le théorème, ou plutôt précise sa signification.

En particulier, soit D le sous-ensemble de l'espace des phases correspondant à un domaine étroit d'énergie :

$$E_0 - \delta E \leq H(q, p) \leq E_0 + \delta E \tag{2.65}$$

et supposons que l'espace des configurations est borné ; par exemple, considérons un gaz de particules dans une boîte. il est raisonnable dans cette situation de supposer que la région D possède un volume fini dans l'espace des phases, si les interactions entre les particules sont faibles. Dans ce cas le théorème du retour s'applique, ce qui a comme conséquence le paradoxe suivant : si un contenant comporte deux parties égales séparées par une cloison amovible, et qu'un gaz est initialement compris dans l'une des deux parties (l'autre étant initialement vide), alors après avoir enlevé la cloison, toutes les molécules du gaz vont éventuellement revenir dans la moitié initiale du contenant. Le temps nécessaire serait vraisemblablement plus long que l'âge de l'Univers, de sorte que ce paradoxe n'est pas vraiment un, mais constitue une conséquence intéressante du théorème de Liouville.

Invariants intégraux de Poincaré

Le théorème de Liouville est un cas particulier d'un résultat plus général démontré par Poincaré, concernant la préservation par évolution temporelle d'intégrales de dimensions paires de la forme suivante :

$$\int \sum_{\alpha} dq_{\alpha} dp_{\alpha} \tag{2 variables conjuguées}$$

$$\int \sum_{\alpha,\beta} dq_\alpha dp_\alpha dq_\beta dp_\beta \quad (4 \text{ variables conjuguées}) \quad (2.66)$$

etc.

2.B.3 Lois de conservation et symétries

Les différentes lois de conservation de la physique s'expriment de manière particulièrement simple dans la formulation hamiltonienne de la mécanique.

Commençons par la conservation de l'énergie. Cette règle s'applique lorsque le hamiltonien H ne dépend pas explicitement du temps. En effet, la dérivée de l'énergie par rapport au temps est

$$\begin{aligned} \frac{dH}{dt} &= \frac{\partial H}{\partial q} \cdot \dot{q} + \frac{\partial H}{\partial p} \cdot \dot{p} + \frac{\partial H}{\partial t} \\ &= [H, H] + \frac{\partial H}{\partial t} \\ &= \frac{\partial H}{\partial t} \end{aligned} \quad (2.67)$$

où on a utilisé l'antisymétrie du crochet de Poisson : $[H, H] = 0$. Ceci s'annule si le hamiltonien ne dépend pas explicitement du temps.

Plus généralement, une coordonnée q_α est qualifiée de *cyclique* si elle n'apparaît pas dans le hamiltonien. Dans ce cas, $\partial H / \partial q_\alpha = 0$ et donc le moment conjugué correspondant p_α est conservé : sa dérivée temporelle s'annule en vertu des équations de Hamilton. Le crochet de Poisson $[p_\alpha, H]$ s'annule. Un exemple de coordonnée cyclique est l'angle azimutal φ dans un potentiel central (voir l'exemple 2.3 ci-haut).

Flots de symétrie

On dit que le flot temporel est généré par le hamiltonien, car la variation d'une fonction F au cours de cet écoulement est donnée par le crochet de Poisson avec H :

$$\frac{dF}{dt} = [F, H] \quad (2.68)$$

(nous avons supposé que la fonction F ne dépend pas explicitement du temps). On peut également considérer d'autres flots, générés par d'autres fonctions que le hamiltonien. Ces flots ont les mêmes propriétés que le flot temporel, en particulier le théorème de Liouville s'applique encore, car sa démonstration ne repose pas sur la fonction particulière qui est utilisée pour générer le flot. Afin d'éviter toute confusion, nous n'utiliserons pas la variable t pour dénoter la variable de flot, mais plutôt le symbole s . Encore une fois, chaque point $\eta = (q, p)$ de l'espace des phases se voit associer une courbe unique $\eta(s)$ qui passe par ce point, et un champ de vitesses $d\eta/ds$ existe dans l'espace des phases, qui décrit ce flot. Ce champ de vitesses est sans divergence et le flot est incompressible.

Par exemple, on pourrait considérer le flot généré par l'un des moments conjugués, disons p_α . Dans ce cas particulièrement simple,

$$\frac{dq_\beta}{ds} = [q_\beta, p_\alpha] = \delta_{\alpha\beta} \quad \frac{dp_\beta}{ds} = [p_\beta, p_\alpha] = 0 \quad (2.69)$$

Ceci signifie que la variable s coïncide en fait avec la coordonnée généralisée correspondante q_α . Autrement dit, au cours de ce flot, la coordonnée q_α change de manière uniforme. Ce flot décrit donc une translation de la coordonnée q_α , et on dit que le moment conjugué à une coordonnée généralisée génère les translations de cette coordonnée. L'effet d'une telle translation sur une fonction $F(q, p)$ se calcule de manière différentielle par la définition du flot :

$$\frac{dF}{ds} = [F, p_\alpha] = \frac{\partial F}{\partial q_\alpha} \quad (2.70)$$

En particulier, si la fonction F ne dépend pas de la coordonnée q_α , alors sa valeur est invariante sous le flot, c'est-à-dire la même à tous les points le long de la trajectoire du flot.

Symétries et lois de conservation

Un système mécanique possède une symétrie lorsque la valeur du hamiltonien H demeure la même lorsqu'on effectue une certaine transformation des coordonnées. Par exemple, si le hamiltonien ne dépend pas d'une certaine coordonnée généralisée q_α (cette coordonnée est alors cyclique), alors le hamiltonien est invariant sous la transformation $q_\alpha \rightarrow q_\alpha + a$ et donc est invariant par le flot généré par le moment conjugué correspondant à cette coordonnée, ce qui se traduit par la condition $[p_\alpha, H] = 0$. Cette condition, en retour, entraîne que la valeur du *générateur* du flot, p_α , est constante dans le temps. Autrement dit et plus généralement, l'annulation du crochet de Poisson $[F, H] = 0$ entraîne que le hamiltonien H est invariant par rapport au flot généré par F et, pareillement, que F est constant dans le temps, c'est-à-dire invariant par rapport au flot généré par H . Par exemple :

1. Les *translations* dans la direction x sont *générées* par le moment conjugué p_x . Si le système possède une symétrie de translation le long de l'axe des x , l'énergie potentielle ne dépendra pas de x , le hamiltonien non plus, et alors $[p_x, H] = 0$. Ceci entraîne en retour la conservation de la composante en p_x de la quantité de mouvement.
2. Les *rotations* par rapport à l'axe des z sont décrites par l'angle φ en coordonnées sphériques (ou cylindriques). Ces rotations sont générées par le moment conjugué à φ , soit la composante en z du moment cinétique ($p_\varphi = L_z$, voir éq. (2.34)). Si le hamiltonien ne dépend pas de φ , alors $[p_\varphi, H] = 0$ et cette composante du moment cinétique est conservée. Bien sûr, si le hamiltonien est invariant par rapport à toute rotation (autour d'un axe quelconque), alors toutes les composantes du moment cinétique sont conservées.
3. Un cas particulier de symétrie est l'*homogénéité dans le temps*, c'est-à-dire le fait que rien ne dépende explicitement du temps dans le système considéré. Ceci entraîne que H ne dépend pas explicitement du temps. La conséquence de cela est la conservation de l'énergie, comme vu plus haut.

C Transformations canoniques

2.C.1 Définitions et propriétés générales

Dans la mécanique de Lagrange, on peut en principe effectuer un changement de coordonnées généralisées $q_\alpha \rightarrow q'_\alpha(q)$ dans le but de simplifier le problème étudié. Par exemple, on peut passer des coordonnées cartésiennes aux coordonnées sphériques pour étudier le mouvement d'un objet dans un champ de force central. Tout changement de coordonnées bien défini (c'est-à-dire biunivoque sauf peut-être en quelques points isolés) est admissible.

La mécanique de Hamilton permet d'effectuer des changements de variables plus généraux, qui s'appliquent dans l'espace des phases :

$$q_\alpha \rightarrow q'_\alpha(q, p) \quad p_\alpha \rightarrow p'_\alpha(q, p) \quad \text{ou encore} \quad \eta \rightarrow \eta'(\eta) \quad (2.71)$$

Cependant, tous ces changements ne sont pas utiles : seuls ceux qui préservent la forme des équations de Hamilton le sont. Cela impose une restriction importante sur les transformations possibles : Notons M la matrice jacobienne de la transformation :

$$M = \frac{\partial \eta'}{\partial \eta} \quad \text{ou} \quad M_{ij} = \frac{\partial \eta'_i}{\partial \eta_j} \quad (2.72)$$

Alors il s'ensuit que $\dot{\eta}' = M\dot{\eta}$ et donc que les équations de Hamilton sont, dans les nouvelles coordonnées,

$$\dot{\eta}' = MJ \frac{\partial H}{\partial \eta} \quad \text{où} \quad \frac{\partial H}{\partial \eta_i} = \sum_j \frac{\partial H}{\partial \eta'_j} \frac{\partial \eta'_j}{\partial \eta_i} = \sum_j \frac{\partial H}{\partial \eta'_j} M_{ji} \quad (2.73)$$

cette dernière expression pouvant s'écrire $\tilde{M} \partial H / \partial \eta'$. Donc on obtient

$$\dot{\eta}' = MJ\tilde{M} \frac{\partial H}{\partial \eta'} \quad (2.74)$$

Pour que cette forme soit la même que celle des équations de Hamilton originales, la matrice jacobienne doit obéir à la contrainte suivante :

$$MJ\tilde{M} = J \quad (2.75)$$

Une matrice M respectant cette condition est dite *symplectique*. On montre en particulier que le produit de deux matrices symplectiques est aussi symplectique, ainsi que l'est l'inverse d'une matrice symplectique.² En conclusion : un changement de coordonnées dans l'espace des phases préserve la forme des équations de Hamilton si le jacobien $M(\eta)$ de la transformation est une matrice symplectique partout dans l'espace des phases.

La propriété de groupe des matrices symplectiques nous permet d'affirmer directement que si deux transformations sont canoniques, alors la composition de ces deux transformations l'est aussi. De même, l'inverse d'une transformation canonique est aussi canonique.

2. Ceci nous permet d'affirmer que les matrices symplectiques forment un *groupe* au sens mathématique du terme, dit le *groupe symplectique*.

Crochets de Poisson et transformations canoniques

Les crochets de Poisson ne sont pas affectés par les transformations canoniques : ils sont invariants. Pour démontrer cette assertion, rappelons qu'on peut écrire le crochet de Poisson comme

$$[F, G] = \widetilde{\frac{\partial F}{\partial \eta}} J \frac{\partial G}{\partial \eta} \quad (2.76)$$

Lors d'une transformation canonique $\eta \rightarrow \eta'$, on cherche à exprimer les quantités impliquées en fonction de η' . Comme

$$\frac{\partial F}{\partial \eta} = \widetilde{M} \frac{\partial F}{\partial \eta'} \quad (2.77)$$

on trouve

$$[F, G] = \widetilde{\widetilde{M} \frac{\partial F}{\partial \eta'} J \widetilde{M} \frac{\partial G}{\partial \eta'}} = \widetilde{\frac{\partial F}{\partial \eta'}} M J \widetilde{M} \frac{\partial G}{\partial \eta'} = \widetilde{\frac{\partial F}{\partial \eta'}} J \frac{\partial G}{\partial \eta'} \quad (2.78)$$

ce qui signifie que le crochet de Poisson, maintenant exprimé en fonction des dérivées par rapport aux nouvelles variables η'_i , a exactement la même forme que celui qui s'exprime en fonction des anciennes variables η_i . Autrement dit, le crochet de Poisson est un *invariant canonique*. En particulier, les nouvelles variables canoniques respectent encore la condition

$$[q'_\alpha, p'_\beta] = \delta_{\alpha\beta} \quad [q'_\alpha, q'_\beta] = 0 \quad [p'_\alpha, p'_\beta] = 0 \quad (2.79)$$

2.C.2 Fonctions génératrices

Le principe de la moindre action s'applique également à tous les systèmes de coordonnées canoniques. Cela signifie que la variation de l'action est nulle, indépendamment du système de coordonnées canoniques utilisé, et donc

$$\delta \int (\mathbf{p} \cdot d\mathbf{q} - H(\mathbf{q}, \mathbf{p}, t) dt) = \delta \int (\mathbf{p}' \cdot d\mathbf{q}' - H'(\mathbf{q}', \mathbf{p}', t) dt) \quad (2.80)$$

où nous avons laissé ouverte la possibilité que le Hamiltonien approprié dans le nouveau système de coordonnées (H'), puisse différer du hamiltonien original. Étant donné que l'intégrand de l'action est fixe aux bornes de l'intégrale, cette condition est équivalente à la suivante :

$$\mathbf{p} \cdot d\mathbf{q} - H(\mathbf{q}, \mathbf{p}, t) dt = \mathbf{p}' \cdot d\mathbf{q}' - H'(\mathbf{q}', \mathbf{p}', t) dt + dF \quad (2.81)$$

où F est une fonction dans l'espace des phases (avec accessoirement une dépendance explicite dans le temps), qui donc possède la même intégrale entre les deux extrémités quel que soit le système de coordonnées utilisé. Cette relation peut aussi s'écrire

$$dF = \mathbf{p} \cdot d\mathbf{q} - \mathbf{p}' \cdot d\mathbf{q}' + (H' - H) dt \quad (2.82)$$

La fonction F peut être considérée comme fonction d'un ensemble de $2N$ coordonnées sur l'espace des phases et du temps, mais pas nécessairement les ensembles $(\mathbf{q}, \mathbf{p}, t)$ ou $(\mathbf{q}', \mathbf{p}', t)$. On pourrait par

exemple l'exprimer en fonction des variables (q, q', t) . Sous cette forme, nous la désignerons par F_1 . D'après l'expression ci-dessus, cette dépendance en (q, q') entraîne que

$$p = \frac{\partial F_1}{\partial q} \quad p' = -\frac{\partial F_1}{\partial q'} \quad H' = H + \frac{\partial F_1}{\partial t} \quad (2.83)$$

Donc : si on spécifie la fonction $F = F_1(q, q', t)$, les relations ci-haut nous donnent une expression pour p et p' et nous permettent donc de définir complètement la transformation. Pour cette raison, la fonction F_1 est appelée le *générateur* de la transformation.

Exemple 2.5

Supposons que $F_1 = q \cdot q'$. Dans ce cas, nous avons les équations de transformation suivantes :

$$p = \frac{\partial F_1}{\partial q} = q' \quad , \quad p' = -\frac{\partial F_1}{\partial q'} = -q \quad \text{et} \quad H' = H \quad (2.84)$$

Cette transformation plutôt triviale revient à échanger les coordonnées généralisées avec les moments généralisés. Elle souligne en fait que la distinction entre coordonnées et moments est purement accidentelle dans la formulation hamiltonienne.

Il est aussi possible de générer une transformation canonique à partir d'un générateur $F_2 = F_1 + p' \cdot q'$ exprimé en fonction de q et p' :

$$F_2(q, p') = p' \cdot q' + F_1(q, q') \implies dF_2 = q' \cdot dp' + p \cdot dq + (H' - H) dt \quad (2.85)$$

où donc

$$p = \frac{\partial F_2}{\partial q} \quad q' = \frac{\partial F_2}{\partial p'} \quad (2.86)$$

De même, on peut définir des générateurs $F_3(q', p)$ et $F_4(p, p')$, dont les définitions et propriétés sont résumées dans le tableau 2.1.

Exemple 2.6 Oscillateur harmonique

Le hamiltonien de l'oscillateur harmonique est

$$H = \frac{1}{2m} (p^2 + m^2 \omega^2 x^2) \quad (2.87)$$

Comme il s'agit d'une somme de carrés, on serait tenté d'utiliser des variables $p' = N$ et $q' = \varphi$ qui permettrait de mettre cette forme à profit pour éliminer la variable φ , en supposant une transformation canonique de la forme

$$p = f(N) \cos \varphi \quad x = \frac{f(N)}{m\omega} \sin \varphi \quad (2.88)$$

$f(N)$ étant une fonction à déterminer, de sorte que

$$H = H' = \frac{f(N)^2}{2m} (\cos^2 \varphi + \sin^2 \varphi) = \frac{f(N)^2}{2m} \quad (2.89)$$

Fonction génératrice	dérivées	exemple trivial
$F = F_1(q, q', t)$	$p = \frac{\partial F_1}{\partial q} \quad p' = -\frac{\partial F_1}{\partial q'}$	$F_1 = q \cdot q' \quad q' = p \quad p' = -q$
$F = F_2(q, p', t) - q' \cdot p'$	$p = \frac{\partial F_2}{\partial q} \quad q' = \frac{\partial F_2}{\partial p'}$	$F_2 = q \cdot p' \quad q' = q \quad p' = p$
$F = F_3(q', p, t) + q \cdot p$	$q = -\frac{\partial F_3}{\partial p} \quad p' = -\frac{\partial F_3}{\partial q'}$	$F_3 = q' \cdot p \quad q' = -q \quad p' = -p$
$F = F_4(p, p', t) + q \cdot p - q' \cdot p'$	$q = -\frac{\partial F_4}{\partial p} \quad q' = \frac{\partial F_4}{\partial p'}$	$F_4 = p \cdot p' \quad q' = p \quad p' = -q$

TABLE 2.1
Fonctions génératrices des transformations canoniques.

La fonction f doit être choisie de manière à ce que la transformation soit canonique. Si la manœuvre réussit, alors la coordonnée φ est cyclique et le moment conjugué N est alors constant. En utilisant un générateur de la forme $F_1(x, \varphi)$, on devrait trouver

$$N = -\frac{\partial F_1}{\partial \varphi} \quad \text{et} \quad p = \frac{\partial F_1}{\partial x} = f(N) \cos \varphi \quad (2.90)$$

sachant que $x = f(N) \sin \varphi / m\omega$, cela entraîne

$$m\omega x \cot \varphi = \frac{\partial F_1}{\partial x} \implies F_1 = \frac{m\omega x^2}{2} \cot \varphi \quad (2.91)$$

et donc

$$N = -\frac{\partial F_1}{\partial \varphi} = \frac{m\omega x^2}{2 \sin^2 \varphi} = \frac{f(N)^2}{2m\omega} \implies f(N) = \sqrt{2m\omega N} \quad (2.92)$$

La forme précise de la transformation canonique est donc

$$p = \sqrt{2m\omega N} \cos \varphi \quad x = \sqrt{\frac{2N}{m\omega}} \sin \varphi \quad (2.93)$$

et le hamiltonien s'exprime comme

$$H = \frac{f(N)^2}{2m} = \omega N \quad (2.94)$$

N est donc une constante du mouvement proportionnelle à l'énergie, ce qui entraîne que

$$\dot{\varphi} = \frac{\partial H}{\partial N} = \omega \implies \varphi(t) = \varphi_0 + \omega t \quad (2.95)$$

La solution des équations du mouvement est donc

$$p = \sqrt{2m\omega N} \cos(\omega t + \varphi_0) \quad x = \sqrt{\frac{2N}{m\omega}} \sin(\omega t + \varphi_0) \quad (2.96)$$

Il peut être intéressant de vérifier que la transformation (2.93) respecte effectivement la condition (2.75). Le caractère canonique de cette transformation découlant d'un principe variationnel, la chose n'est pas immédiatement évidente. Le jacobien M ne peut pas être calculé directement, car la forme (2.93) exprime les anciennes coordonnées en fonction des nouvelles, et non l'inverse. On peut cependant calculer le jacobien inverse :

$$M^{-1} = \frac{\partial \eta}{\partial \eta'} = \begin{pmatrix} \sqrt{(2N/m\omega)} \cos \varphi & (1/\sqrt{2Nm\omega}) \sin \varphi \\ -\sqrt{2Nm\omega} \sin \varphi & \sqrt{(m\omega/2N)} \cos \varphi \end{pmatrix} \quad (2.97)$$

L'inverse de la condition (2.75) s'exprime simplement comme

$$\tilde{M}^{-1} J M^{-1} = J \quad (2.98)$$

ce qui se traduit concrètement par

$$\begin{pmatrix} \sqrt{(2N/m\omega)} \cos \varphi & -\sqrt{2Nm\omega} \sin \varphi \\ (1/\sqrt{2Nm\omega}) \sin \varphi & \sqrt{(m\omega/2N)} \cos \varphi \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{(2N/m\omega)} \cos \varphi & (1/\sqrt{2Nm\omega}) \sin \varphi \\ -\sqrt{2Nm\omega} \sin \varphi & \sqrt{(m\omega/2N)} \cos \varphi \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (2.99)$$

Cette équation se vérifie simplement.

2.C.3 Transformations canoniques infinitésimales

Considérons une transformation canonique $\eta \rightarrow \eta'$ arbitrairement proche de la transformation identité, qu'on pourrait écrire généralement comme

$$\eta' = \eta + \delta \eta \quad (2.100)$$

où $\delta \eta(\eta)$ est une fonction infinitésimale dans l'espace des phase. Nous allons maintenant démontrer que si $\delta \eta$ est obtenu par un écoulement généré par le crochet de Poisson avec une fonction G , alors la transformation (2.100) est canonique.

Supposons donc que $\delta \eta$ a la forme suivante :

$$\delta \eta_i = \varepsilon [\eta_i, G] \quad (2.101)$$

où $G(\eta)$ est la fonction génératrice de l'écoulement et où ε est une quantité infinitésimale. Cette relation peut aussi s'écrire, d'après la forme symplectique (2.76) du crochet de Poisson,

$$\delta \eta_i = \varepsilon \frac{\partial \eta_i}{\partial \eta_j} \frac{\partial G}{\partial \eta_l} J_{jl} = \varepsilon \frac{\partial G}{\partial \eta_l} J_{il} \quad (2.102)$$

La condition pour qu'un changement de variables dans l'espace des phases soit canonique est donnée en (2.75). Dans le cas de la transformation (2.100), la matrice jacobienne M est

$$M_{ij} = \delta_{ij} + \frac{\partial \delta \eta_i}{\partial \eta_j} \quad (2.103)$$

et la condition (2.75) devient, au premier ordre en ε ,

$$\begin{aligned} 0 &= \frac{\partial \delta \eta_i}{\partial \eta_k} J_{kj} + J_{ik} \frac{\partial \delta \eta_j}{\partial \eta_k} \\ &= \varepsilon \frac{\partial^2 G}{\partial \eta_k \partial \eta_l} (J_{il} J_{kj} + J_{jl} J_{ik}) \\ &= \varepsilon \frac{\partial^2 G}{\partial \eta_k \partial \eta_l} (J_{il} J_{kj} + J_{jk} J_{il}) \end{aligned} \quad (2.104)$$

Dans la dernière équation, nous avons échangé les indices k et l dans le deuxième terme, en utilisant l'égalité des dérivées deuxièmes croisées dans le préfacteur. Sous cette forme, il est clair que l'expression s'annule bel et bien, car le deuxième terme compense exactement le premier en raison de l'antisymétrie de la matrice J : $J_{jk} = -J_{kj}$. Conclusion : les transformations infinitésimales générées par le crochet de Poisson avec une fonction G sont canoniques.

Par conséquent, les transformations finies (non infinitésimales) qui résultent d'une succession de transformations infinitésimales sont aussi canoniques. Plus précisément, une fonction génératrice G génère un écoulement en vertu de l'équation

$$\frac{d\eta_i}{ds} = [\eta_i, G] \quad (2.105)$$

Si on suit cet écoulement sur un intervalle donné de la variable s , on obtient une transformation canonique finie. En particulier, l'évolution temporelle est une transformation canonique, car elle est générée par le hamiltonien. Ainsi, la transformation

$$\eta(t) \rightarrow \eta(t + \tau) \quad (2.106)$$

qui nous fait passer du temps t au temps $t + \tau$ est canonique.

D Théorie de Hamilton-Jacobi

2.D.1 Équation de Hamilton-Jacobi

Les transformations canoniques peuvent être utilisées pour simplifier la solution d'un problème, en nous permettant d'utiliser des variables en fonction desquelles les équations de Hamilton sont plus simples à résoudre. Manifestement, le cas extrême d'une telle simplification se produit si le nouveau hamiltonien H' est identiquement nul : $H' \equiv 0$. Dans ce cas, toutes les coordonnées et les impulsions sont des constantes dans le temps :

$$\dot{p}' = 0 \quad \dot{q}' = 0 \quad (2.107)$$

Appelons $S(q, p')$ la fonction génératrice de type F_2 (voir 2.1) qui mènerait à une telle transformation. Demander $H' = 0$ revient à dire que

$$H(q, p, t) + \frac{\partial S}{\partial t} = 0 \quad (2.108)$$

Comme $p = \partial S / \partial q$, cela se traduit par une équation différentielle aux dérivées partielles :

$$H\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0 \quad (2.109)$$

Il s'agit de l'équation de Hamilton-Jacobi.³ L'approche de Jacobi remplace la solution d'un ensemble d'équations différentielles couplées par une seule équation aux dérivées partielles du premier ordre, en $N+1$ variables (temps inclus). À première vue, le gain n'est pas évident, mais en réalité cette équation permet de résoudre des problèmes difficiles d'accès autrement, et est particulièrement utile dans l'étude des orbites. Cette approche a aussi le mérite de décrire particulièrement bien la mécanique classique comme cas limite de la mécanique quantique (voir section 2.D.3).

Exemple 2.7

Le hamiltonien d'une particule libre en une dimension est $H = p^2/2m$. L'équation de Hamilton-Jacobi correspondante est donc

$$\frac{1}{2m} \left(\frac{\partial S}{\partial x} \right)^2 + \frac{\partial S}{\partial t} = 0 \quad (2.110)$$

Dans le cas d'un oscillateur harmonique, il faut ajouter un potentiel $V = \frac{1}{2}m\omega^2 x^2$ et l'équation devient

$$\frac{1}{2m} \left(\frac{\partial S}{\partial x} \right)^2 + \frac{1}{2}m\omega^2 x^2 + \frac{\partial S}{\partial t} = 0 \quad (2.111)$$

3. La méthode de solution des problèmes mécaniques basée sur la solution de cette équation a été proposée par C. Jacobi en 1837 : *Note sur l'intégration des équations différentielles de la dynamique*, Comptes rendus hebdomadaires des séances de l'Académie des sciences de Paris, 1837.

Pour un problème tridimensionnel en coordonnées sphériques, avec un potentiel $V(r, \theta)$, l'équation aurait plutôt la forme suivante :

$$\frac{1}{2m} \left(\frac{\partial S}{\partial r} \right)^2 + \frac{1}{2mr^2} \left(\frac{\partial S}{\partial \theta} \right)^2 + \frac{1}{2mr^2 \sin^2 \theta} \left(\frac{\partial S}{\partial \varphi} \right)^2 + V(r, \theta) + \frac{\partial S}{\partial t} = 0 \quad (2.112)$$

La solution de l'équation de Hamilton-Jacobi fait nécessairement apparaître $N + 1$ constantes d'intégration, dont l'une est une simple constante additive qui n'a aucun effet sur la solution, car les équations de transformation canonique ne font intervenir que les dérivées de S . En somme, N constantes d'intégration sont non triviales. Désignons ces constantes par $\alpha_1, \alpha_2, \dots, \alpha_N$. La solution complète de l'équation prendrait donc la forme

$$S = S(q_1, \dots, q_N; \alpha_1, \dots, \alpha_N; t) \quad (2.113)$$

Nous allons identifier ces N constantes aux N moments transformés (p'_1, \dots, p'_N) . En fait, nous avons une grande liberté dans le choix des nouvelles impulsions généralisées, et le fait d'identifier les nouvelles impulsions aux constantes d'intégrations de l'équation de Hamilton-Jacobi est parfaitement légitime, même s'il est arbitraire. La fonction S ainsi définie est appelée *fonction principale de Hamilton*. Suite à ce choix, les nouvelles coordonnées q' sont données par les N relations

$$q'_\gamma := \beta_\gamma = \frac{\partial S}{\partial \alpha_\gamma} \quad (\gamma = 1, \dots, N) \quad (2.114)$$

Ces nouvelles coordonnées sont constantes dans le temps, bien sûr. Ces N relations peuvent être inversées pour exprimer les anciennes coordonnées q_β en fonction des $2N$ constantes $(\alpha_\gamma, \beta_\gamma)$ et du temps, ce qui constitue alors la solution complète du problème mécanique.

Exemple 2.8

Particule libre en une dimension L'équation de Hamilton-Jacobi dans ce cas prend la forme

$$\frac{1}{2m} \left(\frac{\partial S}{\partial x} \right)^2 + \frac{\partial S}{\partial t} = 0 \quad (2.115)$$

Nous allons tenter de résoudre cette équation en posant que la fonction S est la somme d'une fonction de t et d'une fonction de x : $S(x, t) = S_0(t) + W(x)$ (on dit alors que la solution est *séparable*). En substituant cette forme dans l'équation, on trouve

$$\frac{1}{2m} \left(\frac{dW}{dx} \right)^2 + \frac{dS_0}{dt} = 0 \quad (2.116)$$

Comme les deux termes de cette équation dépendent de variables différentes, ils doivent nécessairement être constants et opposés, sinon l'égalité ne pourrait être vraie pour toutes les valeurs de x et de t . Appelons cette constante α :

$$\frac{1}{2m} \left(\frac{dW}{dx} \right)^2 = \alpha \quad \frac{dS_0}{dt} = -\alpha \quad (2.117)$$

La solution à ces équations, modulo une constante additive, est

$$S_0(t) = -\alpha t \quad W(x) = \sqrt{2m\alpha} x \quad (2.118)$$

et donc au total on obtient la solution complète

$$S(x, \alpha, t) = \sqrt{2m\alpha}x - \alpha t \quad (2.119)$$

La nouvelle coordonnée constante est alors

$$\beta = \frac{\partial S}{\partial \alpha} = x\sqrt{\frac{m}{2\alpha}} - t \quad (2.120)$$

et l'inversion de cette équation nous permet d'isoler x :

$$x = \sqrt{\frac{2\alpha}{m}}(t + \beta) \quad (2.121)$$

l'impulsion p est, quant à elle, donnée par

$$p = \frac{\partial S}{\partial x} = \sqrt{2m\alpha} \quad (2.122)$$

On voit de l'équation (2.109) que la constante α n'est rien d'autre que l'énergie de la particule, et donc $\sqrt{2m\alpha}$ est la quantité de mouvement correspondante et $\sqrt{2\alpha/m}$ la vitesse constante associée.

Exemple 2.9 Oscillateur harmonique

L'équation de Hamilton-Jacobi dans ce cas prend la forme

$$\frac{1}{2m} \left(\frac{\partial S}{\partial x} \right)^2 + \frac{1}{2}m\omega^2 x^2 + \frac{\partial S}{\partial t} = 0 \quad (2.123)$$

Nous allons encore supposer que la solution est séparable. Pour les mêmes raisons que ci-haut, on trouve plutôt les équations suivantes :

$$\frac{1}{2m} \left(\frac{dW}{dx} \right)^2 + \frac{1}{2}m\omega^2 x^2 = \alpha \quad \frac{dS_0}{dt} = -\alpha \quad (2.124)$$

En isolant $\partial W/\partial x$ et en intégrant, on trouve

$$S(x, \alpha, t) = \int dx \sqrt{2m\alpha - m^2\omega^2 x^2} - \alpha t \quad (2.125)$$

Il n'est pas nécessaire d'intégrer à ce stade, car ce sont les dérivées de S qui nous intéressent. En effet, la nouvelle coordonnée constante est alors

$$\begin{aligned} \beta &= \frac{\partial S}{\partial \alpha} = m \int dx \frac{1}{\sqrt{2m\alpha - m^2\omega^2 x^2}} - t \\ &= \sqrt{\frac{m}{2\alpha}} \int \frac{dx}{\sqrt{1 - (m\omega^2 x^2/2\alpha)}} - t \\ &= \frac{1}{\omega} \arcsin \left(\sqrt{\frac{m}{2\alpha}} \omega x \right) - t \end{aligned} \quad (2.126)$$

et l'inversion de cette relation donne

$$x = \sqrt{\frac{2\alpha}{m\omega^2}} \sin(\omega(t + \beta)) \quad (2.127)$$

ce qui est la solution bien connue de l'oscillateur harmonique. Notons en plus que l'impulsion conjuguée est

$$p = \frac{\partial S}{\partial x} = \sqrt{2m\alpha - m^2\omega^2 x^2} = \sqrt{2m\alpha(1 - \sin^2(\omega t + \omega\beta))} = \sqrt{2m\alpha} \cos(\omega t + \omega\beta) \quad (2.128)$$

Encore une fois, α représente l'énergie totale du système.

2.D.2 Fonction caractéristique de Hamilton

Dans tous les cas où H est indépendant du temps, on pourra au moins séparer la dépendance temporelle de l'équation de Hamilton-Jacobi comme ci-haut, en écrivant $S = W - Et$. On peut alors se concentrer sur l'équation de Hamilton-Jacobi réduite :

$$H\left(q, \frac{\partial W}{\partial q}\right) = E \quad (2.129)$$

où $E = \alpha_1$ est l'énergie du système. La fonction W est appelée *fonction caractéristique de Hamilton*. Elle dépend de N constantes α_i qui sont les nouvelles impulsions généralisées. La première des nouvelles coordonnées généralisées constantes, β_1 , est alors

$$\beta_1 = \frac{\partial S}{\partial E} = -t + \frac{\partial W}{\partial E} \quad \text{ou encore} \quad t + \beta_1 = \frac{\partial W}{\partial E} \quad (2.130)$$

Cette équation nous donne la dépendance temporelle du problème. La constante β_1 reflète l'arbitraire dans l'origine des temps. Pour les systèmes à plus d'un degré de liberté, l'équation de Hamilton-Jacobi est utile en pratique en autant qu'elle soit séparable, c'est-à-dire que la fonction W puisse être écrite comme une somme de fonctions séparées des coordonnées :

$$W(q) = W_1(q_1; \mathbf{a}) + W_2(q_2; \mathbf{a}) + \dots + W_N(q_N; \mathbf{a}) \quad (2.131)$$

Dans ce cas, les $N - 1$ constantes $\alpha_2, \dots, \alpha_N$ peuvent être les constantes de séparation et les $N - 1$ équations

$$\beta_\gamma = \frac{\partial W}{\partial \alpha_\gamma} \quad (\gamma = 2, 3, \dots, N) \quad (2.132)$$

nous donnent les nouvelles coordonnées constantes.

Fonction caractéristique de Hamilton comme fonction génératrice

On peut utiliser la fonction caractéristique de Hamilton W différemment : on peut l'adopter comme fonction génératrice au lieu de $S = W - Et$. Cette fonction ne dépendant pas du temps, le nouvel hamiltonien H' est identique à l'ancien, et égal à la constante $\alpha_1 = E$. Les nouvelles impulsions p'_γ peuvent être n'importe quelle combinaison des constantes α_γ , pas nécessairement ces constantes

directement. On a donc un hamiltonien $H'(p')$ qui ne dépend que des nouvelles impulsions, c'est-à-dire que toutes les coordonnées sont cycliques. Ces nouvelles coordonnées obéissent aux équations différentielles simples

$$\dot{q}'_\gamma = \frac{\partial H'}{\partial p'_\gamma} = v_\gamma \quad (2.133)$$

où les « vitesses » v_γ sont constantes (car fonctions des impulsions constantes p'). La solution du mouvement est alors

$$q'_\gamma = v_\gamma t + \beta_\gamma \quad (2.134)$$

où les constantes β_γ sont déterminées par les conditions initiales, tout comme les valeurs de φ'_γ .

2.D.3 La mécanique classique comme limite de la mécanique ondulatoire

L'équation de Hamilton-Jacobi nous permet de faire le lien entre la mécanique quantique (ou mécanique ondulatoire) et la mécanique classique. Nous allons montrer que l'équation de Schrödinger, qui décrit le comportement des ondes de matière, mène naturellement à l'équation de Hamilton-Jacobi dans la limite où l'énergie potentielle varie peu sur une distance de l'ordre de la longueur d'onde de de Broglie. Ainsi, la mécanique classique est l'équivalent pour la mécanique ondulatoire de l'optique géométrique pour l'optique ondulatoire.

Commençons par l'équation de Schrödinger pour une particule en dimension 3, dans un potentiel $V(\mathbf{r})$:

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(\mathbf{r})\psi \quad (2.135)$$

Écrivons maintenant la fonction d'onde ψ comme suit :

$$\psi(\mathbf{r}, t) = \psi_0 \exp \frac{iS(\mathbf{r}, t)}{\hbar} \quad (2.136)$$

où la fonction S représente surtout la phase de l'onde, mais est en général complexe, et ψ_0 est une constante. En substituant cette forme dans l'équation de Schrödinger, on trouve

$$\begin{aligned} -\psi \frac{\partial S}{\partial t} &= -i \frac{\hbar}{2m} \nabla \cdot (\psi \nabla S) + V\psi \\ &= \frac{1}{2m} \psi \nabla S \cdot \nabla S - i \frac{\hbar}{2m} \psi \nabla^2 S + V\psi \end{aligned} \quad (2.137)$$

ou encore, en simplifiant ψ ,

$$\frac{1}{2m} (\nabla S)^2 + V(\mathbf{r}) + \frac{\partial S}{\partial t} = \frac{i\hbar}{2m} \nabla^2 S \quad (2.138)$$

Cette équation est strictement équivalente à l'équation de Schrödinger : aucune approximation n'a été faite, seul un changement de variable $\psi \rightarrow S$.

On remarque immédiatement qu'on retrouve l'équation de Hamilton-Jacobi si le membre de droite s'annule. D'ailleurs, la constante de Planck ne se retrouve que dans le membre de droite, ce qui indique déjà qu'on retrouve l'équation de Hamilton-Jacobi dans la limite classique ($\hbar \rightarrow 0$).

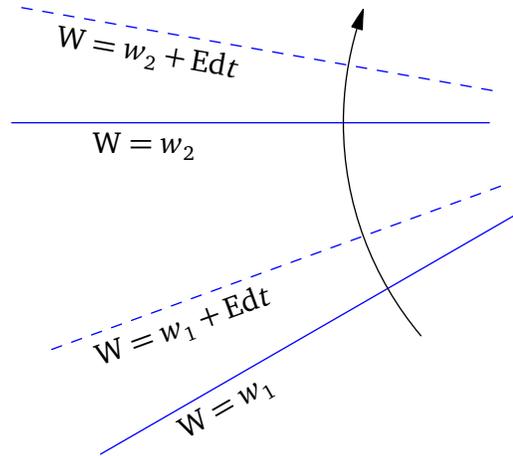


FIGURE 2.2

Cependant, la limite classique ne peut être simplement décrite par une opération sur une constante comme \hbar , dont la valeur numérique peut être grande ou petite selon le système d'unités choisi ; elle doit plutôt être décrite en fonction d'un rapport de quantités physiques qui tend vers zéro. Voyons comment. Si nous supposons que la fonction S est la fonction principale de Hamilton, alors $\nabla S = \mathbf{p}$, la quantité de mouvement de la particule. Cette quantité de mouvement varie d'un point à l'autre de l'espace (S étant une fonction de \mathbf{r}). Cette variation de la quantité de mouvement est bien-sûr causée par la présence du potentiel $V(\mathbf{r})$. Le membre de droite de l'équation (2.138) peut être négligé si

$$\frac{i\hbar}{2m} \nabla^2 S \ll \frac{1}{2m} (\nabla S)^2 \quad (2.139)$$

ou encore

$$\frac{i\hbar}{2m} \nabla \cdot \mathbf{p} \ll \frac{1}{2m} \mathbf{p}^2 \quad (2.140)$$

Dans le cas unidimensionnel, cette condition se réduit à

$$\frac{\hbar}{p^2} \frac{dp}{dx} \ll 1 \quad \text{ou encore} \quad \lambda \frac{1}{p} \frac{dp}{dx} \ll 1 \quad (2.141)$$

où

$$\lambda = \frac{\hbar}{p} = \frac{1}{2\pi} \frac{h}{p} = \frac{\lambda}{2\pi} \quad (2.142)$$

est la longueur d'onde de de Broglie réduite. Cette condition signifie justement que la variation relative de la quantité de mouvement sur une distance de l'ordre de la longueur d'onde de de Broglie doit être petite. Or la variation de p est liée à la variation de $V(\mathbf{r})$, car $|\mathbf{p}| = \sqrt{2m(E - V(\mathbf{r}))}$.

Intuitivement, on doit imaginer les surfaces à S constant comme des fronts d'onde qui avancent en fonction du temps dans la direction de la vitesse de la particule (voir figure 2.2). En effet, la quantité de mouvement \mathbf{p} étant le gradient de $W = S + Et$, elle est perpendiculaire aux surfaces à W constant : la particule se propage perpendiculairement aux fronts d'onde. Comme $S = W - Et$, en un temps dt , la fonction caractéristique changera par $dW = Edt$ pour un front d'onde donné (c'est-à-dire une valeur donnée de S). Par contre, $dW = \mathbf{p} \cdot d\mathbf{r}$, ce qui mène à la relation

$$\mathbf{p} \cdot d\mathbf{r} = Edt \quad \text{ou} \quad \mathbf{p} \cdot \mathbf{u} = E \quad \mathbf{u} := \text{vitesse du front d'onde} \quad (2.143)$$

Cette dernière relation nous donne la vitesse \mathbf{u} du front d'onde à un endroit donné, en fonction de la quantité de mouvement de la particule. Comme l'énergie E est constante, on voit que la vitesse

du front d'onde varie inversement avec la vitesse $\mathbf{v} = \mathbf{p}/m$ de la particule. Les fronts d'onde sont incurvés par les variations d'énergie potentielle (les forces). Ainsi, en pénétrant dans une région où V diminue, la vitesse v de la particule augmente et, en contrepartie, $|\mathbf{u}|$ diminue, ce qui incurve les fronts d'onde autour de la région en question (voir figure 2.2). La vitesse \mathbf{u} est la vitesse de groupe de l'onde ; la relation entre \mathbf{p} , \mathbf{u} et l'énergie E est conforme à la définition habituelle de la vitesse de groupe :

$$\mathbf{u} = \frac{\partial E}{\partial \mathbf{p}} \quad (2.144)$$

Cette situation est entièrement analogue à celle d'une onde lumineuse se propageant dans un milieu dans lequel l'indice de réfraction varie en fonction de la position : $n = n(\mathbf{r})$. La fonction $n(\mathbf{r})$ joue alors le rôle de l'énergie potentielle $V(\mathbf{r})$ et les rayons optiques sont alors déviés par les variations de n . La mécanique classique est en quelque sorte la limite «géométrique» de la mécanique ondulatoire.

Problèmes

Problème 2.1 Film de savon

Un film de savon est étiré entre deux anneaux de rayon a séparés par une distance l (voir figure). On peut montrer que la forme du film est dictée par le principe de l'énergie minimum, qui se traduit ici par un principe de surface minimum (l'énergie du film étant donnée par la tension superficielle multipliée par la surface). Les deux anneaux partageant l'axe de symétrie x et étant centrés sur cet axe à des positions $x = \pm l/2$, le film adopte une surface de révolution générée par une courbe $y(x)$ (en rouge sur la figure) bornée par les valeurs $y(\pm l/2) = a$.

A L'aire du film est donnée par une fonctionnelle de $y(x)$:

$$S[y] = \int dx F(y, y') \quad y' = \frac{dy}{dx} \quad (2.145)$$

Trouvez l'expression de la fonction F .

B En construisant l'analogie du hamiltonien pour ce problème, par transformation de Legendre, montrez que

$$\frac{y}{\sqrt{1 + (y')^2}} \quad (2.146)$$

est indépendant de x .

C En partant de ce fait, montrez que

$$y(x) = \gamma \cosh \frac{x}{\gamma} \quad \text{où } \gamma \text{ est une constante telle que } a = \gamma \cosh \frac{l}{2\gamma} \quad (2.147)$$

Le cosinus hyperbolique de révolution porte le nom de *caténoïde*.

D Calculez l'aire en fonction de γ et l . Quel est le rapport l/a maximum au-delà duquel cette surface n'est plus stable en principe? Indice : comparez à l'aire d'un film de savon qui serait constitué de deux surfaces planes bornées par chacun des deux anneaux.

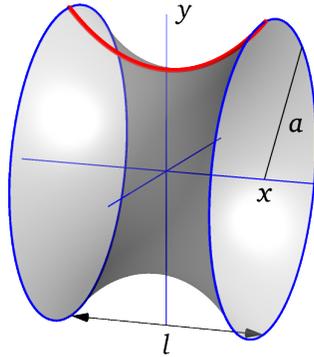
Problème 2.2 Pendule avec tige élastique

Un pendule simple de masse m est suspendu à un pivot situé à l'origine, par une tige élastique dont la longueur à l'équilibre est ℓ et la fréquence naturelle d'oscillation ω . Le pendule est libre de se mouvoir dans toutes les directions. Utilisez les coordonnées sphériques, avec l'origine au pivot et l'axe polaire dirigé vers le bas, dans la même direction que l'accélération gravitationnelle de grandeur g .

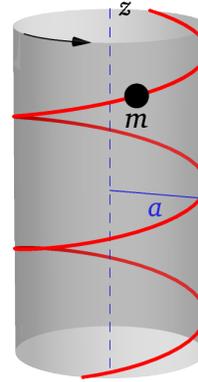
A Écrivez le hamiltonien de ce système. Basez-vous sur la forme générale du hamiltonien en

coordonnées sphériques.

- B** Quelles sont les quantités conservées dans ce problème ? Si $g = 0$, y a-t-il une quantité conservée supplémentaire ?
- C** Écrivez les équations du mouvement de Hamilton.



Prob. 2.1



Prob. 2.4

Problème 2.3 Pendule à pivot tournant : hamiltonien

- A** Obtenez le Hamiltonien associé au lagrangien du problème 1.3A.
- B** Écrivez les équations de Hamilton et vérifiez qu'elles se combinent bel et bien pour donner l'équation du mouvement d'un pendule simple.
- C** Effectuons maintenant une transformation canonique

$$p_\varphi \rightarrow P_\varphi = p_\varphi + \alpha(\varphi) \quad \varphi \rightarrow \varphi \quad (2.148)$$

où $\alpha(\varphi)$ est une fonction de φ choisie de manière à ce que H ne contienne pas de terme linéaire en P_φ . Démontrez que cette transformation est bel et bien canonique. Ensuite trouvez la forme appropriée de $\alpha(\varphi)$ et exprimez H en fonction de P_φ et de φ . À quel système simple a-t-on affaire ?

Problème 2.4 Masse glissant le long d'une hélice libre de tourner

Une bille de masse m est contrainte de suivre une courbe hélicoïdale fixée sur la périphérie d'un cylindre de moment d'inertie I . La bille descend, sans frottement, sous l'influence de la gravité. Le cylindre est libre de tourner autour de l'axe z . Le rayon du cylindre est a et la distance verticale entre deux pas de l'hélice est $2\pi h$.

- A** Écrivez le lagrangien de ce système en fonction des coordonnées cylindriques φ et z de la bille. Procédez par élimination explicite de la contrainte.
- B** Écrivez le hamiltonien correspondant et identifiez un moment conjugué conservé, en plus de

l'énergie.

C Obtenez une expression pour l'accélération constante \ddot{z} de la bille vers le bas.

Problème 2.5 Masse glissant le long d'une hélice

Refaites le problème 2.4, cette fois en supposant que le cylindre tourne sur lui-même à une fréquence angulaire ω , sans égard à la bille.

Problème 2.6 Coordonnées elliptiques

Les coordonnées elliptiques (ξ, η, φ) sont définies de la manière suivante en fonction des coordonnées cartésiennes :

$$\begin{aligned} x &= c\sqrt{(\xi^2 - 1)(1 - \eta^2)} \cos \varphi \\ y &= c\sqrt{(\xi^2 - 1)(1 - \eta^2)} \sin \varphi \\ z &= c\xi\eta \end{aligned} \tag{2.149}$$

où le domaine de variation de ξ est $[1, \infty)$ et celui de η est $[-1, 1]$. L'angle φ a la même signification qu'en coordonnées sphériques ou cylindriques.

A Montrez que, dans le plan $\varphi = 0$, les courbes à ξ constant sont des ellipses dont les foyers sont situés à $\pm c$ le long de l'axe des z , et que les courbes à η constant sont des hyperboles avec les mêmes foyers. Rappel : la somme des distances d'un point aux deux foyers d'une ellipse est constante. Pour l'hyperbole, c'est la différence qui est constante. Il est donc utile de commencer par trouver une expression pour les distances r_1 et r_2 entre un point quelconque et les deux points $(0, 0, \pm c)$.

B Une particule de masse m se déplace dans le champ gravitationnel de deux objets fixes, de masses m_1 et m_2 , situés aux deux foyers $(0, 0, \pm c)$. Écrivez le lagrangien de ce système dans le système des coordonnées elliptiques. L'utilisation de *Mathematica* ou d'un autre logiciel de calcul symbolique est conseillée.

C Obtenez l'expression du hamiltonien. Identifiez deux quantités conservées.

Problème 2.7 Crochets de Poisson et moment cinétique

A Montrez que les crochets de Poisson des composantes L_i du moment cinétique d'une particule sont les suivants :

$$[L_i, L_j] = \varepsilon_{ijk} L_k \tag{2.150}$$

où ε_{ijk} est le symbole complètement antisymétrique et où la convention de sommation sur les indices répétés est suivie.

B Soit $f(\mathbf{q}, \mathbf{p})$ une fonction de l'espace des phases à une particule qui est invariante par rotation. Montrez que $[f, \mathbf{L}] = 0$, c'est-à-dire que le crochet de Poisson de cette fonction avec toute composante du moment cinétique s'annule. Vous devez pour cela identifier quelles combinaisons

de \mathbf{p} et de \mathbf{q} peuvent apparaître dans cette fonction.

C Soit $\mathbf{A}(\mathbf{q}, \mathbf{p})$ une fonction vectorielle dans l'espace des phases à une particule. Montrez que

$$[\mathbf{A}, \mathbf{L} \cdot \mathbf{n}] = \mathbf{n} \wedge \mathbf{A} \quad (2.151)$$

où \mathbf{n} est un vecteur quelconque (on peut le choisir normé).

Problème 2.8 Transformations canoniques

A Montrez que la transformation suivante en un degré de liberté :

$$q' = q \quad p' = p + f(q) \quad (2.152)$$

où la fonction f quelconque, est canonique.

B Cela reste-t-il vrai pour une transformation à plus d'un degré de liberté, c'est-à-dire de la forme

$$q'_\alpha = q_\alpha \quad p'_\alpha = p_\alpha + f_\alpha(q) \quad ? \quad (2.153)$$

À quelle condition cette transformation est-elle canonique ?

C On désire effectuer un changement de variable dans l'espace des configurations :

$$q'_\alpha = f_\alpha(q) \quad (2.154)$$

où les N fonctions f_α sont quelconques. Trouvez une fonction génératrice F qui puisse générer ce changement de variables sous la forme d'une transformation canonique.

Problème 2.9 Équation de Hamilton-Jacobi pour une force constante

Une particule de masse m se déplace en une dimension sous l'influence d'une force constante f .

A Écrivez le hamiltonien de ce système en fonction de la coordonnée x et de l'impulsion conjuguée p . Écrivez ensuite l'équation de Hamilton-Jacobi correspondante pour la fonction caractéristique de Hamilton W .

B Résolvez l'équation de Hamilton-Jacobi et obtenez une expression de $x(t, \beta, E)$ et $p(t, \beta, E)$, où β est la nouvelle coordonnée constante et E l'énergie de la particule. Exprimez ensuite les mêmes quantités en fonction des valeurs initiales x_0 et p_0 . Note : ce problème est très simple du point de vue physique ou calculatoire. il ne vise qu'à réviser les procédures de solution de l'équation de Hamilton-Jacobi.

Problème 2.10 Équation de Hamilton-Jacobi en coordonnées elliptiques

Continuez le problème 2.6 en démontrant que l'équation de Hamilton-Jacobi est séparable en coordonnées elliptiques. Obtenez une expression intégrale pour la fonction caractéristique de

Hamilton dans ce cas.

CHAPITRE 3

Mouvement des corps rigides

A Cinématique des corps rigides

Un corps rigide est par définition composé d'une multitude de particules, ou points matériels, dont les distances mutuelles sont constantes dans le temps. Nous allons montrer qu'un tel système possède 6 degrés de liberté au total, soit 3 degrés de liberté de translation et 3 de rotation. Aux fins de l'argument, supposons que les différentes particules qui forment le corps sont numérotées de 1 à N . Il faut manifestement 3 variables pour spécifier la position de la première particule du système. Une fois cette position fixée, il faut deux angles (par exemple une longitude et une latitude) pour spécifier la direction dans laquelle se trouve la deuxième particule du système. Une fois la deuxième particule fixée, il reste encore la possibilité d'effectuer une rotation par un angle quelconque par rapport à l'axe qui relie les deux premières particules. Une fois cette rotation effectuée, il ne reste manifestement plus de liberté à l'objet. Cela fait donc 6 degrés de liberté au total. Remarquons que trois particules suffisent à déterminer un corps solide¹ : les particules numérotées de 4 à N voient leurs positions déterminées par celles des trois premières particules.

3.A.1 Matrices de Rotation

Supposons donc que l'une des particules du corps soit fixée, par exemple à l'origine des coordonnées, et attardons-nous aux trois degrés de liberté de rotation. Considérons un repère cartésien S , dont les axes sont fixes dans l'espace et les vecteurs de base sont notés $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Un deuxième repère cartésien, S' , a ses axes fixes par rapport à l'objet et ses vecteurs de base sont notés $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$; ce deuxième repère est qualifié de *mobile*. Le deuxième repère peut être exprimé en fonction du premier, à l'aide d'une matrice R , dont les composantes sont notées R_{ij} :

$$\mathbf{e}'_i = R_{i1}\mathbf{e}_1 + R_{i2}\mathbf{e}_2 + R_{i3}\mathbf{e}_3 = R_{ij}\mathbf{e}_j \quad (3.1)$$

(nous adopterons la *convention de sommation*, par laquelle on somme sur les indices répétés). La position \mathbf{r} d'un point matériel peut alors être exprimée dans l'une ou l'autre base :

$$\mathbf{r} = x_i\mathbf{e}_i = x'_i\mathbf{e}'_i = x'_iR_{ij}\mathbf{e}_j = x'_jR_{ji}\mathbf{e}_i \quad (3.2)$$

1. À condition qu'elles ne soient pas situées sur une même droite, bien sûr.

En faisant correspondre les coefficients de e_i de part et d'autre – ce qui est légitime puisque les e_i forment une base – on trouve

$$x_i = x'_j R_{ji} \quad (3.3)$$

Désignons par x le vecteur colonne abritant les composantes x_i (et de même pour x'). On écrit alors

$$x = \tilde{R}x' \quad \text{et donc} \quad x' = \tilde{R}^{-1}x \quad (3.4)$$

où \tilde{R} désigne la transposée de la matrice R .

Notation

Nous allons utiliser des linéales pour désigner les vecteurs colonnes et les matrices (ex. x , R , A), mais des italiques pour leurs composantes (ex. x_i , R_{ij}). Les caractères gras seront réservés aux vecteurs géométriques (ex. r). La distinction entre r et x est importante : r est un objet géométrique indépendant de la base utilisée, alors que x est un ensemble de composantes dans une base donnée.

La relation entre les deux repères S et S' est une rotation, et doit préserver la distance entre deux points, en particulier la norme du vecteur r . Cette norme au carré peut s'exprimer, en notation matricielle, comme $\tilde{x}x$ et donc une condition essentielle sur la matrice R est

$$\tilde{x}'x' = \tilde{x}x = \widetilde{(\tilde{R}x')} \tilde{R}x' = \tilde{x}'R\tilde{R}x' \quad (3.5)$$

Comme cette relation doit s'appliquer sur tous les vecteurs x' pour une matrice R donnée, on doit conclure que $R\tilde{R} = \mathbb{I}$: l'inverse de la matrice R est sa transposée. On peut également écrire cette condition comme $\tilde{R}R = \mathbb{I}$ Sachant cela, la relation ci-haut entre x et x' devient

$$x' = Rx \quad \text{et} \quad x = \tilde{R}x' \quad (3.6)$$

Une matrice dont l'inverse coïncide avec sa transposée est dite *orthogonale*. Ces matrices ont les propriétés suivantes :

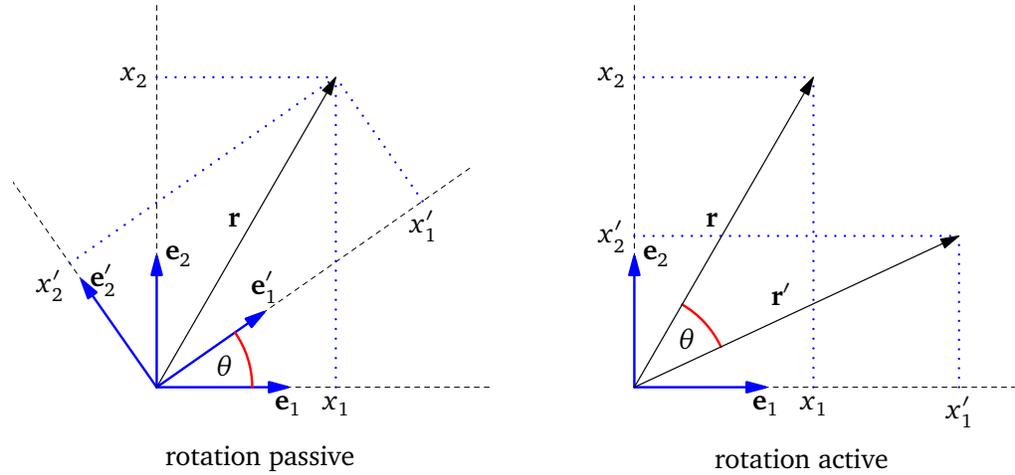
1. La relation $\tilde{R}R = \mathbb{I}$ signifie que les colonnes de R forment une base orthonormée.
2. Le produit de deux matrices orthogonales est encore orthogonal. En effet, si $\tilde{R}_1R_1 = \mathbb{I}$ et $\tilde{R}_2R_2 = \mathbb{I}$, alors $\widetilde{(R_1R_2)}(R_1R_2) = \tilde{R}_2\tilde{R}_1R_1R_2 = \tilde{R}_2R_2 = \mathbb{I}$.
3. L'ensemble des matrices orthogonales forme ce qu'on appelle un *groupe*, qu'on note $O(n)$ dans l'espace à n dimensions (ici on s'intéresse à $O(3)$). Pour qu'un ensemble de matrices forment un groupe, il faut que :
 - le produit g_1g_2 de deux éléments de l'ensemble soit encore un élément de l'ensemble.
 - chaque élément g comporte un inverse g^{-1} dans l'ensemble (donc la matrice identité doit appartenir à l'ensemble).
4. Le déterminant d'une matrice étant égal à celui de sa transposée, on a la relation $\det(\tilde{R}R) = (\det R)^2 = 1$, c'est-à-dire que le déterminant d'une matrice orthogonale est ± 1 . Si le déterminant est $+1$, alors la matrice R est qualifiée de *directe*, sinon elle est *indirecte*.

La matrice qui effectue une rotation d'angle θ autour de l'axe des z est

$$R(z, \theta) = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.7)$$

FIGURE 3.1

À gauche : rotation passive d'un angle θ par rapport à l'axe des z . À droite, rotation active correspondante.



Cette matrice, appliquée à un vecteur x , produit les composantes du vecteur x' qui représente le même vecteur position r , mais cette fois dans le repère S' en rotation par rapport au repère S (voir partie gauche de la figure 3.1).

On montre que, pour un axe de rotation général n , les composantes de la matrice de rotation sont données par l'expression suivante :

$$R_{ij} = n_i n_j (1 - \cos \theta) + \delta_{ij} \cos \theta + \varepsilon_{ijk} n_k \sin \theta \quad (3.8)$$

ou encore

$$R(\mathbf{n}, \theta) = \begin{pmatrix} n_x^2(1 - \cos \theta) + \cos \theta & n_x n_y(1 - \cos \theta) + n_z \sin \theta & n_x n_z(1 - \cos \theta) - n_y \sin \theta \\ n_y n_x(1 - \cos \theta) - n_z \sin \theta & n_y^2(1 - \cos \theta) + \cos \theta & n_y n_z(1 - \cos \theta) + n_x \sin \theta \\ n_z n_x(1 - \cos \theta) + n_y \sin \theta & n_z n_y(1 - \cos \theta) - n_x \sin \theta & n_z^2(1 - \cos \theta) + \cos \theta \end{pmatrix} \quad (3.9)$$

Cette représentation est générale : toute matrice de rotation directe (donc de déterminant unité) peut être caractérisée par un axe de rotation n et un angle de rotation θ .

Théorème 3.1 Théorème d'Euler

Toute matrice de rotation directe R peut être caractérisée par un axe de rotation n et un angle de rotation θ autour de cet axe.

Preuve: Considérons l'identité suivante, valable pour une matrice orthogonale :

$$(\mathbb{I} - R)\tilde{R} = (\tilde{R} - \mathbb{I}) \quad (3.10)$$

Prenons le déterminant de chaque membre de cette équation et appliquons la propriété $\det A = \det \tilde{A}$, ainsi que la propriété $\det R = \det \tilde{R} = 1$ pour une matrice orthogonale directe. On trouve alors

$$\det(\mathbb{I} - R) = \det(\tilde{R} - \mathbb{I}) = \det(R - \mathbb{I}) = -\det(\mathbb{I} - R) \quad (3.11)$$

la dernière égalité s'appliquant à toute matrice 3×3 . On conclut que $\det(R - \mathbb{I}) = 0$, ce qui signifie que 1 est toujours une valeur propre de R . Appelons n le vecteur propre correspondant à

cette valeur propre : $R\mathbf{n} = \mathbf{n}$. Ce vecteur n'est donc pas affecté par la rotation et détermine l'axe de rotation : tout multiple de ce vecteur restera aussi invariant par application de la rotation, et la droite $\mathbf{x} = \alpha\mathbf{n}$ ($\alpha \in \mathbb{R}$) constitue l'axe de rotation.

D'autre part, la matrice R est forcément diagonalisable, car elle est normale.^a Elle possède donc trois valeurs propres $\lambda_{1,2,3}$. On sait par ailleurs que le déterminant d'une matrice est le produit de ses valeurs propres, et que sa trace est la somme de ses valeurs propres :

$$\det R = \lambda_1\lambda_2\lambda_3 \quad \text{et} \quad \text{tr} R = \lambda_1 + \lambda_2 + \lambda_3 \quad (3.12)$$

Comme $\det R = 1$, on a donc $\lambda_1\lambda_2\lambda_3 = 1$. Comme on vient de le voir, on peut choisir $\lambda_3 = 1$, ce qui nous laisse avec $\lambda_1\lambda_2 = 1$. De plus, comme une matrice orthogonale est réelle, son polynôme caractéristique est à coefficients réels et ses racines sont soit réelles, soit existent en paires complexes conjuguées. Nous avons donc deux possibilités :

1. $\lambda_1 = \lambda_2^{-1}$ est réel. Dans ce cas, les vecteurs propres de R seraient également réels. Soit \mathbf{x} un vecteur propre associé à la valeur propre λ_1 . Comme $R\mathbf{x} = \lambda_1\mathbf{x}$, on voit que la rotation ne préserve pas la longueur du vecteur \mathbf{x} , à moins que $\lambda_1 = \lambda_2 = 1$ ou $\lambda_1 = \lambda_2 = -1$. Dans le premier cas, R est la matrice identité. Dans le deuxième cas, R décrit une rotation de π autour de l'axe \mathbf{n} .
2. $\lambda_1 = \lambda_2^*$ et λ_1 est unimodulaire : on peut l'écrire sous la forme $e^{i\theta}$. Dans ce cas, $\text{tr} R = 1 + 2\cos\theta$ et θ , en comparant avec l'expression (3.7), est bien l'angle de rotation.

Dans tous les cas, la direction du vecteur propre associé à la valeur propre $\lambda_3 = 1$ est la direction de l'axe de rotation, car un vecteur dans cette direction n'est pas affecté par la rotation : $R\mathbf{x} = \mathbf{x}$.

a. Rappelons qu'une matrice A est dite *normale* si $AA^\dagger = A^\dagger A$, où $A^\dagger \equiv \tilde{A}^*$. Une matrice normale est diagonalisable.

3.A.2 Rotations passives et actives

Les rotations que nous avons décrites ci-haut sont qualifiées de *passives*, au sens que le vecteur \mathbf{r} n'est pas affecté par la rotation : seules ses composantes le sont, car la rotation n'est qu'un changement de base. On peut aussi considérer des rotations *actives*, qui n'ont rien à voir avec la relation entre deux repères, mais décrivent plutôt un changement réel du vecteur \mathbf{r} , toujours exprimé dans la même base. Dans le cas actif, le vecteur colonne \mathbf{x}' contient les composantes du nouveau vecteur \mathbf{r}' ayant subi la rotation :

$$\mathbf{r}' = x'_i \mathbf{e}_i = R_{ij}^{(\text{act})} x_j \mathbf{e}_i \quad \text{ou} \quad \mathbf{x}' = R^{(\text{act})} \mathbf{x} \quad (3.13)$$

Notez que la relation entre les nouvelles et les anciennes composantes du vecteur est la même quel que soit le point de vue adopté. Il est assez évident, en considérant la figure 3.1, que, pour des \mathbf{x} et \mathbf{x}' donnés, la rotation active $\mathcal{R}^{(\text{act})}$ est géométriquement l'inverse de la rotation passive $\mathcal{R}^{(\text{pass})}$: elle s'exerce dans l'autre sens et donc, au sens géométrique, $\mathcal{R}^{(\text{act})} = \mathcal{R}^{(\text{pass})^{-1}}$. En effet, le passage de \mathbf{x} vers \mathbf{x}' illustré sur la partie gauche de la figure peut être vu comme la conséquence d'un changement de base effectué par la matrice $R(\mathbf{n}, \theta)$, ou encore comme l'effet d'une rotation réelle du vecteur \mathbf{r} par un angle $-\theta$, les vecteurs de base étant constants (partie droite de la figure). Par contre, la matrice de rotation qui relie \mathbf{x}' à \mathbf{x} est la même dans les deux cas $R^{(\text{act})} = R^{(\text{pass})}$. En fonction de la matrice $R(\mathbf{n}, \theta)$ définie en (3.8), on a la correspondance

$$\mathcal{R}^{(\text{pass})}(\mathbf{n}, \theta) \rightarrow R(\mathbf{n}, \theta) \quad \text{et} \quad \mathcal{R}^{(\text{act})}(\mathbf{n}, \theta) \rightarrow R(\mathbf{n}, -\theta) \quad (3.14)$$

Nous devons spécifier au besoin lequel des deux modes, actif ou passif est utilisé.

B Vitesse angulaire

Un mouvement de rotation peut être décrit comme une matrice de rotation $R(t)$ qui dépend du temps de manière continue. Cette matrice relie le repère fixe dans l'espace au repère mobile (fixe à l'objet). On peut convenir, par exemple, que $R(0) = \mathbb{I}$, la matrice identité.

3.B.1 Rotations infinitésimales

Considérons une matrice de rotation infinitésimale, c'est-à-dire très proche de la matrice identité. On peut écrire une telle matrice comme $R = \mathbb{I} + \delta M$, δM étant une matrice infinitésimale. La condition $\tilde{R}R = \mathbb{I}$ revient alors à

$$\mathbb{I} = (\mathbb{I} + \delta \tilde{M})(\mathbb{I} + \delta M) = \mathbb{I} + \delta \tilde{M} + \delta M + \delta \tilde{M} \delta M \quad (3.15)$$

On peut négliger le produit $\delta \tilde{M} \delta M$, qui est du deuxième ordre de petitesse, et on trouve la condition

$$\delta \tilde{M} + \delta M = 0 \quad (3.16)$$

ce qui signifie que δM est une matrice antisymétrique.

Considérons maintenant une matrice de rotation dépendant du temps, $R(t)$, et sa dérivée par rapport au temps, $\dot{R}(t)$. Par définition, $R(t + dt) = R(t) + \dot{R}(t) dt$. D'un autre côté, la rotation $R(t + dt)$ ne différant de $R(t)$ que par une quantité infinitésimale, elle peut s'obtenir de $R(t)$ en appliquant cette dernière sur une rotation infinitésimale, c'est-à-dire que

$$R(t + dt) = R(t)(\mathbb{I} + \delta M) \quad (3.17)$$

Cette matrice δM doit être proportionnelle à dt , et en plus être antisymétrique. Écrivons-la sous la forme $\Omega(t) dt$, où $\Omega(t)$ est une matrice finie, antisymétrique et dépendant du temps en général. Cela nous donne

$$R(t) + \dot{R}(t) dt = R(t)(\mathbb{I} + \Omega(t) dt) \implies \dot{R}(t) = R(t)\Omega \quad \text{ou} \quad \Omega(t) = \tilde{R}(t)\dot{R}(t) \quad (3.18)$$

La matrice Ω étant antisymétrique, elle comporte trois composantes indépendantes et peut s'écrire sous la forme suivante :

$$\Omega = \begin{pmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{pmatrix} \quad \text{ou} \quad \Omega_{ij} = \varepsilon_{ijk} \omega_k \quad (3.19)$$

où ε_{ijk} est le symbole complètement antisymétrique de Levi-Civita et où nous avons défini un objet à trois composantes $\omega = (\omega_x, \omega_y, \omega_z)$ que nous désignerons par *vitesse angulaire* associée au mouvement $R(t)$. Nous n'avons pas démontré que cet objet est un vecteur, c'est-à-dire qu'il se transforme comme un vecteur lors d'un changement de repère ; c'est l'objet de la section suivante.

3.B.2 Caractère vectoriel de la vitesse angulaire

Démontrons maintenant que la vitesse angulaire est un vecteur, c'est-à-dire que ses composantes cartésiennes se comportent comme un vecteur lors d'un changement de repère (rotation passive). Nous devons pour cela démontrer comment la matrice de rotation active R diffère d'un repère à l'autre. La question doit être comprise précisément ici : nous allons considérer une matrice de rotation passive P qui effectue un changement de repère de S à S' , les deux repères étant cette fois fixes dans l'espace, mais différents. La matrice de rotation active n'est pas la même dans les deux repères, bien évidemment. Soyons prudents avec la notation ici et appelons x_R les composantes d'un vecteur ayant subi une rotation active, exprimées dans le repère S , et x'_R les composantes du même vecteur, cette fois dans le repère S' :

$$x_R = Rx \quad x'_R = R'x' \quad x' = Px \quad x'_R = Px_R \quad (3.20)$$

En combinant ces différentes équations, on trouve

$$x'_R = Px_R = PRx = PR\tilde{P}x' \implies R' = PR\tilde{P} \quad (3.21)$$

Autrement dit, la transformation d'une matrice en passant d'un repère $S \rightarrow S'$ est

$$R' = PR\tilde{P} \quad \text{ou} \quad R = \tilde{P}R'P \quad (3.22)$$

En conséquence,

$$\Omega' = \tilde{R}'R' = PR\tilde{P}PR\tilde{P} = P\tilde{R}\tilde{R}P = P\Omega\tilde{P} \quad (3.23)$$

ce qui s'exprime, en composantes,

$$\varepsilon_{ijk}\omega'_k = P_{in}\varepsilon_{nml}\omega_l P_{jm} \quad (3.24)$$

Supposons maintenant que ω est effectivement un vecteur et vérifions que la relation ci-dessus est satisfaite dans ce cas. On doit donc, d'après (3.6), poser $\omega_l = P_{kl}\omega'_k$ et substituer cette relation dans le membre de droite, qui devient

$$\varepsilon_{nml}P_{in}P_{jm}P_{kl}\omega'_k \quad (3.25)$$

Par définition du déterminant, on a précisément

$$\varepsilon_{nml}P_{in}P_{jm}P_{kl} = \varepsilon_{ijk} \det P \quad (3.26)$$

Comme $\det P = 1$ pour une rotation directe, l'expression (3.25) n'est autre que $\varepsilon_{ijk}\omega'_k$ et on vérifie donc le résultat annoncé : la vitesse angulaire est un vecteur.

3.B.3 Composition des vitesses

Supposons qu'un repère cartésien S' soit en rotation par rapport à un autre repère cartésien S , de sorte qu'on puisse écrire, symboliquement, $S' = R(t)S$. Considérons ensuite un vecteur \mathbf{a} qui dépend du temps, dont les composantes forment le vecteur colonne \mathbf{a} dans le repère S et le vecteur colonne

\mathbf{a}' dans le repère S' , reliés par $\mathbf{a}' = R\mathbf{a}$. Les dérivées de ces vecteurs, telles qu'observées dans les deux repères, sont $\dot{\mathbf{a}}$ et $\dot{\mathbf{a}}'$ et sont donc reliées par

$$\dot{\mathbf{a}}' = \dot{R}\mathbf{a} + R\dot{\mathbf{a}} = R\Omega\mathbf{a} + R\dot{\mathbf{a}} = R(\Omega\mathbf{a} + \dot{\mathbf{a}}) \quad (3.27)$$

Appelons $\dot{\mathbf{a}}_R$ le vecteur géométrique formé par les composantes \dot{a}'_i ; il représente la dérivée de \mathbf{a} telle qu'observée dans le repère tournant. En développant explicitement ses composantes dans le repère S , qui forment le vecteur colonne $\Omega\mathbf{a} + \dot{\mathbf{a}}$, on trouve

$$\dot{a}'_{Ri} = \varepsilon_{ijk}\omega_k a_j + \dot{a}_i \quad (3.28)$$

Ce qui s'exprime, vectoriellement, par

$$\dot{\mathbf{a}}_R = -\boldsymbol{\omega} \wedge \mathbf{a} + \dot{\mathbf{a}} \quad \text{ou} \quad \dot{\mathbf{a}} = \dot{\mathbf{a}}_R + \boldsymbol{\omega} \wedge \mathbf{a} \quad (3.29)$$

Appliquons cette relation au vecteur position ($\mathbf{a} \rightarrow \mathbf{r}$). On obtient alors une relation entre la vitesse $\mathbf{v} = \dot{\mathbf{r}}$ observée dans le repère S et la vitesse $\mathbf{v}_R = \dot{\mathbf{r}}_R$ observée dans le repère tournant S' :

$$\mathbf{v} = \mathbf{v}_R + \boldsymbol{\omega} \wedge \mathbf{r} \quad (3.30)$$

En particulier, si S' est un système d'axes liés à l'objet, la vitesse d'une particule de l'objet serait nulle dans S' ($\mathbf{v}_R = 0$) et donc donnée par

$$\mathbf{v} = \boldsymbol{\omega} \wedge \mathbf{r} \quad (3.31)$$

Voyons ensuite comment des rotations composées mènent à une addition des vitesses angulaires correspondantes. En particulier, considérons trois repères cartésiens, que nous noterons S_1 , S_2 et S_3 . S_1 peut être considéré comme fixe dans l'espace, alors que S_2 est obtenu de S_1 par une rotation $R_1(t)$ et S_3 est obtenu de S_2 par une rotation $R_2(t)$. La rotation qui nous fait passer directement de S_1 à S_3 est alors $R_2(t)R_1(t)$ (notez l'ordre des matrices : il est important). Montrons que la vitesse angulaire $\boldsymbol{\omega}$ associée à cette rotation directe est la somme des vitesses angulaires associées aux deux rotations successives :

$$\boldsymbol{\omega} = \boldsymbol{\omega}_1 + \boldsymbol{\omega}_2 \quad (3.32)$$

À cette fin, remarquons que

$$\begin{aligned} \Omega &= \widetilde{R}\dot{R} \\ &= \widetilde{(R_2R_1)}(\dot{R}_2R_1 + R_2\dot{R}_1) \\ &= \widetilde{R}_1\widetilde{R}_2\dot{R}_2R_1 + \widetilde{R}_1\widetilde{R}_2R_2\dot{R}_1 \\ &= \widetilde{R}_1\Omega_2R_1 + \Omega_1 \end{aligned} \quad (3.33)$$

Le deuxième terme représente la vitesse angulaire $\boldsymbol{\omega}_1$. Par contre, Ω_2 représente la vitesse angulaire $\boldsymbol{\omega}_2$, mais exprimée dans le repère S_2 , car c'est dans ce repère que s'applique la rotation R_2 . Dans le repère S_1 , cette matrice aurait précisément la forme $\widetilde{R}_1\Omega_2R_1$, d'après la relation (3.22), ce qui correspond aux composantes de $\boldsymbol{\omega}_2$ dans le repère S_1 , comme établi à la section 3.B.2. Ceci démontre l'équation (3.32).

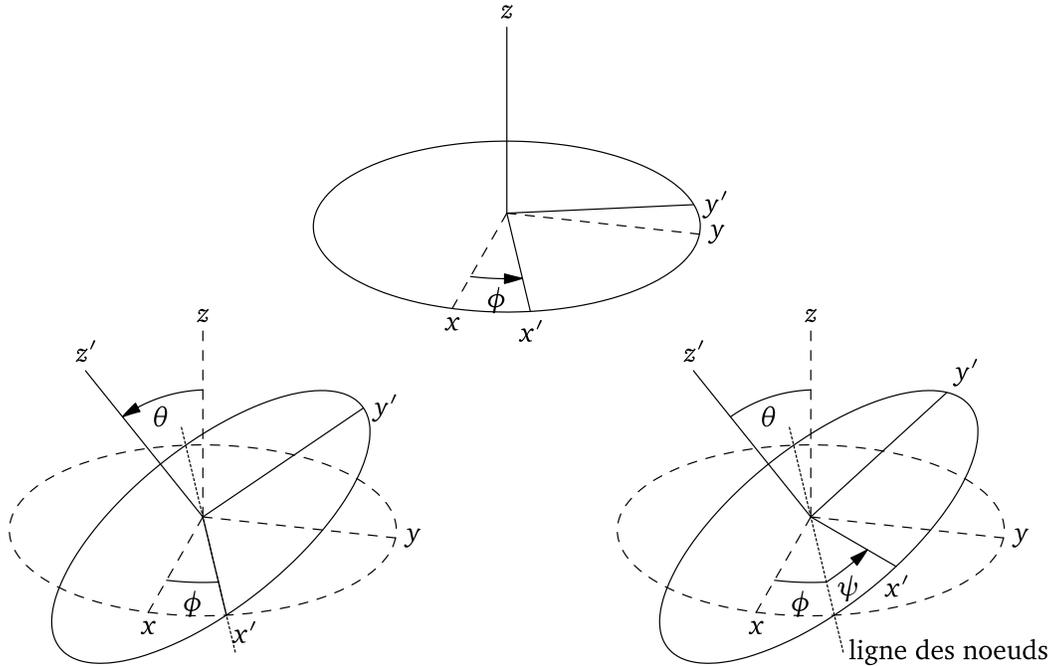


FIGURE 3.2
Définition des angles d'Euler

3.B.4 Angles d'Euler

Afin de représenter le mouvement de rotation en fonction du temps, il est commode et habituel d'utiliser les *angles d'Euler*, définis à la figure 3.2. Ces angles définissent une séquence de trois rotations successives qui nous font passer d'un repère S , fixe dans l'espace, à un repère S' fixe à l'objet. Les trois rotations successives sont :

1. Une rotation $R_1 = R(\mathbf{z}, \phi)$ qui nous amène du repère S au repère S_1 , en rotation de ϕ par rapport à l'axe des z .
2. Une rotation $R_2 = R(\mathbf{x}, \theta)$ d'un angle θ par rapport à l'axe x' de S_1 , qui nous amène vers un repère S_2 .
3. Une rotation $R_3 = R(\mathbf{z}, \psi)$ d'un angle ψ par rapport à l'axe z' de S_2 , qui nous amène vers le repère mobile $S_3 = S'$.

Schématiquement, on écrit :

$$S \xrightarrow{R_1} S_1 \xrightarrow{R_2} S_2 \xrightarrow{R_3} S_3 = S' \quad (3.34)$$

Il faut bien comprendre que chaque matrice de rotation est définie dans le repère sur lequel elle s'applique, c'est-à-dire que R_1 est définie dans le repère S , R_2 dans le repère S_1 et R_3 dans le repère S_2 .

La rotation nette est donc le produit des trois matrices (notez l'ordre des opérations)

$$R(t) = R_3 R_2 R_1 = R(\mathbf{z}, \psi) R(\mathbf{x}, \theta) R(\mathbf{z}, \phi) \quad (3.35)$$

ou, sous forme matricielle,

$$R(t) = \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.36)$$

Trouvons maintenant une expression pour la vitesse angulaire ω de l'objet, en se basant sur les relations (3.32) et (3.33). La vitesse angulaire nette est la somme

$$\omega = \omega_\phi + \omega_\theta + \omega_\psi \quad (3.37)$$

des trois vitesses angulaires associées à chacune des trois rotations. La difficulté est d'exprimer ces trois vitesses dans le même repère. Nous allons en fait les exprimer dans le repère mobile S' , fixe à l'objet. Dans le repère S , on a évidemment $\omega_\phi = (0, 0, \dot{\phi})$. Lors des rotations successives qui mènent à S' , ces composantes se transforment successivement comme

$$(0, 0, \dot{\phi}) \xrightarrow{R_1} (0, 0, \dot{\phi}) \xrightarrow{R_2} (0, \dot{\phi} \sin \theta, \dot{\phi} \cos \theta) \xrightarrow{R_3} (\dot{\phi} \sin \psi \sin \theta, \dot{\phi} \cos \psi \sin \theta, \dot{\phi} \cos \theta) \quad (3.38)$$

Dans le repère S_1 , on a évidemment $\omega_\theta = (\dot{\theta}, 0, 0)$. Lors des rotations successives qui mènent à S' , ces composantes se transforment successivement comme

$$(\dot{\theta}, 0, 0) \xrightarrow{R_2} (\dot{\theta}, 0, 0) \xrightarrow{R_3} (\dot{\theta} \cos \psi, -\dot{\theta} \sin \psi, 0) \quad (3.39)$$

Enfin, dans le repère S' , on a bien évidemment $\omega_\psi = (0, 0, \dot{\psi})$. Donc, au total, les composantes de la vitesse angulaire exprimées en fonction des angles d'Euler sont

$$\omega = \begin{pmatrix} \dot{\phi} \sin \psi \sin \theta + \dot{\theta} \cos \psi \\ \dot{\phi} \cos \psi \sin \theta - \dot{\theta} \sin \psi \\ \dot{\phi} \cos \theta + \dot{\psi} \end{pmatrix} \quad (3.40)$$

C Théorème du moment cinétique

Le théorème du moment cinétique stipule que la dérivée par rapport au temps du moment cinétique d'un objet est donnée par le couple total agissant sur l'objet :

$$\frac{d\mathbf{L}}{dt} = \mathbf{N} \quad (3.41)$$

Le couple \mathbf{N} est défini par

$$\mathbf{N} = \sum_a \mathbf{r}_a \wedge \mathbf{f}_a \quad (3.42)$$

où \mathbf{f}_a est la force appliquée sur la particule a de l'objet. La démonstration est élémentaire :

$$\begin{aligned} \frac{d\mathbf{L}}{dt} &= \frac{d}{dt} \sum_a (\mathbf{r}_a \wedge \mathbf{p}_a) \\ &= \sum_a (\dot{\mathbf{r}}_a \wedge \mathbf{p}_a + \mathbf{r}_a \wedge \dot{\mathbf{p}}_a) \end{aligned} \quad (3.43)$$

Comme la quantité de mouvement \mathbf{p}_a de la particule a est parallèle à sa vitesse $\dot{\mathbf{r}}_a$ et que la dérivée de la quantité de mouvement \mathbf{p}_a est précisément la force \mathbf{f}_a , le théorème s'ensuit.

Par contre, ce théorème a une portée plus grande qu'il ne paraît à première vue. Rappelons à cet effet le théorème de König pour le moment cinétique.

Théorème 3.2 Théorème de König pour le moment cinétique

le moment cinétique d'un objet est la somme du moment cinétique intrinsèque \mathbf{L}_{cm} , calculé par rapport au centre de masse de l'objet, et d'une contribution $\mathbf{r}_{\text{cm}} \wedge \mathbf{p}$, appelée moment cinétique *orbital*, qui serait le moment cinétique du système si toute la masse de celui-ci était concentrée en son centre de masse (\mathbf{p} est la quantité de mouvement totale de l'objet) :

$$\mathbf{L} = \mathbf{L}_{\text{cm}} + \mathbf{r}_{\text{cm}} \wedge \mathbf{p} \quad (3.44)$$

Preuve: Considérons un point \mathbf{r}_0 en mouvement et le référentiel S' attaché à ce point. Ce référentiel peut en général être accéléré, c'est-à-dire non inertielle. La position de la particule a dans S' est $\mathbf{r}'_a = \mathbf{r}_a - \mathbf{r}_0$. Donc la vitesse de la même particule dans S' est $\mathbf{v}'_a = \mathbf{v}_a - \mathbf{v}_0$. Le moment

cinétique évalué à \mathbf{r}_0 dans le référentiel S' est alors

$$\begin{aligned}
 \mathbf{L}' &= \sum_{a=1}^N m_a (\mathbf{r}_a - \mathbf{r}_0) \wedge (\mathbf{v}_a - \mathbf{v}_0) \\
 &= \mathbf{L} - \sum_{a=1}^N m_a \mathbf{r}_0 \wedge \mathbf{v}_a - \sum_a m_a \mathbf{r}_a \wedge \mathbf{v}_0 + \sum_a m_a \mathbf{r}_0 \wedge \mathbf{v}_0 \\
 &= \mathbf{L} - \mathbf{r}_0 \wedge \left(\sum_a m_a \mathbf{v}_a \right) - \left(\sum_a m_a \mathbf{r}_a \right) \wedge \mathbf{v}_0 + m \mathbf{r}_0 \wedge \mathbf{v}_0 \\
 &= \mathbf{L} - \mathbf{r}_0 \wedge \mathbf{p} - m (\mathbf{r}_{\text{cm}} - \mathbf{r}_0) \wedge \mathbf{v}_0
 \end{aligned} \tag{3.45}$$

où \mathbf{p} est l'impulsion totale du système et m sa masse totale. En particulier, si $\mathbf{r}_0 = \mathbf{r}_{\text{cm}}$, on retrouve la relation $\mathbf{L}_{\text{cm}} = \mathbf{L} - \mathbf{r}_{\text{cm}} \wedge \mathbf{p}$, ce qui démontre le théorème.

Le second théorème de König s'applique aussi au couple total :

$$\mathbf{N} = \mathbf{N}_{\text{cm}} + \mathbf{r}_{\text{cm}} \wedge \mathbf{f} \tag{3.46}$$

où \mathbf{f} est la force totale agissant sur le système. La même démonstration que ci-haut peut être reprise, en la modifiant légèrement :

$$\begin{aligned}
 \mathbf{N}' &= \sum_a (\mathbf{r}_a - \mathbf{r}_0) \wedge \mathbf{f}_a \\
 &= \mathbf{N} - \mathbf{r}_0 \wedge \sum_a \mathbf{f}_a \\
 &= \mathbf{N} - \mathbf{r}_0 \wedge \mathbf{f}
 \end{aligned} \tag{3.47}$$

Si $\mathbf{r}_0 = \mathbf{r}_{\text{cm}}$, on retrouve la relation $\mathbf{N}_{\text{cm}} = \mathbf{N} - \mathbf{r}_{\text{cm}} \wedge \mathbf{f}$. Notons ici que la force \mathbf{f}_a est indépendante du référentiel.

Le second théorème de König ne serait pas très utile si le théorème du moment cinétique ne s'appliquait pas séparément au moment cinétique intrinsèque et au couple intrinsèque :

$$\mathbf{N}_{\text{cm}} = \frac{d\mathbf{L}_{\text{cm}}}{dt} \tag{3.48}$$

Cette relation n'est pas immédiatement évidente, car le théorème du moment cinétique découle de la deuxième loi de Newton ($\mathbf{f} = m\mathbf{a}$) et celle-ci n'est strictement valable que dans un référentiel d'inertie, alors que la relation (3.48) reste valable même si le centre de masse de l'objet est accéléré. Pour démontrer la relation (3.48), substituons les relations (3.44) et (3.46) dans le théorème du moment cinétique (3.41) :

$$\frac{d\mathbf{L}_{\text{cm}}}{dt} + \frac{d}{dt} (\mathbf{r}_{\text{cm}} \wedge \mathbf{p}) = \mathbf{N}_{\text{cm}} + \mathbf{r}_{\text{cm}} \wedge \mathbf{f} \tag{3.49}$$

Or,

$$\begin{aligned}
 \frac{d}{dt} (\mathbf{r}_{\text{cm}} \wedge \mathbf{p}) &= \frac{d\mathbf{r}_{\text{cm}}}{dt} \wedge \mathbf{p} + \mathbf{r}_{\text{cm}} \wedge \frac{d\mathbf{p}}{dt} \\
 &= \mathbf{v}_{\text{cm}} \wedge \mathbf{p} + \mathbf{r}_{\text{cm}} \wedge \mathbf{f}
 \end{aligned} \tag{3.50}$$

Le premier terme s'annule, car $\mathbf{p} = m\mathbf{v}_{\text{cm}}$. On peut donc affirmer, par différence avec l'équation précédente, que $\mathbf{N}_{\text{cm}} = d\mathbf{L}_{\text{cm}}/dt$, c'est-à-dire que le théorème du moment cinétique s'applique aux parties orbitales et intrinsèques séparément.

D Tenseur d'inertie

3.D.1 Relation entre vitesse angulaire et moment cinétique

Passons maintenant à une description de la dynamique des corps rigides. Désignons par \mathbf{r}_a les positions des différentes particules formant l'objet rigide, et par m_a leurs masses. Nous avons vu à l'éq. (3.31) que la vitesse d'une particule liée à l'objet est donnée par $\mathbf{v}_a = \boldsymbol{\omega} \wedge \mathbf{r}_a$. Cette équation est valable dans tout repère, parce qu'elle est exprimée en fonction de vecteurs et non de composantes particulières.

Le moment cinétique d'une particule étant défini comme $\mathbf{L} = \mathbf{r} \wedge \mathbf{p} = m\mathbf{r} \wedge \mathbf{v}$, le moment cinétique total de l'objet a l'expression suivante :

$$\mathbf{L} = \sum_a m_a \mathbf{r}_a \wedge \mathbf{v}_a = \sum_a m_a \mathbf{r}_a \wedge (\boldsymbol{\omega} \wedge \mathbf{r}_a) = \sum_a m_a [\boldsymbol{\omega} r_a^2 - \mathbf{r}_a (\boldsymbol{\omega} \cdot \mathbf{r}_a)] \quad (3.51)$$

ou, en fonction des composantes,

$$L_i = \sum_a m_a [\omega_i r_a^2 - \omega_j x_{ai} x_{aj}] = I_{ij} \omega_j \quad (3.52)$$

où la matrice \mathbf{I} est définie par ses composantes :

$$I_{ij} = \sum_a m_a [r_a^2 \delta_{ij} - x_{ai} x_{aj}] \quad (3.53)$$

ou, explicitement,

$$\mathbf{I} = \sum_a m_a \begin{pmatrix} y_a^2 + z_a^2 & -x_a y_a & -x_a z_a \\ -x_a y_a & x_a^2 + z_a^2 & -y_a z_a \\ -x_a z_a & -y_a z_a & x_a^2 + y_a^2 \end{pmatrix} \quad (3.54)$$

Cette matrice se calcule dans un repère d'axes fixes à l'objet, sinon elle change en fonction du temps ! On l'appelle la *matrice d'inertie* de l'objet. Plus précisément, il s'agit d'un *tenseur*, dont les composantes prennent la forme d'une matrice. En notation matricielle, la relation entre vitesse angulaire et moment cinétique peut s'écrire de manière concise comme

$$\mathbf{L} = \mathbf{I} \boldsymbol{\omega} \quad (3.55)$$

L'énergie cinétique associée à la rotation de l'objet est, quant à elle,

$$\begin{aligned} T &= \frac{1}{2} \sum_a m_a \mathbf{v}_a^2 = \frac{1}{2} \sum_a m_a \mathbf{v}_a \cdot (\boldsymbol{\omega} \wedge \mathbf{r}_a) = \frac{1}{2} \sum_a m_a \boldsymbol{\omega} \cdot (\mathbf{r}_a \wedge \mathbf{v}_a) \\ &= \frac{1}{2} \boldsymbol{\omega} \cdot \sum_a m_a (\mathbf{r}_a \wedge \mathbf{v}_a) = \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{L} = \frac{1}{2} \tilde{\boldsymbol{\omega}} \mathbf{I} \boldsymbol{\omega} \end{aligned} \quad (3.56)$$

3.D.2 Tenseurs et axes principaux

Un ensemble de trois quantités sous forme de vecteur colonne \times forment un *vecteur* dans l'espace si, lors du passage $S \rightarrow S'$ entre deux repères cartésiens, elles se transforment comme indiqué en (7.30) ou, plus explicitement,

$$x'_i = R_{ij}x_j \quad (3.57)$$

L'idée essentielle est que les composantes du vecteur changent d'un repère à l'autre, mais le vecteur géométrique reste le même et existe indépendamment des repères utilisés pour en décrire les composantes.

Un *tenseur de rang 2* est une quantité géométrique qui est représentée, dans un repère donné, par une matrice T . Les composantes de cette matrice changent d'un repère à l'autre en suivant la règle :

$$T'_{ij} = R_{ik}R_{jl}T_{kl} \quad \text{ou} \quad T' = RT\tilde{R} \quad (3.58)$$

Encore une fois, l'objet physique qu'on appelle tenseur est indépendant du repère, mais ses composantes ne le sont pas. Notons que la matrice antisymétrique Ω représentant la vitesse angulaire se transforme précisément de cette manière, selon l'éq. (3.23). On peut donc affirmer que Ω constitue un tenseur antisymétrique. Il se trouve qu'en trois dimensions, tout tenseur de rang 2 antisymétrique possède trois composantes indépendantes et peut être représenté par un vecteur, comme le vecteur ω représente le tenseur Ω .

De même, un *tenseur de rang 3* est un objet dont les composantes sont affublées de trois indices : T_{ijk} , et ces composantes se transforment de la manière suivante :

$$T'_{ijk} = R_{il}R_{jm}R_{kn}T_{lmn} \quad (3.59)$$

Dans ce cas, on n'utilise pas de notation particulière, comme la notation matricielle, pour désigner l'ensemble des composantes. Un vecteur n'est rien d'autre qu'un tenseur de rang 1.

La notion de tenseur est très importante en géométrie riemannienne, et donc en relativité générale. Nous l'avons définie ici de la manière la plus simple possible, mais il faut généralement utiliser des repères qui ne sont pas orthonormés et qui mènent à la distinction entre indices covariants et indices contravariants. Nous ne ferons pas cette distinction ici.

Revenons au tenseur d'inertie. Par construction, la matrice qui le représente est toujours symétrique : $\tilde{T} = T$. Cette propriété du tenseur T est valable dans tous les repères, car

$$\tilde{T}' = \widetilde{RT\tilde{R}} = R\tilde{T}\tilde{R} = RT\tilde{R} = T' \quad (3.60)$$

Puisqu'elle est symétrique, la matrice T peut être diagonalisée à l'aide d'une matrice orthogonale, c'est-à-dire une matrice de rotation. Autrement dit, il existe une matrice de rotation R telle que, dans le repère S' , la matrice $T' = RT\tilde{R}$ est diagonale. Il est avantageux d'adopter non seulement des axes cartésiens fixes à l'objet, mais des axes en fonction desquels la matrice d'inertie est diagonale :

$$I = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \quad (3.61)$$

On appelle ces axes les *axes principaux* de l'objet. Chacune des trois composantes diagonales est le *moment d'inertie* selon l'axe correspondant :

$$I_1 = \sum_a m_a(x_2^2 + x_3^2) \quad I_2 = \sum_a m_a(x_1^2 + x_3^2) \quad I_3 = \sum_a m_a(x_1^2 + x_2^2) \quad (3.62)$$

L'énergie cinétique prend alors la forme simple

$$T = \frac{1}{2}I_1\omega_1^2 + \frac{1}{2}I_2\omega_2^2 + \frac{1}{2}I_3\omega_3^2 \quad (3.63)$$

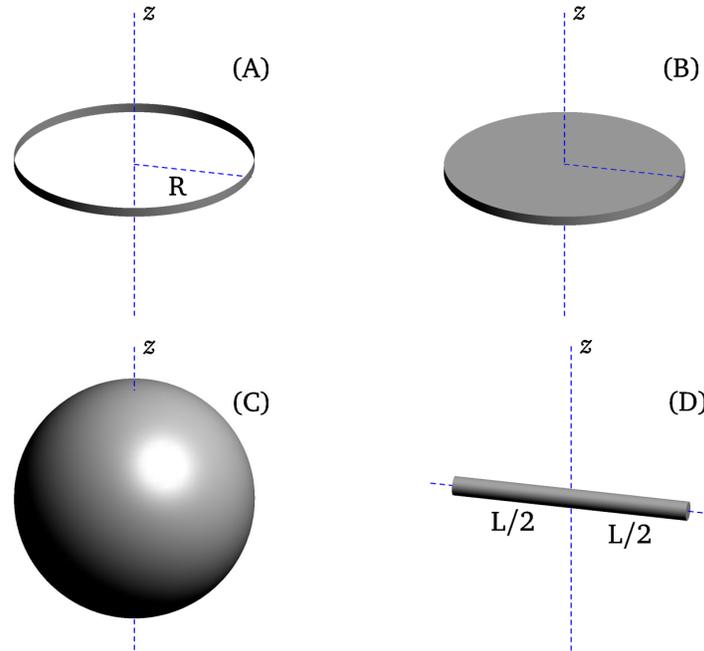


FIGURE 3.3

Exemple 3.1 Moment d'inertie d'un anneau

Considérons un anneau de rayon R et de masse M (fig. 3.3A). Calculons le moment d'inertie I par rapport à l'axe de l'anneau. Comme tous les points de l'anneau sont situés à égale distance de l'axe de rotation (on néglige ici la largeur de l'anneau), la définition (3.62) donne simplement

$$I = \sum_i m_i \rho_i^2 = R^2 \sum_i m_i = MR^2 \quad (3.64)$$

La même formule s'applique à un cylindre creux d'épaisseur négligeable par rapport à son rayon.

Exemple 3.2 Moment d'inertie d'un disque ou d'un cylindre

Considérons un disque plein de rayon R et de masse M (fig. 3.3B). On peut diviser ce disque en une série d'anneaux concentriques de rayon r (r va de 0 à R). Chacun de ces anneaux a une

largeur dr et possède une masse $2\pi\sigma r dr$, où σ est la densité superficielle (constante) du disque. Le moment d'inertie de l'anneau de rayon r est donc $dI = 2\pi\sigma r^3 dr$ et le moment d'inertie total est

$$I = \int_0^R dr 2\pi\sigma r^3 = \frac{1}{2}\pi\sigma R^4 = \frac{1}{2}MR^2 \quad (M = \pi R^2\sigma) \quad (3.65)$$

Remarques:

- ◆ Cette expression vaut aussi pour un cylindre de hauteur h . En fait, la distribution de la masse en fonction de la coordonnée z parallèle à l'axe de calcul peut être quelconque, car cette coordonnée n'intervient pas dans le calcul du moment d'inertie.
- ◆ Le principe de superposition permet de calculer le moment d'inertie d'un objet pouvant être considéré comme la 'différence' de deux objets. Par exemple, un cylindre creux de rayon intérieur a et de rayon extérieur b peut être considéré comme un cylindre plein de rayon b duquel on a soustrait un cylindre plein de rayon $a < b$. Le moment d'inertie d'un tel objet est alors

$$I = \frac{1}{2}\pi\sigma(b^4 - a^4) = \frac{1}{2}\pi\sigma(b^2 - a^2)(b^2 + a^2) = \frac{1}{2}M(b^2 + a^2) \quad (3.66)$$

Exemple 3.3 Moment d'inertie d'une sphère

Comme troisième exemple, considérons une sphère pleine de rayon R et de masse M (fig. 3.3C). On peut diviser cette sphère en une série de disques superposés. Un disque situé à une hauteur z par rapport au centre (z va de $-R$ à R) possède une épaisseur dz , un rayon $r = \sqrt{R^2 - z^2}$ et une masse $dM = \mu\pi r^2 dz$ où μ est la densité volumique (constante) de la sphère. Le moment d'inertie de chaque disque est

$$dI = \frac{1}{2}dMr^2 = \frac{1}{2}\mu\pi r^4 dz = \frac{1}{2}\mu\pi(R^2 - z^2)^2 dz \quad (3.67)$$

et le moment d'inertie total est obtenu en intégrant sur z de $-R$ à R :

$$I = \frac{1}{2}\mu\pi \int_{-R}^R dz (R^2 - z^2)^2 = \mu\pi \frac{8}{15}R^5 = \frac{2}{5}MR^2 \quad (M = \frac{4}{3}\pi R^3\mu) \quad (3.68)$$

Exemple 3.4 Moment d'inertie d'une tige

Comme dernier exemple, considérons une tige de longueur L et de masse M (fig. 3.3D). On subdivise cette tige en parcelles de longueur dx et de masse λdx , λ étant la densité linéaire (ou masse par unité de longueur) de la tige. Chaque parcelle possède un moment d'inertie $dI = \lambda x^2 dx$, où x est la distance par rapport à l'axe. Le moment d'inertie total est

$$I = \int_{-L/2}^{L/2} dx \lambda x^2 = \frac{1}{12}\lambda L^3 = \frac{1}{12}ML^2 \quad (M = \lambda L) \quad (3.69)$$

E Mouvement libre d'un objet rigide

3.E.1 Équations d'Euler

Considérons un objet rigide en rotation libre, c'est-à-dire en l'absence de couple appliqué. Plus précisément, supposons que le couple appliqué au centre de masse de l'objet, \mathbf{N}_{cm} , est nul. Nous nous placerons dans le référentiel du centre de masse de l'objet. Un objet qu'on lance dans les airs est un exemple de cette situation, si on néglige le couple produit par la résistance de l'air. Dans ce cas, le théorème du moment cinétique (3.48) stipule que le moment cinétique \mathbf{L} de l'objet est constant²

Même si le moment cinétique \mathbf{L} de l'objet est constant, les composantes de \mathbf{L} dans le repère S' dont les axes sont liés à l'objet ne sont pas constantes. D'après la relation (3.29), elles obéissent plutôt à la relation suivante :

$$\frac{d\mathbf{L}}{dt} = \left(\frac{d\mathbf{L}}{dt} \right)_{\mathbf{R}} + \boldsymbol{\omega} \wedge \mathbf{L} = 0 \quad (3.70)$$

Dans le système des axes principaux, la relation entre le moment cinétique et la vitesse angulaire est

$$\mathbf{L} = I_1 \omega_1 \mathbf{e}_1 + I_2 \omega_2 \mathbf{e}_2 + I_3 \omega_3 \mathbf{e}_3 \quad (3.71)$$

Les moments d'inertie I_a étant constants, la relation (3.70) équivaut au système d'équations suivant, appelé *équations d'Euler* :

$$\begin{cases} I_1 \dot{\omega}_1 + (I_3 - I_2) \omega_2 \omega_3 = 0 \\ I_2 \dot{\omega}_2 + (I_1 - I_3) \omega_3 \omega_1 = 0 \\ I_3 \dot{\omega}_3 + (I_2 - I_1) \omega_1 \omega_2 = 0 \end{cases} \quad (3.72)$$

3.E.2 Rotation libre d'un objet symétrique

Supposons maintenant que deux des moments d'inertie de l'objet soient identiques, disons $I_1 = I_2 = I$. C'est le cas si l'objet a une symétrie de rotation par rapport à l'axe \mathbf{e}_3 , mais cette condition n'est pas nécessaire ; par exemple, un parallélepède à base carrée respecte cette condition. On dira cependant qu'un tel objet rigide est *symétrique*. Les équations d'Euler se simplifient ainsi pour un objet symétrique :

$$\begin{cases} \dot{\omega}_1 + \gamma \omega_2 \omega_3 = 0 \\ \dot{\omega}_2 - \gamma \omega_3 \omega_1 = 0 \\ \dot{\omega}_3 = 0 \end{cases} \quad \text{où } \gamma = \frac{I_3 - I}{I} \quad (3.73)$$

Donc, dans ce cas, ω_3 est constant. D'autre part, en différentiant la première équation et en substituant dans la deuxième, on trouve

$$\ddot{\omega}_1 + (\gamma \omega_3)^2 \omega_1 = 0 \quad (3.74)$$

2. Nous ne ferons plus la distinction entre \mathbf{L} et \mathbf{L}_{cm} dans cette section, car il est entendu que nous plaçons l'origine au centre de masse de l'objet.

La solution de cette équation nous est bien connue :

$$\omega_1(t) = A \cos(\Gamma t + \phi) \quad \Gamma = \gamma \omega_3 \quad (3.75)$$

où A et ϕ sont des constantes déterminées par les conditions initiales. De même, ω_2 oscille avec la même fréquence et s'obtient directement de la première des équations (3.73) :

$$\omega_2(t) = A \sin(\Gamma t + \phi) \quad (3.76)$$

Le vecteur $\omega_{\perp} = \omega_1 \mathbf{e}_1 + \omega_2 \mathbf{e}_2$ tourne donc dans le temps dans le repère lié à l'objet, à une fréquence $\Gamma = \gamma \omega_3$. La grandeur ω_{\perp} de cette projection est fixe dans ce référentiel, et donc dans le référentiel inertiel également (la grandeur d'un vecteur ne change pas d'un référentiel inertiel à un référentiel en rotation). D'autre part, comme ω_3 est constant, la grandeur du vecteur ω est fixe dans les deux référentiels, mais ce vecteur précesse autour de l'axe \mathbf{e}_3 dans le référentiel fixe à l'objet. Ce mouvement de rotation est appelé *nutaton*.

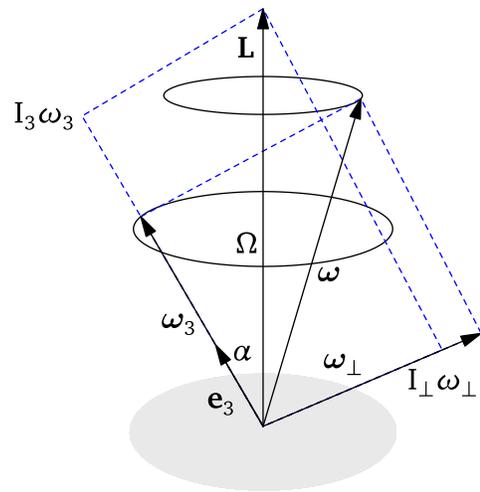


FIGURE 3.4
Orientation relative des vecteurs \mathbf{L} et ω dans le problème de la nutation d'un objet libre.

Considérons maintenant la figure 3.4. Le vecteur \mathbf{L} est fixe dans l'espace (disons, dans la direction z). Nous venons de voir aussi que le vecteur ω tourne autour de l'axe \mathbf{e}_3 ; la fréquence de cette rotation, vue du référentiel lié à l'objet, est Γ . Comme ω_3 est constant dans le temps, $I_3 \omega_3 = L_3$ l'est aussi : la composante de \mathbf{L} le long de l'axe \mathbf{e}_3 est constante dans le temps et donc l'angle α entre \mathbf{e}_3 et \mathbf{L} est constant. Enfin, la projection de ω sur \mathbf{L} est elle aussi constante dans le temps, car l'énergie cinétique de rotation, qui est constante, est $\frac{1}{2} \mathbf{L} \cdot \omega$; les grandeurs de \mathbf{L} et de ω étant constantes, l'angle entre \mathbf{L} et ω l'est donc aussi. Donc les trois vecteurs \mathbf{e}_3 , ω et \mathbf{L} sont rigides les uns par rapport aux autres. Appelons S_3 le référentiel dans lequel ces trois vecteurs sont fixes. Ce référentiel, donc l'axe \mathbf{e}_3 et le vecteur ω , tournent autour de \mathbf{L} à une fréquence Ω , la véritable fréquence de nutation observée, qui n'est pas identique à Γ , comme nous allons le voir.

Le référentiel S_3 est en rotation par rapport aux axes liés à l'objet. La vitesse angulaire de cette rotation est précisément $\Gamma \mathbf{e}_3$: c'est la vitesse angulaire de ω par rapport aux axes fixes, tel que démontré par les équations d'Euler. Or, la rotation de S_3 par rapport à l'espace – de vitesse angulaire $\Omega \mathbf{z}$ – est la *composition* de la rotation de S_3 par rapport à l'objet – de vitesse angulaire $\Gamma \mathbf{e}_3$ – et de la rotation de l'objet par rapport à l'espace – de vitesse angulaire ω . Autrement dit, la rotation de S_3 par rapport à l'espace s'obtient en ajoutant deux mouvements de rotation – et donc deux vitesses angulaires, comme suit :

$$\Omega \mathbf{z} = \Gamma \mathbf{e}_3 + \omega \quad (3.77)$$

En projetant cette relation sur le vecteur \mathbf{e}_3 , on trouve

$$\Omega \cos \alpha = \Gamma + \omega_3 = \frac{I_3}{I} \omega_3 \quad (3.78)$$

La fréquence de nutation est donc

$$\Omega = \frac{I_3}{I} \frac{\omega_3}{\cos \alpha} \quad (3.79)$$

Par exemple, dans le cas d'un disque plat de rayon a , on montre que $I = \frac{1}{4}ma^2$ et on sait que $I_3 = \frac{1}{2}ma^2$. Si la nutation est de petite amplitude (α petit), alors $\cos \alpha \approx 1$ et la fréquence de nutation observée est environ deux fois la vitesse angulaire de rotation ω_3 de l'objet. Si une flèche était dessinée sur le disque de son centre vers le bord, alors deux périodes de nutation se produiraient dans le temps que la flèche effectue une rotation complète.

Nutation de la Terre

Il est tentant d'appliquer ce calcul à la rotation de la Terre. La Terre étant légèrement aplatie mais symétrique par rapport à son axe, on a $I_1 = I_2 < I_3$. On évalue le rapport γ à

$$\gamma = \frac{I_3 - I}{I} \approx 0,00327 \quad (3.80)$$

En théorie, la période de nutation $2\pi\Gamma^{-1}$ de la Terre (par rapport aux axes fixes à la Terre) devrait être de 306 jours. Cependant, la réalité est plus compliquée, car la Terre n'est pas un objet parfaitement rigide. Un mouvement de nutation de faible amplitude (environ 15 m d'amplitude au pôle nord) est observé, mais il n'est pas régulier, c'est-à-dire qu'il est la superposition de différents mouvements de fréquences différentes. Il comporte une composante annuelle attribuée aux saisons (reliée au mouvement des masses d'air, etc.) et une autre composante importante d'une période de 420 jours, qui correspond plus à la nutation prédite. Cette période plus longue que prévu (420 jours au lieu de 306) peut être attribuée à l'élasticité interne de la Terre.

3.E.3 Rotation libre d'un objet asymétrique

Supposons maintenant que les trois moments d'inertie de l'objet sont différents : $I_1 > I_2 > I_3$. Nous pouvons qualitativement comprendre le mouvement de l'objet sans avoir recours aux équations d'Euler. En fonction des composantes (L_1, L_2, L_3) du moment cinétique dans le repère des axes principaux, la conservation de l'énergie cinétique T définit une surface, un ellipsoïde appelé *ellipsoïde de Binet*, spécifié par l'équation suivante :

$$\frac{L_1^2}{2I_1} + \frac{L_2^2}{2I_2} + \frac{L_3^2}{2I_3} = T \quad (3.81)$$

Les longueurs des demi-axes de cet ellipsoïde sont respectivement données par $\sqrt{2I_1T}$, $\sqrt{2I_2T}$ et $\sqrt{2I_3T}$. D'autre part, la conservation du moment cinétique dans l'espace se traduit par un vecteur (L_1, L_2, L_3) de longueur constante, dont l'extrémité se déplace sur la surface d'une sphère de rayon $|\mathbf{L}|$ dans le repère fixe à l'objet. L'équation de cette sphère est

$$L_1^2 + L_2^2 + L_3^2 = |\mathbf{L}|^2 \quad (3.82)$$

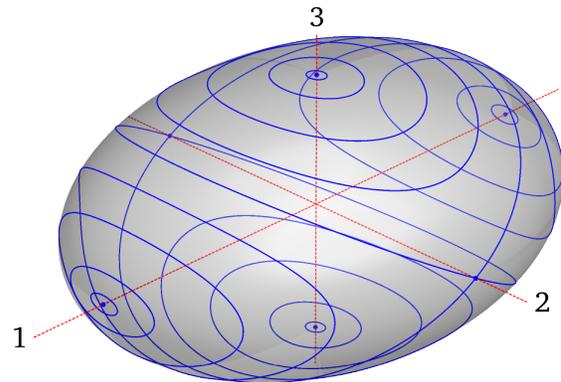
Au cours du temps, le vecteur (L_1, L_2, L_3) se déplace donc sur l'intersection d'une sphère et d'un ellipsoïde.

Cette vision géométrique aide à comprendre pourquoi la rotation d'un objet par rapport à un axe principal est stable dans le cas du moment d'inertie minimum (I_3) et du moment d'inertie maximum (I_1), mais instable dans le cas du moment d'inertie moyen (I_2). Ceci est illustré à la Fig. 3.5. Si la rotation est initiée exactement par rapport à l'axe 1, alors le moment cinétique demeure parallèle à la vitesse angulaire et $L_2 = L_3 = 0$ en tout temps. Ce mouvement correspond au point indiqué à l'intersection de l'axe 1 et de l'ellipsoïde de Binet. Si la rotation dévie légèrement de l'axe 1, alors les composantes L_2 et L_3 resteront petites au cours du temps; ce mouvement correspond à une courbe fermée qui entoure la même intersection. La même chose se produit si la rotation est initiée par rapport à l'axe 3. La rotation par rapport à ces axes est donc *stable* : une petite déviation par rapport à l'axe reste petite au cours du temps.

Par contre, un phénomène intéressant se produit si la rotation est initiée par rapport à l'axe 2, c'est-à-dire l'axe moyen, dont le moment d'inertie I_2 a une valeur intermédiaire entre I_1 et I_3 . Dans ce cas, la rotation exacte par rapport à l'axe 2, même si elle est mathématiquement possible, est une solution instable des équations d'Euler : la moindre variation par rapport à cet axe fait suivre au vecteur (L_1, L_2, L_3) une grande courbe fermée qui croise l'intersection de l'axe 2 avec l'ellipsoïde. En pratique, cette instabilité est facilement visible si on essaie de mettre un objet possédant trois axes clairement différents (par exemple un livre dont les couverts ont été préalablement attachés) en rotation par rapport à son axe moyen.

FIGURE 3.5

L'ellipsoïde de Binet et le mouvement du moment cinétique dans le repère des axes fixes à l'objet. Les trois axes principaux (1, 2 et 3) sont indiqués. Les courbes bleues sont les intersections de l'ellipsoïde avec des sphères de rayons différents.



F Mouvement d'une toupie symétrique

3.F.1 Lagrangien, hamiltonien et problème effectif à une variable

Considérons maintenant le problème d'une toupie symétrique avec un point fixe, tel qu'illustré à la figure 3.6 : le pivot de la toupie est à l'origine, et son centre de masse est situé à une distance h du pivot, le long de son axe de symétrie.

Décrivons le mouvement de la toupie en spécifiant les angles d'Euler requis pour amener le repère fixe dans l'espace (axe z vertical) vers les axes principaux de la toupie à un moment donné. Sur la figure 3.2, les axes primés (x', y', z') sont les axes principaux de l'objet. Donc les dérivées temporelles des angles d'Euler ont l'interprétation suivante :

1. $\dot{\psi}$: rotation de la toupie par rapport à son axe de symétrie.
2. $\dot{\phi}$: précession de la toupie par rapport à la verticale
3. $\dot{\theta}$: nutation de l'axe de la toupie du haut vers le bas par rapport à la verticale.

Cependant, d'après la relation (3.40), la composante de la vitesse angulaire le long de l'axe de symétrie de la toupie est

$$\omega_3 = \dot{\phi} \cos \theta + \dot{\psi} \quad (3.83)$$

et non simplement $\dot{\psi}$. Ceci s'explique par la contribution du mouvement de précession, donc de $\dot{\phi}$, à la rotation par rapport à l'axe de symétrie. En particulier, lorsque $\theta = 0$, il est clair que $\omega_3 = \dot{\phi} + \dot{\psi}$ ³

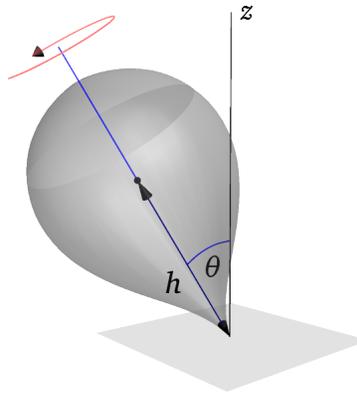


FIGURE 3.6

La toupie est dans un champ gravitationnel constant, et son énergie potentielle est $V = mgh \cos \theta$ ($h \cos \theta$ est donc la hauteur z du centre de masse par rapport au sol).

D'après les équations (3.63) et (3.40), la fonction de Lagrange correspondante est

$$L = \frac{1}{2}I \{(\dot{\phi} \sin \psi \sin \theta + \dot{\theta} \cos \psi)^2 + (\dot{\phi} \cos \psi \sin \theta - \dot{\theta} \sin \psi)^2\} + \frac{1}{2}I_3(\dot{\phi} \cos \theta + \dot{\psi})^2 - mgh \cos \theta \quad (3.84)$$

3. Il y a d'ailleurs, pour cette valeur de θ , ambiguïté dans la définition des angles d'Euler, c'est-à-dire que plusieurs combinaisons de ϕ et ψ correspondent à la même rotation.

En simplifiant, ceci se réduit à

$$L = \frac{1}{2}I(\dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2) + \frac{1}{2}I_3(\dot{\phi} \cos \theta + \dot{\psi})^2 - mgh \cos \theta \quad (3.85)$$

Équations du mouvement

Les moments conjugués aux trois angles d'Euler sont les suivants :

$$p_\psi = \frac{\partial L}{\partial \dot{\psi}} = I_3(\dot{\phi} \cos \theta + \dot{\psi}) = I_3 \omega_3 \quad (3.86)$$

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = I \sin^2 \theta \dot{\phi} + I_3 \cos \theta (\dot{\phi} \cos \theta + \dot{\psi}) = I \sin^2 \theta \dot{\phi} + p_\psi \cos \theta \quad (3.87)$$

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = I \dot{\theta} \quad (3.88)$$

Comme le lagrangien ne dépend explicitement ni de ϕ , ni de ψ , les moments conjugués p_ϕ et p_ψ sont conservés :

$$\dot{p}_\phi = \frac{\partial L}{\partial \phi} = 0 \quad \dot{p}_\psi = \frac{\partial L}{\partial \psi} = 0 \quad (3.89)$$

C'est là tout l'avantage d'utiliser les angles d'Euler comme coordonnées généralisées. En particulier, le taux instantané de précession $\dot{\phi}$ s'exprime en fonction de ces moments conjugués et de θ :

$$\dot{\phi} = \frac{p_\phi - p_\psi \cos \theta}{I \sin^2 \theta} \quad (3.90)$$

L'équation de Lagrange pour θ , la seule qui soit non triviale, est

$$\dot{p}_\theta = \frac{\partial L}{\partial \theta} \quad (3.91)$$

ou, en développant explicitement,

$$I\ddot{\theta} = I \sin \theta \cos \theta \dot{\phi}^2 - p_\psi \dot{\phi} \sin \theta + mgh \sin \theta \quad (3.92)$$

Lorsqu'on l'exprime en fonction de p_ϕ au lieu de $\dot{\phi}$, cette équation devient

$$I\ddot{\theta} = \frac{(p_\phi - p_\psi \cos \theta)(p_\phi \cos \theta - p_\psi)}{I \sin^3 \theta} + mgh \sin \theta \quad (3.93)$$

Hamiltonien

Construisons maintenant le hamiltonien, d'après le lagrangien (3.85) et les moments conjugués (3.86-3.88) :

$$\begin{aligned} H &= p_\psi \dot{\psi} + p_\phi \dot{\phi} + p_\theta \dot{\theta} - L \\ &= \frac{p_\psi^2}{2I_3} + \frac{p_\theta^2}{2I} + \frac{(p_\phi - p_\psi \cos \theta)^2}{2I \sin^2 \theta} + mgh \cos \theta \end{aligned} \quad (3.94)$$

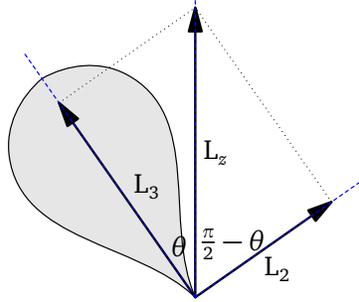


FIGURE 3.7

On peut aussi comprendre cette expression directement, sachant que l'énergie cinétique en fonction des composantes du moment cinétique est

$$T = \frac{L_3^2}{2I_3} + \frac{L_1^2}{2I} + \frac{L_2^2}{2I} \quad (3.95)$$

En posant que l'axe \mathbf{e}_1 est dans la direction de la ligne des noeuds, l'axe \mathbf{e}_2 est alors perpendiculaire à la fois à l'axe de symétrie \mathbf{e}_3 et à \mathbf{e}_1 . D'après la définition des angles d'Euler, les moments conjugués $(p_\psi, p_\phi, p_\theta)$ sont respectivement les composantes de \mathbf{L} le long de l'axe \mathbf{e}_3 , \mathbf{z} et \mathbf{e}_1 (i.e. la ligne des noeuds). Les deux premiers termes de (3.94) sont donc évidents. Par contre, comme l'axe \mathbf{z} fait un angle θ par rapport à \mathbf{e}_3 et $\pi/2 - \theta$ par rapport à \mathbf{e}_2 (voir fig. 3.7), on a

$$p_\phi = L_z = L_3 \cos \theta + L_2 \sin \theta = p_\psi \cos \theta + L_2 \sin \theta \implies L_2 = \frac{p_\phi - p_\psi \cos \theta}{\sin \theta} \quad (3.96)$$

ce qui mène directement au troisième terme de (3.94).

On constate que, le Hamiltonien ne dépendant ni de ψ , ni de ϕ , les moments p_ψ et p_ϕ peuvent être traités comme des constantes. Autrement dit, le problème se réduit à une seule coordonnée, θ , évoluant dans une énergie potentielle effective donnée par

$$V_{\text{eff}}(\theta) = \frac{(p_\phi - p_\psi \cos \theta)^2}{2I \sin^2 \theta} + mgh \cos \theta \quad (3.97)$$

(le premier terme de (3.94) étant, de ce point de vue, une constante). La conservation de l'énergie s'écrit alors comme $E = T + V$, ou encore

$$E = \frac{p_\psi^2}{2I_3} + \frac{I\dot{\theta}^2}{2} + \frac{(p_\phi - p_\psi \cos \theta)^2}{2I \sin^2 \theta} + mgh \cos \theta \quad (3.98)$$

3.F.2 Précession uniforme

Considérons une solution particulière du mouvement de la toupie, dans laquelle l'angle θ est constant. Dans ce cas, il est clair, d'après l'éq. (3.90), que $\dot{\phi}$ est constant lui aussi, donc que la précession est uniforme. L'équation du mouvement (3.92), appliquée au cas θ constant, équivaut

à la condition que le système se situe au minimum du potentiel effectif. Décrivons ce potentiel en fonction de la variable $u = \cos \theta$:

$$V_{\text{eff}}(u) = \frac{(p_\phi - p_\psi u)^2}{2I(1-u^2)} + mghu \quad (3.99)$$

La condition de minimisation de ce potentiel est

$$\begin{aligned} V'_{\text{eff}}(u) &= mgh - p_\psi \frac{(p_\phi - p_\psi u)}{I(1-u^2)} + u \frac{(p_\phi - p_\psi u)^2}{I(1-u^2)^2} \\ &= mgh + \frac{(p_\phi - p_\psi u)}{I(1-u^2)^2} (-(1-u^2)p_\psi + u(p_\phi - p_\psi u)) \\ &= mgh + \frac{(p_\phi - p_\psi u)(up_\phi - p_\psi)}{I(1-u^2)^2} = 0 \end{aligned} \quad (3.100)$$

Exprimons cette condition en fonction de $\dot{\phi}$ et non de p_ϕ , en utilisant l'expression (3.90) :

$$\dot{\phi} = \frac{p_\phi - p_\psi u}{I(1-u^2)} \implies p_\phi = I\dot{\phi}(1-u^2) + up_\psi \quad (3.101)$$

ce qui entraîne

$$\begin{aligned} V'_{\text{eff}}(u) &= mgh + \frac{\dot{\phi}(up_\phi - p_\psi)}{(1-u^2)} \\ &= mgh + \frac{\dot{\phi}(-(1-u^2)p_\psi + uI\dot{\phi}(1-u^2))}{(1-u^2)} \\ &= mgh + \dot{\phi}(-p_\psi + uI\dot{\phi}) \\ &= mgh - \dot{\phi}p_\psi + I\dot{\phi}^2 \cos \theta = 0 \end{aligned} \quad (3.102)$$

(nous avons exclu la possibilité $\theta = 0$, que nous traiterons à la section 3.F4). Il s'agit d'une équation quadratique en $\dot{\phi}$, dont la solution est

$$\dot{\phi} = \frac{p_\psi}{2I \cos \theta} \left\{ 1 - \sqrt{1 - \frac{4Imgh \cos \theta}{p_\psi^2}} \right\} \quad (3.103)$$

Le signe négatif de la racine a été choisi, de sorte que le taux de précession s'annule dans la limite $g \rightarrow 0$ (absence de couple).

On note que cette solution est impossible si le discriminant est négatif. Il en résulte une condition élémentaire pour l'existence d'une précession uniforme :

$$4Imgh \cos \theta < p_\psi^2 = I_3^2 \omega_3^2 \quad (3.104)$$

Autrement dit, il faut que la rotation de la toupie autour de son axe de symétrie soit suffisamment rapide.

Le régime d'une *toupie rapide* est défini par l'approximation

$$\frac{1}{2}I_3 \omega_3^2 \gg 2mgh \frac{I}{I_3} \quad (\text{toupie rapide}) \quad (3.105)$$

et décrit la situation où l'énergie cinétique de rotation autour de l'axe de symétrie est beaucoup plus grande que la variation possible de l'énergie potentielle. L'expression de la fréquence de précession se simplifie alors, en utilisant le développement du binôme sur l'éq. (3.103) :

$$\dot{\phi} \approx \frac{I_3 \omega_3}{2I \cos \theta} \left\{ 1 - \left(1 - \frac{2I m g h \cos \theta}{I_3^2 \omega_3^2} \right) \right\} = \frac{m g h}{I_3 \omega_3} \quad (3.106)$$

3.F.3 Solution générale

Afin d'étudier le mouvement général de la toupie symétrique, en utilisant le plus possible les lois de conservation, il est pratique de définir les paramètres a , b , α et β :

$$p_\psi = I a \quad p_\phi = I b \quad \alpha = \frac{2E - I_3 \omega_3^2}{I} \quad \beta = \frac{2m g h}{I} \quad (3.107)$$

Les paramètres a et b ont les dimensions d'une vitesse angulaire, alors que α et β ont les dimensions d'une vitesse angulaire au carré. En fonction de ces paramètres, l'équation de conservation de l'énergie (3.98) prend la forme suivante :

$$\alpha = \dot{\theta}^2 + \frac{(b - a \cos \theta)^2}{\sin^2 \theta} + \beta \cos \theta \quad (3.108)$$

alors que la vitesse de précession $\dot{\phi}$ s'exprime comme suit :

$$\dot{\phi} = \frac{b - a \cos \theta}{\sin^2 \theta} \quad (3.109)$$

Procédons ensuite au changement de variable suivant : $u = \cos \theta$. En fonction de u , l'équation (3.108) devient

$$\dot{u}^2 = (1 - u^2)(\alpha - \beta u) - (b - au)^2 := f(u) \quad (3.110)$$

On peut facilement séparer les variables et obtenir une expression intégrale qui, en principe, nous donne l'évolution dans le temps de u , et donc de θ :

$$t = \int \frac{du}{\sqrt{(1 - u^2)(\alpha - \beta u) - (b - au)^2}} \quad (3.111)$$

Cette intégrale peut être réduite à des fonctions elliptiques, ou simplement être calculée numériquement. Donc en principe le problème est résolu, puisque connaissant $u(t)$, et donc $\theta(t)$, on peut en déduire $\phi(t)$ et $\psi(t)$ en utilisant l'expression des moments conjugués conservés p_ψ et p_ϕ .

Cependant, quoiqu'il soit rassurant de savoir comment calculer exactement le mouvement de la toupie, une description qualitative du mouvement de la toupie serait plus instructive. À cette fin, rappelons-nous que la fonction

$$f(u) := (1 - u^2)(\alpha - \beta u) - (b - au)^2 = \beta u^3 - u^2(a^2 + \alpha) + (2ab - \beta)u + \alpha - b^2 \quad (3.112)$$

est à tout moment égale à \dot{u}^2 . Les racines de $f(u)$ correspondent donc aux points de rebroussement du mouvement de la toupie, c'est-à-dire aux angles θ auxquels $\dot{\theta}$ change de signe, donc entre lesquels la toupie évolue.

Le polynôme $f(u)$ étant cubique, il possède au moins une racine réelle. Dans un cas particulier, il possède une racine réelle double, et dans un cas plus général trois racines réelles. Gardons en tête que la constante β est toujours positive ; donc $f(u)$ est négatif quand $u \rightarrow -\infty$ et positif quand $u \rightarrow \infty$. D'autre part, quand $u = \pm 1$, on voit de la définition (3.112) que $f(u) < 0$. Donc au moins une racine de $f(u)$ doit être plus grande que 1, alors que cette valeur ne correspond pas à un angle réel ($1 < u < -1$). Le mouvement physique de la toupie doit donc correspondre au cas où deux racines réelles de $f(u)$ — désignons-les par u_1 et u_2 — sont situées dans l'intervalle $[-1, 1]$ (voir figure 3.8). Le cas d'une précession uniforme correspond en particulier à l'existence d'une racine double $u_1 = u_2$.

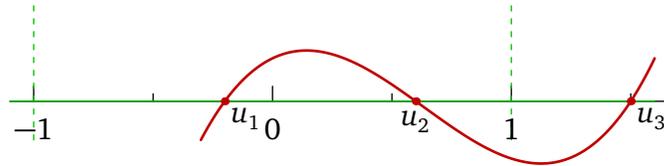


FIGURE 3.8

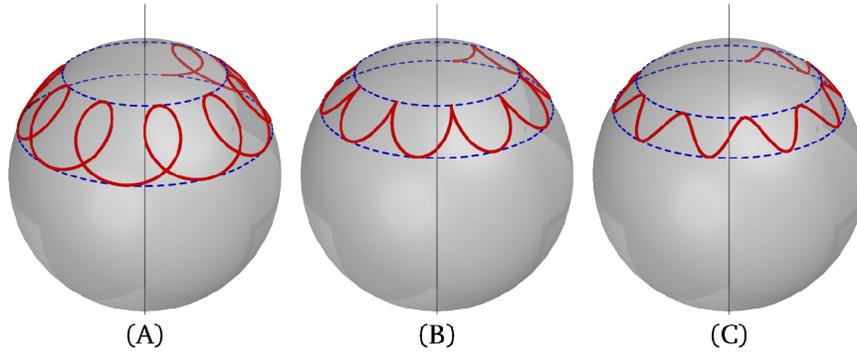


FIGURE 3.9

Les deux racines $u_1 = \cos \theta_1$ et $u_2 = \cos \theta_2$ correspondent à deux angles $\theta_2 < \theta_1$ qui bornent le mouvement de la toupie. Pour préciser les types de mouvement possibles, définissons $u' = b/a$ et, pour fixer les idées, supposons que $a > 0$. Nous avons alors les trois cas suivants :

1. Si $u' > u_2$, alors $b - a \cos \theta_2 > 0$ et donc $b - a \cos \theta > 0$ en tout temps, car $\cos \theta < \cos \theta_2$. Cela entraîne, d'après l'éq. (3.109), que $\dot{\phi}$ est toujours du même signe (positif) et donc que le mouvement de l'axe de la toupie décrit une courbe qualitativement semblable à la figure 3.9C.
2. Si, en revanche, $u_1 < u' < u_2$, alors le signe de $\dot{\phi}$ n'est le même sur les deux bornes θ_1 et θ_2 et l'axe de la toupie décrit plutôt une courbe semblable à la figure 3.9A.
3. Enfin, le cas limite entre les deux se produit lorsque $u' = u_2$. Dans ce cas, $\dot{\phi} = 0$ lorsque $\theta = \theta_2$ et des pointes apparaissent dans la courbe décrite par l'axe de la toupie, comme illustré à la figure 3.9B.

Ce dernier cas correspond en fait aux conditions initiales les plus courantes lorsqu'on expérimente avec une toupie (ou un gyroscope pesant). Supposons en effet que les conditions initiales du mouvement sont $\theta = \theta_2$ et $\dot{\theta} = \dot{\phi} = 0$. Dans ce cas, on voit que

$$b = a \cos \theta_2 = au_2 \quad \text{et} \quad \alpha = \beta \cos \theta_2 = \beta u_2 \quad (3.113)$$

et donc

$$f(u) = (1 - u^2)\beta(u_2 - u) - a(u_2 - u)^2 = (u - u_2)[\beta(1 - u^2) - a(u_2 - u)] \quad (3.114)$$

On voit que, dans ce cas, $u' = b/a = u_2$ est l'une des racines de $f(u)$ et donc correspond au cas illustré à la figure 3.9B.

3.F.4 Toupie dormante

Examinons maintenant le cas particulier d'une toupie qui tourne rapidement à la verticale, et étudions la stabilité d'un tel mouvement. Les conditions initiales correspondent dans ce cas à $u_1 = u_2 = 1$ et donc, d'après les équations (3.113), on a $a = b$ et $\alpha = \beta$. L'équation (3.110) se simplifie alors comme suit :

$$\dot{u}^2 = \beta(1 - u^2)(1 - u) - a^2(1 - u)^2 = (1 - u)^2 [\beta(1 + u) - a^2] \quad (3.115)$$

Il est donc manifeste que si $u = 1$ est une racine, cette racine est nécessairement double, c'est-à-dire $u_1 = u_2 = 1$. La troisième racine, u_3 , est donnée par

$$u_3 = \frac{a^2}{\beta} - 1 \quad (3.116)$$

Cette solution est compatible avec la solution générale trouvée précédemment si $u_3 > 1$, ce qui mène à la condition

$$\frac{a^2}{\beta} > 2 \quad \text{ou} \quad \left(\frac{I_3 \omega_3}{I} \right)^2 > \frac{4mgh}{I} \quad \text{ou enfin} \quad \boxed{\omega_3^2 > \frac{4mghI}{I_3^2}} \quad (3.117)$$

Cette condition correspond à une toupie dite *dormante*. Dans une situation réelle, le frottement du pivot réduisant progressivement la valeur de ω_3 , une toupie dormante 'décrochera' quand sa vitesse angulaire ω_3 tombera en deçà de cette valeur critique.

Une autre possibilité de solution est que u_3 soit compris entre -1 et 1 . Dans ce cas, il suffit de renommer les racines et de supposer plutôt que $-1 < u_1 < 1$ et $u_2 = u_3 = 1$. Cette situation correspond en fait à une toupie en nutation, mais qui atteint la verticale dans son mouvement de nutation ($\theta_2 = 0$).

G Précession des équinoxes

La précession des équinoxes est le mouvement de précession de la Terre sous l'influence gravitationnelle de la Lune et du Soleil. Le phénomène est aisément décrit par les relations démontrées à la section précédente, sauf qu'il faut remplacer le potentiel gravitationnel $mgh \cos \theta$ par une expression décrivant l'énergie potentielle de la Terre dans le champ gravitationnel de la Lune et du Soleil.

3.G.1 Énergie potentielle d'un objet plongé dans un champ gravitationnel

Considérons un objet constitué de N particules de masses m_a à des positions \mathbf{r}_a , en présence d'un champ gravitationnel $\mathbf{g}(\mathbf{r})$ créé par une source lointaine située en \mathbf{r} , qu'on considérera ponctuelle aux fins de l'argument. L'énergie potentielle de l'objet est

$$V = -GM \sum_a \frac{m_a}{|\mathbf{r}_a - \mathbf{r}|} \quad (3.118)$$

Supposons que l'origine des coordonnées est placée quelque part dans l'objet, par exemple en son centre de masse. Si ψ_a désigne l'angle entre les vecteurs \mathbf{r} et \mathbf{r}_a , on a

$$\frac{1}{|\mathbf{r}_a - \mathbf{r}|} = \frac{1}{\sqrt{r^2 + r_a^2 - 2rr_a \cos \psi_a}} = \frac{1}{r} \frac{1}{\sqrt{1 + ((r_a/r))^2 - 2(r_a/r) \cos \psi_a}} \quad (3.119)$$

Vu que la source est lointaine (donc $r \gg r_a$), on peut procéder au développement en série suivant :

$$\frac{1}{|\mathbf{r}_a - \mathbf{r}|} = \frac{1}{r} \sum_{n=0}^{\infty} \left(\frac{r_a}{r}\right)^n P_n(\cos \psi_a) \quad (3.120)$$

où P_n désigne le polynôme de Legendre de degré n .

Rappelons que les polynômes de Legendre sont précisément définis par la fonction génératrice

$$G(x, t) = \frac{1}{\sqrt{1 - 2xt + x^2}} = \sum_{l=0}^{\infty} x^l P_l(t) \quad (3.121)$$

et qu'ils sont obtenus en effectuant correctement le développement binomial. Les premiers polynômes sont

$$\begin{aligned} P_0(t) &= 1 \\ P_1(t) &= t \\ P_2(t) &= \frac{1}{2}(3t^2 - 1) \\ P_3(t) &= \frac{1}{2}(5t^3 - 3t) \end{aligned}$$

Une formule explicite peut être obtenue pour P_l :

$$P_l(t) = \frac{(2l-1)!!}{l!} \left[t^l - \frac{l(l-1)}{2(2l-1)} t^{l-2} + \frac{l(l-1)(l-2)(l-3)}{2 \cdot 4 \cdots (2l-3)(2l-1)} t^{l-4} + \dots \right] \quad (3.122)$$

Les polynômes de Legendre sont habituellement définis entre -1 et 1 , et ont la propriété d'orthogonalité sur cet intervalle :

$$\int_{-1}^1 dt P_n(t) P_m(t) = \frac{2}{2l+1} \delta_{mn} \quad (3.123)$$

En plus, on a les valeurs $P_n(1) = 1$ et $P_n(-1) = (-1)^n$, qui se démontre aisément en posant $t = \pm 1$ dans la fonction génératrice $G(x, t)$, qui mène alors au développement géométrique simple de $1/(1 \mp x)$. Les cinq premiers polynômes de Legendre sont illustrés à la figure 3.10.

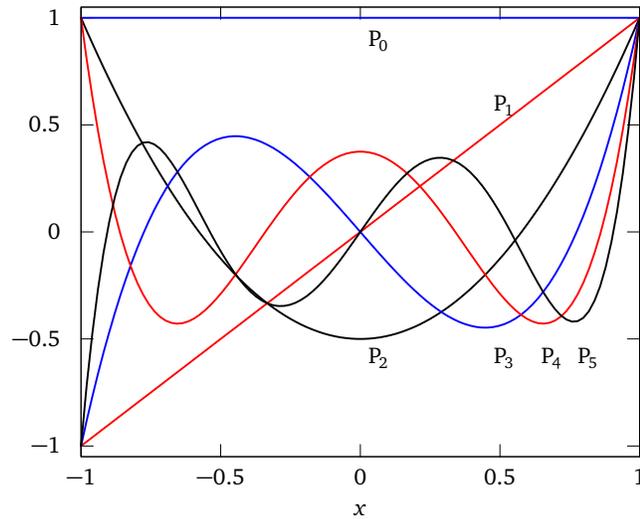


FIGURE 3.10
Les 5 premiers polynômes de Legendre.

En limitant le développement aux trois premiers termes, l'énergie potentielle gravitationnelle a l'expression suivante :

$$V = -\frac{GM}{r} \left(\sum_a m_a \right) - \frac{GM}{r^3} \mathbf{r} \cdot \left(\sum_a m_a \mathbf{r}_a \right) - \frac{GM}{2r^5} \left(\sum_a m_a [3(\mathbf{r} \cdot \mathbf{r}_a)^2 - r^2 r_a^2] \right) - \dots \quad (3.124)$$

Désignons par m la masse totale de l'objet. Le premier terme du développement ci-dessus n'est autre que

$$V_0 = -\frac{GMm}{r} \quad (3.125)$$

Le deuxième terme n'est autre que

$$V_1 = -\frac{GMm}{r^3} \mathbf{r} \cdot \mathbf{r}_{\text{cm}} \quad (3.126)$$

où \mathbf{r}_{cm} est la position du centre de masse de l'objet. Donc ce terme s'annule si on place l'origine des coordonnées au centre de masse, ce que nous supposons par la suite.

Enfin, le troisième terme est

$$V_2 = -\frac{GM}{r^5} \left(\frac{3}{2} x_i x_j - \frac{1}{2} r^2 \delta_{ij} \right) \sum_a m_a x_{a,i} x_{a,j} \quad (3.127)$$

On peut exprimer ce terme en fonction du tenseur d'inertie \mathbf{I} défini en (3.53) :

$$\sum_a m_a x_{a,i} x_{a,j} = -I_{ij} + \delta_{ij} \sum_a m_a r_a^2 \quad (3.128)$$

En prenant la trace de cette expression, c'est-à-dire en posant $i = j$ et en sommant sur i , on constate que

$$\sum_a m_a r_a^2 = -\text{tr } \mathbf{I} + 3 \sum_a m_a r_a^2 \implies \sum_a m_a r_a^2 = \frac{1}{2} \text{tr } \mathbf{I} \quad (3.129)$$

Donc

$$\sum_a m_a x_{a,i} x_{a,j} = -I_{ij} + \frac{1}{2} \delta_{ij} \text{tr } \mathbf{I} \quad (3.130)$$

Il s'ensuit que

$$\begin{aligned} V_2 &= -\frac{\text{GM}}{r^5} \left(\frac{3}{2} x_i x_j - \frac{1}{2} r^2 \delta_{ij} \right) (-I_{ij} + \frac{1}{2} \delta_{ij} \text{tr } \mathbf{I}) \\ &= \frac{\text{GM}}{r^5} \left(\frac{3}{2} \tilde{x} \mathbf{I} \tilde{x} - \frac{1}{2} \tilde{x} \tilde{x} \text{tr } \mathbf{I} \right) \quad \tilde{x} \tilde{x} = r^2 \end{aligned} \quad (3.131)$$

Jusqu'ici nous n'avons pas imposé de contraintes à la géométrie de l'objet. Supposons maintenant que l'objet a une symétrie cylindrique, de sorte que $I_1 = I_2 = I \neq I_3$. Soit γ l'angle que fait l'axe de symétrie de l'objet avec la direction de la source située en \mathbf{r} . Dans le repère des axes principaux de l'objet, la direction du vecteur \mathbf{x} peut être décrite par l'angle polaire γ et un angle azimutal φ , de sorte que

$$\mathbf{x} = (r \sin \gamma \cos \varphi, r \sin \gamma \sin \varphi, r \cos \gamma) \quad (3.132)$$

et donc

$$\tilde{x} \mathbf{I} \tilde{x} = r^2 (I \sin^2 \gamma + I_3 \cos^2 \gamma) = r^2 [I + (I_3 - I) \cos^2 \gamma] \quad (3.133)$$

L'énergie potentielle prend donc la forme suivante, en fonction de r et γ :

$$V = -\frac{\text{GM}m}{r} + \frac{\text{GM}}{r^3} (I_3 - I) P_2(\cos \gamma) \quad (3.134)$$

Le premier terme est l'énergie potentielle d'un objet de masse m entièrement concentré en son centre de masse et n'implique pas du tout l'orientation de l'objet. Ce terme n'aura donc aucun rôle à jouer dans la précession des équinoxes. Nous allons donc nous concentrer sur le deuxième terme.

Lors du mouvement orbital de la Terre autour de la source (ou vice-versa, si la source est la Lune) l'angle γ va changer. Cependant, la précession des équinoxes est un phénomène très lent, de sorte qu'en première approximation, l'orientation de l'axe de la Terre est toujours la même, alors que le vecteur \mathbf{r} effectue une révolution complète. Nous pouvons donc calculer la moyenne du potentiel V_2 le long d'une orbite complète et utiliser cette valeur moyenne dans le traitement de la précession. Nous allons supposer que l'orbite de la Terre autour du Soleil est circulaire, c'est-à-dire que nous allons négliger la faible excentricité de l'orbite terrestre. Soit θ l'angle que fait l'axe de symétrie de la Terre par rapport à l'axe polaire de l'orbite (l'axe perpendiculaire au plan de l'orbite) et η l'angle azimutal de la Terre le long de son orbite circulaire (voir figure 3.11). Le cosinus de l'angle γ entre l'axe de la Terre et sa position le long de son orbite est donné par le produit scalaire

$$\cos \gamma = (0, -\sin \theta, \cos \theta) \cdot (\cos \eta, \sin \eta, 0) = -\sin \theta \sin \eta \quad (3.135)$$

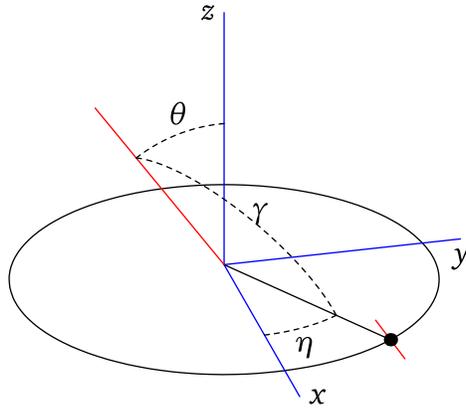


FIGURE 3.11

La moyenne du carré $\langle \cos^2 \gamma \rangle$ se calcule facilement :

$$\langle \cos^2 \gamma \rangle = \sin^2 \theta \langle \sin^2 \eta \rangle = \frac{1}{2} \sin^2 \theta = \frac{1}{2} - \frac{1}{2} \cos^2 \theta \quad (3.136)$$

Donc

$$\langle P_2(\cos \gamma) \rangle = \frac{3}{2} \left(\frac{1}{2} - \frac{1}{2} \cos^2 \theta \right) - \frac{1}{2} = -\frac{1}{2} \left(\frac{3}{2} \cos^2 \theta - \frac{1}{2} \right) = -\frac{1}{2} P_2(\cos \theta) \quad (3.137)$$

Le potentiel moyen que nous allons utiliser pour étudier le mouvement de précession est donc

$$V(\theta) = -\frac{GM}{2r^3} (I_3 - I) P_2(\cos \theta) \quad (3.138)$$

3.G.2 Fréquence de précession

Le lagrangien décrivant la rotation de la Terre sur elle-même est une modification triviale du lagrangien (3.85) de la toupie symétrique, où le potentiel $V(\theta)$ est maintenant donné par l'éq. (3.138) :

$$L = \frac{1}{2} I (\dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2) + \frac{1}{2} I_3 (\dot{\phi} \cos \theta + \dot{\psi})^2 - V(\theta) \quad (3.139)$$

L'équation du mouvement (3.92) est alors remplacée par

$$\begin{aligned} I \ddot{\theta} &= I \sin \theta \cos \theta \dot{\phi}^2 - p_\psi \dot{\phi} \sin \theta - \frac{\partial V}{\partial \theta} \\ &= \sin \theta \left\{ -I \cos \theta \dot{\phi}^2 - p_\psi \dot{\phi} + \frac{\partial V}{\partial u} \right\} \quad (u = \cos \theta) \end{aligned} \quad (3.140)$$

La fréquence de précession uniforme s'obtient en posant $\ddot{\theta} = 0$:

$$-I \cos \theta \dot{\phi}^2 - p_\psi \dot{\phi} + \frac{\partial V}{\partial u} = 0 \quad (3.141)$$

La solution se trouve en trouvant les racines de cette équation quadratique. Un raccourci vers une réponse utile se prend en reconnaissant que la rotation de la Terre est rapide en comparaison de sa

précession, de sorte que le deuxième terme de l'équation ci-haut domine complètement le premier. Dans ce cas, on trouve immédiatement

$$\begin{aligned}
 \dot{\phi} &= \frac{1}{p_\psi} \frac{\partial V}{\partial u} \\
 &= -\frac{GM}{2r^3} \frac{I_3 - I}{I_3 \omega_3} P_2'(\cos \theta) \\
 &= -\frac{3GM}{2\omega_3 r^3} \frac{I_3 - I}{I_3} \cos \theta
 \end{aligned} \tag{3.142}$$

Dans le cas de la Terre,

$$\frac{I_3 - I}{I_3} = 0,00327 \tag{3.143}$$

et l'angle θ vaut $23^\circ 27'$. Une évaluation numérique mène à une période de précession de 81 000 ans. La période mesurée est de 25 771 ans, ce qui correspond à un déplacement angulaire de $50,26''$ par année. La différence est bien sûr attribuable à l'influence de la Lune, qui est plus forte que celle du Soleil. Même si la masse M de la Lune est considérablement plus petite que celle du Soleil, le rayon de son orbite est aussi beaucoup plus petit, et le rayon intervient au cube dans la formule (3.142). Un calcul correct doit ajouter l'effet de la Lune à celui du Soleil, et tenir compte du fait que l'orbite de la Lune autour de la Terre n'est pas exactement dans le même plan que celle de la Terre autour du Soleil. Le résultat de ces effets est une valeur théorique compatible avec la valeur observée.

Problèmes

Problème 3.1 Matrice de rotation générale

Nous allons vérifier dans ce problème que les éléments de la matrice de rotation $R(\mathbf{n}, \theta)$, où \mathbf{n} est un vecteur unitaire, sont donnés explicitement par l'expression suivante :

$$R_{ij} = n_i n_j (1 - \cos \theta) + \delta_{ij} \cos \theta + \varepsilon_{ijk} n_k \sin \theta \quad (3.144)$$

- A** Montrez que cette expression définit une matrice orthogonale.
- B** Montrez que l'application de cette matrice sur le vecteur \mathbf{n} le laisse invariant.
- C** Vérifiez que la trace de cette matrice est ce qu'elle devrait être pour une rotation d'angle θ .
- D** Montrez qu'on retrouve la forme connue de la matrice de rotation dans le cas $\mathbf{n} = \mathbf{z}$.
- E** Montrez qu'en langage vectoriel, le vecteur $\mathbf{x}' = R\mathbf{x}$ s'exprime comme

$$\mathbf{r}' = \mathbf{n} \cdot \mathbf{r} (1 - \cos \theta) \mathbf{n} + \mathbf{r} \cos \theta + \mathbf{r} \wedge \mathbf{n} \sin \theta \quad (3.145)$$

Problème 3.2 Rotations et matrices de Pauli

Dans ce problème nous allons explorer une représentation des rotations en fonction de matrices unitaires 2×2 . Commençons par définir les trois matrices suivantes, dites *matrices de Pauli* :

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.146)$$

Ces trois matrices sont hermitiennes : $\sigma_i^\dagger = \sigma_i$.

- A** Montrez que ces matrices obéissent aux relations suivantes :

$$\sigma_i \sigma_j = \delta_{ij} \mathbb{I} + i \varepsilon_{ijk} \sigma_k \quad (3.147)$$

- B** Si \mathbf{n} est un vecteur unitaire quelconque, montrez que

$$e^{i\mathbf{a}\mathbf{n}\cdot\boldsymbol{\sigma}} = \mathbb{I} \cos \alpha + i\boldsymbol{\sigma} \cdot \mathbf{n} \sin \alpha \quad (3.148)$$

où α est un réel quelconque. Nous avons introduit un 'vecteur de matrices' $\boldsymbol{\sigma}$, dont les composantes sont les matrices de Pauli. Ainsi, $\mathbf{n} \cdot \boldsymbol{\sigma} = n_1 \sigma_1 + n_2 \sigma_2 + n_3 \sigma_3$ est une matrice hermitienne 2×2 . Indice : pour calculer l'exponentielle, utilisez son développement en série de Taylor, ainsi que la propriété $(\boldsymbol{\sigma} \cdot \mathbf{n})^2 = \mathbb{I}$, que vous devrez démontrer.

- C** Une matrice U est qualifiée d'*unitaire* si son inverse est donné par son conjugué hermitien : $U^\dagger U = 1$. Montrez que la matrice $\exp(i\mathbf{a}\mathbf{n} \cdot \boldsymbol{\sigma})$ est unitaire.

D Étant donné un vecteur \mathbf{r} , on peut former la matrice $\mathbf{r} \cdot \boldsymbol{\sigma}$. On peut donc faire correspondre une matrice hermitienne 2×2 à un vecteur \mathbf{r} . Montrez que si on définit un vecteur \mathbf{r}' par la relation

$$\mathbf{r}' \cdot \boldsymbol{\sigma} = U(\mathbf{r} \cdot \boldsymbol{\sigma})U^\dagger \quad (3.149)$$

où U est une matrice unitaire, alors $|\mathbf{r}'| = |\mathbf{r}|$. La matrice U effectue donc une rotation sur le vecteur \mathbf{r} , mais via une application plus complexe qu'une simple multiplication matrice \times vecteur.

E Montrez que la matrice unitaire

$$U(\mathbf{n}, \theta) = e^{i\theta \mathbf{n} \cdot \boldsymbol{\sigma} / 2} = \mathbb{I} \cos \theta / 2 + i \boldsymbol{\sigma} \cdot \mathbf{n} \sin \theta / 2 \quad (3.150)$$

produit sur la matrice $\mathbf{r} \cdot \boldsymbol{\sigma}$ l'effet correspondant à la matrice de rotation $R(\mathbf{n}, \theta)$ sur le vecteur \mathbf{r} . Autrement dit, montrez que

$$U(\mathbf{n}, \theta)(\mathbf{r} \cdot \boldsymbol{\sigma})U^\dagger(\mathbf{n}, \theta) = \mathbf{r}' \cdot \boldsymbol{\sigma} \quad (3.151)$$

où

$$\mathbf{r}' = \mathbf{n} \cdot \mathbf{r}(1 - \cos \theta)\mathbf{n} + \mathbf{r} \cos \theta + \mathbf{r} \wedge \mathbf{n} \sin \theta \quad (3.152)$$

Vous devez pour cela utiliser les propriétés des matrices de Pauli démontrées plus haut.

Problème 3.3 Matrices de rotations

Vérifiez que les matrices de rotation $R_1 = R(\mathbf{z}, \frac{\pi}{2})$ et $R_2 = R(\mathbf{x}, \frac{\pi}{2})$ ne commutent pas. Exprimez les produits des deux matrices, dans un ordre et dans l'autre (c'est-à-dire $R_1 R_2$ et $R_2 R_1$), en fonction d'une matrice de rotation simple. Cela nécessite de déterminer l'axe et l'angle de la rotation associés à chacun des produits. Note : l'utilisation de Mathematica, notamment de la fonction `RotationMatrix`, est permise, et peut être utile, surtout pour la dernière partie.

Problème 3.4 Vitesse angulaire dans le repère fixe

Exprimez les composantes de la vitesse angulaire $\boldsymbol{\omega}$ dans le repère fixe dans l'espace S – au lieu du repère fixe à l'objet – en fonction des angles d'Euler et de leurs dérivées. Comment peut-on décrire la différence entre l'expression obtenue et celle qui s'applique au repère S' ?

Problème 3.5 Tenseur d'inertie de la molécule de méthane

La molécule de méthane (CH_4) a la structure géométrique d'un tétraèdre régulier, avec le carbone au centre et les 4 hydrogènes sur les sommets. Exprimez la matrice d'inertie d'une molécule de méthane en fonction de la distance a entre le noyau de carbone et chacun des noyaux d'hydrogène, et de la masse m du proton (on néglige la masse de l'électron).

Problème 3.6 Moment d'inertie d'un disque mince

Calculez le moment d'inertie d'un disque de rayon a , de masse m et d'épaisseur négligeable, par rapport à un axe qui le traverse le long d'un diamètre.

Problème 3.7 Moments d'inertie d'un ellipsoïde

A Calculez les moments d'inertie $I_1 = I_2 = I$ et I_3 d'un ellipsoïde de révolution de densité uniforme et de masse m dont les demi-axes sont a dans le plan xy et c le long de l'axe des z .

B Dans le cas où c est très proche de a , trouvez une approximation simple pour le rapport $(I_3 - I)/I$

Problème 3.8 Angles d'Euler et rotation libre

Utilisez les angles d'Euler pour décrire la rotation libre d'un objet. Adoptez la direction du moment cinétique conservé L comme axe des z . Démontrez que (i) θ est constant et (ii) la grandeur du moment cinétique est $L = I\dot{\phi}$.

Problème 3.9 Toupie dormante et potentiel effectif

A En général, le potentiel effectif d'une toupie symétrique diverge quand $\theta \rightarrow 0$. À quelle condition cette divergence est-elle absente ?

B Dans l'hypothèse où cette condition est satisfaite, sous quelle condition supplémentaire la solution $\theta = 0$ est-elle stable, c'est-à-dire la toupie est-elle dormante ? Obtenez cette condition uniquement en considérant un développement de Taylor du potentiel effectif au deuxième ordre en θ .

Problème 3.10 Toupie reposant sur une plaque sans frottement

A Au lieu de supposer qu'une toupie symétrique est liée à un pivot fixe, supposons au contraire que sa pointe repose sur une surface sur laquelle elle est libre de glisser, sans frottement. Écrivez le lagrangien de ce système, en tenant compte maintenant des degrés de liberté associés au mouvement du centre de masse de la toupie (translation), en plus des degrés de liberté de rotation décrits par les angles d'Euler. La contrainte que la pointe de la toupie repose sur la plaque doit être décrite par un paramètre de Lagrange. Écrivez les équations de Lagrange correspondantes.

B En supposant que la hauteur z du centre de masse de la toupie est constante, ce qui est un cas particulier, l'expression de la fréquence de précession est-elle différente de celle associée à une toupie à pivot fixe ? Pour une forme et une masse donnée de la toupie, la valeur de cette fréquence de précession serait-elle la même ?

C Le cas d'un hauteur z non constante est-il plus difficile à résoudre que pour une toupie à pivot fixe ? Pourquoi ?

Problème 3.11 Précession rapide

L'équation quadratique que respecte la fréquence de précession $\dot{\phi}$ dans le cas où θ est constant s'écrit

$$mgh - \dot{\phi} p_{\psi} + I \dot{\phi}^2 \cos \theta = 0 \quad (3.153)$$

Cette équation possède deux racines :

$$\dot{\phi} = \frac{p_{\psi}}{2I \cos \theta} \left\{ 1 \mp \sqrt{1 - \frac{4Imgh \cos \theta}{p_{\psi}^2}} \right\} \quad (3.154)$$

Nous avons déjà étudié le sens de la première racine (signe $-$).

- A** Expliquez le sens de l'autre racine (signe $+$) dans le cas $g = 0$.
- B** Que devient cette deuxième racine dans le cas où g est non nul, mais où la toupie est rapide (c'est-à-dire quand le discriminant est proche de, mais inférieur à 1) ?

Problème 3.12 Rotation libre d'un objet asymétrique

Considérez un objet asymétrique, avec trois moments d'inertie bien distincts $I_1 > I_2 > I_3$, en rotation libre dans l'espace.

- A** Montrez qu'un mouvement de rotation uniforme autour de chacun des axes principaux est toujours une solution aux équations d'Euler.
- B** Montrez que les rotations sont stables autour des deux axes extrêmes \mathbf{e}_1 et \mathbf{e}_3 , mais que la rotation autour de l'axe médian est instable. Pour ce faire, supposez que les deux autres composantes de la vitesse angulaire, au lieu d'être nulles, sont très petites, de sorte que leur produit peut être négligé, et résolvez les équations d'Euler correspondantes.

Problème 3.13 Mouvement de billes magnétiques

Une bille de masse m et rayon a possède un moment magnétique permanent \mathbf{M} . Ce moment magnétique, en interaction avec le champ magnétique terrestre \mathbf{B} , qu'on considère uniforme et constant, cause un couple $\mathbf{M} \wedge \mathbf{B}$ sur la bille. La bille se trouve sur une surface plane ; on note \mathbf{v} la vitesse de son centre de masse, nécessairement dans le plan xy . La seule force horizontale agissant sur la bille est la force de frottement \mathbf{f} causée par le glissement du point de contact de la bille avec le sol. Nous allons adopter le modèle suivant pour cette force : si \mathbf{v}_r dénote la vitesse du point de la bille qui touche le sol, la force sera constante en grandeur et opposée à \mathbf{v}_r . Par contre, la force sera exactement nulle si $\mathbf{v}_r = 0$. Comme cette dépendance mathématique en fonction de v_r est singulière quand $v_r \rightarrow 0$, nous allons plutôt la remplacer par la fonction suivante :

$$\mathbf{f} = -f \frac{\mathbf{v}_r}{|\mathbf{v}_r| + \eta} \quad \text{où } \eta \text{ est une quantité petite} \quad (3.155)$$

Le champ \mathbf{B} peut-être dans une direction quelconque.

- A** Obtenez une expression pour \mathbf{v}_r en fonction de \mathbf{v} et de $\boldsymbol{\omega}$, la vitesse angulaire de la bille.
- B** Écrivez les équations du mouvement donnant les dérivées temporelles de \mathbf{r} , de \mathbf{v} , de $\boldsymbol{\omega}$ et de \mathbf{M} . Indice : le vecteur \mathbf{M} est constant dans le repère lié à l'objet, mais dépend du temps dans le repère de l'observateur, celui qu'on utilise ici. Les équations peuvent être laissées sous forme vectorielle, mais doivent être équivalentes à 10 équations indépendantes (en supposant qu'on ne s'occupe pas de la dimension z pour la position et la vitesse).
- C** Programmez ce système d'équations différentielles dans Mathematica, avec les valeurs suivantes des paramètres : $m = 1$, $a = 1$, $f = 1$, $\boldsymbol{\omega}(0) = (-1, 0, 0)$, $\mathbf{v}(0) = (0, 1, 0)$, $\mathbf{r}(0) = (0, 0, 0)$, $\mathbf{M}(0) = (1, 0, 0)$, $\mathbf{B} = (0, 0.1/\sqrt{2}, 0.1/\sqrt{2})$, $\eta = 10^{-3}$. Résolvez ces équations numériquement de $t = 0$ à $t = 400$ et portez la trajectoire en graphique. Les fonctions suivantes de Mathematica seront utiles à des degrés divers : `NDSolve`, `ParametricPlot`, `Cross`, `Norm`, `Evaluate`.

CHAPITRE 4

Forces centrales

A Solution générale

4.A.1 Réduction du problème à deux corps

Soit deux corps de masses m_1 et m_2 et de positions \mathbf{r}_1 et \mathbf{r}_2 , évoluant sous l'influence d'un potentiel central $V(|\mathbf{r}_1 - \mathbf{r}_2|)$ qui ne dépend que de la distance entre les deux corps. La fonction de Lagrange est

$$L = \frac{1}{2}m_1\dot{\mathbf{r}}_1^2 + \frac{1}{2}m_2\dot{\mathbf{r}}_2^2 - V(|\mathbf{r}_1 - \mathbf{r}_2|) \quad (4.1)$$

Nous allons utiliser les coordonnées généralisées \mathbf{r} et \mathbf{r}_{cm} définies ci-dessous, au lieu de \mathbf{r}_1 et \mathbf{r}_2 :

$$\begin{cases} \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \\ \mathbf{r}_{\text{cm}} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} \end{cases} \quad \text{ou, à l'inverse} \quad \begin{cases} \mathbf{r}_1 = \mathbf{r}_{\text{cm}} + \frac{m_2\mathbf{r}}{m_1 + m_2} \\ \mathbf{r}_2 = \mathbf{r}_{\text{cm}} - \frac{m_1\mathbf{r}}{m_1 + m_2} \end{cases} \quad (4.2)$$

Il s'ensuit que

$$\dot{\mathbf{r}}_1 = \mathbf{v}_{\text{cm}} + \frac{m_2\dot{\mathbf{r}}}{m_1 + m_2} \quad \dot{\mathbf{r}}_2 = \mathbf{v}_{\text{cm}} - \frac{m_1\dot{\mathbf{r}}}{m_1 + m_2} \quad (4.3)$$

et que

$$\begin{aligned} \frac{1}{2}m_1\dot{\mathbf{r}}_1^2 &= \frac{1}{2}m_1\mathbf{v}_{\text{cm}}^2 + \frac{1}{2}\frac{m_1m_2^2\dot{\mathbf{r}}^2}{(m_1 + m_2)^2} + \frac{m_1m_2}{m_1 + m_2}\dot{\mathbf{r}}\mathbf{v}_{\text{cm}} \\ \frac{1}{2}m_2\dot{\mathbf{r}}_2^2 &= \frac{1}{2}m_2\mathbf{v}_{\text{cm}}^2 + \frac{1}{2}\frac{m_2m_1^2\dot{\mathbf{r}}^2}{(m_1 + m_2)^2} - \frac{m_1m_2}{m_1 + m_2}\dot{\mathbf{r}}\mathbf{v}_{\text{cm}} \end{aligned} \quad (4.4)$$

dont la somme est

$$T(\mathbf{r}, \mathbf{r}_{\text{cm}}) = \frac{1}{2}(m_1 + m_2)\mathbf{v}_{\text{cm}}^2 + \frac{1}{2}\frac{m_1m_2}{m_1 + m_2}\dot{\mathbf{r}}^2 \quad (4.5)$$

On définit ensuite la masse totale M et la *masse réduite* m ainsi :

$$M = m_1 + m_2 \quad m = \frac{m_1m_2}{m_1 + m_2} \quad \text{ou} \quad \frac{1}{m} = \frac{1}{m_1} + \frac{1}{m_2} \quad (4.6)$$

De sorte que le lagrangien a la forme finale suivante :

$$L = \frac{1}{2}M\mathbf{v}_{\text{cm}}^2 + \frac{1}{2}m\dot{\mathbf{r}}^2 - V(r) \quad (r := |\mathbf{r}|) \quad (4.7)$$

Le lagrangien est la somme de deux parties qui impliquent des variables différentes : le premier terme ci-haut représente l'énergie cinétique associée au mouvement du centre de masse des deux corps, sans énergie potentielle associée. Les deux autres termes forment le lagrangien d'une particule effective de masse m se déplaçant dans un potentiel central $V(r)$. Le problème à deux corps a donc été réduit à un problème à un corps, qu'on peut résoudre séparément du mouvement du centre de masse, qui se produit alors à vitesse constante.

Remarques:

- ◆ En fonction de la masse réduite, on exprime les positions $\mathbf{r}_{1,2}$ comme suit :

$$\mathbf{r}_1 = \mathbf{r}_{\text{cm}} + \frac{m}{m_1}\mathbf{r} \quad \mathbf{r}_2 = \mathbf{r}_{\text{cm}} - \frac{m}{m_2}\mathbf{r} \quad (4.8)$$

- ◆ Si, en plus du potentiel central $V(r)$, les deux corps évoluent sous l'influence d'un troisième corps, alors cette réduction n'est plus valable. Cependant, si l'interaction en cause est gravitationnelle, alors l'énergie potentielle supplémentaire associée au troisième corps prend la forme

$$m_1\Phi(\mathbf{r}_1) + m_2\Phi(\mathbf{r}_2) \quad (4.9)$$

où Φ est le potentiel gravitationnel créé par le troisième corps. Si ce dernier est très éloigné, alors on peut supposer que Φ varie lentement sur une échelle de distance de l'ordre de $|\mathbf{r}|$ et procéder au développement limité suivant autour de \mathbf{r}_{cm} :

$$\Phi(\mathbf{r}_{\text{cm}} + \delta\mathbf{r}) \approx \Phi(\mathbf{r}_{\text{cm}}) + \delta x_i \partial_i \Phi(\mathbf{r}_{\text{cm}}) + \frac{1}{2} \delta x_i \delta x_j \partial_i \partial_j \Phi(\mathbf{r}_{\text{cm}}) + \dots \quad (4.10)$$

(on somme sur les indices répétés). En appliquant ce développement à $\delta\mathbf{r} = (m/m_1)\mathbf{r}$ et $\delta\mathbf{r} = -(m/m_2)\mathbf{r}$ et en sommant, on trouve, à cet ordre d'approximation :

$$\begin{aligned} & m_1\Phi(\mathbf{r}_1) + m_2\Phi(\mathbf{r}_2) \\ &= M\Phi + mx_i \partial_i \Phi - mx_i \partial_i \Phi + \frac{1}{2} \frac{m^2}{m_1} x_i x_j \partial_i \partial_j \Phi + \frac{1}{2} \frac{m^2}{m_2} x_i x_j \partial_i \partial_j \Phi + \dots \\ &= M\Phi + \frac{1}{2} mx_i x_j \partial_i \partial_j \Phi + \dots \end{aligned} \quad (4.11)$$

où Φ et ses dérivées sont évaluées à \mathbf{r}_{cm} . Le centre de masse évolue dans le potentiel gravitationnel $\Phi(\mathbf{r}_{\text{cm}})$, alors que la variable relative \mathbf{r} est sujette à la force de marée $\partial_i \partial_j \Phi(\mathbf{r}_{\text{cm}})$ causée par le troisième corps. Celle-ci dépend de \mathbf{r}_{cm} , ce qui viole la séparation entre les variables relatives et de centre de masse que nous avons jusqu'ici; cependant, si le problème du mouvement du centre de masse est résolu dans un premier temps, alors la solution retenue peut être utilisée pour raffiner le mouvement relatif en tenant compte de la force de marée.

- ◆ Cette réduction exacte du problème à deux corps vers le problème à un corps ne s'étend malheureusement pas au problème à trois corps ou au-delà.

4.A.2 Problème radial effectif

Concentrons-nous maintenant sur le problème à un corps dans un potentiel central, dont le lagrangien s'écrit

$$L = \frac{1}{2}m\mathbf{v}^2 - V(r) \quad (4.12)$$

Le lagrangien est invariant par toute rotation des coordonnées autour de l'origine (la source du potentiel). Cela entraîne que le moment cinétique $\mathbf{L} = \mathbf{r} \wedge \mathbf{p}$ est conservé en tout temps, en grandeur et en direction. La trajectoire de l'objet est toujours contenue dans le plan perpendiculaire au moment cinétique, comme l'indique clairement la définition de \mathbf{L} . Adoptons comme coordonnées généralisées les coordonnées polaires obtenues en alignant l'axe des z avec le moment cinétique orbital. Le lagrangien s'exprime alors comme suit (voir l'exemple 2.3) :

$$L = \frac{1}{2}m \{ \dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\varphi}^2 \} - V(r) \quad (4.13)$$

Les moments conjugués sont :

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r} \quad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2 \dot{\theta} \quad p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = mr^2 \sin^2 \theta \dot{\varphi} \quad (4.14)$$

et le hamiltonien est

$$H = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\varphi^2}{2mr^2 \sin^2 \theta} + V(r) \quad (4.15)$$

Avec l'orientation choisie de l'axe z , $\theta = \pi/2$ et donc $p_\theta = 0$ et le Hamiltonien se réduit à un problème bidimensionnel, dans le plan xy (ou $\theta = \pi/2$) :

$$H = \frac{p_r^2}{2m} + \frac{\ell^2}{2mr^2} + V(r) \quad (4.16)$$

où nous avons substitué $p_\varphi = \ell$, la composante en z du moment cinétique, constante dans le temps. La conservation de ℓ nous permet donc de réduire le problème à un problème effectif dans une seule dimension décrite par la coordonnée radiale r . Dans ce problème effectif, l'énergie potentielle est donnée par

$$V_{\text{eff}} = \frac{\ell^2}{2mr^2} + V(r) \quad (4.17)$$

et la conservation de l'énergie s'exprime comme

$$E = \frac{1}{2}m\dot{r}^2 + \frac{\ell^2}{2mr^2} + V(r) \quad (4.18)$$

Le deuxième terme ci-dessous porte le nom de *potentiel centrifuge*.

Le caractère de la trajectoire peut être qualitativement déduit de la forme du potentiel effectif. La figure 4.1 illustre trois exemples de potentiels effectifs. Considérons premièrement le cas illustré sur la partie gauche de la figure, qui correspond au potentiel gravitationnel causé par une source ponctuelle à l'origine. Le potentiel effectif tend vers $+\infty$ quand $r \rightarrow 0$ parce que le potentiel centrifuge, en $1/r^2$, croît plus rapidement que $1/r$ dans cette limite. À l'inverse, il décroît plus rapidement que $1/r$ quand $r \rightarrow \infty$, et donc $V(r)$ domine V_{eff} dans cette limite. Ceci entraîne l'existence d'un minimum de V_{eff} , situé à $r = r_0$. Il est clair que la valeur minimale E_0 de E est obtenue quand $\dot{r} = 0$ en

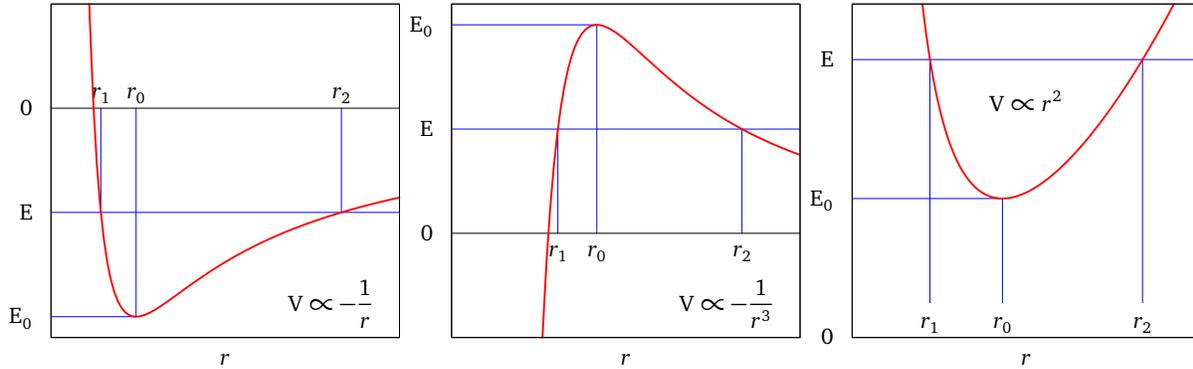


FIGURE 4.1
Trois exemples de potentiels effectifs.

tout temps, et est alors égale au minimum du potentiel effectif. La trajectoire est alors circulaire. Si $0 > E > E_0$, la trajectoire est bornée par les valeurs r_1 et r_2 , qui sont les *points de rebroussement* : lorsque la particule atteint ces valeurs de r , la dérivée \dot{r} s'annule pour ensuite changer de signe. La trajectoire est donc bornée entre ces deux valeurs, menant à une orbite stable. On montrera plus bas que cette trajectoire est une ellipse dont l'un des foyers est à l'origine. Si $E > 0$, il n'existe qu'un point de rebroussement r_1 et l'orbite n'est plus bornée : la particule va s'échapper à l'infini, après s'être approchée à une distance minimale $r = r_1$ du centre d'attraction. On montre que la trajectoire est alors hyperbolique.

Le cas illustré sur la partie centrale de la figure 4.1 correspond à un potentiel attractif en $1/r^3$ qui, à l'inverse du précédent, domine V_{eff} quand $r \rightarrow 0$, alors que le potentiel centrifuge domine quand $r \rightarrow \infty$. Il en résulte un potentiel effectif de forme inversée par rapport au cas précédent, donc un potentiel sans borne inférieure. Même si une orbite circulaire est possible, lorsque $E = E_0$ et $r = r_0$, cette orbite est instable, car elle correspond au maximum local du potentiel effectif. Comme la barrière centrifuge est trop faible pour le potentiel $V(r)$, la particule va passer par l'origine $r = 0$ si son énergie est plus grande que E_0 , ou si elle se trouve à un moment donné à une distance $r < r_0$ de l'origine. Dans ce cas, la vitesse de la particule devient infinie à l'origine. La particule ne peut éviter la catastrophe que si elle vient de l'infini et que son énergie E est inférieure à E_0 .

Enfin, le cas illustré sur la partie droite de la figure 4.1 correspond à un potentiel harmonique, c'est-à-dire en r^2 . Dans ce cas, la particule ne peut jamais s'échapper à l'infini et les points de rebroussement sont toujours bien définis. On montre que la trajectoire est une ellipse centrée à l'origine, qui dégénère en un cercle lorsque $E = E_0$.

Loi des aires

La réduction à un problème radial effectif ne doit pas nous faire oublier la coordonnée angulaire φ . Celle-ci a une dépendance temporelle déterminée par la conservation du moment cinétique :

$$\ell = mr^2\dot{\varphi} \quad (\theta = \pi/2) \quad (4.19)$$

Dans un temps dt , le rayon vecteur \mathbf{r} balaie une aire $\frac{1}{2}r^2 d\varphi$, de sorte que l'aire balayée par unité de temps, ou *vitesse aréolaire*, est

$$\frac{dS}{dt} = \frac{1}{2}r^2\dot{\varphi} = \frac{\ell}{2m} = \text{cte} \quad (4.20)$$

La constance dans le temps de la vitesse aréolaire constitue la *loi des aires*, ou deuxième loi de Kepler. Elle permet de relier la période τ de l'orbite (si orbite il y a) au moment cinétique et à la trajectoire $r(\varphi)$:

$$\frac{\ell}{2m}\tau = \frac{1}{2} \int_0^{2\pi} r^2 d\varphi \implies \tau = \frac{m}{\ell} \int_0^{2\pi} r^2 d\varphi = \frac{2m}{\ell} S \quad (4.21)$$

où S est l'aire totale de l'orbite sur un tour complet.

4.A.3 Intégrale de la trajectoire

En appliquant la conservation de l'énergie et du moment cinétique, il est possible d'obtenir une expression intégrale qui permet en principe de déterminer la trajectoire de la particule, c'est-à-dire la relation entre φ et r . En effet, à partir de l'expression de l'énergie

$$E = \frac{1}{2}m\dot{r}^2 + \frac{\ell^2}{2mr^2} + V(r) \quad (4.22)$$

on isole \dot{r} :

$$\frac{dr}{dt} = \sqrt{\frac{2E}{m} - \frac{\ell^2}{m^2r^2} - \frac{2V(r)}{m}} \quad (4.23)$$

Comme

$$\ell = mr^2 \frac{d\varphi}{dt} \quad \text{et donc} \quad \frac{d\varphi}{dt} = \frac{\ell}{mr^2} \quad (4.24)$$

on obtient la relation différentielle suivante entre r et φ :

$$\frac{mr^2}{\ell} d\varphi = \frac{dr}{\sqrt{(2E/m) - (\ell^2/m^2r^2) - (2V(r)/m)}} \quad (4.25)$$

Effectuons maintenant le changement de variable $u = \ell/r$, tel que $du = -\ell dr/r^2$. On trouve alors la relation

$$d\varphi = -\frac{du}{\sqrt{2mE - u^2 - 2mV(\ell/u)}} \quad (4.26)$$

qu'il suffit en principe d'intégrer :

$$\varphi = - \int \frac{du}{\sqrt{2mE - u^2 - 2mV(\ell/u)}} + \text{cte} \quad (4.27)$$

Le reste dépend de la forme explicite de $V(r)$. Bien sûr, ceci suppose que $\ell \neq 0$. On sait par contre que φ est constant si $\ell = 0$.

4.A.4 Théorème du viriel

Considérons un ensemble de N particules de positions \mathbf{r}_i et d'impulsions \mathbf{p}_i . On définit le *viriel* G comme suit :

$$G = \sum_i \mathbf{r}_i \cdot \mathbf{p}_i \quad (4.28)$$

Calculons la dérivée par rapport au temps du viriel :

$$\begin{aligned} \frac{dG}{dt} &= \sum_i \dot{\mathbf{r}}_i \cdot \mathbf{p}_i + \sum_i \mathbf{r}_i \cdot \dot{\mathbf{p}}_i \\ &= \sum_i m_i \mathbf{v}_i^2 + \sum_i \mathbf{r}_i \cdot \mathbf{f}_i \\ &= 2T + \sum_i \mathbf{r}_i \cdot \mathbf{f}_i \end{aligned} \quad (4.29)$$

où T est l'énergie cinétique du système. Calculons ensuite la moyenne de cette relation sur un temps τ très long :

$$2\langle T \rangle + \sum_i \langle \mathbf{r}_i \cdot \mathbf{f}_i \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt \frac{dG}{dt} = \lim_{\tau \rightarrow \infty} \frac{G(\tau) - G(0)}{\tau} \quad (4.30)$$

Si les particules du système sont confinées dans un espace fini, soit en raison d'une contrainte (par exemple un gaz dans un contenant) ou parce que le système est lié par une attraction (par exemple le système solaire), alors le viriel G ne peut pas excéder une certaine borne en tout temps, ce qui entraîne que le membre de droite de l'équation ci-dessus est nul dans la limite $\tau \rightarrow \infty$. On peut donc affirmer la relation suivante entre les valeurs moyennes :

$$\langle T \rangle = -\frac{1}{2} \sum_i \langle \mathbf{r}_i \cdot \mathbf{f}_i \rangle \quad (4.31)$$

Ceci est la forme générale du *théorème du viriel*.

Supposons maintenant, comme cas particulier, que les particules du système interagissent via une force centrale en loi de puissance, tirée d'un potentiel $V(r) = kr^\alpha$.

$$\mathbf{f}_i = -k \frac{\partial}{\partial \mathbf{r}_i} \sum_{\substack{j \\ (j \neq i)}} |\mathbf{r}_i - \mathbf{r}_j|^\alpha = -k\alpha \sum_{\substack{j \\ (j \neq i)}} (\mathbf{r}_i - \mathbf{r}_j) |\mathbf{r}_i - \mathbf{r}_j|^{\alpha-2} \quad (4.32)$$

Dans ce cas,

$$\sum_i \mathbf{r}_i \cdot \mathbf{f}_i = -k\alpha \sum_{i \neq j} \mathbf{r}_i \cdot (\mathbf{r}_i - \mathbf{r}_j) |\mathbf{r}_i - \mathbf{r}_j|^{\alpha-2} = k\alpha \sum_{i \neq j} \mathbf{r}_j \cdot (\mathbf{r}_i - \mathbf{r}_j) |\mathbf{r}_i - \mathbf{r}_j|^{\alpha-2} \quad (4.33)$$

Après la dernière égalité, nous avons échangé les rôles de i et j . En prenant la moyenne des deux membres de cette égalité, on retrouve la même quantité sous la forme suivante :

$$-\frac{1}{2} k\alpha \sum_{i \neq j} (\mathbf{r}_i - \mathbf{r}_j)^2 |\mathbf{r}_i - \mathbf{r}_j|^{\alpha-2} = -\frac{1}{2} k\alpha \sum_{i \neq j} |\mathbf{r}_i - \mathbf{r}_j|^\alpha \quad (4.34)$$

Cette dernière expression n'est rien d'autre que $-\alpha V$, où V est maintenant l'énergie potentielle totale :

$$V = \frac{1}{2} \sum_{i \neq j} k |\mathbf{r}_i - \mathbf{r}_j|^\alpha \quad (4.35)$$

Le facteur $\frac{1}{2}$ provient de ce qu'on doit compter chaque paire de particules une seule fois. Le théorème du viriel peut donc s'écrire comme suit dans ce cas :

$$\langle T \rangle = \frac{1}{2} \alpha \langle V \rangle \quad (4.36)$$

Dans le cas encore plus particulier du potentiel gravitationnel, $\alpha = -1$, et

$$\langle T \rangle = -\frac{1}{2} \langle V \rangle \quad (\text{potentiel en } 1/r) \quad (4.37)$$

Comme l'énergie totale est conservée, $E = \langle E \rangle = \langle T \rangle + \langle V \rangle$ et on peut récrire ce résultat comme

$$E = \frac{1}{2} \langle V \rangle = -\langle T \rangle \quad (4.38)$$

Ce résultat s'applique bien sûr au cas particulier d'une seule paire de particules. Dans ce cas, il suffit de prendre la moyenne sur une seule période de l'orbite puisque le mouvement est périodique.

B Le problème de Kepler

4.B.1 Sections coniques

Poursuivons notre analyse de la trajectoire dans le cas d'un potentiel de la forme

$$V(r) = -\frac{k}{r} \quad (4.39)$$

où, par exemple, $k = Gm_1m_2$ dans le cas du problème à deux corps en gravitation. On calcule sans peine que la position du minimum du potentiel effectif, ainsi que sa valeur, sont données par

$$r_0 = \frac{\ell^2}{mk} \quad E_0 = -\frac{mk^2}{2\ell^2} \quad (4.40)$$

On définit aussi la valeur correspondante de u :

$$u_0 = \frac{\ell}{r_0} = \frac{mk}{\ell} \quad (4.41)$$

D'après l'éq. (4.27), la relation entre φ et r est

$$\varphi = - \int \frac{du}{\sqrt{2mE - u^2 + 2uu_0}} + \text{cte} \quad (4.42)$$

Complétons le carré dans la racine :

$$2mE - u^2 + 2uu_0 = 2mE + u_0^2 - (u - u_0)^2 = (eu_0)^2 - (u - u_0)^2 \quad (4.43)$$

où on a défini

$$e = \sqrt{1 + \frac{2mE}{u_0^2}} = \sqrt{1 + \frac{2E\ell^2}{mk^2}} = \sqrt{1 - \frac{E}{E_0}} \quad (4.44)$$

L'intégrale est alors

$$\varphi = - \int \frac{du}{\sqrt{(eu_0)^2 - (u - u_0)^2}} = - \arccos \frac{u - u_0}{eu_0} + \text{cte} \quad (4.45)$$

Le choix de la constante d'intégration dépend de l'origine des φ (c'est-à-dire la direction de l'axe des x). On peut toujours choisir cet axe de manière à ce que cette constante s'annule, de sorte que

$$e \cos \varphi = \frac{u}{u_0} - 1 \implies u = u_0 (e \cos \varphi + 1) \quad (4.46)$$

ou encore

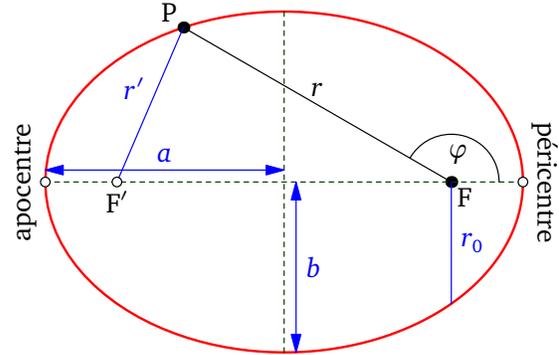
$$r(\varphi) = \frac{r_0}{1 + e \cos \varphi} \quad (4.47)$$

Il s'agit de l'équation d'une section conique d'excentricité e dont l'un des foyers est à l'origine.

Les types de trajectoires suivants sont possibles, selon la valeur de e :

1. $E = E_0$: l'excentricité est nulle. L'orbite est circulaire.
2. $E_0 < E < 0$: l'excentricité est comprise entre 0 et 1. L'orbite est elliptique.
3. $E = 0$: l'excentricité est $e = 1$. L'orbite est parabolique.
4. $E > 0$: l'excentricité est supérieure à 1. L'orbite est hyperbolique.

FIGURE 4.2
Paramètres géométriques de l'ellipse.



Orbite elliptique

L'ellipse est souvent définie comme l'ensemble des points P tels que la somme des distances entre P et deux points fixes appelés foyers (F et F') est constante. Voyons à cet effet la figure 4.2. Cette définition s'écrit alors

$$r + r' = 2a \quad (4.48)$$

où $2a$ est la largeur de l'ellipse (a est appelé le demi-grand axe). Cette définition de l'ellipse entraîne que la courbe est symétrique par rapport à une inversion gauche-droite, car les deux foyers F et F' y jouent des rôles identiques. Appelons c la distance entre chaque foyer et le centre de l'ellipse ; la distance entre F et F' est donc $2c$. La loi des cosinus nous permet d'exprimer r' en fonction de r et de φ :

$$r'^2 = r^2 + 4c^2 - 4rc \cos(\pi - \varphi) = r^2 + 4c^2 + 4rc \cos \varphi \quad (4.49)$$

Comme $r' = 2a - r$, on obtient alors

$$(2a - r)^2 = 4a^2 - 4ar + r^2 = r^2 + 4c^2 + 4rc \cos \varphi \implies r(c \cos \varphi + a) = a^2 - c^2 \quad (4.50)$$

ou encore

$$r = \frac{a^2 - c^2}{a + c \cos \varphi} \quad (4.51)$$

On constate que cette relation est exactement de la forme (4.47), avec la correspondance suivante :

$$c = ea \quad r_0 = \frac{a^2 - c^2}{a} = a(1 - e^2) \quad (4.52)$$

Le demi-petit axe b de l'ellipse est défini comme la moitié de sa hauteur en son centre, et s'obtient immédiatement du théorème de Pythagore quand $r = r' = a$:

$$r^2 = a^2 = c^2 + b^2 \implies b = \sqrt{a^2 - c^2} = a\sqrt{1 - e^2} \quad (4.53)$$

Notons que r_0 est la distance entre le foyer principal de l'ellipse et la position de l'objet en quadrature (c.-à-d. à $\varphi = \pm\pi/2$). D'autre part, comme $a > c$, l'excentricité e est nécessairement comprise entre

0 et 1. On retrouve manifestement un cercle quand $F = F'$ (et donc $c = 0$), ce qui correspond à une excentricité nulle ($e = 0$). À l'autre extrême, une ellipse infiniment aplatie correspond à $b = 0$, donc à $e = 1$.

En fonction des paramètres physiques E et ℓ , les paramètres géométriques de l'ellipse s'expriment ainsi :

$$\begin{aligned} e &= \sqrt{1 + \frac{2E\ell^2}{mk^2}} \\ a &= \frac{r_0}{1-e^2} = -\frac{k}{2E} \\ b &= a\sqrt{1-e^2} = \frac{\ell}{\sqrt{-2mE}} \end{aligned} \quad (4.54)$$

La période de l'orbite elliptique s'obtient directement de la formule (4.21) :

$$\tau = \frac{2m}{\ell} S = \frac{2m}{\ell} \pi ab = \pi \frac{2m}{\ell} \frac{k}{2|E|} \frac{\ell}{\sqrt{2m|E|}} \quad (4.55)$$

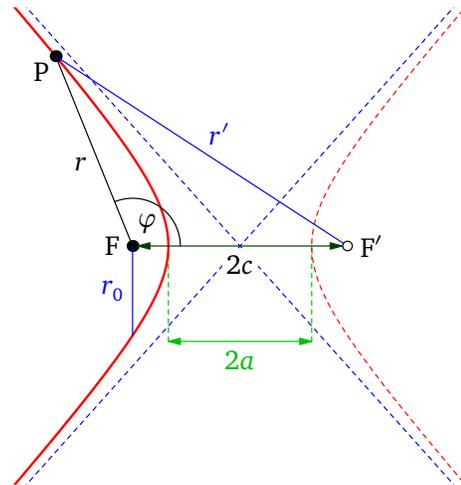
ou enfin

$$\tau = \pi k \sqrt{\frac{m}{2|E|^3}} \quad (4.56)$$

Dans le cas du problème gravitationnel, où $k = GMm$, on obtient la troisième loi de Kepler :

$$\tau = 2\pi \sqrt{\frac{a^3}{GM}} \quad \text{ou} \quad \tau^2 = \frac{(2\pi)^2}{GM} a^3 \quad (4.57)$$

FIGURE 4.3
Paramètres d'une orbite hyperbolique.



Orbite hyperbolique

L'hyperbole est définie comme l'ensemble des points P tels que la *différence* des distances entre P et deux points fixes appelés foyers (F et F') est constante, que nous appellerons $2a$ et que nous supposerons positive. Voyons à cet effet la figure 4.3. Cette définition s'écrit alors

$$r' - r = 2a \quad (4.58)$$

Appelons toujours $2c$ la distance entre les deux foyers. Notons premièrement que la droite verticale bissectrice entre les deux foyers étant l'ensemble des points tels que $r' = r$, la courbe ne peut pas traverser cette droite. La loi des cosinus nous permet encore d'exprimer r' en fonction de r et de φ :

$$r'^2 = r^2 + 4c^2 - 4rc \cos \varphi \quad (4.59)$$

Comme $r' = 2a + r$, on obtient alors

$$(2a + r)^2 = 4a^2 + 4ar + r^2 = r^2 + 4c^2 - 4rc \cos \varphi \implies r(c \cos \varphi + a) = c^2 - a^2 \quad (4.60)$$

ou encore

$$r = \frac{c^2 - a^2}{a + c \cos \varphi} \quad (4.61)$$

Cette fois encore cette relation est exactement de la forme (4.47), avec cette fois la correspondance suivante :

$$c = ea \quad r_0 = \frac{c^2 - a^2}{a} = a(e^2 - 1) \quad (4.62)$$

On constate que $e > 1$. Dans le cas extrême $e \rightarrow 1$ (ou $c \rightarrow a$), l'hyperbole est infiniment aplatie et située le long de l'axe horizontal à gauche du foyer principal. L'autre extrême ($e \rightarrow \infty$) correspond à $a/c \rightarrow 0$, donc à une courbe qui tend vers la droite bissectrice.

Notons qu'en supposant que a est négatif (donc $r' < r$), on obtiendrait plutôt la courbe pointillée sur la figure, soit l'autre branche de l'hyperbole, obtenue en échangeant les rôles des deux foyers. Cette branche correspond formellement à une valeur négative de r_0 , donc une valeur négative de k , donc à un potentiel répulsif, comme nous le verrons plus bas.

4.B.2 Équation de Kepler

Nous avons démontré que la trajectoire d'un objet soumis à un potentiel attractif $-k/r$ est une ellipse ou une hyperbole. Mais nous n'avons pas obtenu d'information temporelle sur le mouvement, c'est-à-dire une expression précise du mouvement $r(t)$ et $\varphi(t)$. Kepler a trouvé une relation mathématique particulièrement ingénieuse qui permet d'obtenir l'angle φ en fonction du temps écoulé dans une orbite elliptique ; il s'agit de la célèbre équation de Kepler, que nous allons maintenant expliquer et démontrer.

On commence par définir l'*anomalie excentrique* E , selon la figure 4.4 : Le centre d'attraction étant au foyer F et l'objet au point P , traçons un cercle de rayon a concentrique à l'ellipse et traçons une droite RQ parallèle au petit axe de l'ellipse et passant par le point P . Cette droite croise le cercle au point Q . L'anomalie excentrique E est l'angle $\angle RCQ$.

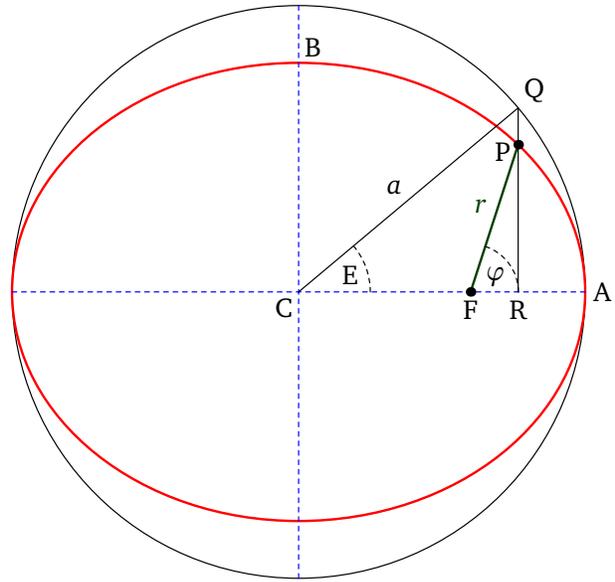
Démontrons la relation suivante entre l'anomalie excentrique et l'angle φ (φ est appelé *anomalie vraie* dans le contexte astronomique) :

$$\tan \frac{E}{2} = \sqrt{\frac{1-e}{1+e}} \tan \frac{\varphi}{2} \quad (4.63)$$

La démonstration est géométrique, comme au temps de Kepler :

$$FR = CR - CF = a \cos E - ae \quad (4.64)$$

FIGURE 4.4
Diagramme associé à l'équation de Kepler.



mais $FR = r \cos \varphi$ et donc

$$r \cos \varphi = a(\cos E - e) \quad (4.65)$$

D'autre part, en vertu d'une propriété simple de l'ellipse et de son cercle concentrique,

$$\frac{PR}{QR} = \frac{b}{a} = \sqrt{1 - e^2} \quad (4.66)$$

donc

$$\frac{r \sin \varphi}{a \sin E} = \sqrt{1 - e^2} \implies r \sin \varphi = a \sqrt{1 - e^2} \sin E \quad (4.67)$$

En mettant au carré et en additionnant le carré de (4.65), on trouve

$$r^2 = a^2(\cos^2 E + e^2 - 2e \cos E + [1 - e^2] \sin^2 E) \implies r = a(1 - e \cos E) \quad (4.68)$$

Par une relation trigonométrique standard, et en utilisant (4.65) et (4.68), on trouve

$$\begin{aligned} 2r \sin^2(\varphi/2) &= r(1 - \cos \varphi) = a(1 - e \cos E) - a(\cos E - e) \\ &= a(1 + e)(1 - \cos E) = 2a(1 + e) \sin^2(E/2) \end{aligned} \quad (4.69)$$

De même,

$$\begin{aligned} 2r \cos^2(\varphi/2) &= r(1 + \cos \varphi) = a(1 - e \cos E) + a(\cos E - e) \\ &= a(1 - e)(1 + \cos E) = 2a(1 - e) \cos^2(E/2) \end{aligned} \quad (4.70)$$

La racine carrée du rapport des deux dernières équations donne bien

$$\tan \frac{\varphi}{2} = \sqrt{\frac{1 + e}{1 - e}} \tan \frac{E}{2} \quad (4.71)$$

Nous allons maintenant démontrer l'équation de Kepler proprement dite. Par la deuxième loi de Kepler (la loi des aires),

$$\frac{\text{aire(FPA)}}{\text{aire de l'ellipse}} = \frac{t}{\tau} \quad (4.72)$$

où t est le temps écoulé depuis le passage au périhélie et τ la période de l'orbite. Exprimé autrement,

$$\text{aire(FPA)} = \frac{\pi abt}{\tau} = \frac{1}{2} abM \quad (4.73)$$

où M est l'*anomalie moyenne*, le temps écoulé depuis le passage au périhélie, exprimé de manière angulaire, c'est-à-dire en fraction de période ($\times 2\pi$).

L'aire FPA peut se calculer ainsi :

$$\text{aire(FPA)} = \text{aire(FRP)} + \text{aire(ARP)} \quad (4.74)$$

D'autre part,

$$\text{aire(ARP)} = \frac{b}{a} \text{aire(ARQ)} \quad (4.75)$$

ce qui se démontre en divisant l'aire en languettes verticales infinitésimales.¹ Aussi,

$$\text{aire(ARQ)} = \text{aire(ACQ)} - \text{aire(RCQ)} = \frac{1}{2} a^2 E - \frac{1}{2} a^2 \sin E \cos E \quad (4.76)$$

Comme $\text{aire(FRP)} = \frac{1}{2} r^2 \sin \varphi \cos \varphi$, on trouve

$$\begin{aligned} \text{aire(FPA)} &= \frac{1}{2} r^2 \sin \varphi \cos \varphi + \frac{1}{2} ab(E - \sin E \cos E) \\ &= \frac{1}{2} a^2 \sqrt{1-e^2} (\cos E - e) \sin E + \frac{1}{2} ab(E - \sin E \cos E) \\ &= \frac{1}{2} ab(E - e \sin E) \\ &= \frac{1}{2} abM \end{aligned} \quad (4.77)$$

et donc

$$M = E - e \sin E \quad (4.78)$$

Il s'agit de l'équation de Kepler, qui nous permet, en fonction de M (donc du temps t) de calculer E et ensuite φ (par l'éq. (4.63)).

Notons que, dans le cas d'une orbite circulaire, les trois anomalies (vraie, excentrique et moyenne) se confondent, mais elles sont toutes les trois différentes dans une orbite elliptique (*anomalie* est simplement un synonyme d'*angle* dans ce contexte).

L'équation de Kepler (4.78) est transcendante et ne peut être résolue que numériquement. Dans le cas de petites excentricités, la méthode de récurrence est la plus indiquée. On récrit l'équation de Kepler ainsi :

$$E = M + e \sin E \quad (4.79)$$

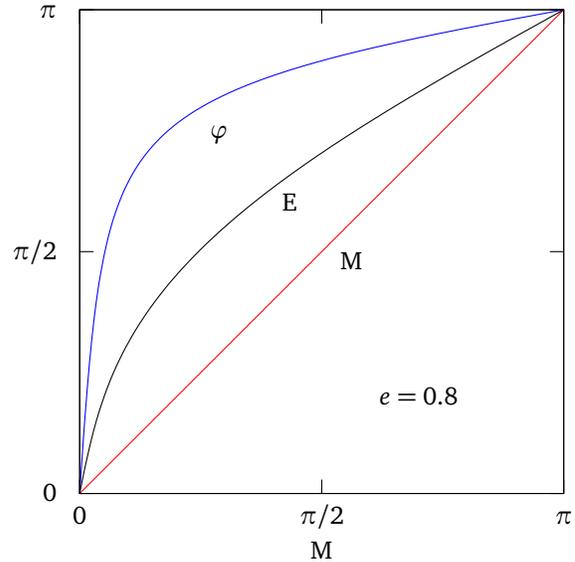
On procède ensuite par approximations successives, en posant comme première solution approchée $E_0 = M$. Ensuite, on substitue cette solution dans l'équation et on recommence le processus. On obtient ainsi une suite de solutions approchées, qui converge vers la solution exacte :

$$\begin{aligned} E_0 &= M \\ E_1 &= M + e \sin E_0 = M + e \sin M \\ E_2 &= M + e \sin E_1 = M + e \sin(M + e \sin M) \\ &\text{etc.} \end{aligned} \quad (4.80)$$

1. Cette technique fait de Kepler l'un des précurseurs du calcul intégral.

FIGURE 4.5

Solution numérique de l'équation de Kepler pour une excentricité $e = 0.8$, présentée pour $M \in [0, \pi]$ (la solution pour $M \in [-\pi, 0]$ s'obtient par réflexion : les fonctions $\varphi(M)$ et $E(M)$ sont impaires). On remarque que φ varie rapidement près du périhélie ($M = 0$) et lentement près de l'apogée ($M = \pi$).



Cet algorithme, par sa nature itérative, est particulièrement adapté à une solution par ordinateur. Un graphique illustrant les valeurs relatives de M , E et φ est présenté à la figure 4.5.

Une expression de l'anomalie vraie φ en fonction de l'anomalie moyenne M existe aussi sous la forme d'une série de Fourier dont les coefficients sont des fonctions de Bessel. On démontre que

$$\cos \varphi = -e + \frac{2(1-e^2)}{e} \sum_{k=1}^{\infty} J_k(ke) \cos(kM) \quad (4.81)$$

4.B.3 Repérage des orbites dans l'espace

Il est généralement nécessaire de repérer les orbites dans l'espace, sans que le moment cinétique orbital ne pointe dans la direction z , ou que le périhélie de l'orbite soit dans la direction x . Pour ce faire, le plan de l'orbite, et la direction du périhélie dans ce plan sont repérés dans l'espace par des angles d'Euler. Ces angles ne sont cependant pas notés par les mêmes symboles que dans le contexte du mouvement des corps rigides. Voir à cet effet la figure 4.6. On a plutôt la correspondance suivante :

$$\theta \rightarrow i \quad \phi \rightarrow \Omega \quad \psi \rightarrow \omega \quad (4.82)$$

L'angle entre le plan de l'orbite et le plan équatorial est l'*inclinaison* i de l'orbite. La *ligne des noeuds* est l'intersection de ces deux plans. La *longitude du noeud ascendant* Ω est l'angle entre une direction de référence sur le plan équatorial (par exemple le point vernal dans le cas des orbites autour du soleil) et la ligne des noeuds, plus précisément le point où l'orbite traverse le plan équatorial vers le haut. L'angle ω entre la ligne des noeuds et le périhélie de l'orbite est appelé *argument du périhélie*. L'angle φ entre la position réelle de l'objet et le rayon vecteur du périhélie est l'*anomalie vraie*. Pour trouver la position précise de l'objet à un moment donné, il faut également spécifier l'origine des temps, soit le moment précis τ_p où l'objet est passé au périhélie. L'ensemble des six quantités

$$i, \Omega, \omega, a, e, \tau_p \quad (4.83)$$

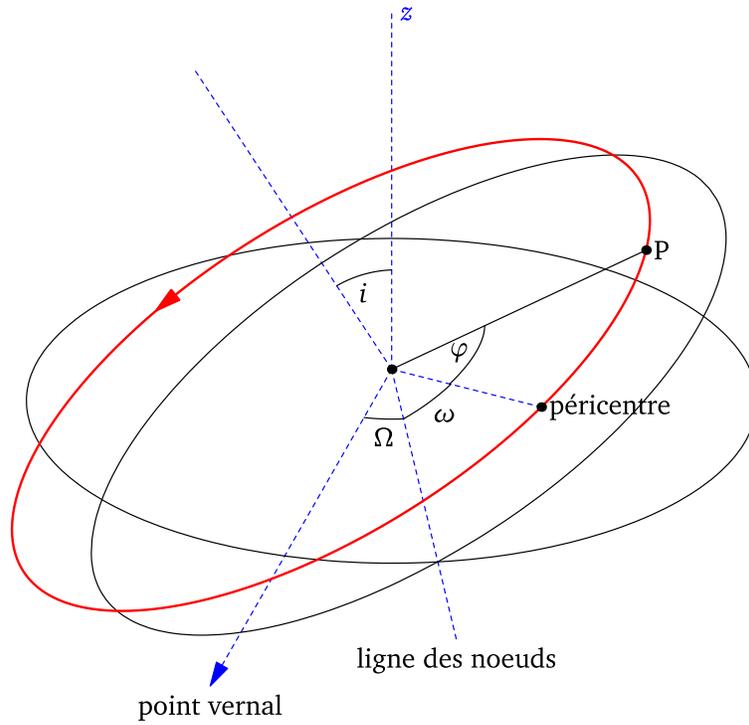


FIGURE 4.6

forment ce qu'on appelle les *éléments* de l'orbite elliptique et permettent en principe de trouver la position précise d'un objet (planète, astéroïde, satellite, etc.) dans l'espace, à tout instant.

C Théorie classique de la diffusion par un potentiel central

4.C.1 Section efficace

Considérons maintenant la situation générale d'une collision entre deux corps résultant d'une force centrale. Nous travaillerons dans le cadre du problème à un corps équivalent, donc dans celui de l'interaction d'une particule ayant la masse réduite m se déplaçant dans un potentiel central $V(r)$. Supposons que la particule se dirige initialement dans la direction \mathbf{z} . Après la collision, la particule est déviée dans une direction spécifiée par les angles polaires (φ, θ) . Ce processus de collision est généralement appelé *diffusion*, en raison de la dualité onde-corpuscule : la déviation d'une particule par un obstacle correspond, pour une onde, à une diffusion progressive. L'origine est alors appelée le *centre de diffusion*, et l'angle de déviation θ est appelé *angle de diffusion*.

La distance b entre la particule incidente et l'axe des z est appelé *paramètre d'impact*. L'angle de diffusion θ est une fonction de b , déterminée par la forme précise du potentiel de diffusion $V(r)$.

Un concept très important dans la physique des particules élémentaires et la physique nucléaire est celui de *section efficace*, que nous allons maintenant définir. Supposons qu'un flux quasi continu de projectiles, tous de même quantité de mouvement, soit incident sur le centre de diffusion (ou *cible*). On supposera que les particules dans ce flux sont distribuées de manière uniforme dans le plan perpendiculaire à l'axe z , très loin de la cible. On notera Φ le nombre de particules par unité de temps qui traversent une unité de surface dans ce plan. Le nombre de particules incidentes sur la cible par unité de temps, avec un paramètre d'impact situé entre b et $b + db$, est donc $2\pi b\Phi db$.

La *section différentielle de diffusion* est définie comme le nombre de particules diffusées dans la direction (θ, φ) , par unité de temps et d'angle solide, divisé par le flux incident Φ :

$$\frac{d\sigma}{d\Omega} = \frac{\# \text{ de particules diffusées vers } (\theta, \varphi) \text{ par sec.}}{\text{flux incident} \times \text{angle solide } d\Omega} \quad (4.84)$$

La section différentielle est une fonction des angles (θ, φ) et possède les unités d'une surface. La *section efficace* σ est alors le nombre de particules déviées par unité de temps, divisé par le flux incident, soit l'intégrale sur les angles de la section différentielle :

$$\sigma = \int d\Omega \frac{d\sigma}{d\Omega} \quad (d\Omega = |\sin \theta d\theta d\varphi| = |d \cos \theta, d\varphi|) \quad (4.85)$$

Autrement dit, le nombre de particules déviées par unité de temps est $\sigma\Phi$.

La relation entre le paramètre d'impact et l'angle θ mène directement à une expression de la section différentielle : Considérons l'ensemble des particules qui ont un paramètre d'impact compris entre b et $b + db$. Le nombre de particules dans cette situation, par unité de temps, est $\Phi 2\pi b db$, soit l'aire d'un anneau de rayon b et de largeur db , fois le flux incident de particules. Toutes ces particules seront déviées dans un anneau solide situé entre θ et $\theta + d\theta$, correspondant à un angle solide total

de $2\pi \sin \theta |d\theta|$ (la valeur absolue est nécessaire, car $d\theta < 0$ si $db > 0$). Par définition de la section différentielle,

$$\Phi 2\pi b |db| = \Phi \frac{d\sigma}{d\Omega} 2\pi \sin \theta |d\theta| \implies \boxed{\frac{d\sigma}{d\Omega} = \frac{b}{\sin \theta} \left| \frac{db}{d\theta} \right|} \quad (4.86)$$

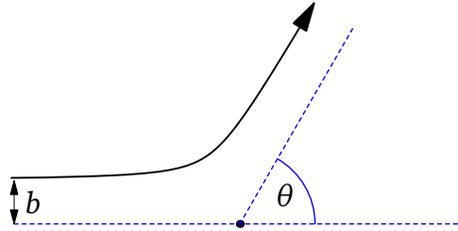


FIGURE 4.7

Pour comprendre l'origine de l'expression *section efficace*, considérons le potentiel d'une *sphère dure*, c'est-à-dire

$$V(r) = \begin{cases} \infty & (r < a) \\ 0 & (r > a) \end{cases} \quad (\text{sphère dure}) \quad (4.87)$$

C'est le potentiel associé à une répulsion de contact avec une sphère de rayon a . Dans ce cas, le projectile n'entre en collision que si le paramètre d'impact est inférieur à a , et donc le nombre de collisions par unité de temps est $\pi a^2 \Phi$. La section efficace est donc $\sigma = \pi a^2$, égale à la section géométrique de la cible. Dans un problème plus général (sans interaction de contact), la section efficace nous indique la capacité d'un diffuseur à dévier les particules incidentes, en donnant la superficie équivalente d'un objet qui diffuserait uniquement par contact.

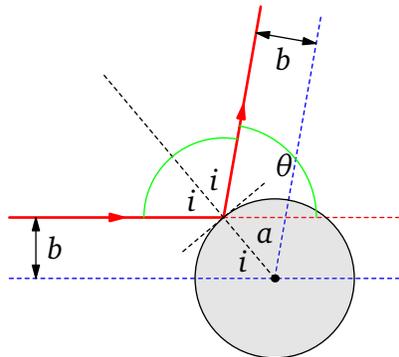


FIGURE 4.8

Exemple 4.1 Diffusion par une sphère dure

Calculons la section différentielle de diffusion dans le cas du potentiel (4.87). La réflexion de la particule sur la surface de la sphère se fait de telle manière que les deux angles marqués i sur la figure 4.8 sont égaux. Cette propriété est la seule façon de conserver le moment cinétique, car

elle s'assure que le paramètre d'impact est le même avant et après la collision ou, dit autrement, que la collision inverse est également possible. Il s'ensuit que l'angle de diffusion est $\theta = \pi - 2i$. D'autre part, la relation entre b et i s'obtient par trigonométrie élémentaire :

$$b = a \sin i = a \sin(\pi/2 - \theta/2) = a \cos(\theta/2) \quad (4.88)$$

Donc

$$\left| \frac{db}{d\theta} \right| = \frac{1}{2} a \sin(\theta/2) \quad (4.89)$$

et, d'après la relation (4.86),

$$\frac{d\sigma}{d\Omega} = \frac{b}{\sin \theta} \left| \frac{db}{d\theta} \right| = \frac{1}{2} a^2 \frac{\cos(\theta/2) \sin(\theta/2)}{\sin \theta} = \frac{1}{4} a^2, \quad (4.90)$$

soit une constante indépendante des angles. La section efficace est alors donnée par l'angle solide total (4π) multiplié par la section différentielle, soit $\sigma = \pi a^2$, c'est-à-dire la section géométrique.

4.C.2 Formule de Rutherford

Considérons maintenant le cas d'une particule de charge e_1 diffusant sur une particule de charge e_2 . Le potentiel $V(r)$ prend alors la forme

$$V(r) = \frac{e_1 e_2}{4\pi \epsilon_0 r} = \frac{k}{r} \quad (4.91)$$

La trajectoire du problème effectif à une particule s'obtient comme dans le cas du potentiel gravitationnel attractif, et l'équation (4.45) peut encore être appliquée, si on garde en tête que le paramètre u_0 est maintenant $u_0 = -mk/\ell$, car le signe du potentiel central est différent. Donc on a

$$\varphi = - \int \frac{du}{\sqrt{(eu_0)^2 - (u - u_0)^2}} = - \arccos \frac{u - u_0}{e|u_0|} + \text{cte} \quad (4.92)$$

ou encore, en adoptant une constante d'intégration nulle,

$$e \cos \varphi = \frac{u - u_0}{|u_0|} \implies r(\varphi) = \frac{\ell^2/mk}{e \cos \varphi - 1} \quad (4.93)$$

Cette forme diffère de l'éq. (4.47), mais on retrouve la forme précédente si on remplace φ par $\varphi + \pi$ et qu'on change le signe de k . Notons que, dans ce cas, l'excentricité (4.44) est toujours supérieure à un, l'énergie étant toujours positive. La trajectoire est donc hyperbolique, sauf qu'elle suit plutôt le tracé pointillé de la figure 4.3, correspondant à l'autre branche de l'hyperbole.

Le paramètre d'impact b est la distance entre le foyer de l'hyperbole et l'asymptote. Si v_∞ désigne la vitesse du projectile à l'infini, le moment cinétique s'exprime en fonction du paramètre d'impact et de v_∞ comme suit :

$$\ell = m v_\infty b \quad (4.94)$$

C. Théorie classique de la diffusion par un potentiel central

L'énergie totale étant donnée par $E = \frac{1}{2}mv_{\infty}^2$, l'excentricité e peut alors s'exprimer comme

$$e = \sqrt{1 + \frac{2E\ell^2}{mk^2}} = \sqrt{1 + \left(\frac{2Eb}{k}\right)^2} \quad \text{ou} \quad \frac{2Eb}{k} = \sqrt{e^2 - 1} \quad (4.95)$$

Soit ensuite φ_{∞} l'angle associé à l'asymptote de l'hyperbole. D'après l'éq. (4.93), cet angle est $\varphi_{\infty} = \arccos(1/e)$. L'angle de diffusion θ est l'angle entre les deux asymptotes et s'obtient de φ_{∞} par la relation $2\varphi_{\infty} + \theta = \pi$, ou encore $\varphi_{\infty} = \pi/2 - \theta/2$. Ceci nous permet de relier θ à b :

$$\cos \varphi_{\infty} = \sin(\theta/2) = \frac{1}{e} \implies \sqrt{e^2 - 1} = \sqrt{\frac{1}{\sin^2(\theta/2)} - 1} = \cot(\theta/2) \quad (4.96)$$

ou enfin

$$\cot(\theta/2) = \frac{2Eb}{k} \quad (4.97)$$

La section différentielle de diffusion s'obtient alors directement de la formule (4.86) :

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{b}{\sin \theta} \left| \frac{db}{d\theta} \right| \\ &= \left(\frac{k}{2E}\right)^2 \frac{1}{2} \frac{\cot(\theta/2)}{\sin \theta} \frac{1}{\sin^2(\theta/2)} \\ &= \left(\frac{k}{4E}\right)^2 \frac{1}{\sin^4(\theta/2)} \end{aligned} \quad (4.98)$$

Comme $k = e_1 e_2 / 4\pi\epsilon_0$, on peut enfin écrire la formule de Rutherford :

$$\frac{d\sigma}{d\Omega} = \left(\frac{e_1 e_2}{4\pi\epsilon_0}\right)^2 \frac{1}{16E^2} \operatorname{cosec}^4 \frac{\theta}{2} \quad (4.99)$$

D Forces centrales et équation de Hamilton-Jacobi

4.D.1 Équation de Hamilton-Jacobi dans le cas d'un potentiel azimutal

Voyons maintenant comment l'équation de la trajectoire peut être obtenue par la méthode de Hamilton-Jacobi. Considérons pour commencer un problème plus général : celui d'une particule se déplaçant dans un potentiel de la forme

$$V(r, \theta) = U(r) + \frac{g(\theta)}{r^2} \quad (4.100)$$

en trois dimensions, où r et θ sont les coordonnées sphériques. On suppose que le potentiel est indépendant de φ , de sorte que le problème possède une symétrie de rotation par rapport à l'axe polaire. L'équation de Hamilton-Jacobi réduite (2.129) prend alors la forme

$$\frac{1}{2m} \left(\frac{\partial W}{\partial r} \right)^2 + \frac{1}{2mr^2} \left(\frac{\partial W}{\partial \theta} \right)^2 + \frac{1}{2mr^2 \sin^2 \theta} \left(\frac{\partial W}{\partial \varphi} \right)^2 + U(r) + \frac{g(\theta)}{r^2} = E \quad (4.101)$$

Nous allons poser encore une fois que la solution est séparable, c'est-à-dire la somme de trois fonctions indépendantes :

$$W = W_1(r) + W_2(\theta) + W_3(\varphi) \quad (4.102)$$

L'équation devient alors

$$\frac{1}{2m} \left(\frac{dW_1}{dr} \right)^2 + \frac{1}{2mr^2} \left(\frac{dW_2}{d\theta} \right)^2 + \frac{1}{2mr^2 \sin^2 \theta} \left(\frac{dW_3}{d\varphi} \right)^2 + U(r) + \frac{g(\theta)}{r^2} = E \quad (4.103)$$

En multipliant par $2mr^2 \sin^2 \theta$, on trouve

$$r^2 \sin^2 \theta \left(\frac{dW_1}{dr} \right)^2 + \sin^2 \theta \left(\frac{dW_2}{d\theta} \right)^2 + \left(\frac{dW_3}{d\varphi} \right)^2 + 2mr^2 \sin^2 \theta \left(U(r) + \frac{g(\theta)}{r^2} - E \right) = 0 \quad (4.104)$$

Le terme en W_3 étant le seul qui dépend de φ , ce terme doit nécessairement être constant afin que l'équation soit satisfaite pour toutes les valeurs de (r, θ, φ) . On pose donc

$$\frac{dW_3}{d\varphi} = p_\varphi = \ell_z = \text{cte} \quad (4.105)$$

Sachant cela, l'équation (4.103) devient, en la multipliant par $2mr^2$,

$$r^2 \left(\frac{dW_1}{dr} \right)^2 + \left\{ \left(\frac{dW_2}{d\theta} \right)^2 + \frac{\ell_z^2}{\sin^2 \theta} + 2mg(\theta) \right\} + 2mr^2 U(r) - 2mr^2 E = 0 \quad (4.106)$$

Cette fois, ce sont les termes entre accolades qui doivent former une constante, car la variable θ n'apparaît pas ailleurs. On peut donc poser

$$\left(\frac{dW_2}{d\theta} \right)^2 + \frac{\ell_z^2}{\sin^2 \theta} + 2mg(\theta) = \alpha_2^2 \quad (4.107)$$

et il reste alors l'équation suivante pour W_1 :

$$\left(\frac{dW_1}{dr}\right)^2 + \frac{\alpha_2^2}{r^2} + 2mU(r) = 2mE \quad (4.108)$$

Cette dernière équation peut aussi s'écrire ainsi, en fonction de la composante radiale de l'impulsion, p_r :

$$\frac{p_r^2}{2m} + \frac{\alpha_2^2}{2mr^2} + U(r) = E \quad (4.109)$$

ce qui nous permet d'identifier la constante α_2 avec la grandeur ℓ du moment cinétique orbital.

Nous avons donc réussi à séparer l'équation différentielle aux dérivées partielles en trois équations différentielles ordinaires, dont une triviale pour $W_3(\varphi)$. Formellement, nous avons les solutions suivantes :

$$\begin{aligned} W_1(r) &= \int dr \sqrt{2m(E - U(r)) - \frac{\ell^2}{r^2}} \\ W_2(\theta) &= \int d\theta \sqrt{\ell^2 - 2mg(\theta) - \frac{\ell_z^2}{\sin^2 \theta}} \end{aligned} \quad (4.110)$$

et la fonction caractéristique de Hamilton devient

$$W(r, \theta, \varphi) = \int dr \sqrt{2m(E - U(r)) - \frac{\ell^2}{r^2}} + \int d\theta \sqrt{\ell^2 - 2mg(\theta) - \frac{\ell_z^2}{\sin^2 \theta}} + \ell_z \varphi \quad (4.111)$$

Les nouvelles coordonnées constantes sont alors

$$\begin{aligned} t + \beta_1 &= \frac{\partial W}{\partial E} = m \int \frac{dr}{\sqrt{2m(E - U(r)) - (\ell^2/r^2)}} \\ \beta_2 &= \frac{\partial W}{\partial \ell} = - \int \frac{\ell dr}{r^2 \sqrt{2m(E - U(r)) - (\ell^2/r^2)}} + \int \frac{\ell d\theta}{\sqrt{\ell^2 - 2mg(\theta) - (\ell_z^2/\sin^2 \theta)}} \\ \beta_3 &= \frac{\partial W}{\partial \ell_z} = \varphi - \int \frac{d\theta \ell_z}{\sin^2 \theta \sqrt{\ell^2 - 2mg(\theta) - (\ell_z^2/\sin^2 \theta)}} \end{aligned} \quad (4.112)$$

On affirme alors que le problème a été réduit à des quadratures, c'est-à-dire à des intégrales qui peuvent être calculées, numériquement s'il le faut. En fait, ces relations sont utiles surtout si les intégrales sont possibles analytiquement, car il faut ensuite inverser ces relations pour trouver une expression de (r, θ, φ) en fonction des constantes E , ℓ et $\ell_z = p_\varphi$.

En supposant qu'on puisse procéder à cette inversion, nous aurions trois expressions

$$r = F_r(t + \beta_1, E, \ell) \quad \theta = F_\theta(r, \beta_2, E, \ell, \ell_z) \quad \varphi = \beta_3 + F_\varphi(\theta, \ell, \ell_z) \quad (4.113)$$

La première équation nous renseigne sur la dépendance temporelle de la coordonnée radiale. La deuxième sur la relation entre l'angle polaire θ et r . La troisième, enfin, donne φ en fonction des deux autres coordonnées. Les deux dernières équations donnent donc la forme de la trajectoire dans l'espace.

4.D.2 Cas d'un potentiel central

Supposons maintenant, pour simplifier le problème, que $g(\theta) = 0$, et donc que $V(r) = U(r)$. Dans ce cas, le problème étant purement central, les trois composantes du moment cinétique sont conservées. Il est alors judicieux de supposer que le mouvement est confiné au plan xy , et donc que le système ne comporte effectivement que deux degrés de liberté, avec les coordonnées (r, φ) , θ étant égal à $\pi/2$. L'équation d'Hamilton-Jacobi réduite est alors

$$\frac{1}{2m} \left(\frac{\partial W}{\partial r} \right)^2 + \frac{1}{2mr^2} \left(\frac{\partial W}{\partial \varphi} \right)^2 + V(r) = E \quad (4.114)$$

Comme avant, on procède par séparation des variables : $W = W_1(r) + W_3(\varphi)$ et la fonction W_3 doit être particulièrement simple : $W_3(\varphi) = \ell\varphi$. Il reste l'expression suivante pour W_1 :

$$W_1(r) = \int dr \sqrt{2m(E - V(r)) - \frac{\ell^2}{r^2}} \quad (4.115)$$

Les nouvelles coordonnées constantes sont alors

$$\begin{aligned} t + \beta_1 &= \frac{\partial W}{\partial E} = m \int \frac{dr}{\sqrt{2m(E - V(r)) - (\ell^2/r^2)}} \\ \beta_3 &= \frac{\partial W}{\partial \ell} = \varphi - \int \frac{\ell dr}{r^2 \sqrt{2m(E - V(r)) - (\ell^2/r^2)}} \end{aligned} \quad (4.116)$$

En supposant qu'on puisse calculer ces intégrales indéfinies, nous pourrions en principe inverser ces relations et obtenir deux expressions du type

$$r = F_r(t + \beta_1, E, \ell) \quad \varphi = \beta_3 + F_\varphi(r, E, \ell) \quad (4.117)$$

La dernière équation, en particulier, nous donne l'information sur l'orbite $r(\varphi)$, sans égard au temps écoulé. Elle est équivalente à l'équation (4.25) trouvée plus haut, si on pose la condition initiale $\beta_3 = 0$.

4.D.3 Variables d'action-angle

Considérons un système séparable dont chacune des coordonnées q_i varie dans un intervalle fini ou périodique. C'est le cas du problème général traité dans les sous-sections ci-dessus : la coordonnée radiale r varie entre deux points de rebroussement, et les autres coordonnées sont des angles dont le domaine de définition est périodique (de 0 à 2π). Comme le problème est séparable, la fonction caractéristique de Hamilton a la forme

$$W(\mathbf{q}, \boldsymbol{\alpha}) = \sum_i W_i(q_i, \alpha_i) \quad (4.118)$$

et les impulsions correspondantes sont $p_i = \partial W_i / \partial q_i$. Formellement, les termes de la fonction caractéristique sont les intégrales indéfinies suivantes :

$$W_i = \int p_i dq_i \quad (4.119)$$

Nous allons maintenant définir les *variables d'action* J_i comme suit :

$$J_i = \oint p_i dq_i = \oint \frac{\partial W_i}{\partial q_i} dq_i \quad (4.120)$$

où le symbole \oint dénote une intégrale *définie* prise sur une période complète de la variable q_i . Comme la fonction $W_i(q_i, \alpha_i)$ dépend aussi des constantes de séparation α_i , la variable d'action J_i sera exprimée en fonction des α_i , mais ne dépendra pas des coordonnées q_i , car celles-ci sont les variables d'intégration d'une intégrale définie. Notons qu'il doit être possible d'identifier clairement les bornes de l'intégrale sur q_i d'après la forme connue de la fonction W_i ; le mouvement de la coordonnée correspondante est soit une rotation ou une libration (c'est-à-dire une oscillation). Comme toute combinaison des constantes de séparation α_i peut faire office de nouvelle impulsion, les N variables d'action seront ici désignées comme les nouvelles impulsions du problème. En retour, les termes W_i de la fonction caractéristique pourront être exprimés en fonction des q_i et des J_i : $W_i(q, J)$.

Les coordonnées conjuguées aux variables d'action, notées w_i , sont formellement données par

$$w_i = \frac{\partial W}{\partial J_i} = \sum_j \frac{\partial W_j}{\partial J_i} \quad (4.121)$$

Comme les variables d'action sont des constantes du mouvement, la dépendance temporelle de w_i est très simple :

$$\dot{w}_i = \frac{\partial H}{\partial J_i} := \nu_i(J) = \text{cte} \quad \text{et donc} \quad w_i = \nu_i t + \beta_i \quad (4.122)$$

où les β_i sont des constantes d'intégration fixées par les conditions initiales. Les variables w_i sont appelées *variables d'angle*. Cette nomenclature vient du fait que le changement de w_i au cours d'une période de la coordonnée q_i est de 1. On peut démontrer cette assertion de la manière suivante :

$$\Delta w_i = \oint \frac{\partial w_i}{\partial q_i} dq_i = \oint \frac{\partial^2 W_i}{\partial J_i \partial q_i} dq_i = \frac{\partial}{\partial J_i} \oint \frac{\partial W_i}{\partial q_i} dq_i = \frac{\partial J_i}{\partial J_i} = 1 \quad (4.123)$$

Remarquons que J_i est constant dans l'intégrand, ce qui permet justement de sortir l'opérateur de différentiation du signe d'intégration. La variable w_i est donc analogue à un angle divisé par 2π . Les constantes ν_i sont alors les fréquences (en Hz) des mouvements correspondants, et les périodes sont $1/\nu_i$.

L'un des avantages des variables d'action-angle est qu'elles donnent accès aux fréquences des mouvements sans qu'une solution complète des équations de Hamilton soit nécessaire : l'intégrale qui définit J_i peut souvent être calculée en tant qu'intégrale définie sans qu'il soit possible de calculer la même intégrale pour un intervalle quelconque. Mais la principale utilité de ces variables apparaît dans le théorie des perturbations.

Exemple 4.2 Oscillateur harmonique en deux dimensions

Le hamiltonien est

$$H = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{1}{2}m\omega_x^2 x^2 + \frac{1}{2}m\omega_y^2 y^2 \quad (4.124)$$

La fonction caractéristique est manifestement séparable, $W = W_x + W_y$, chacun des deux termes obéissant à l'équation suivante :

$$\frac{1}{2m} \left(\frac{\partial W_x}{\partial x} \right)^2 + \frac{1}{2} m \omega_x^2 x^2 = \alpha_x \quad \frac{1}{2m} \left(\frac{\partial W_y}{\partial y} \right)^2 + \frac{1}{2} m \omega_y^2 y^2 = \alpha_y \quad \alpha_x + \alpha_y = E \quad (4.125)$$

Donc

$$p_x = (2m\alpha_x - m^2\omega_x^2 x^2)^{1/2} \quad p_y = (2m\alpha_y - m^2\omega_y^2 y^2)^{1/2} \quad (4.126)$$

La variable d'action J_x est

$$J_x = \oint p_x dx = \oint dx (2m\alpha_x - m^2\omega_x^2 x^2)^{1/2} \quad (4.127)$$

Le domaine d'intégration doit être pris entre les valeurs de x qui annulent la racine carrée, dans une direction, puis dans l'autre. On procède à cette fin au changement de variable d'intégration $x = \sqrt{2\alpha_x/m\omega_x^2} \sin \theta$, où θ doit être intégré de 0 à 2π . On trouve enfin

$$J_x = \frac{2\alpha_x}{\omega_x} \int_0^{2\pi} \cos^2 \theta d\theta = \frac{2\pi\alpha_x}{\omega_x} \quad \text{et pareillement} \quad J_y = \frac{2\pi\alpha_y}{\omega_y} \quad (4.128)$$

Le Hamiltonien s'exprime alors ainsi

$$H = E = \alpha_x + \alpha_y = \frac{\omega_x J_x}{2\pi} + \frac{\omega_y J_y}{2\pi} \quad (4.129)$$

et les fréquences associées sont

$$v_x = \frac{\partial H}{\partial J_x} = \frac{\omega_x}{2\pi} \quad v_y = \frac{\partial H}{\partial J_y} = \frac{\omega_y}{2\pi} \quad (4.130)$$

Comme on s'y attendait.

4.D.4 Variables d'action-angle dans le problème de Kepler

Les termes de la fonction caractéristique de Hamilton, dans le cas du potentiel de Kepler, mais sans présumer de l'orientation du plan de l'orbite, sont, d'après l'éq. (4.111),

$$W_r = \int dr \sqrt{2mE + \frac{2mk}{r} - \frac{\ell^2}{r^2}} \quad W_\theta = \int d\theta \sqrt{\ell^2 - \frac{\ell_z^2}{\sin^2 \theta}} \quad W_\varphi = \ell_z \varphi \quad (4.131)$$

Ceci mène aux expressions suivantes pour les variables d'action :

$$J_r = \oint dr \sqrt{2mE + \frac{2mk}{r} - \frac{\ell^2}{r^2}} \quad J_\theta = \oint d\theta \sqrt{\ell^2 - \frac{\ell_z^2}{\sin^2 \theta}} \quad J_\varphi = 2\pi \ell_z \quad (4.132)$$

L'intégrale définissant J_θ peut se faire par une série de substitutions. Si i est l'inclinaison du plan de l'orbite par rapport au plan xy , alors l'angle θ est compris entre θ_0 et $\pi - \theta_0$ tel que $\sin \theta_0 =$

$\cos i$, comme on voit en annulant la racine carrée figurant dans l'intégrant. On procède ensuite au changement de variable $\theta \rightarrow \psi$ où $\cos \theta = \sin i \sin \psi$ pour trouver

$$J_\theta = 4\ell \sin^2 i \int_0^{\pi/2} d\psi \frac{\cos^2 \psi}{1 - \sin^2 i \sin^2 \psi} \quad (4.133)$$

Un changement de variable supplémentaire $u = \tan \psi$ nous mène à la forme suivante :

$$J_\theta = 4\ell \int_0^\infty du \frac{1}{(1+u^2)(1+u^2 \cos^2 i)} \quad (4.134)$$

et cette dernière forme se calcule en la décomposant en fractions partielles, ce qui donne

$$J_\theta = 2\pi\ell(1 - \cos i) = 2\pi(\ell - \ell_z) \quad \text{car} \quad \cos i = \frac{\ell_z}{\ell} \quad (4.135)$$

Nous aurions pu obtenir ce résultat plus simplement, sachant que l'énergie cinétique s'exprime comme suit :

$$2T = p_r \dot{r} + p_\theta \dot{\theta} + p_\varphi \dot{\varphi} \quad (4.136)$$

ou encore, de manière équivalente, en utilisant des coordonnées polaires (r, ψ) dans le plan de l'orbite,

$$2T = p_r \dot{r} + p_\psi \dot{\psi} \quad (4.137)$$

Ceci entraîne que

$$\oint p_\theta d\theta = \oint p_\psi d\psi - \oint p_\varphi d\varphi = 2\pi\ell - 2\pi\ell_z \quad (4.138)$$

En effet, le premier terme du membre de droite n'est rien d'autre que $2\pi\ell'_z$ dans le système de coordonnées où le plan de l'orbite est précisément le plan xy , et donc $\ell'_z = \ell$, la grandeur du moment cinétique.

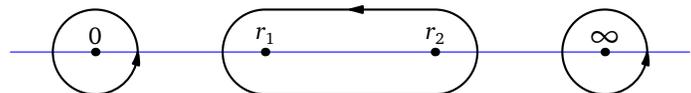
Il reste à calculer la valeur de J_r :

$$J_r = \oint dr \sqrt{2mE + \frac{2mk}{r} - \frac{\ell^2}{r^2}} = - \oint \frac{dz}{z^2} \sqrt{2mE + 2mkz - \ell^2 z^2} \quad z := \frac{1}{r} \quad (4.139)$$

où l'intégrale est prise entre les deux points de rebroussement r_1 et r_2 caractérisés par l'annulation de la racine carrée. Cette intégrale peut se faire à l'aide de la méthode des résidus, avec le résultat suivant :

$$J_r = -2\pi\ell + \pi k \sqrt{\frac{2m}{-E}} \quad (4.140)$$

Voyons dans le détail comment faire cette intégrale par la méthode des résidus. L'intégrant, en fonction de r , possède des points singuliers à $r = 0$, $r = \infty$ et deux points de branchement à r_1 et r_2 , qui sont les points de rebroussement, où la racine carrée s'annule. Le contour utilisé est illustré ci-dessous :



$$(4.141)$$

Ce contour est l'union de trois contours de même orientation, chacun entourant une singularité de l'intégrant. La somme des intégrales autour de ces contours est donc nulle, de par le théorème des résidus : l'extérieur

des trois contours ne contenant aucune singularité. La contribution du contour entourant $r = 0$ s'obtient en considérant que l'intégrand, autour de cette valeur, est essentiellement $i\ell/r$, et donc le résidu est $i\ell$. La contribution correspondante est donc $-2\pi\ell$. La contribution du contour entourant $r = \infty$ s'obtient en passant à la variable d'intégration $z = 1/r$ et en évaluant l'opposé du résidu de l'intégrand autour de $z = 0$. Il faut pour cela développer la racine carrée autour de $z = 0$ pour trouver

$$\begin{aligned} \frac{1}{z^2} \sqrt{2mE + 2mkz - \ell^2 z^2} &= \frac{\sqrt{-2mE}}{z^2} \sqrt{-1 + \frac{kz}{E} - \frac{\ell^2 z^2}{2mE}} = i \frac{\sqrt{-2mE}}{z^2} \sqrt{1 - \frac{kz}{E} + \frac{\ell^2 z^2}{2mE}} \\ &= i \frac{\sqrt{-2mE}}{z^2} \left(1 - \frac{kz}{2E} + \dots \right) = i \frac{\sqrt{-2mE}}{z^2} + ik \sqrt{\frac{m}{-2E}} \frac{1}{z} + \dots \end{aligned} \quad (4.142)$$

ce qui donne un résidu $ik\sqrt{-m/2E}$ et donc une contribution $\pi k\sqrt{-2m/E}$ en multipliant par $-2\pi i$. Enfin, le contour qui entoure les points de branchement donne précisément l'intégrale J_r recherchée, sauf pour le signe qui est opposé. La somme des trois contributions étant nulle, on trouve bien le résultat annoncé.

Ce résultat permet d'exprimer l'énergie E en fonction des trois variables d'action. Le résultat est particulièrement simple dans ce cas :

$$H = E = -\frac{2\pi^2 mk^2}{(J_r + J_\theta + J_\varphi)^2} \quad (4.143)$$

Les fréquences associées sont toutes égales :

$$\nu_r = \nu_\theta = \nu_\varphi = \frac{4\pi^2 mk^2}{(J_r + J_\theta + J_\varphi)^3} \quad (4.144)$$

Parce que ces trois fréquences sont égales, on dit que le problème de Kepler est *complètement dégénéré*. Physiquement, cela signifie que l'orbite est une courbe fermée dans l'espace (une ellipse) : la période est la même pour la variation des trois coordonnées.

Il est alors avantageux de procéder à une transformation canonique additionnelle parmi les J_i et les w_i , c'est-à-dire d'introduire trois autres variables d'action J_1, J_2 et J_3 , ainsi que les variables d'angle correspondantes, par le biais de la fonction génératrice suivante :

$$F = (w_\varphi - w_\theta)J_1 + (w_\theta - w_r)J_2 + w_r J_3 \quad (4.145)$$

avec comme résultat

$$\begin{aligned} w_1 &= \frac{\partial F}{\partial J_1} = w_\varphi - w_\theta & J_r &= \frac{\partial F}{\partial w_r} = J_3 - J_2 \\ w_2 &= \frac{\partial F}{\partial J_2} = w_\theta - w_r & J_\theta &= \frac{\partial F}{\partial w_\theta} = J_2 - J_1 \\ w_3 &= \frac{\partial F}{\partial J_3} = w_r & J_\varphi &= \frac{\partial F}{\partial w_\varphi} = J_1 \end{aligned} \quad (4.146)$$

En fonction de ces nouvelles variables, le hamiltonien est

$$H = -\frac{2\pi^2 mk^2}{J_3^2} \quad (4.147)$$

et les fréquences ν_1 et ν_2 sont nulles, c'est-à-dire que les variables d'angle w_1 et w_2 sont en fait des constantes, tout comme les trois variables d'action. Quant à la variable w_3 , elle progresse linéairement dans le temps, avec une valeur initiale w_3^0 .

Or ces cinq constantes, plus la valeur initiale w_3^0 , doivent pouvoir être exprimées en fonction des six éléments d'une orbite énumérés à l'éq. (4.83). On sait déjà que l'énergie totale E s'exprime à la fois en fonction de J_3 et du demi-grand axe a :

$$E = -\frac{k}{2a} = -\frac{2\pi^2 mk^2}{J_3^2} \quad \text{et donc} \quad a = \frac{J_3^2}{4\pi^2 mk} \quad (4.148)$$

D'autre part, l'excentricité de l'orbite est reliée au moment cinétique $\ell = (J_\theta + J_\varphi)/2\pi = J_2/2\pi$:

$$e = \sqrt{1 + \frac{2E\ell^2}{mk^2}} = \sqrt{1 - \frac{J_2^2}{J_3^2}} \quad (4.149)$$

Nous avons vu ci-dessus que l'inclinaison de l'orbite i s'exprime en fonction de ℓ_z et ℓ :

$$\cos i = \frac{\ell_z}{\ell} = \frac{J_\varphi}{J_\theta + J_\varphi} = \frac{J_1}{J_2} \quad (4.150)$$

Le temps de passage au péricentre est relié à l'origine des temps, tout comme la valeur initiale w_3^0 . Il reste donc à exprimer la longitude du noeud ascendant Ω et l'argument du péricentre ω en fonction des variables restantes, w_1 et w_2 . On sait que $J_1 = J_\varphi = 2\pi\ell_z$. Donc J_1 génère les rotations dans le plan xy . La variable conjuguée à J_1 doit donc être un angle défini dans ce plan (divisé par 2π) ; cet angle n'est défini qu'à une constante additive près, et doit être constant dans le problème de Kepler. On peut donc, sans perte de généralité, l'identifier à la longitude du noeud ascendant Ω . D'autre part, $J_2 = 2\pi\ell$ et donc génère les rotations dans le plan de l'orbite (le plan perpendiculaire au moment cinétique). La variable conjuguée à J_2 est donc un angle dans ce plan (divisé par 2π) qui est constant dans le problème de Kepler. Cet angle peut être choisi comme l'argument du péricentre ω . En somme :

$$\Omega = 2\pi w_1 \quad \text{et} \quad \omega = 2\pi w_2 \quad (4.151)$$

Résumons donc la relation entre les nouvelles variables canoniques ($J_{1,2,3}, w_{1,2,3}$) et les éléments de l'orbite :

$J_1 = 2\pi\ell_z$	$w_1 = \frac{\Omega}{2\pi}$	(4.152)
$J_2 = 2\pi\ell$	$w_2 = \frac{\omega}{2\pi}$	
$J_3 = 2\pi\sqrt{mka}$	$w_3 = \frac{t}{\tau}$	

L'expression pour w_3 vient du fait que cette variable varie uniformément dans le temps, et qu'elle change par $\Delta w_3 = 1$ au cours de la période τ . Ces variables (sauf w_3) sont constantes dans le temps dans le cas du problème de Kepler, mais si le hamiltonien est modifié par une perturbation, elles ne seront plus constantes et affecteront un changement graduel dans le temps qui constitue une précession du plan de l'orbite dans le cas de w_1 , et une précession de l'orbite elliptique dans son plan dans le cas de w_2 . L'important est que les variables J_i et w_i forment un ensemble de variables canoniques pour tout hamiltonien, pas seulement pour le problème de Kepler.

E Introduction à la théorie des perturbations

Très peu de problèmes mécaniques peuvent être résolus analytiquement, en particulier très peu de problèmes utiles ! De nos jours, on pourrait se contenter, semble-t-il, de résoudre numériquement les équations du mouvement quand une solution analytique est impossible. Cependant, il est utile d'avoir une connaissance analytique, même approximative, des solutions, car une solution numérique ne permet pas facilement d'entrevoir la totalité des solutions possibles.

Il existe une méthode puissante, analytique cette fois, pour obtenir une solution approchée au problème, lorsque le hamiltonien est proche d'un autre hamiltonien dont on connaît la solution. Il s'agit de la *théorie des perturbations*. Cette méthode suppose qu'on puisse écrire le hamiltonien comme la somme de deux termes : $H = H_0 + H_1$, et que la solution au problème défini par H_0 est connue ; elle suppose en outre que H_1 , appelée la *perturbation*, est petit en comparaison de H_0 . Nous allons expliquer sommairement ici la méthode des perturbations dites *dépendantes du temps*, ou méthode de *variation des constantes*, et nous allons l'appliquer à une perturbation des orbites elliptiques de Kepler causée par un potentiel supplémentaire en loi de puissance.

4.E.1 Méthode de variation des constantes

Supposons que le hamiltonien du système considéré puisse s'écrire ainsi :

$$H(q, p, t) = H_0(q, p, t) + H_1(q, p, t) \quad (4.153)$$

où la fonction principale de Hamilton $S_0(q, \alpha, t)$ associée à H_0 nous est connue. Nous pouvons utiliser α et β comme nouvelles variables canoniques et le hamiltonien transformé est alors

$$H'(\beta, \alpha, t) = H_0 + H_1 + \frac{\partial S_0}{\partial t} = H_1(\beta, \alpha) \quad (4.154)$$

car S_0 obéit à l'équation de Hamilton-Jacobi (2.109) pour le hamiltonien H_0 . La transformation $(q, p) \rightarrow (\beta, \alpha)$ est canonique, quel que soit le hamiltonien utilisé (le fait qu'une transformation soit ou non canonique ne dépend pas du hamiltonien).

Les variables β et α sont constantes dans le temps si H_0 est le hamiltonien complet, car alors $H' = 0$. Mais la perturbation H_1 cause une variation dans le temps de ces « constantes » (d'où l'expression « variation des constantes ») qui obéit aux équations de Hamilton, avec hamiltonien H_1 :

$$\dot{\alpha} = -\frac{\partial H_1}{\partial \beta} \quad \dot{\beta} = \frac{\partial H_1}{\partial \alpha} \quad (4.155)$$

Comme H_1 est supposé petit, les variables (α, β) évoluent lentement.

Exemple 4.3 Oscillateur harmonique en tant que perturbation d'une particule libre
Considérons le hamiltonien de l'oscillateur harmonique comme la somme d'une partie non per-

turbée H_0 et d'une perturbation H_1 :

$$H_0 = \frac{p^2}{2m} \quad H_1 = \frac{1}{2}m\omega^2 x^2 \quad (4.156)$$

D'après l'exemple 2.8 (p. 60), la transformation nous amenant vers des variables canoniques (β, α) constantes sous H_0 est

$$x = \sqrt{\frac{2\alpha}{m}}(t + \beta) \quad p = \sqrt{2m\alpha} \quad (4.157)$$

Donc la perturbation s'exprime ainsi en fonction de ces variables :

$$H_1(\beta, \alpha) = \frac{1}{2}m\omega^2 \frac{2\alpha}{m}(t + \beta)^2 = \alpha\omega^2(t + \beta)^2 \quad (4.158)$$

Les équations de Hamilton pour (β, α) sont donc

$$\dot{\alpha} = -\frac{\partial H_1}{\partial \beta} = -2\alpha\omega^2(t + \beta) \quad \dot{\beta} = \frac{\partial H_1}{\partial \alpha} = \omega^2(t + \beta)^2 \quad (4.159)$$

On pourrait évidemment résoudre ces équations exactement, mais là n'est pas le but de l'exemple. En pratique, si on ne savait comment résoudre ces équations, on supposerait que ω est petit et qu'en première approximation, pour une durée limitée, il est possible de supposer que les valeurs de (β, α) dans les membres de droite de ces équations sont les valeurs initiales $(\beta_0 = 0, \alpha_0)$, et donc qu'on a la solution approchée, dite de « premier ordre » :

$$\alpha_1 = \alpha_0 - \alpha_0\omega^2 t^2 \quad \beta_1 = \frac{1}{3}\omega^2 t^3 \quad (4.160)$$

La coordonnée x devient alors, à cet ordre d'approximation,

$$\begin{aligned} x(t) &= \sqrt{\frac{2\alpha_1}{m}}(t + \beta_1) \\ &= \sqrt{\frac{2\alpha_0}{m}} \sqrt{1 - \omega^2 t^2} \left(t + \frac{1}{3}\omega^2 t^3 \right) \\ &\approx \sqrt{\frac{2\alpha_0}{m}} \left(1 - \frac{1}{2}\omega^2 t^2 \right) \left(1 + \frac{1}{3}\omega^2 t^2 \right) t \\ &\approx \sqrt{\frac{2\alpha_0}{m}} \frac{1}{\omega} \left(\omega t - \frac{1}{6}\omega^3 t^3 \right) \end{aligned} \quad (4.161)$$

On notera que cette expression constitue en fait les deux premiers termes du développement en série de la solution exacte, soit

$$x(t) = \sqrt{\frac{2\alpha_0}{m}} \frac{1}{\omega} \sin(\omega t) \quad (4.162)$$

où α_0 est l'énergie de l'oscillateur

4.E.2 Précession des orbites

Nous allons maintenant appliquer la théorie des perturbations au problème d'un objet se déplaçant dans un potentiel de la forme

$$V(r) = -\frac{k}{r} - \frac{h}{r^n} \quad (4.163)$$

où le deuxième terme sera considéré comme une perturbation. Nous allons utiliser dans ce problème les variables d'action-angle $J_{1,2,3}$ et $w_{1,2,3}$ définies à la section 4.D.4. Le hamiltonien H_0 ne dépend alors que de J_3 :

$$H_0 = -\frac{2\pi^2 mk^2}{J_3^2} \quad (4.164)$$

ce qui entraîne que les coordonnées w_2 et w_3 sont constantes dans le temps si $h = 0$.

Au lieu d'utiliser la valeur instantanée de H_1 dans ce calcul, nous allons utiliser sa valeur moyenne sur une période τ de l'orbite non perturbée. Si la perturbation est petite, son effet sera lent de toute manière et ce qui nous intéresse généralement est l'effet accumulé sur plusieurs périodes. L'utilisation de la valeur moyenne est alors pleinement justifiée. Appelons ψ la coordonnée angulaire dans le plan de l'orbite (qui ne coïncide pas nécessairement avec le plan $\theta = \pi/2$) : On définit donc la perturbation moyenne

$$\begin{aligned} \overline{H_1} &= -\frac{1}{\tau} \int_0^\tau dt \frac{h}{r^n(t)} \\ &= -\frac{1}{\tau} \int_0^{2\pi} d\psi \frac{1}{\dot{\psi}} \frac{h}{r^n(t)} \\ &= -\frac{h}{\tau} \int_0^{2\pi} d\psi \frac{mr^2}{\ell} \frac{1}{r^n} \\ &= -\frac{mh}{\ell\tau} \left(\frac{mk}{\ell^2}\right)^{n-2} \int_0^{2\pi} d\psi [1 + e \cos \psi]^{n-2} \end{aligned} \quad (4.165)$$

Nous avons utilisé l'expression de r en fonction de l'angle ψ :

$$r = \frac{\ell^2/mk}{1 + e \cos \psi} \quad (4.166)$$

Perturbation en inverse du carré de la distance

Dans le cas $n = 2$, on trouve

$$\overline{H_1} = -\frac{2\pi mh}{\ell\tau} \quad (n = 2) \quad (4.167)$$

La période τ s'exprime en fonction de J_3 (elle est l'inverse de $v_3 = \partial H_0 / \partial J_3$) et le moment cinétique ℓ en fonction de J_2 :

$$\overline{H_1}(J_2, J_3) = -\frac{4\pi^2 mh}{J_2 \tau} \quad \frac{1}{\tau} = v_3 = \frac{\partial H_0}{\partial J_3} = \frac{4\pi^2 mk^2}{J_3^3} \quad (4.168)$$

La perturbation a donc un effet sur la variable d'angle $w_2 = \omega/2\pi$, qui n'est plus constante. L'argument du péricentre avance donc dans le temps, avec un taux de précession

$$\dot{\omega} = 2\pi\nu_2 = 2\pi \frac{\partial \overline{H}_1}{\partial J_2} = \frac{8\pi^3 m h}{J_2^2 \tau} = \frac{2\pi}{(1-e^2)\tau} \frac{h}{ka} \quad (4.169)$$

où, dans cette dernière expression, nous avons utilisé les correspondances

$$J_3 = 2\pi \sqrt{mka} \quad \text{et} \quad 1-e^2 = \frac{J_2^2}{J_3^2} \quad (4.170)$$

Notons que le rapport h/ka est sans unités et représente la grandeur relative de la perturbation au potentiel initial $-k/r$. La présence de la période τ dans cette formule nous permet d'estimer immédiatement le changement de ω au bout d'une période :

$$\Delta\omega = \dot{\omega}\tau = \frac{2\pi}{(1-e^2)} \frac{h}{ka} \quad (4.171)$$

Ce changement est proportionnel à la perturbation, et d'autant plus important que l'excentricité est grande.

Perturbation en inverse du cube de la distance

Dans le cas $n = 3$, l'intégrale sur ψ figurant dans (4.165) donne encore 1, et

$$\overline{H}_1 = -\frac{2\pi m^2 h k}{\ell^3 \tau} = -\frac{(2\pi)^4 m^2 h k}{J_2^3 \tau} \quad (n = 3) \quad (4.172)$$

de sorte que le taux de précession moyen est

$$\dot{\omega} = 2\pi \frac{\partial \overline{H}_1}{\partial J_2} = 3 \frac{(2\pi)^5 m^2 h k}{J_2^4 \tau} = \frac{6\pi h}{ka^2(1-e^2)^2 \tau} \quad (4.173)$$

Une perturbation de ce genre intervient dans la théorie de la relativité générale, où on peut montrer que le potentiel newtonien est augmenté d'une perturbation

$$H_1 = -\frac{k\ell^2}{m^2 c^2 r^3} \implies h = \frac{k\ell^2}{m^2 c^2} = \quad (4.174)$$

où c est la vitesse de la lumière. Le taux de précession moyen est alors²

$$\dot{\omega} = \frac{6\pi}{\tau(1-e^2)} \frac{R}{a} \quad (4.175)$$

où $R := GM/c^2$ est le rayon gravitationnel du centre d'attraction.

Dans le cas du système solaire, cet effet est maximum pour la planète ayant la plus petite valeur de a , soit Mercure. L'avance observée du périhélie de Mercure est de 5 600 secondes d'arc par siècle.

2. Même si le moment cinétique ℓ figure dans la constante h , cette dépendance additionnelle en ℓ ne doit pas être prise en compte lors du calcul (4.173) de la dérivée. En effet, le rôle de ℓ dans la valeur de h n'est pas celui d'une variable canonique, mais plutôt celui d'un paramètre externe.

Chapitre 4. Forces centrales

La précession des équinoxes de la Terre et l'effet des autres planètes parvient à expliquer la majeure partie de cette précession, mais il reste une avance résiduelle de 43 secondes d'arc par siècle, qui ne peut pas être expliquée par les lois de Newton.³ L'effet de la relativité générale exprimé ci-haut est de 42.98 secondes d'arc par siècle et explique parfaitement la précession résiduelle de Mercure. Il s'agit de l'une des confirmations classiques de la théorie de la relativité générale.

3. Au 19^e siècle, l'astronome Le Verrier voulut expliquer cette précession résiduelle par l'influence gravitationnelle d'une planète jusqu'alors inobservée, suivant une orbite encore plus rapprochée du soleil que Mercure ; il l'avait même baptisée *Vulcain*. Mais une telle planète ne fut jamais observée.

Problèmes

Problème 4.1 Diffusion de Rutherford : cas attractif

Comment doit-on modifier la formule de la section différentielle de Rutherford dans le cas d'une force électrique attractive au lieu de répulsive ?

Problème 4.2 Diffusion par un potentiel en $1/r^2$

Calculez la section différentielle de diffusion pour un projectile subissant de la part de la cible une force de répulsion dérivant du potentiel $V(r) = k/r^2$. La section efficace est-elle finie ?

Problème 4.3 Diffusion par une sphère molle

Dans ce problème nous allons étudier la diffusion de particules par le potentiel central suivant :

$$V(r) = \begin{cases} V_0 & (r < a) \\ 0 & (r > a) \end{cases} \quad (\text{sphère molle}) \quad (4.176)$$

Nous allons supposer que $V_0 > 0$. Quand l'énergie E de la particule incidente est inférieure à V_0 , le potentiel est impénétrable et se comporte comme celui d'une sphère dure. Les choses sont différentes si $E > V_0$.

A Démontrez la relation suivante entre le paramètre d'impact b et l'angle de diffusion θ :

$$\cos \frac{\theta}{2} = \frac{1}{\gamma a^2} \left(b^2 + \sqrt{(a^2 - b^2)(\gamma^2 a^2 - b^2)} \right) \quad \gamma^2 := 1 - \frac{V_0}{E} \quad (4.177)$$

Indice : à l'intérieur de chacune des régions $r < a$ et $r > a$, la force est nulle. En passant d'une région à l'autre, l'énergie totale et le moment cinétique doivent être conservés. La preuve est géométrique.

B Montrez qu'il y a un angle de diffusion maximum θ_0 (donnez-en une expression), et qu'en deçà de cet angle, deux valeurs du paramètre d'impact b sont possibles pour chaque valeur de θ . Comment doit-on calculer la section différentielle de diffusion, dans ce cas ?

C À l'aide de Mathematica, tracez la section différentielle de diffusion en fonction de θ pour $\gamma = 0.5$.

Problème 4.4 Théorie des perturbations pour l'oscillateur anharmonique

Considérez un oscillateur harmonique de masse m et fréquence ω , dont la coordonnée canonique

est x , auquel on ajoute une perturbation de la forme $H_1 = \lambda x^4$.

A En vous basant sur la solution de l'équation de Hamilton-Jacobi pour l'oscillateur harmonique, obtenez une expression pour la perturbation moyenne \bar{H}_1 en fonction de la nouvelle coordonnée β et du nouveau moment α , la moyenne étant effectuée sur une période de l'oscillation.

B Quel est le taux de variation moyen de β ? Expliquez comment ce taux peut être interprété comme un changement relatif de la période de l'oscillateur dû à la perturbation.

C Considérez maintenant un pendule simple et développez l'énergie potentielle en fonction de l'angle φ jusqu'à l'ordre φ^4 inclus. En vous servant du résultat obtenu en **B**, donnez une expression de la correction relative à la période du pendule ($\Delta\tau/\tau$) en fonction de l'amplitude φ_0 de l'oscillation.

Problème 4.5 Perturbation d'une orbite par une force constante

Considérons un satellite artificiel en orbite rapide autour de la Terre. Supposons que la force de marée causée par la lune (ou le soleil) agisse sur ce satellite, et supposons qu'on puisse considérer approximativement que cette force est constante en grandeur et en direction. Alternativement, on peut penser à l'influence de Jupiter sur une planète comme Mercure; comme Jupiter se déplace lentement et qu'elle est très éloignée, la force qu'elle exerce sur Mercure est approximativement constante sur le cours de quelques périodes. Utilisons un système de coordonnées dans lequel l'orbite de l'objet est contenue dans le plan xy ($\theta = \pi/2$) et notons l'argument du péricentre par ω . Cette force constante de grandeur f est dirigée dans une direction définie par l'angle polaire θ_0 et l'angle azimutal φ_0 .

A Écrivez la forme de cette perturbation, et calculez sa valeur moyenne dans une période.

B Quelles sont les quantités affectées par cette perturbation et comment sont-elles affectées?

Problème 4.6 Précession de l'orbite d'un satellite

Dans ce problème, nous allons étudier comment la forme aplatie de la Terre influence les paramètres de l'orbite elliptique d'un satellite artificiel. Les paramètres de cette orbite sont : l'excentricité e , le demi-grand axe a , l'inclinaison i du plan de l'orbite par rapport à l'axe polaire terrestre, la longitude Ω du noeud ascendant, et l'argument du péricentre ω . Les moments d'inertie de la Terre sont notés I_3 et I , et sa masse M . L'énergie potentielle d'un satellite de masse m en présence de la Terre aplatie est approximativement

$$V(r) = -\frac{k}{r} + \frac{k}{M} \frac{I_3 - I}{r^3} P_2(\cos \theta) \quad k := GMm \quad (4.178)$$

A Démontrez que, le long de l'orbite du satellite décrite par l'angle φ , la valeur de l'angle polaire θ est donnée par

$$\cos \theta = \sin i \sin(\varphi + \omega) \quad (4.179)$$

B Démontrez que la perturbation H_1 causée par l'aplatissement de la Terre est, moyennée sur

une période de l'orbite,

$$\bar{H}_1 = \frac{\pi m^2 k^2 (I_3 - I_1)}{2M\ell^3 \tau} (1 - 3 \cos^2 i) \quad (4.180)$$

où ℓ est le moment cinétique du satellite et τ sa période.

C Démontrez que le plan de l'orbite du satellite précède autour de l'axe polaire terrestre, et que la fraction de tour effectuée à chaque période est

$$-\frac{3}{2} \frac{I_3 - I_1}{Ma^2(1 - e^2)^2} \cos i \quad (4.181)$$

Pour un satellite en orbite circulaire à une altitude de 1 000 km et d'inclinaison $i = \pi/4$, quel est ce changement relatif, en pourcentage ?

D Démontrez que le péricentre de l'orbite précède dans son propre plan, et que la fraction de tour effectuée à chaque période est

$$\frac{3}{4} \frac{I_3 - I_1}{Ma^2(1 - e^2)^2} (5 \cos^2 i - 1) \quad (4.182)$$

E Cet effet s'applique également à l'orbite d'une planète autour du soleil. Quel devrait être l'aplatissement du soleil pour qu'il soit aussi important sur l'orbite de Mercure que celui prédit par la relativité générale ?

CHAPITRE 5

Relativité et électromagnétisme

A Théorie de la relativité restreinte

5.A.1 Espace-temps

Considérons deux référentiels S et S' se déplaçant l'un par rapport à l'autre à une vitesse $V = \beta c$ le long de l'axe des x , comme illustré à la figure 5.1. Comme démontré dans toute introduction à la théorie de la relativité, les coordonnées et le temps dans les deux référentiels sont reliés par la *transformation de Lorentz* :

$$\begin{aligned}x' &= \gamma(x - \beta ct) & y' &= y \\ ct' &= \gamma(ct - \beta x) & z' &= z\end{aligned}\tag{5.1}$$

où $\gamma = 1/\sqrt{1 - \beta^2}$. On introduit aussi la *rapidité* η définie par la relation

$$\tanh \eta = \beta\tag{5.2}$$

La transformation de Lorentz prend alors la forme

$$\begin{aligned}x' &= x \cosh \eta - ct \sinh \eta & y' &= y \\ ct' &= ct \cosh \eta - x \sinh \eta & z' &= z\end{aligned}\tag{5.3}$$

L'avantage de la rapidité est que (1) la transformation de Lorentz est formellement similaire à une rotation dans l'espace, où les fonctions circulaires ont été remplacées par des fonctions hyperboliques et (2) la composition de deux transformations de Lorentz successives dans la même direction se fait par simple addition des rapidités.

L'espace euclidien tridimensionnel, auquel on ajoute la dimension temporelle, forme ce qu'on appelle l'*espace-temps*. Spécifier une position (x, y, z) et un temps t définit un *événement* x dans l'espace-temps. On introduit alors un ensemble de 4 coordonnées d'espace-temps (ct, x, y, z) , numérotées de 0 à 3 :

$$x^0 = ct \quad x^1 = x \quad x^2 = y \quad x^3 = z\tag{5.4}$$

L'équivalent de la distance dans l'espace à trois dimensions est l'*intervalle* s^2 :

$$s^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2\tag{5.5}$$

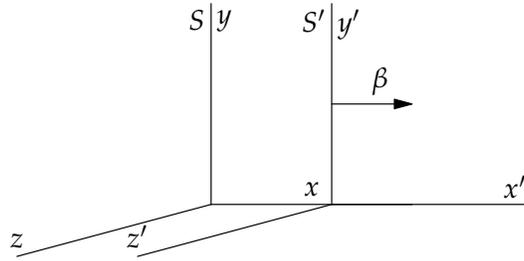


FIGURE 5.1

La transformation (5.1) préserve l'intervalle, qui est le même dans tous les référentiels.

Les concepts introduits à l'annexe B sont également utiles dans l'espace-temps : on peut introduire dans l'espace-temps une base notée e_μ ($\mu = 0, 1, 2, 3$) où le vecteur e_0 est un vecteur unitaire dans la direction temporelle. Un changement de référentiel est en fait une transformation qui nous amène vers une nouvelle base e'_μ . Il existe un produit scalaire bien défini sur l'espace-temps, défini alors par le tenseur métrique suivant :

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (5.6)$$

En fonction des coordonnées $(ct, x, y, z) = (x^0, x^1, x^2, x^3)$, qui sont les composantes contravariantes du quadrivecteur position de l'événement $x = x^\mu e_\mu$, la norme de x s'exprime comme

$$x^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \quad (5.7)$$

et porte le nom d'*intervalle* (sous-entendu : intervalle entre l'événement x et l'origine).

La transformation de Lorentz (5.1) permet de relier entre elles les composantes décrivant le même événement dans deux référentiels différents. Lors d'une transformation de Lorentz générale, la coordonnée spatio-temporelle x^μ se transforme comme suit :

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \quad (5.8)$$

où $\Lambda^\mu{}_\nu$ est une matrice de transformation. L'indice de gauche numérote les rangées, l'indice de droite les colonnes et les indices répétés sont sommés. Par exemple, pour la transformation (5.1) la matrice de transformation est

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cosh \eta & -\sinh \eta & 0 & 0 \\ -\sinh \eta & \cosh \eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5.9)$$

mais elle serait différente pour un changement de référentiel accompagné d'une rotation des axes, ou suivant un axe quelconque. La matrice Λ est l'équivalent de la matrice S décrite à l'annexe B.

Une matrice Λ appartiendra, par définition, au *groupe de Lorentz* si elle préserve la forme (5.7) de l'intervalle. Cela revient à dire que les composantes du tenseur métrique 5.6 (ou de son inverse $g^{\mu\nu}$) doivent être les mêmes dans tous les référentiels. Autrement dit,

$$g^{\mu\nu} = \Lambda^\mu_\alpha \Lambda^\nu_\beta g^{\alpha\beta} \quad (5.10)$$

Cette condition doit être respectée par une matrice Λ pour qu'elle appartienne au groupe de Lorentz. L'équation ci-haut représente 10 conditions indépendantes, car elle est symétrique en (μ, ν) . Comme la matrice Λ comporte a priori 16 composantes indépendantes, cela laisse six paramètres libres qui peuvent spécifier un élément du groupe de Lorentz. Ces six paramètres correspondent aux trois composantes de la vitesse relative entre deux référentiels, ainsi qu'à trois paramètres de rotation entre les axes cartésiens des deux référentiels (trois angles d'Euler, par exemple).

Les composantes covariantes se transforment à l'aide de la transposée de la matrice inverse de Λ , soit

$$A'_\mu = (\Lambda^{-1})^\nu_\mu A_\nu \quad (5.11)$$

La matrice inverse de (5.9) s'obtient simplement en changeant le signe de β .

5.A.2 Exemples d'invariants et de quadrivecteurs

Quadri gradient

Examinons maintenant comment se transforme l'opérateur gradient augmenté de la dérivée par rapport au temps, noté

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{c\partial t}, \nabla \right) \quad (5.12)$$

Sachant que $x^\nu = (\Lambda^{-1})^\nu_\mu x'^\mu$ (transformation inverse), on voit que

$$\partial'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} = (\Lambda^{-1})^\nu_\mu \partial_\nu \quad (5.13)$$

et donc que le quadri gradient se transforme comme les composantes covariantes d'un quadrivecteur, ce qui justifie la notation utilisée.

Temps propre

Le *temps propre* τ d'un objet en mouvement est simplement le temps tel qu'il s'écoule dans le référentiel de l'objet. Dans le référentiel inertiel S' qui se déplace avec l'objet à un instant donné, la différentielle de temps propre coïncide avec dt' . Cependant, l'objet peut en général être accéléré de sorte que le référentiel S' n'est pas le même à tous les instants. Quel que soit le référentiel dans lequel on observe l'objet, la différentielle de temps propre peut toujours être exprimée comme suit :

$$c d\tau = \sqrt{dx^\mu dx_\mu} = c dt \sqrt{1 - \beta^2} = \frac{c dt}{\gamma} \quad (5.14)$$

où βc est la vitesse de l'objet, qui peut dépendre du temps. En effet, dans le référentiel S' , $dx^\mu dx_\mu = c^2 d\tau^2$ et de plus l'expression ci-haut est invariante, puisque $dx^\mu dx_\mu$ est un invariant. Pour l'obtenir le temps propre total écoulé sur la trajectoire de l'objet, on intègre $d\tau$:

$$\tau = \int d\tau = \int \frac{dt}{\gamma} = \int dt \sqrt{1 - \beta^2} \quad (5.15)$$

Quadri vitesse

Considérons un objet se déplaçant à une vitesse instantanée $\mathbf{v} = \beta c$. On définit sa *quadri vitesse* u^μ comme

$$u^\mu = \frac{dx^\mu}{d\tau} \quad (5.16)$$

La différentielle $d\tau$ est invariante, alors que dx^μ se transforme comme les coordonnées ; donc u^μ est une quadrivecteur. En fonction de la vitesse β , les composantes explicites sont

$$u^\mu : (\gamma c, \gamma \mathbf{v}) \quad (5.17)$$

Par contraction on obtient l'invariant $u^\mu u_\mu = c^2$. Le temps propre écoulé le long d'une trajectoire peut alors s'écrire comme

$$\tau = \int d\tau = \frac{1}{c^2} \int u_\mu dx^\mu \quad (5.18)$$

Quadri accélération

On définit aussi la quadri accélération

$$a^\mu = \frac{du^\mu}{d\tau} = \gamma \frac{du^\mu}{dt} \quad (5.19)$$

L'expression explicite (5.17) de la quadri vitesse donne

$$a^\mu = \left(\frac{(\boldsymbol{\beta} \cdot \mathbf{a})}{(1 - \beta^2)^2}, \frac{\mathbf{a}}{1 - \beta^2} + \frac{(\boldsymbol{\beta} \cdot \mathbf{a})\boldsymbol{\beta}}{(1 - \beta^2)^2} \right) \quad (5.20)$$

À partir de ce quadrivecteur, on peut construire certains invariants. Par exemple,

$$\begin{aligned} u_\mu a^\mu &= u_\mu \frac{du^\mu}{d\tau} \\ &= \frac{1}{2} \frac{d}{d\tau} (u_\mu u^\mu) \\ &= 0 \end{aligned} \quad (5.21)$$

En utilisant les expressions explicites pour u^μ et a^μ en fonction des composantes temporelles et spatiales, la condition $u_\mu a^\mu = 0$ ressort de manière triviale. Un autre invariant à considérer est $a^\mu a_\mu$. On calcule que

$$a^\mu a_\mu = -\gamma^4 (\mathbf{a}^2 + \gamma^2 (\boldsymbol{\beta} \cdot \mathbf{a})^2) \quad (5.22)$$

Cette quantité est la même dans tous les référentiels, même si les valeurs de β et de \mathbf{a} changent.

Quadricourant

Un autre quadrivecteur est formé par la densité de courant et la densité de charge :

$$\mathbf{J}^\mu : (c\rho, \mathbf{J}) \quad (5.23)$$

Pour s'en convaincre, il suffit d'écrire l'équation de continuité comme

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = \partial_\mu \mathbf{J}^\mu = 0 \quad (5.24)$$

Comme la conservation de la charge doit être valide dans tous les référentiels et que ∂_μ est un quadrivecteur, le quadricourant J^μ en est forcément un lui aussi. Une autre façon de se convaincre que J^μ est un quadrivecteur est de considérer l'expression de la densité de courant associée à un ensemble de charges ponctuelles en mouvement :

$$\mathbf{J}(\mathbf{r}) = \sum_i q^{(i)} \mathbf{v}^{(i)} \delta(\mathbf{r} - \mathbf{r}^{(i)}) \quad (5.25)$$

Ici $q^{(i)}$, $\mathbf{r}^{(i)}$ et $\mathbf{v}^{(i)}$ sont respectivement la charge, la position et la vitesse de la i^e particule. La fonction delta dans la formule ci-haut n'est pas un invariant de Lorentz, puisqu'elle est définie par la relation

$$\int d^3 r^{(i)} \delta(\mathbf{r} - \mathbf{r}^{(i)}) = 1 \quad (5.26)$$

valide dans tous les référentiels et que $d^3 r$ n'est pas invariant. Cependant, $\gamma^{(i)} d^3 r^{(i)}$ est un invariant, où $\gamma^{(i)} = 1/\sqrt{1 - (\beta^{(i)})^2}$. La fonction $(1/\gamma^{(i)})\delta(\mathbf{r} - \mathbf{r}^{(i)})$ est donc aussi un invariant. Écrivons donc la densité de courant et la densité de charge comme

$$\mathbf{J}(\mathbf{r}) = \sum_i q^{(i)} \gamma^{(i)} \mathbf{v}^{(i)} [\delta(\mathbf{r} - \mathbf{r}^{(i)})/\gamma^{(i)}] \quad (5.27)$$

$$\rho(\mathbf{r}) = \sum_i q^{(i)} \gamma^{(i)} [\delta(\mathbf{r} - \mathbf{r}^{(i)})/\gamma^{(i)}] \quad (5.28)$$

En fonction de la quadri vitesse de chaque particule, on a

$$J^\mu = \sum_i q^{(i)} u^{(i)\mu} [\delta(\mathbf{r} - \mathbf{r}^{(i)})/\gamma^{(i)}] \quad (5.29)$$

Comme $q^{(i)}$ et la fonction delta modifiée sont des invariants, on conclut que J^μ est bel et bien un quadrivecteur.

B Dynamique relativiste

5.B.1 Action d'une particule libre

En relativité, les équations du mouvement doivent être les mêmes dans tous les référentiels inertiels. Pour cela il faut que l'action soit invariante par rapport aux transformations de Lorentz, c'est-à-dire qu'elle ait la même forme dans tous les référentiels. Ainsi, si elle est stationnaire dans un référentiel, elle le sera dans tous les référentiels. Guidés par ce principe, essayons d'obtenir l'action d'une particule libre.

Soit $x^\mu(\tau)$ la trajectoire de la particule dans l'espace-temps, en fonction du temps propre τ . La quantité invariante la plus simple qu'on peut construire avec une telle trajectoire est sa longueur propre, c'est-à-dire le temps écoulé dans le référentiel (non inertielle, en général) de la particule. Nous supposons donc que l'action d'une particule libre est simplement son temps propre. Cette quantité doit être multipliée par une constante qui donne à l'action ses unités naturelles, soit celles d'une énergie fois un temps. La seule constante relative à la particule ayant les unités d'une énergie est sa masse m , fois c^2 . On suppose donc que l'action est

$$S = -mc^2 \int d\tau \quad (5.30)$$

Puisque $d\tau = dt/\gamma$, on peut écrire

$$S = -mc^2 \int dt \sqrt{1 - \beta^2} \quad (5.31)$$

Le lagrangien d'une particule libre est donc

$$L = -mc^2 \sqrt{1 - \beta^2} \quad (5.32)$$

Dans l'approximation non relativiste ($\beta \ll 1$) on a $\sqrt{1 - \beta^2} \approx 1 - \beta^2/2$ et donc

$$L \approx -mc^2 + \frac{1}{2}mv^2 \quad (5.33)$$

On retrouve donc le lagrangien habituel, plus une constante ($-mc^2$) qui n'affecte pas les équations du mouvement.

L'impulsion généralisée \mathbf{p} est définie comme

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1 - \beta^2}} \quad (5.34)$$

Le hamiltonien qui correspond à ce lagrangien est

$$\begin{aligned} H &= \mathbf{p} \cdot \mathbf{v} - L \\ &= \frac{mv^2}{\sqrt{1 - \beta^2}} + mc^2 \sqrt{1 - \beta^2} \end{aligned}$$

$$= \frac{mc^2}{\sqrt{1-\beta^2}} \quad (5.35)$$

Le hamiltonien est dans ce cas égal à l'énergie cinétique de la particule, plus une constante appelée *énergie de masse*. Dans l'approximation non relativiste, H devient

$$H \approx mc^2 + \frac{1}{2}mv^2 \quad (5.36)$$

On reconnaît l'énergie cinétique usuelle, plus une constante : l'énergie au repos mc^2 .

Le hamiltonien doit en principe être exprimé en fonction des moments conjugués. On vérifie facilement dans ce cas-ci que H a l'expression suivante en fonction de \mathbf{p} :

$$H = \sqrt{\mathbf{p}^2c^2 + m^2c^4} \quad (5.37)$$

5.B.2 Particule chargée dans un champ électromagnétique

Considérons maintenant non pas une particule libre, mais une particule de charge e dans un champ électromagnétique. Nous allons supposer que le champ électromagnétique peut être décrit par un quadrivecteur $A^\mu(x)$, fonction de la position et du temps. Nous constaterons plus tard que ce quadrivecteur est constitué du potentiel électrique Φ comme composante temporelle, et du potentiel vecteur \mathbf{A} comme partie spatiale : $A^\mu = (\Phi/c, \mathbf{A})$, mais cette identification n'est pas nécessaire pour le moment. Nous allons cependant supposer que le produit eA^μ a les unités de l'énergie, ce qui compatible avec cette identification.¹

À l'action de la particule libre – appelons-la S_0 – on doit ajouter une action S_1 décrivant l'interaction de cette particule avec le champ électromagnétique. Cette action doit bien sûr être invariante. L'invariant le plus simple qu'on peut construire à l'aide du quadrivecteur A^μ et du mouvement de la particule est l'intégrale

$$S_1 = -e \int A_\mu dx^\mu \quad (5.38)$$

où nous avons multiplié par la charge e de la particule, car l'interaction entre une particule chargée et le champ électromagnétique doit nécessairement être proportionnelle à sa charge. Le facteur $1/c$ assure que l'action a les bonnes unités. L'intégrale est effectuée sur la trajectoire de la particule. Comme l'intégrand est la contraction de deux quadrivecteurs, il est forcément invariant. Cette action peut aussi s'exprimer ainsi :

$$\begin{aligned} S_1 &= e \int (\mathbf{A} \cdot d\mathbf{r} - \Phi dt) \\ &= e \int dt(\mathbf{A} \cdot \mathbf{v} - \Phi) \end{aligned} \quad (5.39)$$

L'action à considérer est l'action totale $S_0 + S_1$, ou le lagrangien correspondant :

$$L = -mc^2 \sqrt{1 - \frac{v^2}{c^2}} + e\mathbf{A} \cdot \mathbf{v} - e\Phi \quad (5.40)$$

1. Nous allons utiliser le système international (SI) pour les unités électromagnétiques, en dépit qu'il ne soit pas idéal pour décrire l'électromagnétisme dans un contexte relativiste. Le système de Heaviside-Lorentz lui est certainement préférable, mais est moins familier.

Nous allons maintenant obtenir les équations du mouvement et s'assurer qu'elles coïncident avec la formule connue de la force de Lorentz. Les équations de Lagrange sont, dans ce cas,

$$\frac{d}{dt} \frac{\partial L}{\partial \mathbf{v}} - \frac{\partial L}{\partial \mathbf{r}} = 0 = \frac{d}{dt} \left(\frac{m\mathbf{v}}{\sqrt{1-v^2/c^2}} \right) + e\dot{\mathbf{A}} - ev_i(\nabla A_i) + e\nabla\Phi \quad (5.41)$$

Le premier terme du membre de droite est par définition la force \mathbf{F} (la dérivée de la quantité de mouvement par rapport au temps). On peut donc écrire

$$\mathbf{F} = -e\nabla\Phi - e\dot{\mathbf{A}} + ev_j\nabla A_j \quad (5.42)$$

Notons que la dérivée temporelle $\dot{\mathbf{A}}$ est une dérivée *totale* par rapport au temps et non une dérivée partielle. Cette dérivée est non seulement due à la variation temporelle de \mathbf{A} à une position donnée, mais aussi au déplacement de la particule dans le temps d'un point à un autre :

$$\dot{\mathbf{A}} = \frac{\partial \mathbf{A}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{A} \quad (5.43)$$

La composante i de la force totale peut donc s'exprimer comme

$$F_i = -e\partial_i\Phi - e\frac{\partial A_i}{\partial t} - ev_j\partial_j A_i + ev_j\partial_i A_j \quad (5.44)$$

mais

$$-v_j\partial_j A_i + v_j\partial_i A_j = \varepsilon_{ijk}\varepsilon_{krs}v_j\partial_r A_s = (\mathbf{v} \wedge (\nabla \wedge \mathbf{A}))_i \quad (5.45)$$

Donc on obtient

$$\mathbf{F} = -e\nabla\Phi - e\frac{\partial \mathbf{A}}{\partial t} + e\mathbf{v} \wedge (\nabla \wedge \mathbf{A}) \quad (5.46)$$

En fonction du champ électrique \mathbf{E} et du champ magnétique \mathbf{B} définis par

$$\mathbf{E} = -\nabla\Phi - \frac{\partial \mathbf{A}}{\partial t} \quad \mathbf{B} = \nabla \wedge \mathbf{A} \quad (5.47)$$

on écrit finalement

$$\mathbf{F} = e\mathbf{E} + e\mathbf{v} \wedge \mathbf{B} \quad (5.48)$$

qu'on appelle la *force de Lorentz*.

Construisons maintenant le hamiltonien. Le moment généralisé est

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1-\beta^2}} + e\mathbf{A} \quad (5.49)$$

Le moment généralisé est, dans ce cas, distinct de la quantité de mouvement $m\gamma\mathbf{v}$. Le hamiltonien correspondant est

$$H = \mathbf{v} \cdot \mathbf{p} - L = \frac{mc^2}{\sqrt{1-\beta^2}} + e\Phi \quad (5.50)$$

où il est compris que la vitesse \mathbf{v} doit être remplacée par $(\mathbf{p} - e\mathbf{A})/m\gamma$. Si on définit la quantité de mouvement

$$\boldsymbol{\pi} := \frac{m\mathbf{v}}{\sqrt{1-\beta^2}} = \mathbf{p} - e\mathbf{A} \quad (5.51)$$

alors on a la relation suivante, conséquence immédiate de la relation connue entre l'énergie et la quantité de mouvement d'une particule :

$$H = \sqrt{m^2c^4 + \pi^2c^2} + e\Phi = \sqrt{m^2c^4 + (\mathbf{p} - e\mathbf{A})^2c^2} + e\Phi \quad (5.52)$$

Ceci est le hamiltonien relativiste d'une particule chargée dans un champ électromagnétique quelconque.

Dans l'approximation non relativiste, le hamiltonien devient

$$H \approx mc^2 + \frac{1}{2m} (\mathbf{p} - e\mathbf{A})^2 + e\Phi \quad (5.53)$$

5.B.3 Forme covariante de l'équation du mouvement

Plutôt que de passer par les équations de Lagrange, qui ne sont pas manifestement covariantes, nous pourrions déduire les équations du mouvement sous une forme covariante directement à partir du principe de la moindre action. C'est ce que nous ferons ici. L'action d'une particule chargée dans un champ électromagnétique est

$$\begin{aligned} S &= -mc^2 \int d\tau - e \int A_\mu dx^\mu \\ &= -mc \int \sqrt{dx^\mu dx_\mu} - e \int A_\mu dx^\mu \end{aligned}$$

Nous devons maintenant effectuer une variation de la trajectoire : $x^\mu \rightarrow x^\mu + \delta x^\mu$. La variation correspondante de l'action est, au premier ordre,

$$\begin{aligned} \delta S &= -mc \int \frac{1}{\sqrt{dx^\mu dx_\mu}} dx_\mu d\delta x^\mu - e \int \partial_\nu A_\mu \delta x^\nu dx^\mu - e \int A_\mu d\delta x^\mu \\ &= -m \int \dot{x}_\mu d\delta x^\mu - e \int \partial_\mu A_\nu \delta x^\mu dx^\nu - e \int A_\mu d\delta x^\mu \\ &= -m \int \dot{x}_\mu \delta \dot{x}^\mu d\tau - e \int \partial_\mu A_\nu \delta x^\mu \dot{x}^\nu d\tau - e \int A_\mu \delta \dot{x}^\mu d\tau \\ &= m \int \ddot{x}_\mu \delta x^\mu d\tau - e \int \partial_\mu A_\nu \dot{x}^\nu \delta x^\mu d\tau + e \int \partial_\nu A_\mu \dot{x}^\nu \delta x^\mu d\tau \\ &= m \int \ddot{x}_\mu \delta x^\mu d\tau - e \int F_{\mu\nu} \dot{x}^\nu \delta x^\mu d\tau \end{aligned} \quad (5.54)$$

Nous avons effectué des intégrations par parties dans la quatrième ligne et utilisé la notation

$$\dot{x}^\mu = \frac{dx^\mu}{d\tau} \quad \text{et la relation} \quad dx^\mu = \dot{x}^\mu d\tau \quad (5.55)$$

Nous avons aussi introduit le *tenseur de Faraday*

$$F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu \quad (5.56)$$

Le tenseur de Faraday, de par sa définition, est antisymétrique. On constate, en utilisant la relation (5.47) que ses composantes sont les composantes des champs électrique et magnétique :

$$F_{\mu\nu} : \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix} \quad (5.57)$$

Comme la variation δx^μ est arbitraire, la condition $\delta S = 0$ entraîne l'équation suivante :

$$m\ddot{x}_\mu = eF_{\mu\nu}\dot{x}^\nu \quad (5.58)$$

Il s'agit en fait de la force de Lorentz, exprimée sous une forme covariante. Aucune approximation n'a été effectuée : c'est donc une expression relativiste. C'est un exercice simple de démontrer qu'on retrouve la forme habituelle de cette force en substituant l'expression de la quadri accélération et de la quadri vitesse.

Cette équation s'exprime également sous la forme suivante :

$$m\dot{u}^\mu = eF^\mu{}_\nu u^\nu \quad (5.59)$$

où le tenseur mixte $F^\mu{}_\nu = g^{\mu\lambda}F_{\lambda\nu}$, a les composantes suivantes (les trois dernières rangées changent de signe par rapport à $F_{\mu\nu}$) :

$$F^\mu{}_\nu : \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{pmatrix} \quad (5.60)$$

Exemple 5.1 Charge dans un champ électrique uniforme

Considérons une particule de masse m et de charge e dans un champ électrique constant et uniforme $\mathbf{E} = E\mathbf{x}$. L'éq. (5.59) se réduit alors à

$$\frac{d}{d\tau} \begin{pmatrix} u^0 \\ u^1 \\ u^2 \\ u^3 \end{pmatrix} = \begin{pmatrix} 0 & \varepsilon & 0 & 0 \\ \varepsilon & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u^0 \\ u^1 \\ u^2 \\ u^3 \end{pmatrix} \quad \varepsilon = \frac{eE}{mc} \quad (5.61)$$

ou encore

$$\dot{u}^0 = \varepsilon u^1 \quad \dot{u}^1 = \varepsilon u^0 \quad u^2 = \text{cte} \quad u^3 = \text{cte} \quad (5.62)$$

Les deux premières équations se découplent pour donner

$$\ddot{u}^0 = \varepsilon^2 u^0 \quad (5.63)$$

dont la solution s'exprime par des fonctions hyperboliques :

$$u^0(\tau) = u^0(0) \cosh(\varepsilon\tau) + \frac{\dot{u}^0(0)}{\varepsilon} \sinh(\varepsilon\tau) \quad (5.64)$$

En supposant que la particule est initialement au repos, c'est-à-dire que $u^0(0) = c$ et $\dot{u}^0(0) = 0$, on trouve

$$u^0(\tau) = c \cosh(\varepsilon\tau) \quad \text{et} \quad u^1(\tau) = c \sinh(\varepsilon\tau) \quad (5.65)$$

En intégrant cette solution sur le temps propre, on trouve

$$x^0(\tau) = \frac{c}{\varepsilon} \sinh(\varepsilon\tau) \quad \text{et} \quad x^1(\tau) = \frac{c}{\varepsilon} (\cosh(\varepsilon\tau) - 1) \quad (5.66)$$

où nous avons ajouté une constante d'intégration de manière à respecter les conditions aux limites $x^0(0) = 0$ et $x^1(0) = 0$. La première de ces équations nous donne la relation entre le temps du référentiel $t = x^0/c$ et le temps propre τ :

$$t = \frac{mc}{eE} \sinh \frac{eE\tau}{mc} \quad (5.67)$$

Quand τ est petit, le développement de Taylor de cette expression donne $t \approx \tau$. Il s'agit du mouvement dit *hyperbolique*, associé en relativité à une force constante.

Exemple 5.2 Charge dans un champ magnétique uniforme

Supposons le champ magnétique orienté selon z . L'éq. (5.59) se réduit alors à

$$\frac{d}{d\tau} \begin{pmatrix} u^0 \\ u^1 \\ u^2 \\ u^3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \omega & 0 \\ 0 & -\omega & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u^0 \\ u^1 \\ u^2 \\ u^3 \end{pmatrix} \quad \omega = \frac{eB}{m} \quad (5.68)$$

ou encore

$$u^0 = \text{cte} \quad \dot{u}^1 = \omega u^2 \quad \dot{u}^2 = -\omega u^1 \quad u^3 = \text{cte} \quad (5.69)$$

La grandeur de la vitesse est constante, puisque u^0 (donc γ) est constant. Donc la relation entre le temps du référentiel t et le temps propre est simplement $t = \gamma\tau$. D'autre part, u^3 est constant, c'est-à-dire que la composante de la vitesse parallèle à \mathbf{B} est constante. Les deux équations du centre, quant à elles, se découpent pour donner

$$\ddot{u}^1 = -\omega^2 u^1 \quad (5.70)$$

dont la solution s'exprime par des fonctions circulaires :

$$u^1(\tau) = u^1(0) \cos(\omega\tau) + \frac{\dot{u}^1(0)}{\omega} \sin(\omega\tau) \quad (5.71)$$

En supposant, par exemple, que $\dot{u}^1(0) = 0$ et $u^1(0) = \gamma v$, on trouve

$$u^1(\tau) = \gamma v \cos(\omega\tau) \quad u^2(\tau) = -\gamma v \sin(\omega\tau) \quad (5.72)$$

ou encore

$$v_x(t) = v \cos(\omega t / \gamma) \quad v_y(t) = -v \sin(\omega t / \gamma) \quad (5.73)$$

La vitesse suit une évolution circulaire à une fréquence ω/γ , où ω est appelée la *fréquence cyclotron*.

5.B.4 Tenseur de Faraday et transformation des champs

La version contravariante du tenseur de Faraday est

$$F^{\mu\nu} : \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix} \quad (5.74)$$

Lors d'un changement de référentiel, le tenseur de Faraday se transforme de la manière suivante :

$$F'^{\mu\nu} = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma F^{\rho\sigma} \quad (5.75)$$

En appliquant la transformation de Lorentz (5.9), on obtient explicitement (nous avons multiplié par c)

$$\begin{aligned} \begin{pmatrix} 0 & -E'_x & -E'_y & -E'_z \\ E'_x & 0 & -cB'_z & cB'_y \\ E'_y & cB'_z & 0 & -cB'_x \\ E'_z & -cB'_y & cB'_x & 0 \end{pmatrix} &= \Lambda \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -cB_z & cB_y \\ E_y & cB_z & 0 & -cB_x \\ E_z & -cB_y & cB_x & 0 \end{pmatrix} \Lambda \\ &= \Lambda \begin{pmatrix} \beta\gamma E_x & -\gamma E_x & -E_y & -E_x \\ \gamma E_x & -\beta\gamma E_x & -cB_z & cB_y \\ \gamma(E_y - \beta cB_z) & \gamma(cB_z - \beta E_y) & 0 & -cB_x \\ \gamma(\beta cB_y + E_x) & -\gamma(cB_y + \beta E_x) & cB_x & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -E_x & -\gamma(-\beta cB_z + E_y) & -\gamma(\beta cB_y + E_z) \\ E_x & 0 & \gamma(\beta E_y - cB_z) & \gamma(cB_y + \beta E_x) \\ \gamma(E_y - \beta cB_z) & \gamma(cB_z - \beta E_y) & 0 & -cB_x \\ \gamma(\beta cB_y + E_x) & -\gamma(cB_y + \beta E_x) & cB_x & 0 \end{pmatrix} \end{aligned}$$

Et donc

$$\begin{aligned} E'_x &= E_x & B'_x &= B_x \\ E'_y &= \gamma(E_y - \beta cB_z) & B'_y &= \gamma(B_y + \beta E_z/c) \\ E'_z &= \gamma(E_z + \beta cB_y) & B'_z &= \gamma(B_z - \beta E_y/c) \end{aligned} \quad (5.76)$$

Exprimons ces transformations différemment en appelant respectivement \mathbf{E}_\perp et \mathbf{E}_\parallel les composantes de \mathbf{E} perpendiculaire et parallèle à la direction de la vitesse relative \mathbf{v} des deux référentiels :

$$\begin{aligned} \mathbf{E}'_\perp &= \gamma (\mathbf{E}_\perp + \mathbf{v} \wedge \mathbf{B}_\perp) & \mathbf{E}'_\parallel &= \mathbf{E}_\parallel \\ \mathbf{B}'_\perp &= \gamma \left(\mathbf{B}_\perp - \frac{\mathbf{v}}{c^2} \wedge \mathbf{E}_\perp \right) & \mathbf{B}'_\parallel &= \mathbf{B}_\parallel \end{aligned} \quad (5.77)$$

Remarques:

- ◆ Supposons que $\mathbf{E} = 0$ et $\mathbf{B} \neq 0$ dans S . Par un changement de référentiel, on peut augmenter la composante de \mathbf{B} perpendiculaire à \mathbf{v} et faire apparaître un champ électrique perpendiculaire à \mathbf{v} .
- ◆ De même, si $\mathbf{E} \neq 0$ et $\mathbf{B} = 0$ dans S , un changement de référentiel augmente la composante perpendiculaire de \mathbf{E} et fait apparaître un champ magnétique.
- ◆ Si $\mathbf{B} \neq 0$ et $\mathbf{E} = 0$ dans un référentiel, alors il n'existe aucun référentiel dans lequel $\mathbf{B} = 0$. De même, si $\mathbf{B} = 0$ et $\mathbf{E} \neq 0$ dans un référentiel, alors $\mathbf{E} \neq 0$ dans tous les référentiels.

Invariants du champ

Un invariant du champ est une quantité formée à l'aide des champs électrique et magnétique qui ne change pas d'un référentiel à l'autre. La contraction $F^\mu{}_\mu$, quoiqu'invariante, n'est pas intéressante, car elle s'annule (le tenseur de Faraday étant antisymétrique). Par contre, la contraction $F^{\mu\nu}F_{\mu\nu}$ est un invariant non trivial. Un calcul explicite donne

$$-\frac{1}{2}F^{\mu\nu}F_{\mu\nu} = \frac{\mathbf{E}^2}{c^2} - \mathbf{B}^2 \quad (5.78)$$

Un autre invariant s'obtient en considérant plutôt la contraction $\varepsilon^{\mu\nu\rho\lambda}F_{\mu\nu}F_{\rho\lambda}$. On calcule que

$$\frac{1}{4}\varepsilon^{\mu\nu\rho\lambda}F_{\mu\nu}F_{\rho\lambda} = \frac{1}{c}\mathbf{E} \cdot \mathbf{B} \quad (5.79)$$

Donc le produit scalaire $\mathbf{E} \cdot \mathbf{B}$ est invariant.

Remarques:

- ◆ Si \mathbf{E} et $c\mathbf{B}$ sont de même grandeur dans un référentiel, ils le sont dans tous les référentiels ($\mathbf{E}^2 - c^2\mathbf{B}^2 = 0$).
- ◆ Si $\mathbf{B} = 0$ et $\mathbf{E} \neq 0$ dans un référentiel, alors \mathbf{E} ne pourra être plus petit dans un autre référentiel.
- ◆ Si \mathbf{E} et \mathbf{B} sont perpendiculaires dans un référentiel, ils le seront dans tous les référentiels ($\mathbf{E} \cdot \mathbf{B} = 0$).

C Théorie des circuits électriques

L'objectif de cette section est de montrer comment on peut décrire un circuit électrique dans le formalisme lagrangien et hamiltonien. L'utilité de décrire des circuits dans ce langage vient essentiellement du besoin de les traiter de manière quantique, donc du besoin de définir un hamiltonien et des relations de commutation entre variables conjuguées.²

5.C.1 Description générale des circuits

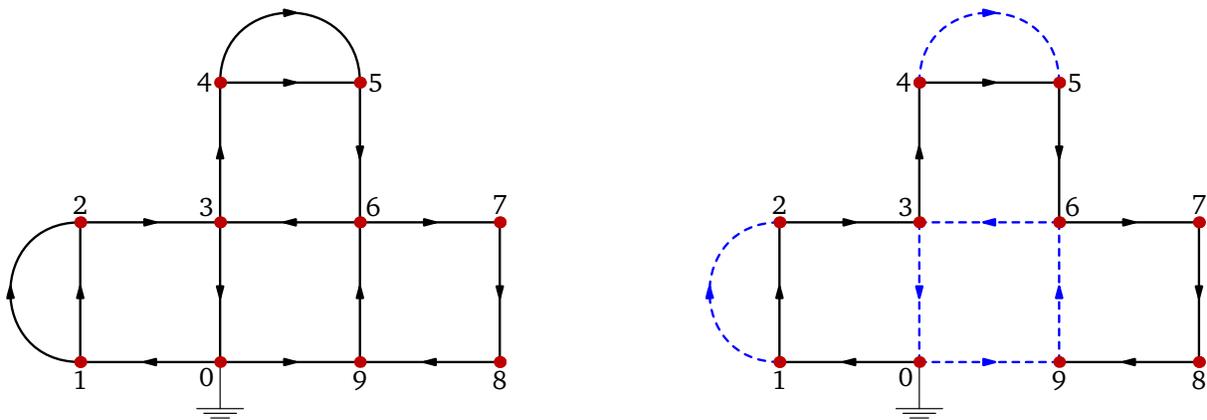


FIGURE 5.2

À gauche : Exemple de circuit. Chaque lien contient un élément de circuit (condensateur, inductance, etc) qui n'est pas représenté ici. En outre, chaque lien possède une orientation conventionnelle. À droite : un arbre fondamental A est indiqué en noir, et la fermeture du circuit en bleu.

Un circuit électrique peut être caractérisé par un graphe, c'est-à-dire par un ensemble de points qu'on appelle les *noeuds* du circuit, reliés par des *éléments de circuit* qui résident sur les *liens* du graphe, qui sont des segments orientés. Considérons à cet effet le graphe illustré sur la figure 5.2. Les noeuds ont été numérotés; les liens auraient aussi pu être numérotés, indépendamment. L'un des sites (numéroté 0) est mis à la masse. Notez que deux noeuds peuvent être reliés par plus d'un lien, par exemple si deux éléments de circuits sont mis en parallèle.

2. Comment un circuit peut-il nécessiter un traitement quantique? Cette question déborde largement le cadre du cours, mais expliquons tout de même sommairement : Un système doit être traité par la mécanique quantique si (i) l'énergie thermique $k_B T$ est inférieure à la séparation des niveaux d'énergie et (ii) les niveaux d'énergie sont bien séparés les uns des autres, c'est-à-dire que l'interaction avec l'environnement ne donne pas à ces niveaux une largeur (ou une incertitude) du même ordre ou supérieure à leur séparation. Dans le cas d'un circuit LC simple, cela entraîne que la fréquence de résonance $\omega = 1/\sqrt{LC}$ doit être telle que $\hbar\omega > k_B T$ et que le bruit qui entre dans le circuit par l'environnement doit être suffisamment faible. Cela est effectivement possible avec des microcircuits, qui possèdent des fréquences caractéristiques de l'ordre du GHz, si on les étudie à des températures inférieures à 100 mK. Ajoutons cependant que le traitement quantique d'un circuit est surtout intéressant si ce circuit comporte des éléments non linéaires, par exemple les *jonctions Josephson*, qui sont des structures composées de matériaux supraconducteurs entre lesquels une mince barrière isolante a été insérée. Des circuits comportant ces éléments nous permettent de construire des systèmes quantiques artificiels dont les hamiltoniens sont contrôlables jusqu'à un certain point. Ceux-ci sont la clef du développement de technologies de l'information quantique.

Il est essentiel, afin de bien définir le voltage sur chaque noeud, de relier chaque noeud de manière unique à la masse ; à cette fin, on divise l'ensemble des liens en deux parties : un *arbre fondamental* A, qui ne contient aucune boucle, indiqué en noir sur la partie droite de la figure 5.2, et une *fermeture* F, qui contient les autres liens. Cette séparation entre A et F n'est pas unique et comporte un degré d'arbitraire, qui n'affecte pas les équations du circuit obtenues à la fin.

Variables de lien

Les liens d'un circuit peuvent être numérotés et seront désignés par un indice l allant de 1 à N_L (le nombre d'éléments de circuit). On définit alors les deux variables suivantes sur chaque lien :

$$\begin{aligned} v_l &= \int \mathbf{E} \cdot d\mathbf{l} \quad (\text{intégré dans le sens du lien}) \\ i_l &= \frac{1}{\mu_0} \oint \mathbf{B} \cdot d\mathbf{l} \quad (\text{intégré autour du lien}) \end{aligned} \quad (5.80)$$

v_l est la tension électrique entre les deux bornes de l'élément de circuit résidant sur le lien l , et i_l le courant circulant dans l'élément de circuit. Ces variables dépendent du temps, naturellement. Le tension de lien est définie comme la force électromotrice, c'est-à-dire l'intégrale du champ électrique, sur un chemin longeant l'extérieur du dispositif. De même, le courant est défini par une circulation du champ magnétique le long d'une boucle entourant le dispositif, sans le pénétrer, et comprend donc la contribution du courant de déplacement. Ces variables seront indépendantes des chemins précis qui sont suivis, pourvu que ceux-ci ne pénètrent pas à l'intérieur du dispositif.

On définit également les variables auxiliaires suivantes, qui sont les intégrales de la tension et du courant sur le temps, à partir d'un moment distant dans le passé où les tensions et courants étaient nuls et le circuit inerte :

$$\begin{aligned} \Phi_l(t) &:= \int_{-\infty}^t v_l(\tau) d\tau \\ Q_l(t) &:= \int_{-\infty}^t i_l(\tau) d\tau \end{aligned} \quad (5.81)$$

Q_l est la charge totale qui est passée au travers de l'élément de circuit depuis le début. Φ_l est appelé la *variable de flux* associée à l'élément de circuit. Pour comprendre cette terminologie, considérons une boucle fermée B formée d'un ensemble de liens $\{l\}$. Cette boucle est orientée, et soit S_{Bl} l'orientation du lien l dans cette boucle : $S_{Bl} = +1$ si l'orientation conventionnelle du lien coïncide avec celle de la boucle, -1 si elle est inversée, et 0 si le lien ne fait pas partie de la boucle. Dans ce cas, la force électromotrice autour de cette boucle est, par définition,

$$\text{force électromotrice} = \sum_l S_{Bl} v_l \quad (5.82)$$

D'après la loi de Faraday, cette force électromotrice est l'opposée de la dérivée du flux magnétique $\tilde{\Phi}_B$ passant au travers de la boucle, avec la convention de signe habituelle (règle de la main droite). Donc

$$\sum_l S_{Bl} v_l = \frac{d}{dt} \sum_l S_{Bl} \Phi_l = -\frac{d}{dt} \tilde{\Phi}_B \quad (5.83)$$

et les variables Φ_l ont donc l'interprétation de « flux partiels » (moyennant un signe) :

$$\sum_l S_{Bl} \Phi_l = -\tilde{\Phi}_B \quad (5.84)$$

(nous avons supposé que la constante d'intégration est nulle, les flux étant tous nuls initialement).

De même, on définit une « charge » \tilde{Q}_n associée à chaque noeud, par la relation

$$\sum_l S_{nl} Q_l = \tilde{Q}_n \quad (5.85)$$

où S_{nl} est 1 si le lien l pointe vers le noeud n , -1 s'il pointe en direction contraire, et 0 s'il n'est pas connecté au noeud n . En dérivant cette équation par rapport au temps, on trouve

$$\sum_l S_{nl} \dot{Q}_l = \frac{d\tilde{Q}_n}{dt} = 0 \quad (5.86)$$

c'est-à-dire que la somme des courants de charge aboutissant à un noeud doit s'annuler, par conservation de la charge électrique. En effet, même si le courant i_l , tel que défini sur le lien, peut comporter un courant de déplacement, ce dernier doit être supporté par un dispositif capacitif. Les courants qui aboutissent aux noeuds sont purement des courants de charge. En général, nous allons aussi supposer que le flux $\tilde{\Phi}_B$ au travers d'une boucle est nul ou constant – par exemple causé par un champ magnétique externe au circuit – de sorte que l'éq. (5.84) est effectivement une contrainte sur les variables Φ_l .

Équations constitutives

Les variables (v_l, i_l) sont reliées entre elles, souvent par l'intermédiaire des variables auxiliaires (Φ_l, Q_l) . Si un élément de circuit est purement résistif, caractérisé par une résistance R , on a la relation $v_l = Ri_l$ avec la convention de signe adoptée ci-haut. S'il est purement capacitif, de capacité C on a la relation $Q_l = Cv_l$. Enfin, s'il est purement inductif, d'inductance L , la relation est plutôt $\Phi_l = Li_l$. En général, les éléments de circuit réels seront une combinaison de ces trois types, ou seront de nature plus complexe, menant à une relation *non linéaire* entre les variables i_l et v_l . Remarquons que si un élément de circuit est une inductance, le flux magnétique à travers cette inductance n'est pas rigoureusement défini, sauf si on ferme le circuit ; la relation $\Phi_l = Li_l$ ne faisant pas référence à un circuit fermé, elle n'est qu'approximative, ou idéalisée.

Variables de noeud

Dans le but d'étudier le circuit à l'aide d'un ensemble de variables indépendantes, non sujettes aux contraintes (5.84) et (5.85), nous allons introduire des variables de noeuds comme suit : Après avoir défini un arbre fondamental A pour le circuit, on définit le *flux* ϕ_n associé au noeud n ainsi :

$$\phi_n = \sum_{l \in A_n} S_{nl} \Phi_l \quad (5.87)$$

où A_n est un ensemble de liens formant un chemin unique pour aller du noeud de base (la masse) au noeud n , le long de l'arbre A . La variable de lien Φ_l est obtenue de manière unique comme

$$\Phi_l = \phi_{n'} - \phi_n \quad (5.88)$$

où n' et n sont les noeuds situés de part et d'autre du lien l sur l'arbre A , la direction de l étant définie de n jusqu'à n' . De même, on définit la tension de noeud V_n comme

$$V_n = \sum_{l \in A_n} S_{nl} v_l \quad (5.89)$$

et on a la relation évidente $V_n = \dot{\phi}_n$. Comme la tension est définie par rapport à la masse, le noeud 0 ne définit pas une variable indépendante, et par définition $V_0 = 0$ et $\phi_0 = 0$.

5.C.2 Lagrangien et hamiltonien d'un circuit

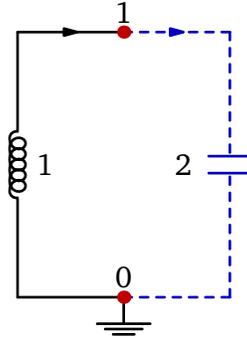


FIGURE 5.3

Circuit LC

Comme exemple fondamental, considérons le circuit LC illustré à la figure 5.3. Ce circuit comporte deux noeuds (indiqués par les points rouges), numérotés 0 et 1. Le noeud 0 est relié à la masse. Les deux liens sont numérotés 1 et 2 et comportent respectivement une inductance L_1 et une capacité C_2 . Par convention, le lien 1 constitue l'arbre fondamental, et le lien 2 la fermeture F. Ce circuit ne comporte qu'une variable indépendante, qu'on peut choisir comme étant ϕ_1 , la variable de flux au noeud 1. Les équations constitutives des éléments présents nous informent que

$$\Phi_1 = L_1 i_1 \quad Q_2 = C_2 v_2 \quad (5.90)$$

Avec les orientations choisies pour les liens, la contrainte de conservation de la charge au noeud 1 s'écrit $i_1 = i_2$, ou encore

$$\frac{\Phi_1}{L_1} = \dot{Q}_2 = C_2 \dot{v}_2 \quad (5.91)$$

Dans ce cas, la variable de noeud ϕ_1 coïncide avec Φ_1 . D'autre part, en supposant que le flux $\tilde{\Phi}$ à travers la boucle est négligeable, on a

$$\Phi_1 + \Phi_2 = 0 \implies \Phi_2 = -\Phi_1 = -\phi_1 \implies v_2 = -\dot{\phi}_1 \quad (5.92)$$

On obtient donc l'équation différentielle suivante :

$$\frac{\phi_1}{L_1 C_2} = \dot{v}_2 = -\ddot{\phi}_1 \implies \ddot{\phi}_1 + \omega^2 \phi_1 = 0 \quad (5.93)$$

où $\omega^2 = (L_1 C_2)^{-1}$. La solution à cette équation est oscillation pure : $\phi_1(t) = \phi(0) \cos(\omega t)$, en supposant que $\dot{\phi}(0) = 0$.

Remarque Une remarque importante s'impose ici, sur l'interprétation des variables de lien Φ_l : On suppose que le flux total $\tilde{\Phi}_B$ passant à travers la boucle unique de ce circuit est négligeable. Pourtant, l'un des éléments du circuit est une inductance, qui doit par définition porter un flux magnétique non négligeable. Comme l'autre élément (le condensateur) ne porte pas de flux magnétique, comme peut-on, au total, supposer que $\tilde{\Phi}_B$ est négligeable ? Rappelons cependant que le

chemin utilisé pour définir la force électromotrice le long de l'inductance passe tout juste à l'extérieur du dispositif, et donc ne contribue pas à porter le flux magnétique de l'inductance. Ainsi, si nous attachons un voltmètre aux bornes 0 et 1 du circuit LC, on se trouve à mesurer la tension de noeud V_1 (égale à la tension de lien v_1), le circuit se refermant dans le voltmètre. Une variation di_1/dt du courant dans l'inductance induit une force contre-électromotrice dans l'inductance, résultant en une tension $V_1 = L_1 di_1/dt$ qui aurait, à elle seule, tendance à contrer la variation de courant. Ceci justifie l'équation constitutive $\Phi_l = L_l i_l$ dans notre convention de signe. Ceci nous fait aussi réaliser que si le chemin d'intégration définissant v_1 passait en plein centre du fil formant une inductance, au lieu de passer à l'extérieur de l'inductance, le résultat obtenu serait l'opposé de v_1 . La définition des variables de lien en fonction de chemins définis à l'extérieur des dispositifs est donc très importante.

L'équation (5.93) pourrait être dérivée du lagrangien suivant :

$$L = \frac{1}{2}C_2\dot{\phi}_1^2 - \frac{1}{2}L_1^{-1}\phi_1^2 \quad (5.94)$$

car l'équation de Lagrange correspondante est

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\phi}_1} - \frac{\partial L}{\partial \phi_1} = C_2\ddot{\phi}_1 + L_1^{-1}\phi_1 = 0 \quad (5.95)$$

ce qui coïncide bien avec l'équation obtenue ci-haut. En fonction du flux ϕ_1 , tout se comporte comme si l'énergie cinétique était l'énergie contenue dans les éléments capacitifs, alors que l'énergie potentielle est contenue dans les éléments inductifs. Ce lagrangien est identique à celui d'un oscillateur harmonique, où la capacité C_2 joue le rôle de la masse (l'inertie) et l'inductance inverse L^{-1} celui de la force de rappel.

On montre, en électromagnétisme, que l'énergie contenue dans le champ électromagnétique prend deux formes : une contribution électrique et une contribution magnétique, données respectivement par

$$E_{\text{el.}} = \frac{1}{2}\epsilon_0 \int d^3r \mathbf{E}^2 \quad E_{\text{mag.}} = \frac{1}{2\mu_0} \int d^3r \mathbf{B}^2 \quad (5.96)$$

Le lagrangien du champ électromagnétique est la différence de ces deux énergies :

$$L_{\text{EM}} = \frac{1}{2} \int d^3r \left(\epsilon_0 \mathbf{E}^2 - \frac{1}{\mu_0} \mathbf{B}^2 \right) \quad (5.97)$$

Cette forme du lagrangien est nécessaire afin que l'action soit invariante de Lorentz (voir l'éq. (5.78)). En fait il s'agit de la seule expression possible pour l'action qui soit à la fois invariante de Lorentz et quadratique en champs :

$$S_{\text{EM}} = \int dt L_{\text{EM}} = \frac{1}{2} \int d^4x \left(\epsilon_0 \mathbf{E}^2 - \frac{1}{\mu_0} \mathbf{B}^2 \right) \quad (5.98)$$

L'intégrant de cette expression est invariant de Lorentz, ainsi que la mesure d'intégration d^4x sur l'espace-temps (le jacobien d'une transformation de Lorentz est l'unité). Il est alors naturel que, dans un circuit, le lagrangien soit la différence entre l'énergie électrique, comme celle emmagasinée dans un condensateur, et l'énergie magnétique, comme celle emmagasinée dans une inductance.

Ces considérations nous mènent à la conclusion que le lagrangien d'un circuit composé de condensateurs et d'inductances doit avoir la forme suivante :

$$L = \sum_l \frac{1}{2} C_l v_l^2 - \sum_l \frac{1}{L_l} \Phi_l^2 \quad (5.99)$$

Il s'agit ensuite, pour un circuit donné, d'exprimer ce lagrangien en fonction des variables de noeud ϕ_n .

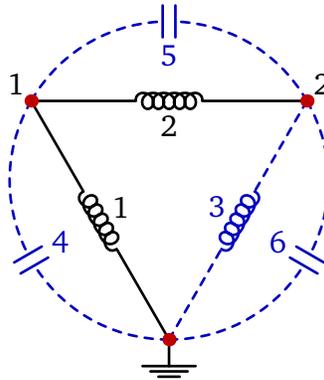
Le Hamiltonien, lui, s'obtient par la procédure habituelle. Dans le cas du circuit LC à une variable considérée plus haut, le moment conjugué à la variable ϕ_1 est défini par

$$q_1 = \frac{\partial L}{\partial \dot{\phi}_1} = C_2 \dot{\phi}_1 = C_2 v_1 \quad (5.100)$$

et n'est autre, dans ce cas-ci, que la charge du condensateur. En général, les variables conjuguées aux variables de flux ϕ_n seront des charges de site q_n , qui n'auront pas toujours une interprétation aussi simple que celle du circuit LC à une boucle. Les variables q_1 et ϕ_1 sont naturellement conjuguées : $[\phi_1, q_1] = 1$.

Exemple 5.3 Circuit LC à deux variables

Considérons le circuit illustré sur la figure, avec l'arbre fondamental et la fermeture indiquées en noir et en bleue pointillé, respectivement. Les noeuds et les liens sont numérotés comme indiqué. Trouvons une expression pour le hamiltonien de ce circuit.



Les deux variables indépendantes sont les flux de noeud ϕ_1 et ϕ_2 . Le lagrangien de ce circuit est

$$L = \frac{1}{2} \left(C_4 v_4^2 + C_5 v_5^2 + C_6 v_6^2 - \frac{\Phi_1^2}{L_1} - \frac{\Phi_2^2}{L_2} - \frac{\Phi_3^2}{L_3} \right) \quad (5.101)$$

En fonction de ϕ_1 et ϕ_2 , et en supposant que les flux de boucle $\tilde{\Phi}_B$ sont nuls, les tensions et flux figurant dans le lagrangien sont

$$\begin{aligned} \Phi_1 &= \phi_1 & v_4 &= \dot{\Phi}_4 = \dot{\phi}_1 \\ \Phi_2 &= \phi_2 - \phi_1 & v_5 &= \dot{\Phi}_5 = \dot{\phi}_2 - \dot{\phi}_1 \\ \Phi_3 &= -\phi_2 & v_6 &= \dot{\Phi}_6 = -\dot{\phi}_2 \end{aligned} \quad (5.102)$$

de sorte que le lagrangien s'exprime comme

$$L = \frac{1}{2} \left(C_4 \dot{\phi}_1^2 + C_5 (\dot{\phi}_2 - \dot{\phi}_1)^2 + C_6 \dot{\phi}_2^2 - \frac{\phi_1^2}{L_1} - \frac{(\phi_2 - \phi_1)^2}{L_2} - \frac{\phi_2^2}{L_3} \right) \quad (5.103)$$

La partie cinétique (i.e. capacitive) s'écrit ainsi sous forme matricielle :

$$T = \frac{1}{2} (\dot{\phi}_1, \dot{\phi}_2) C \begin{pmatrix} \dot{\phi}_1 \\ \dot{\phi}_2 \end{pmatrix} \quad C = \begin{pmatrix} C_4 + C_5 & -C_5 \\ -C_5 & C_5 + C_6 \end{pmatrix} \quad (5.104)$$

Les moments conjugués sont

$$q_1 = (C_4 + C_5) \dot{\phi}_1 - C_5 \dot{\phi}_2 \quad q_2 = (C_5 + C_6) \dot{\phi}_2 - C_5 \dot{\phi}_1 \quad (5.105)$$

Le hamiltonien s'obtient par la procédure habituelle (voir l'éq. (2.18)) :

$$H = \frac{1}{2} (q_1, q_2) C^{-1} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} + \frac{\phi_1^2}{2L_1} + \frac{(\phi_2 - \phi_1)^2}{2L_2} + \frac{\phi_2^2}{2L_3} \quad (5.106)$$

où

$$C^{-1} = \frac{1}{C_4 C_5 + C_5 C_6 + C_4 C_6} \begin{pmatrix} C_5 + C_6 & C_5 \\ C_5 & C_4 + C_5 \end{pmatrix} \quad (5.107)$$

Les moments conjugués et les variables de flux ont les crochets de Poisson suivants :

$$[\phi_i, q_j] = \delta_{ij} \quad (5.108)$$

Un circuit ne comportant que des inductances et des capacités est parfaitement analogue à un système masses-ressorts, tel que décrit à la section D. Il comporte donc des modes propres, chacun avec sa fréquence propre d'oscillation. Les calculs généraux menés à la section D s'appliquent intégralement.

Problèmes

Problème 5.1 Forme générale de la transformation de Lorentz

En partant de la transformation de Lorentz pour une vitesse relative $c\mathbf{u} = V\mathbf{x}$ des deux référentiels, déduisez-en la forme plus générale de la transformation de Lorentz pour une vitesse relative quelconque $c\mathbf{u}$.

Problème 5.2 Tenseur métrique et élément de volume

Supposons que les coordonnées spatiales x^i soient définies relativement à une base quelconque $\{\mathbf{e}_i\}$ ($i = 1, 2, 3$). Montrez que l'élément de volume, indépendant de la base, est donné par

$$d^3x \sqrt{g} \quad (5.109)$$

où $g := \det(g_{ij})$ est le déterminant du tenseur métrique.

Problème 5.3 Effet Compton

Ce problème vise à illustrer comment les lois de conservation de l'énergie et de la quantité de mouvement permettent de déterminer le changement dans l'énergie d'un photon, suite à une collision avec un électron au repos (effet Compton). Le formalisme des quadrivecteurs permet de résoudre ce problème de manière particulièrement efficace et élégante. Rappelons qu'une particule de quantité de mouvement \mathbf{p} est associée à une onde de longueur d'onde $\lambda = h/|\mathbf{p}|$ (h est la constante de Planck).

Un photon de quantité de mouvement \mathbf{p}_1 est incident sur un électron au repos (masse m), et est ensuite dévié d'un angle θ . Exprimez le changement dans la longueur d'onde du photon en fonction de l'angle θ et de la longueur d'onde de Compton de l'électron, définie comme $\lambda_c := h/mc$. Démontrez cette relation en vous servant des quadrivecteurs.

Indice : calculez l'invariant $p_4^2 = (\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{p}_3)^2$, où \mathbf{p}_1 et \mathbf{p}_3 sont respectivement les quadrivecteurs énergie-impulsion des photons incident et diffusé, alors que \mathbf{p}_2 est celui de l'électron qui sert de cible. Effectuez le développement de cette expression le plus loin possible avant de substituer les composantes des quadrivecteurs.

Problème 5.4 Force de Lorentz

Démontrez que l'éq. (5.58) est bel et bien équivalent à l'équation de la force de Lorentz (5.48).

Problème 5.5 Mouvement dans des champs électrique et magnétique croisés

Une particule de charge e est sous l'influence d'un champ électrique $\mathbf{E} = E_y$ et d'un champ magnétique $\mathbf{B} = B_z$, tels que $|E| < c|B|$. On suppose que la particule est initialement au repos et à l'origine des coordonnées.

A Dérivez une expression pour la quadri position $x^\mu(\tau)$ de la particule en fonction de son temps propre. Utilisez pour cela l'équation covariante du mouvement faisant intervenir le tenseur de Faraday et la quadri vitesse. Le problème se ramène à un problème de valeurs propres que vous pouvez résoudre à l'aide de Mathematica. Exprimez votre réponse en fonction du paramètre ξ défini par $\tanh \xi = E/cB$.

B Obtenez maintenant le même résultat, en partant de la solution connue dans le cas où $\mathbf{E} = 0$, et en procédant à une transformation de Lorentz vers un référentiel où le champ électrique est non nul et perpendiculaire à \mathbf{B} .

C Faites un graphique de la trajectoire de l'objet, dans le plan xy , sur une durée assez longue pour constater que la trajectoire se répète. Utilisez pour ce faire la valeur $E = cB/2$.

Problème 5.6 Hamiltonien d'une particule dans un champ magnétique uniforme

Une particule non relativiste de charge e et de masse m se déplace dans un champ magnétique uniforme dans la direction z : $\mathbf{B} = B_z$.

A Écrivez le hamiltonien correspondant en faisant le nécessaire pour qu'il ne dépende explicitement que de x et non pas de y ou z .

B Écrivez les équations de Hamilton correspondantes.

C Montrez que le problème se réduit à celui d'un oscillateur harmonique en raison de la conservation de p_y . Quelle est la fréquence associée ?

Problème 5.7 Problème de Kepler relativiste

Dans ce problème nous allons déterminer la forme de l'orbite d'un objet relativiste de masse m soumis à l'influence d'un potentiel attractif $-k/r$. En supposant, comme avant, que l'orbite est comprise dans le plan xy , le hamiltonien de ce système est, en coordonnées polaires (r, φ) ,

$$H = \sqrt{m^2c^4 + p_r^2c^2 + \frac{p_\varphi^2c^2}{r^2}} - \frac{k}{r} \quad (5.110)$$

A Montrez que l'équation différentielle de la trajectoire est

$$\frac{dr}{d\varphi} = \frac{r^2}{\ell} p_r \quad (5.111)$$

où $\ell = p_\varphi$ est conservé.

B En utilisant la conservation de l'énergie, en particulier en exprimant p_r en fonction de ℓ , de r et de E , montrez que l'orbite de l'objet prend la forme suivante :

$$r(\varphi) = \frac{A}{1 + e \cos(\gamma\varphi)} \quad (5.112)$$

où

$$A = \frac{\ell^2 c^2 - k^2}{kE} \quad (5.113)$$

$$\gamma = \sqrt{1 - \left(\frac{k}{\ell c}\right)^2} \quad (5.114)$$

$$e = \sqrt{1 - \left(1 - \frac{m^2 c^4}{E^2}\right) \left(1 - \frac{\ell^2 c^2}{k^2}\right)} \quad (5.115)$$

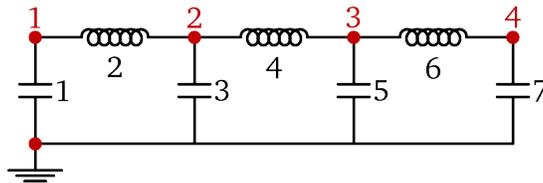
C Montrez que cette formule décrit une ellipse en précession

D Montrez que l'expression de l'excentricité e dans la limite non relativiste est bien celle que nous avons trouvée en traitant le problème de Kepler non relativiste. Posez $E = mc^2 + \epsilon$, où ϵ représente l'énergie totale moins l'énergie de masse, c'est-à-dire la quantité qui est appelée E au chapitre 4. Vous devez pour cela supposer que $\epsilon \ll mc^2$.

E Dans le cas où les effets relativistes sont faibles, mais non nuls, donnez une expression du taux de précession $\dot{\omega}$ du péricentre de l'orbite. Que vaudrait $\dot{\omega}$ pour la planète Mercure, et comment se compare-t-il au taux calculé en tenant compte de la relativité générale ?

Problème 5.8 Ligne de transmission

Considérez le circuit illustré. Tous les condensateurs ont la même capacité C et toutes les inductances ont la même valeur ℓ .



A Donnez une expression du lagrangien de ce circuit en fonction des flux de noeuds ϕ_1, \dots, ϕ_4 . Quel est le hamiltonien correspondant ?

B Écrivez l'équation de Lagrange pour la variable ϕ_3 . Supposons maintenant que le circuit, au lieu de ne contenir que 3 boucle, en compte un nombre infini vers la droite. Quelle serait alors l'équation de Lagrange pour la variable ϕ_n au noeud n ($n > 1$) ?

C Prenez la limite continue de cette équation, en introduisant une variable continue $x = an$, où a est la distance entre deux boucles successives. La fonction $\phi(x)$ est alors définie comme $\phi(an) = \phi_n$. Montrez que l'équation de Lagrange devient alors l'équation d'onde

$$\frac{1}{v^2} \frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} = 0 \quad (5.116)$$

et donnez une expression pour la vitesse de propagation v des ondes décrites par cette équation.

CHAPITRE 6

Introduction aux systèmes chaotiques

A Problème de Hénon-Heiles

6.A.1 Potentiel de Hénon-Heiles

Comme exemple de problème dans lequel le phénomène du chaos se manifeste, considérons le mouvement d'un objet en deux dimensions, dans le potentiel de Hénon-Heiles. Le hamiltonien correspondant est

$$H = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{1}{2}k(x^2 + y^2) + \lambda(x^2y - \frac{1}{3}y^3) \quad (6.1)$$

Le potentiel $V(x, y)$ est la somme d'un potentiel harmonique en deux dimensions et d'une perturbation cubique. Si on exprime ce problème en coordonnées polaires planes, on obtient plutôt la forme suivante :

$$H = \frac{p_r^2}{2m} + \frac{p_\varphi^2}{2mr^2} + \frac{1}{2}kr^2 + \lambda r^3 \sin 3\varphi \quad (6.2)$$

Ce problème est une simplification du problème plus complexe du mouvement d'une étoile par rapport au centre d'une galaxie spirale. L'absence de symétrie de rotation dans le plan xy est alors due aux bras de la galaxie.

Dans le but de simplifier l'analyse du hamiltonien (6.1), les constantes m , k et λ sont choisies de manière à simplifier au maximum les équations : $m = 1$, $k = 1$ et $\lambda = 1$. Il est certainement possible de choisir $m = 1$ sans perte de généralité, car cela revient à choisir des unités particulières pour le temps, c'est-à-dire une certaine unité pour l'énergie. De même, on peut choisir les unités de longueur de manière à fixer $k = 1$ sans perte de généralité. Choisir $\lambda = 1$ représente alors une restriction dans l'espace des paramètres, mais l'effet du terme cubique sera d'autant plus petit que l'énergie totale du système sera petite, ce qui nous permet quand même de juger de l'impact du paramètre λ en étudiant des mouvements d'énergies différentes.

Comme $p_x = \dot{x}$ et $p_y = \dot{y}$, la valeur constante de l'énergie dans ce problème peut s'écrire

$$E = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 + \frac{1}{2}x^2 + \frac{1}{2}y^2 + x^2y - \frac{1}{3}y^3 \quad (6.3)$$

Les équations du mouvement qui découlent de ce hamiltonien simplifié sont

$$\ddot{x} = -x - 2xy \quad \ddot{y} = -y - x^2 + y^2 \quad (6.4)$$

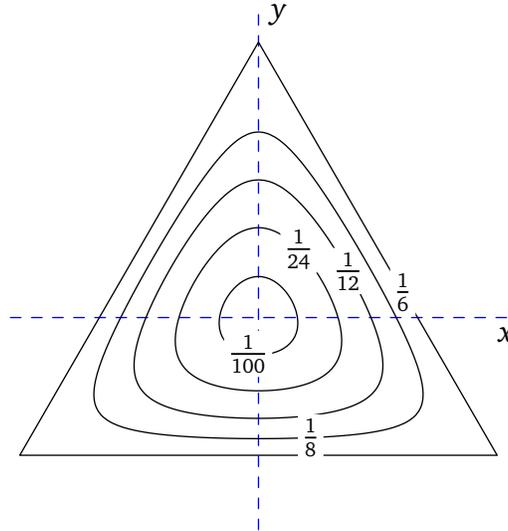


FIGURE 6.1

Ces équations sont impossible à résoudre analytiquement.

La figure 6.1 illustre les courbes du niveau du potentiel de Hénon-Heiles restreint

$$V(x, y) = \frac{1}{2}x^2 + \frac{1}{2}y^2 + x^2y - \frac{1}{3}y^3 = \frac{1}{2}r^2 + r^3 \sin 3\varphi \quad (6.5)$$

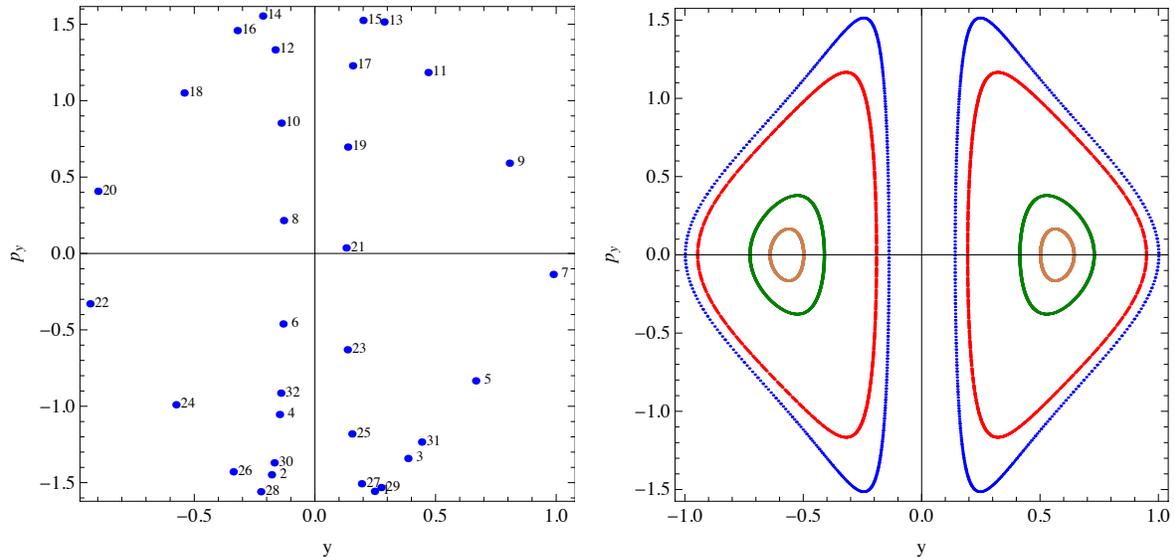
La valeur $\frac{1}{6}$ du potentiel correspond à une inflexion du potentiel au-delà de laquelle le potentiel est instable en raison des termes cubiques. Toutes les valeurs inscrites à l'intérieur du triangle équilatéral associé à la valeur $V = \frac{1}{6}$ sont inférieures à $\frac{1}{6}$. Nous nous bornerons donc à cette région.

Proche de l'origine, les termes cubiques sont petits en comparaison des termes quadratiques et le potentiel devient progressivement isotrope et quadratique. Les courbes d'énergie constantes se rapprochent de cercles et les trajectoires sont des ellipses centrées à l'origine.

6.A.2 Sections de Poincaré

Comme il est impossible de visualiser le mouvement de ce système dans l'espace des phases à 4 dimensions, on visualise plutôt le mouvement du système en sélectionnant une tranche du mouvement en deux dimensions. Une telle tranche est définie comme l'ensemble des points du mouvement $(q(t), p(t))$ qui sont situés sur l'hypersurface d'énergie E et, par exemple, sur le plan $x = 0$. Cet ensemble de points, paramétré par les variables y et $\dot{y} = p_y$, est appelé *section de Poincaré*. Il permet de suivre partiellement l'évolution du système et de caractériser le mouvement comme régulier ou chaotique.

Pour construire numériquement une section de Poincaré, on doit, en même temps que la solution du système d'équations de Hamilton se construit, surveiller cette solution et épinglez au passage, par interpolation, les points tels que $x = 0$ et ajouter les valeurs correspondantes de (y, \dot{y}) à la section. Quant à la valeur associée de \dot{x} , elle sera déterminée par la valeur de l'énergie E et n'est donc pas déterminante.


FIGURE 6.2

Sections de Poincaré pour une particule dans le potentiel (6.6), avec $k = 1$, énergie $E = -0.875$ et une masse unité. À gauche : $g = 0.004$ et une seule valeur des conditions initiales. Quelques points seulement sont illustrés et l'ordre dans lequel ils apparaissent est indiqué. À droite : $g = 0.001$ et différentes valeurs des conditions initiales, indiquées par des couleurs distinctes.

Considérons, pour illustrer le concept, le problème du mouvement d'une particule dans un potentiel képlérien, modifié par une perturbation :

$$V(\mathbf{r}) = -\frac{k}{r} - \frac{g}{3r^3} \quad (6.6)$$

Cette perturbation, comme il est bien connu, va entraîner une précession de l'orbite elliptique de la particule. Les sections de Poincaré illustrées à la figure 6.2 sont obtenues en portant en graphique les valeurs de y et \dot{y} à chaque passage de la particule à $x = 0$. Chaque couleur correspond à un ensemble différent de conditions initiales, donc à un mouvement différent. Chaque section forme un ensemble de deux courbes fermées symétriquement disposées de part et d'autre de $y = 0$, correspondant respectivement au croisement de l'axe $x = 0$ avec $\dot{x} > 0$ et $\dot{x} < 0$. Chaque courbe est en réalité une collection de points qui, à force d'accumulation, forme un ensemble à l'apparence continue. Ceci est le signe d'un mouvement régulier, non chaotique : les points apparaissent, successivement, à des positions différentes le long de la section, de manière régulière et prévisible, comme illustré sur la partie gauche de la figure 6.2.

Considérons ensuite les sections de Poincaré associées au problème de Hénon-Heiles, illustrées à la figure 6.3. Encore une fois, des couleurs différentes correspondent à des conditions initiales différentes. Le premier graphique correspond à une énergie relativement faible ($1/12$) et montre des sections de Poincaré régulières, comme dans le cas du potentiel de Kepler perturbé. En augmentant l'énergie à $E = 1/8$, un phénomène nouveau apparaît. Certaines sections, indiquées en bleu et en vert sur la figure, correspondent encore à des trajectoires régulières, associées à des conditions initiales particulières. Par contre, la section indiquée en rouge remplit un ensemble dense au lieu d'une courbe continue (le graphique ne montre qu'un nombre fini de points en raison de la durée limitée de la simulation, mais en augmentant cette durée, on observe une densité de points crois-

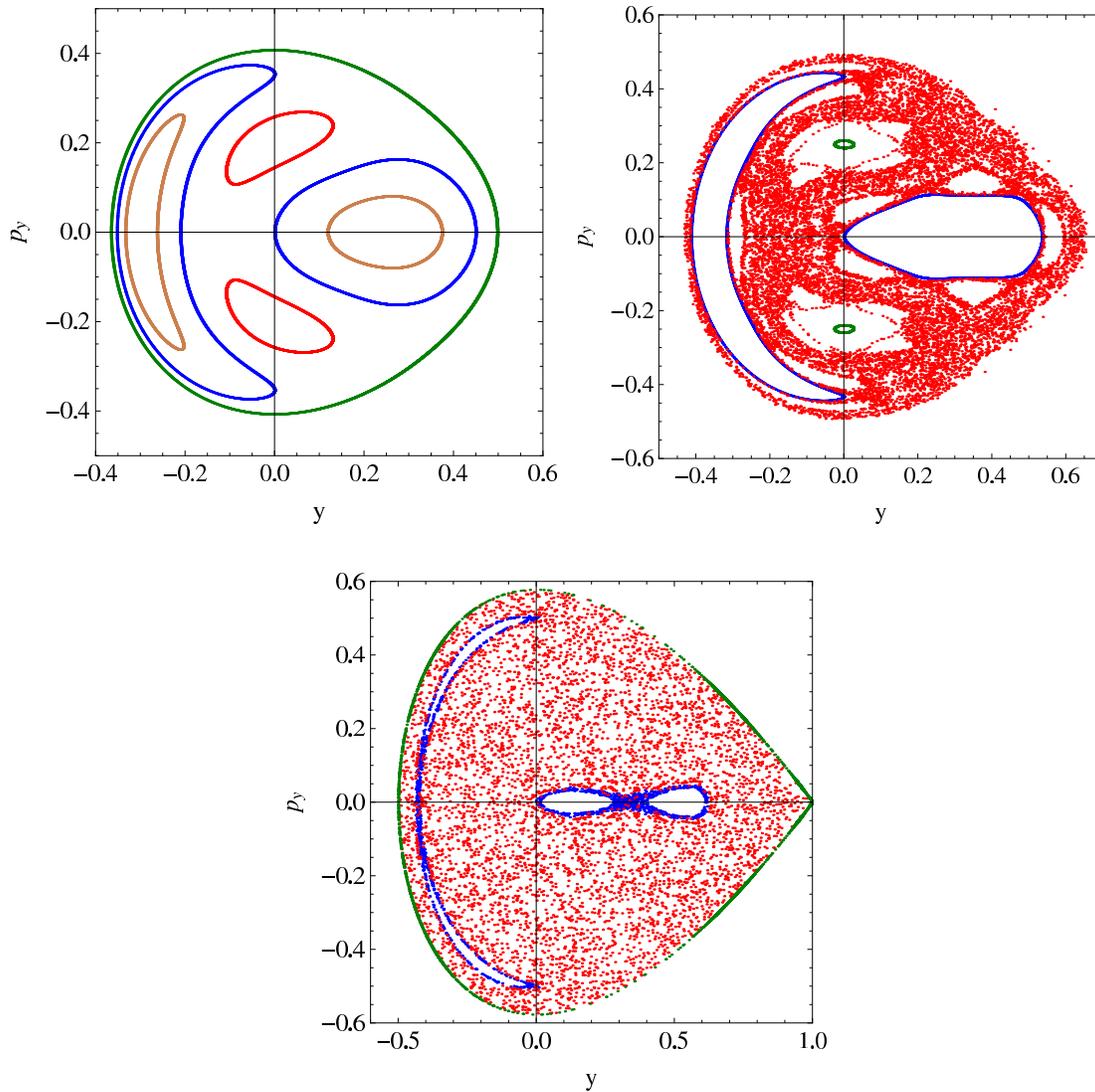


FIGURE 6.3

Sections de Poincaré du système de Hénon-Heiles, à différentes valeurs de l'énergie E . Les différentes couleurs représentent des conditions initiales différentes, mais de mêmes énergies.

En haut à gauche : $E = 1/12$; dans ce cas, on constate que l'orbite coupe la section en suivant des courbes régulières.

En haut à droite : $E = 1/8$; Pour certaines valeurs des conditions initiales, les orbites sont régulières. Pour d'autres, un attracteur étrange est suivi, dont la section remplit une partie dense de la section de Poincaré.

En bas : $E = 1/6$; presque tout l'espace disponible est occupé par l'attracteur étrange.

sante). Le système atteint alors un régime chaotique : les points successifs de la section de Poincaré apparaissent à des positions irrégulières dans un ensemble dense. Ce n'est pas par manque de déterminisme, car la simulation tout comme la solution exacte (non numérique) procède de manière strictement déterministe. C'est plutôt la manifestation d'un patron non répétitif et complexe.

Ce régime chaotique, en pratique, rend les prédictions à long terme impossibles : un léger changement des conditions initiales, ce qui revient à initier la section à un endroit légèrement différent, mène à une séquence de points complètement différents dans laquelle, après un certain nombre d'itérations, la position sur la section serait très différente de celle obtenue après le même nombre d'itérations à partir de la condition initiale originale : c'est la sensibilité aux conditions initiales, qui caractérise généralement un mouvement chaotique.

6.A.3 Stabilité et exposant de Liapounov

La sensibilité aux conditions initiales peut être caractérisée de la manière suivante : Si $\eta(0)$ représente le point initial dans l'espace des phases, alors une légère modification $\delta\eta(0)$ de cette condition initiale va résulter en une modification $\delta\eta(t)$ du mouvement à un temps ultérieur. Pour des changements infinitésimaux, ces variations vont obéir à une loi exponentielle :

$$\delta\eta(t) = \delta\eta(0)e^{\lambda t} \quad (6.7)$$

où λ est l'exposant de Liapounov.

On ne peut pas prouver que le comportement de $\delta\eta(t)$ est exponentiel en général. Par contre, il est facile de voir comment c'est le cas à proximité d'un *point fixe*. Considérons à cet effet un système général d'équations différentielles du premier ordre, pour N variables x_i :

$$\dot{x}_i = F_i(x) \quad (6.8)$$

Les N fonctions F_i sont en général non linéaires et définissent la dynamique considérée. Supposons que ce système d'équations admette un point fixe x_0 , c'est-à-dire tel que $F_i(x_0) = 0$. La solution constante $x_i(t) = x_{0i}$ satisfait évidemment au système (6.8). Considérons maintenant une variation $\delta x_i(t)$ du mouvement provenant d'une variation $\delta x_i(0)$ des conditions initiales. Le mouvement varié est donc $x_i(t) = x_{0i} + \delta x_i(t)$ et obéit aux mêmes équations du mouvement. On peut donc écrire

$$\dot{x}_{0i} + \delta\dot{x}_i = F_i(x_0 + \delta x) \approx F_i(x_0) + \left. \frac{\partial F_i}{\partial x_j} \right|_{x_0} \delta x_j \quad (6.9)$$

ou encore

$$\delta\dot{x}_i = \left. \frac{\partial F_i}{\partial x_j} \right|_{x_0} \delta x_j \quad (6.10)$$

La matrice $\partial_j F_i$ étant indépendante de δx et du temps dans cette approximation. Le problème est redevenu linéaire : nous avons procédé à une *linéarisation* du problème en développant autour d'un point fixe du problème non linéaire. L'équation (6.10) se résout de la manière habituelle, c'est-à-dire par diagonalisation : on définit un ensemble de variables normales δy_i , qui sont les vecteurs propres de la matrice $\partial_j F_i$, avec valeurs propres λ_i . L'évolution temporelle de ces modes propres est alors

$$\delta y_i(t) = \delta y_i(0)e^{\lambda_i t} \quad (6.11)$$

Si toutes les valeurs propres λ_i sont négatives, alors le point fixe x_0 est *stable*. Si l'une au moins des valeurs propres est positive, alors δx va croître avec le temps et le point fixe est instable. L'exposant de Liapounov est dans ce cas la plus grande des valeurs propres λ_i .

B Dynamique discrète et approche du chaos

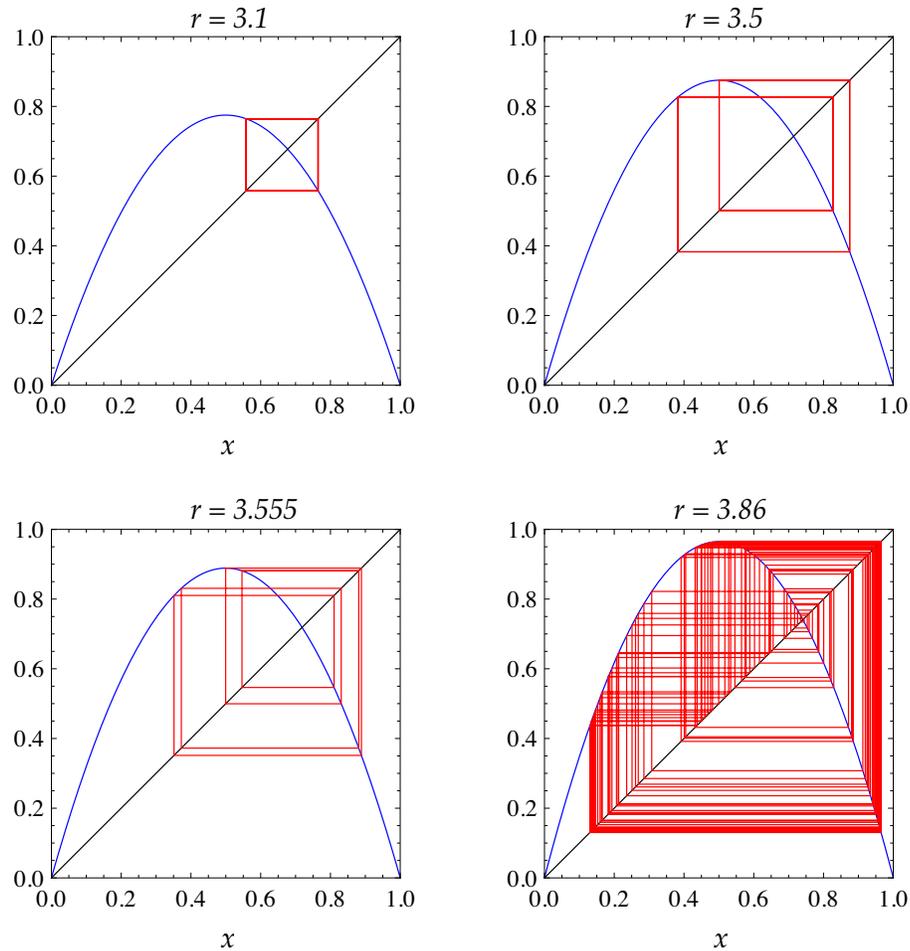


FIGURE 6.4
Exemples d'applications de la carte logistique.

Les sections de Poincaré sont basées sur une dynamique continue, c'est-à-dire sur une résolution d'équations différentielles ; on extrait ensuite de leur solution un ensemble de points discrets formant la section de Poincaré. La différence de temps entre deux points de la section n'est pas nécessairement toujours la même ; en fait, elle est généralement variable.

Une façon simple d'illustrer les propriétés d'une dynamique chaotique est d'étudier carrément une dynamique discrète, c'est-à-dire une règle de récurrence provenant d'une application de l'espace des

phases sur lui-même. Pour simplifier les choses au maximum, nous allons en fait nous concentrer sur une règle de récurrence à une seule variable x , définie dans l'intervalle $[0, 1]$. L'application particulière que nous allons étudier est la *carte logistique*, par laquelle une valeur de x dans l'intervalle $[0, 1]$ est envoyée dans une autre valeur dans cet intervalle :

$$x \mapsto rx(1-x) \quad r \in [0, 4] \quad (6.12)$$

où r est un paramètre fixe compris entre 0 et 4. Cette application, ou carte, prend la forme d'une parabole inversée, qui s'annule à $x = 0$ et $x = 1$, et dont le maximum est situé au point $(0.5, r/4)$.

Nous nous intéressons ici à la dynamique discrète générée par cette application, c'est-à-dire à la séquence $\{x_0, x_1, x_2, \dots\}$ des valeurs de x obtenues en appliquant la carte de manière itérative : $x_{n+1} = rx_n(1-x_n)$. Cette application comporte toujours un *point fixe*, soit une valeur x_f qui n'est pas affectée par l'application :

$$x_f = rx_f(1-x_f) \implies x_f = 0 \quad \text{ou} \quad x_f = \frac{r-1}{r} \quad (6.13)$$

Ces points fixes ne sont pas nécessairement stable. Procédons comme ci-haut, en linéarisant l'équation. Posons $x_n = x_f + \delta x_n$. En développant l'équation de récurrence au premier ordre en δx_n , on trouve

$$x_f + \delta x_{n+1} = rx_f(1-x_f) + r(1-2x_f)\delta x_n \implies \delta x_{n+1} = r(1-2x_f)\delta x_n \quad (6.14)$$

Le point fixe sera manifestement stable si $|r(1-2x_f)| < 1$ et instable si $|r(1-2x_f)| > 1$. En particulier, le point fixe $x_f = 0$ est stable si et seulement si $r < 1$, alors que le point fixe $x_f = (r-1)/r$ est stable si et seulement si

$$|2-r| < 1 \implies \begin{cases} r < 3 & \text{si } r > 2 \\ r > 1 & \text{si } r < 2 \end{cases} \quad (6.15)$$

ce qui revient à dire que le point fixe $x_f = (r-1)/r$ est stable dans l'intervalle $r \in [1, 3]$.

Que se passe-t-il quand $r > 3$? Considérons à cet effet la figure 6.4. Les différents graphiques de cette figure illustrent l'application répétée de la carte logistique en traçant un segment partant de la droite $y = x$ et aboutissant à la parabole inversée, ce qui représente l'application de la carte. Ensuite, à partir de la parabole, un segment horizontal rejoint la droite $y = x$, ce qui représente la substitution $x_{n+1} = f(x_n)$, f étant la carte logistique. On répète ensuite le processus indéfiniment. Les valeurs illustrées sur la figure sont toutes obtenues en appliquant la carte des centaines de fois avant de tracer les segments, de sorte que seul le comportement final (régulier ou non) est illustré.

Le premier graphique de la figure représente un *cycle limite* de longueur 2, c'est-à-dire une séquence de deux valeurs α_1 et α_2 qui se répètent de manière périodique :

$$x_n = x_{n+2} = x_{n+4} = \dots = \alpha_1 \quad x_{n+1} = x_{n+3} = x_{n+5} = \dots = \alpha_2 \quad (6.16)$$

En partant d'une valeur initiale x_0 quelconque, la séquence $\{x_n\}$ va converger vers ce comportement périodique, de période 2. Si on augmente quelque peu la valeur du paramètre de contrôle r , le cycle de période 2 devient éventuellement instable et cède sa place à un cycle de période 4, illustré sur le deuxième graphique de la figure 6.4. Ce doublement de la période du cycle limite porte le nom de *bifurcation*. Le troisième graphique de la figure illustre un cycle de période 8.

Notons r_n la valeur de r à laquelle se produit une bifurcation vers une cycle limite de période 2^n . La figure 6.5 illustre les valeurs de x formant les cycles limites (en ordonnée), pour un grand nombre

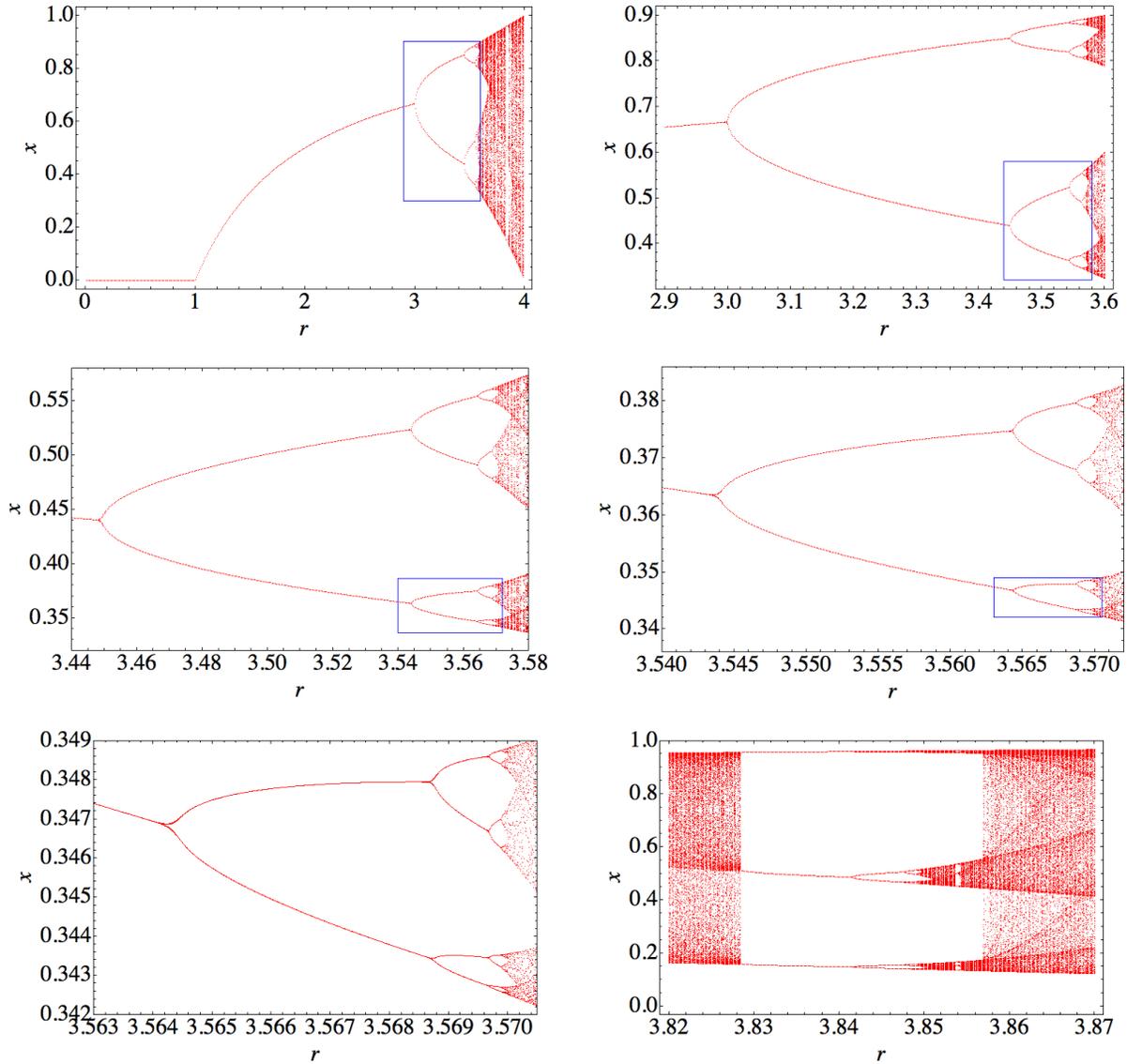


FIGURE 6.5

Diagrammes de Feigenbaum pour la carte logistique. Les diagrammes successifs (de gauche à droite et du haut vers le bas) sont des agrandissements successifs des régions encadrées. Notez l'auto similitude de ces diagrammes. Le dernier diagramme illustre plutôt une région du diagramme où un comportement normal s'observe, au coeur d'une zone chaotique.

de valeurs de r (en abscisse). Le premier graphique de cette figure illustre le domaine complet de r , de 0 à 1. On voit comment le point fixe $x_f = 0$ est stable jusqu'à $r = 1$, comment ensuite le point fixe $x_f = (r - 1)/r$ est stable jusqu'à la première bifurcation, à $r = 3$. Les graphiques suivants, jusqu'à l'avant-dernier, sont des agrandissements successifs des régions marquées en bleu. On constate que ce diagramme comporte une auto similitude évidente, ce qui mène à une structure fractale.

Ces valeurs r_n sont de plus en plus rapprochées, et atteignent une limite r_∞ qui, pour la carte logistique, a la valeur suivante :

$$r_\infty = 3.5699456\dots \quad (6.17)$$

Passée cette valeur de r , le système est dans un régime chaotique : les valeurs de la séquence $\{x_n\}$ ne se répètent jamais, un peu comme les décimales d'un nombre irrationnel. Le dernier graphique de la figure 6.4 illustre justement quelques dizaines d'applications de la carte logistique pour une valeur de r située au-delà de r_∞ , dans la zone chaotique. La zone chaotique apparaîtrait normalement comme une zone complètement remplie de points sur la figure 6.5, si la simulation avait été menée sans interruption. Fait curieux, à l'intérieur même de la zone chaotique se situent des zones normales, avec des cycles de longueur finie, comme illustré dans le dernier graphique de la figure 6.5.

La séquence $\{r_n\}$ des paramètres de bifurcation a un comportement très intéressant : on montre, par expérimentation numérique, que

$$\lim_{n \rightarrow \infty} \frac{r_n - r_{n-1}}{r_{n+1} - r_n} = \delta = 4.66920160910299067185320382\dots \quad (6.18)$$

Autrement dit, les r_n ont un comportement régulier caractérisé par un nombre δ appelé *premier nombre de Feigenbaum*. L'intérêt de ce nombre est son caractère *universel* : le même nombre δ apparaît dans toutes les cartes à une variable possédant un maximum quadratique. Par contre, la valeur limite r_∞ est bien sûr propre à la carte logistique. En somme, si on définit une carte unidimensionnelle avec un maximum localement quadratique, et que cette carte possède un paramètre de contrôle r , la séquence des points de bifurcations va converger vers une limite chaotique avec le même exposant δ . Il y a donc un certain ordre dans l'approche du chaos.

CHAPITRE 7

Annexes

A Rappels d'algèbre linéaire

7.A.1 Notations indicielle et matricielle

Considérons un espace vectoriel \mathcal{V} de dimension n . Étant donné une base dont les éléments sont notés \mathbf{e}_i ($i = 1, 2, \dots, n$), tout élément \mathbf{x} de cet espace peut s'écrire

$$\mathbf{x} = \sum_i x_i \mathbf{e}_i \quad (7.1)$$

où les nombres x_i sont les composantes du vecteur \mathbf{x} dans la base $\{\mathbf{e}_i\}$. Nous allons utiliser la *convention de somme*, qui stipule qu'un signe de somme sur un indice est implicite dans toute expression dans laquelle cet indice est répété, à moins d'avis contraire explicite. Ainsi, l'expression ci-dessus s'écrit simplement $\mathbf{x} = x_i \mathbf{e}_i$. On représente souvent les composantes de \mathbf{x} comme un vecteur colonne \mathbf{x} :

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad (7.2)$$

La distinction entre \mathbf{x} et \mathbf{x} est importante : \mathbf{x} est un vecteur abstrait, un élément de l'espace vectoriel, qui existe indépendamment de la base utilisée, alors que \mathbf{x} est un vecteur colonne qui représente \mathbf{x} dans la base $\{\mathbf{e}_i\}$. La notation \mathbf{x} est introduite uniquement pour pouvoir exprimer des relations entre vecteurs colonne de manière plus concise, sans avoir à les écrire explicitement.

Considérons ensuite un opérateur linéaire sur \mathcal{V} , c'est-à-dire une fonction $\mathcal{M}(\mathbf{x})$ telle que

$$\mathcal{M}(\mathbf{x} + \mathbf{y}) = \mathcal{M}(\mathbf{x}) + \mathcal{M}(\mathbf{y}) \quad \text{et} \quad \mathcal{M}(\lambda \mathbf{x}) = \lambda \mathcal{M}(\mathbf{x}) \quad (7.3)$$

Dans une base $\{\mathbf{e}_i\}$, cet opérateur linéaire peut être représenté par une matrice dont les éléments sont notés M_{ij} (l'indice de rangée à gauche et l'indice de colonne à droite), telle que

$$\mathcal{M}(\mathbf{e}_i) = M_{ji} \mathbf{e}_j \quad (7.4)$$

L'action de l'opérateur linéaire \mathcal{M} sur un vecteur \mathbf{x} est alors

$$\mathbf{y} = \mathcal{M}(\mathbf{x}) = \mathcal{M}(x_j \mathbf{e}_j) = x_j \mathcal{M}(\mathbf{e}_j) = x_j M_{ij} \mathbf{e}_i \implies y_i = M_{ij} x_j \quad (7.5)$$

En notation matricielle, on écrit plutôt $\mathbf{y} = \mathbf{M}\mathbf{x}$, où le symbole \mathbf{M} désigne la matrice

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn} \end{pmatrix} \quad (7.6)$$

La multiplication de deux matrices \mathbf{A} et \mathbf{B} est définie par l'expression

$$(\mathbf{AB})_{ij} = A_{ik} B_{kj} \quad (7.7)$$

et représente l'application successive des deux opérateurs linéaires correspondants. Notez que lors d'un produit de matrices, l'indice sommé est l'indice de colonne du facteur de gauche et l'indice de rangée du facteur de droite.

La *transposée* de la matrice \mathbf{M} , notée $\tilde{\mathbf{M}}$, est obtenue en échangeant rangées et colonnes :

$$\tilde{M}_{ij} = M_{ji} \quad (7.8)$$

La transposée d'un produit est le produit des transposées, dans l'ordre inverse :

$$\widetilde{\mathbf{AB}} = \tilde{\mathbf{B}}\tilde{\mathbf{A}} \quad \text{car} \quad (\widetilde{\mathbf{AB}})_{ij} = A_{jk} B_{ki} = \tilde{B}_{ik} \tilde{A}_{kj} = (\tilde{\mathbf{B}}\tilde{\mathbf{A}})_{ij} \quad (7.9)$$

Une matrice est *symétrique* si elle coïncide avec sa transposée :

$$\mathbf{M} = \tilde{\mathbf{M}} \quad \text{ou} \quad M_{ij} = M_{ji} \quad (\text{matrice symétrique}) \quad (7.10)$$

Elle est *antisymétrique* si elle est l'opposée de sa transposée :

$$\mathbf{M} = -\tilde{\mathbf{M}} \quad \text{ou} \quad M_{ij} = -M_{ji} \quad (\text{matrice antisymétrique}) \quad (7.11)$$

Toute matrice \mathbf{M} peut être exprimée comme la somme d'une partie symétrique et d'une partie antisymétrique :

$$\mathbf{M} = \mathbf{M}^S + \mathbf{M}^A \quad \text{ou} \quad \mathbf{M}^S = \frac{1}{2}(\mathbf{M} + \tilde{\mathbf{M}}) \quad \text{et} \quad \mathbf{M}^A = \frac{1}{2}(\mathbf{M} - \tilde{\mathbf{M}}) \quad (7.12)$$

L'adjoint d'une matrice \mathbf{M} , noté \mathbf{M}^\dagger , est le conjugué complexe de la transposée :

$$\mathbf{M}^\dagger = \tilde{\mathbf{M}}^* \quad \text{ou} \quad (M^\dagger)_{ij} = M_{ji}^* \quad (7.13)$$

Une matrice est dite *hermitienne* si elle coïncide avec son adjoint :

$$\mathbf{M}^\dagger = \mathbf{M} \quad \text{ou} \quad M_{ij} = M_{ji}^* \quad (\text{matrice hermitienne}) \quad (7.14)$$

Les composantes de la matrice identité \mathbf{I} sont notées $I_{ij} = \delta_{ij}$. Le symbole δ_{ij} est appelé *delta de Kronecker*. Lors d'une somme contre un autre symbole, le delta de Kronecker effectue une réduction des indices :

$$\delta_{ik} M_{kj} = M_{ij} \quad (7.15)$$

La trace d'une matrice carrée est la somme de ses éléments diagonaux :

$$\text{Tr } M = M_{ii} \quad (7.16)$$

La trace d'un produit de matrices s'exprime simplement comme

$$\text{Tr}(ABC) = A_{ij}B_{jk}C_{ki} \quad (7.17)$$

Il s'ensuit la propriété cyclique de la trace – la trace d'un produit de plusieurs matrices est la même si on effectue une permutation cyclique des facteurs :

$$\text{Tr}(ABC) = \text{Tr}(BCA) = \text{Tr}(CAB) \quad (7.18)$$

Restreignons maintenant la discussion au cas de l'espace cartésien de dimension 3. Supposons qu'une base orthonormée est utilisée. Le produit scalaire de deux vecteurs \mathbf{A} et \mathbf{B} est alors

$$\mathbf{A} \cdot \mathbf{B} = A_i B_i \quad (7.19)$$

On définit le symbole antisymétrique ε_{ijk} , ou *tenseur de Levi-Civita*, comme suit :

$$\begin{aligned} \varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1 \\ \varepsilon_{132} = \varepsilon_{213} = \varepsilon_{321} = -1 \end{aligned} \quad \text{sinon } \varepsilon_{ijk} = 0 \quad (7.20)$$

Autrement dit, $\varepsilon_{ijk} = 0$ si deux ou trois des indices coïncident. Le symbole ε_{ijk} est invariant lorsqu'on procède à une permutation cyclique de ses indices, mais change de signe lorsqu'on transpose deux indices.

Le déterminant d'une matrice carrée d'ordre 3 s'exprime ainsi en fonction du symbole antisymétrique :

$$\varepsilon_{ijk} \det M = \varepsilon_{lmn} M_{il} M_{jm} M_{kn} \quad (7.21)$$

Le produit vectoriel $\mathbf{A} \wedge \mathbf{B}$ s'exprime

$$\mathbf{A} \wedge \mathbf{B} = \varepsilon_{ijk} A_i B_j \mathbf{e}_k \quad (7.22)$$

On vérifie sans peine que le symbole antisymétrique obéit à la règle suivante :

$$\varepsilon_{ijk} \varepsilon_{lmk} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} \quad (7.23)$$

Cette relation nous permet de démontrer l'identité suivante :

$$\mathbf{A} \wedge (\mathbf{B} \wedge \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}) \quad (7.24)$$

Le gradient d'une fonction s'écrit

$$\nabla f = \frac{\partial f}{\partial x_i} \mathbf{e}_i := \partial_i f \mathbf{e}_i \quad (7.25)$$

La divergence d'un vecteur s'écrit

$$\nabla \cdot \mathbf{A} = \partial_i A_i \quad (7.26)$$

et le rotationnel s'écrit

$$\nabla \wedge \mathbf{A} = \varepsilon_{ijk} \partial_i A_j \mathbf{e}_k \quad (7.27)$$

La notation indicielle permet de facilement démontrer les identités sur les rotationnels et divergences de produits. Par exemple,

$$\nabla \wedge (f \mathbf{A}) = \nabla f \wedge \mathbf{A} + f \nabla \wedge \mathbf{A} \quad \text{car } \varepsilon_{ijk} \partial_i (f A_j) \mathbf{e}_k = \varepsilon_{ijk} \partial_i f A_j \mathbf{e}_k + f \varepsilon_{ijk} \partial_i A_j \mathbf{e}_k \quad (7.28)$$

7.A.2 Changements de base

Un changement de base $\{\mathbf{e}_i\} \rightarrow \{\mathbf{e}'_i\}$ peut être défini par une matrice qui exprime l'ancienne base en fonction de la nouvelle :

$$\mathbf{e}_i = S_{ji} \mathbf{e}'_j \quad (7.29)$$

Les composantes d'un vecteur \mathbf{x} sont affectées par le changement de base de la manière suivante :

$$\mathbf{x} = x'_i \mathbf{e}'_i = x_j \mathbf{e}_j = x_j S_{ij} \mathbf{e}'_i \implies x'_i = S_{ij} x_j \quad \text{ou} \quad \mathbf{x}' = S \mathbf{x} \quad (7.30)$$

Les deux vecteurs colonne \mathbf{x} et \mathbf{x}' contiennent les composantes d'un même vecteur \mathbf{x} dans les bases $\{\mathbf{e}_i\}$ et $\{\mathbf{e}'_i\}$.

Voyons maintenant comment la matrice qui représente un opérateur linéaire \mathcal{M} est affectée par un changement de base. La relation $\mathbf{y} = \mathcal{M}(\mathbf{x})$ peut s'exprimer soit comme $y_i = M_{ij} x_j$, soit comme $y'_i = M'_{ij} x'_j$, selon la base choisie. Cependant, $y'_i = S_{ik} y_k$ et $x_j = S_{jl}^{-1} x'_l$, où la dernière expression désigne l'élément de la matrice inverse S^{-1} , et non l'inverse de l'élément S_{jl} , qu'on noterait plutôt $(S_{jl})^{-1}$ si nécessaire. Donc

$$y'_i = S_{ik} y_k = S_{ik} M_{kj} x_j = S_{ik} M_{kj} S_{jl}^{-1} x'_l \implies M'_{il} = S_{ik} M_{kj} S_{jl}^{-1} \quad (7.31)$$

ou encore

$$\mathbf{M}' = \mathbf{S} \mathbf{M} \mathbf{S}^{-1} \quad (7.32)$$

Le déterminant d'une matrice est le même dans toutes les bases, car

$$\det \mathbf{M}' = \det(\mathbf{S} \mathbf{M} \mathbf{S}^{-1}) = \det \mathbf{S} \det \mathbf{M} \det \mathbf{S}^{-1} = \det \mathbf{M} \quad (7.33)$$

car $\det \mathbf{S}^{-1} = (\det \mathbf{S})^{-1}$. Le déterminant est donc une propriété de l'opérateur linéaire \mathcal{M} lui-même et non des matrices qui le représentent. Il en va de même pour la trace, en vertu de la propriété cyclique :

$$\text{tr} \mathbf{M}' = \text{tr}(\mathbf{S} \mathbf{M} \mathbf{S}^{-1}) = \text{tr}(\mathbf{S}^{-1} \mathbf{S} \mathbf{M}) = \text{tr} \mathbf{M} \quad (7.34)$$

7.A.3 Vecteurs et valeurs propres d'une matrice

On dit que \mathbf{x} est un *vecteur propre* de l'opérateur \mathcal{M} si l'action de \mathcal{M} sur \mathbf{x} donne un résultat proportionnel à \mathbf{x} :

$$\mathcal{M}(\mathbf{x}) = \lambda \mathbf{x} \quad (7.35)$$

La constante λ est la *valeur propre* associée au vecteur propre \mathbf{x} . Cette relation s'écrit de la même manière dans une base donnée, en fonction du vecteur colonne et de la matrice \mathbf{M} qui représente l'action de \mathcal{M} dans cette base :

$$\mathbf{M} \mathbf{x} = \lambda \mathbf{x} \quad (7.36)$$

Si \mathbf{x} est un vecteur propre de \mathcal{M} , alors tout multiple de \mathbf{x} est encore un vecteur propre, avec la même valeur propre. On devrait donc parler de *sous-espace propre*.

L'équation aux valeurs propres peut aussi s'écrire

$$(\mathbf{M} - \lambda \mathbf{I}) \mathbf{x} = \mathbf{0} \quad (7.37)$$

Pour qu'une solution non nulle en x existe, il faut nécessairement que la matrice entre parenthèses soit singulière (c'est-à-dire qu'elle n'ait pas d'inverse), ce qui entraîne que son déterminant est nul :

$$\det |M - \lambda \mathbb{I}| = 0 \quad \text{ou} \quad \begin{vmatrix} M_{11} - \lambda & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} - \lambda & \cdots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn} - \lambda \end{vmatrix} = 0 \quad (7.38)$$

Ce déterminant, en fonction de λ , est un polynôme d'ordre n appelé *polynôme caractéristique*. Il possède n racines (pas nécessairement distinctes) qui sont les valeurs propres possibles de M . Les vecteurs propres correspondants sont ensuite déterminés en résolvant l'équation (7.37) pour chaque valeur propre à tour de rôle.

Si une matrice M possède n vecteurs propres linéairement indépendants, on dit qu'elle est *diagonalisable*, car elle peut être ramenée à une forme diagonale par un changement de base (le caractère diagonalisable est en fait une propriété de l'opérateur linéaire \mathcal{M}). Considérons à cet effet la matrice U dont les colonnes sont n vecteurs propres indépendants de la matrice M . On a donc la relation

$$M_{ik}U_{kj} = \lambda_j U_{ij} \quad (\text{pas de somme sur } j) \quad (7.39)$$

Définissons maintenant une matrice diagonale D dont les éléments sont les valeurs propres λ_i :

$$D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \quad \text{ou} \quad D_{ij} = \delta_{ij} \lambda_j \quad (\text{pas de somme sur } j) \quad (7.40)$$

La relation précédente peut alors s'écrire ainsi :

$$M_{ik}U_{kj} = U_{ik}D_{kj} \quad \text{ou encore} \quad MU = UD \quad (7.41)$$

En post-multipliant et pré multipliant cette relation par l'inverse de U , qui existe nécessairement puisque les colonnes de U sont linéairement indépendantes, on obtient les deux relations suivantes :

$$M = UDU^{-1} \quad \text{ou} \quad D = U^{-1}MU \quad (7.42)$$

D'après la relation (7.32), la matrice $S = U^{-1}$ nous fait passer de la base $\{e_i\}$ à une base $\{e'_i\}$ en fonction de laquelle l'opérateur \mathcal{M} est exprimé par une matrice diagonale D . Les nouveaux vecteurs de base sont exprimés comme suit :

$$e'_i = S_{ji}^{-1} e_j = U_{ji} e_j \quad (7.43)$$

ce qui signifie simplement que les composantes de e'_i sont les éléments de la colonne no i de U , c'est-à-dire les vecteurs propres eux-mêmes.

Une condition suffisante pour qu'une matrice soit diagonalisable est qu'elle soit *normale*, c'est-à-dire qu'elle respecte la condition suivante :

$$M^\dagger M = M M^\dagger \quad (7.44)$$

Si une matrice est diagonalisable, alors son déterminant est le produit de ses valeurs propres, et sa trace est la somme de ses valeurs propres :

$$\det M = \det D = \lambda_1 \lambda_2 \cdots \lambda_n \quad \text{tr } M = \text{tr } D = \lambda_1 + \lambda_2 + \cdots + \lambda_n \quad (7.45)$$

Théorème 7.1

Les valeurs propres d'une matrice hermitienne ($M^\dagger = M$) sont réelles et les vecteurs propres associés à des valeurs propres différentes sont orthogonaux.

Preuve: Considérons en effet deux vecteurs propres x_1 et x_2 associés à deux valeurs propres différentes $\lambda_1 \neq \lambda_2$:

$$Mx_1 = \lambda_1 x_1 \quad Mx_2 = \lambda_2 x_2 \quad (7.46)$$

Le conjugué hermitien de ces relations s'écrit ainsi, puisque M est hermitien :

$$x_1^\dagger M = \lambda_1^* x_1^\dagger \quad x_2^\dagger M = \lambda_2^* x_2^\dagger \quad (7.47)$$

En premier lieu, en combinant les deux équations ci-dessus pour x_1 , on trouve

$$x_1^\dagger Mx_1 = \lambda_1^* x_1^\dagger x_1 = \lambda_1 x_1^\dagger x_1 \quad \text{ou encore} \quad (\lambda_1^* - \lambda_1) x_1^\dagger x_1 = 0 \quad (7.48)$$

Comme $x_1^\dagger x_1 > 0$, il s'ensuit que λ_1 est réel ($\lambda_1^* = \lambda_1$), et cela s'applique bien sûr à toutes les valeurs propres de M . Ensuite, sachant que les valeurs propres sont réelles et combinant les mêmes équations pour x_1 et x_2 , on trouve

$$x_1^\dagger Mx_2 = \lambda_1 x_1^\dagger x_2 = \lambda_2 x_1^\dagger x_2 \implies (\lambda_1 - \lambda_2) x_1^\dagger x_2 = 0 \quad (7.49)$$

On en conclut que si $\lambda_1 \neq \lambda_2$, alors $x_1^\dagger x_2 = 0$, c'est-à-dire que les vecteurs propres sont orthogonaux. Si deux vecteurs propres linéairement indépendants sont associés à la même valeur propre, le résultat tient toujours pour la simple raison qu'on peut toujours choisir les vecteurs de ce sous-espace propre de manière à ce qu'ils soient orthogonaux de toute manière.

On peut donc choisir un ensemble de n vecteurs propres d'une matrice hermitienne afin qu'ils soient mutuellement orthogonaux et, pourquoi pas, de norme unité. Les n vecteurs propres forment alors une base orthonormée. Ces vecteurs propres forment les colonnes d'une matrice U qui diagonalise M . Cette matrice est *unitaire*, c'est-à-dire que

$$U^\dagger U = \mathbb{I} \quad \text{ou} \quad U_{ki}^* U_{kj} = \delta_{ij} \quad (\text{matrice unitaire}) \quad (7.50)$$

Une matrice unitaire O dont tous les éléments sont réels est dite *orthogonale* :

$$\tilde{O}O = \mathbb{I} \quad \text{ou} \quad O_{ki} O_{kj} = \delta_{ij} \quad (\text{matrice orthogonale}) \quad (7.51)$$

Une matrice hermitienne dont les éléments sont réels est tout simplement *symétrique*. Donc :

Toute matrice hermitienne possède des valeurs propres réelles et des vecteurs propres orthogonaux ; elle peut être diagonalisée à l'aide d'une matrice unitaire. De même pour une matrice symétrique, qui peut être diagonalisée à l'aide d'une matrice orthogonale

Vecteurs et valeurs propres généralisées

On dit que λ est une *valeur propre généralisée* des matrices A et B et que x est le vecteur propre correspondant si l'équation suivante est satisfaite :

$$Ax = \lambda Bx \quad (7.52)$$

Ce problème se ramène à la recherche des valeurs et vecteurs propres de la matrice $W = B^{-1}A$ si la matrice B est non singulière. Par contre, même dans le cas où B possède un inverse, certaines propriétés sont plus simples à démontrer à l'aide de la forme (7.52) de l'équation.

Commençons par quelques définitions : une matrice hermitienne B est dite *définie positive* si $x^\dagger Bx > 0$ pour tout vecteur non nul x . Notons premièrement que B doit être hermitienne afin que l'expression $x^\dagger Bx$ soit réelle pour tout vecteur x . Une telle matrice nous permet de définir un produit bilinéaire (ou scalaire) de la manière suivante :

$$\langle x_1 | x_2 \rangle := x_1^\dagger Bx_2 \quad (7.53)$$

Cette définition respecte tous les critères d'un produit bilinéaire : la linéarité est évidente, mais, plus important, le produit d'un vecteur par lui-même est toujours réel et positif :

$$\langle x | x \rangle = x^\dagger Bx > 0 \quad (7.54)$$

Donc, chaque matrice définie positive B permet de définir un produit scalaire *différent*, et on dira que deux vecteurs x_1 et x_2 sont *orthogonaux via la matrice B* si $x_1^\dagger Bx_2 = 0$.

Démontrons maintenant le théorème suivant :

Théorème 7.2

Si les matrices A et B sont hermitiennes, et que la matrice B est définie-positive, alors les valeurs propres généralisées sont réelles et les vecteurs propres associés à des valeurs propres différentes sont orthogonaux via la matrice B.

Preuve: La preuve est presque identique à celle du théorème précédent. Considérons deux vecteurs propres x_1 et x_2 associés à deux valeurs propres différentes $\lambda_1 \neq \lambda_2$:

$$Ax_1 = \lambda_1 Bx_1 \quad Ax_2 = \lambda_2 Bx_2 \quad (7.55)$$

Le conjugué hermitien de ces relations s'écrit ainsi, puisque A et B sont hermitiens :

$$x_1^\dagger A = \lambda_1^* x_1^\dagger B \quad x_2^\dagger A = \lambda_2^* x_2^\dagger B \quad (7.56)$$

En premier lieu, en combinant les deux équations ci-dessus pour x_1 , on trouve

$$x_1^\dagger Ax_1 = \lambda_1^* x_1^\dagger Bx_1 = \lambda_1 x_1^\dagger Bx_1 \quad \text{ou encore} \quad (\lambda_1^* - \lambda_1) x_1^\dagger Bx_1 = 0 \quad (7.57)$$

Comme $x_1^\dagger Bx_1 > 0$, la matrice B étant définie positive, il est résulte que λ_1 est réel ($\lambda_1^* = \lambda_1$), et cela s'applique bien sûr à toutes les valeurs propres. Ensuite, sachant que les valeurs propres

sont réelles et combinant les mêmes équations pour x_1 et x_2 , on trouve

$$x_1^\dagger A x_2 = \lambda_1 x_1^\dagger B x_2 = \lambda_2 x_1^\dagger B x_2 \implies (\lambda_1 - \lambda_2) x_1^\dagger B x_2 = 0 \quad (7.58)$$

On en conclut que si $\lambda_1 \neq \lambda_2$, alors $x_1^\dagger B x_2 = 0$, c'est-à-dire que les vecteurs propres sont orthogonaux via B. Encore une fois, si deux vecteurs propres linéairement indépendants sont associés à la même valeur propre, le résultat tient toujours pour la simple raison qu'on peut toujours choisir les vecteurs de ce sous-espace propre de manière à ce qu'ils soient orthogonaux de toute manière.

Si les matrices A et B sont réelles symétriques, alors le théorème s'applique comme un cas particulier et les vecteurs propres aussi ont des composantes réelles.

B Vecteurs et tenseurs

L'objectif de cette annexe est d'introduire les notions de composantes covariantes et contravariantes de vecteurs et de tenseurs. Pour rendre la discussion plus intuitive, nous travaillerons dans l'espace cartésien à trois dimensions, mais les concepts seront immédiatement généralisables à l'espace-temps quadri-dimensionnel.

Composantes covariantes et contravariantes

Considérons une base $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ dans l'espace tridimensionnel. Cette base n'est pas nécessairement orthonormée. Tout vecteur \mathbf{A} peut être décomposé selon cette base. Nous allons cependant utiliser une notation légèrement différente de celle de l'annexe précédent, en plaçant certains indices en haut :

$$\mathbf{A} = A^1 \mathbf{e}_1 + A^2 \mathbf{e}_2 + A^3 \mathbf{e}_3 = A^i \mathbf{e}_i \quad (7.59)$$

La convention de sommation sur les indices répétés sera appliquée, et ceux-ci apparaîtront toujours l'un en haut, l'autre en bas. Les 3 quantités A^i sont appelées *composantes contravariantes* du vecteur \mathbf{A} . La raison de cette appellation est donnée dans ce qui suit, et est au coeur de la présente annexe. Supposons qu'on procède à un changement de base, c'est-à-dire qu'on définit une nouvelle base \mathbf{e}'_i :

$$\mathbf{e}_i = S^j_i \mathbf{e}'_j \quad (7.60)$$

où les coefficients S^j_i forment une matrice S non singulière (indice de rangée à gauche, indice de colonne à droite). Le vecteur s'exprime également dans l'une ou l'autre base :

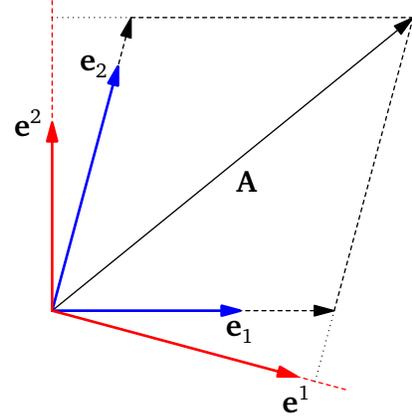
$$\mathbf{A} = A^j \mathbf{e}_j = A'^i \mathbf{e}'_i = A^j S^i_j \mathbf{e}'_i \implies A'^i = S^i_j A^j \quad (7.61)$$

Les composantes A^j doivent s'exprimer en fonction des composantes A'^i par l'intermédiaire de la matrice inverse S^{-1} :

$$A^i = (S^{-1})^i_j A'^j \quad (7.62)$$

Autrement dit, les composantes contravariantes se transforment à l'aide de la matrice inverse transposée de celle utilisée pour transformer les vecteurs de base (d'où l'expression *contravariante*).

FIGURE 7.1



À la base $\{\mathbf{e}_i\}$ on associe une *base duale* $\{\mathbf{e}^i\}$ (l'indice est maintenant en haut) définie par la relation

$$\mathbf{e}^i \cdot \mathbf{e}_j = \delta_j^i \quad (7.63)$$

Comme la relation entre une base et sa duale est la même pour toutes les bases ($\mathbf{e}^i \cdot \mathbf{e}_j = \mathbf{e}'^i \cdot \mathbf{e}'_j = \delta_j^i$) et qu'elle a la même structure que la relation (7.61), on en déduit immédiatement que les vecteurs de la base duale se transforment aussi comme les composantes contravariantes lors d'un changement de base :

$$\mathbf{e}'^i = S^i_j \mathbf{e}^j \quad (7.64)$$

Autrement dit, la base duale est contravariante. Les bases directe et duale sont illustrées sur la figure 7.1.

Tout vecteur peut également être représenté de manière unique sur la base duale, mais cette fois par des composantes dites *covariantes* :

$$\mathbf{A} = A_i \mathbf{e}^i = A'_i \mathbf{e}'^i \quad (7.65)$$

Il est maintenant évident que les composantes covariantes méritent ce nom parce qu'elles se transforment de la même manière que les vecteurs de base :

$$A_i = S^j_i A'_j \quad (7.66)$$

Les composantes covariantes s'obtiennent par projection du vecteur sur la base originale :

$$A_i = \mathbf{A} \cdot \mathbf{e}_i \quad (7.67)$$

Si la base duale coïncide avec la base directe, alors il n'y a pas lieu de distinguer entre composantes covariantes et contravariantes. C'est le cas avec une base orthonormée dans l'espace euclidien à trois dimensions, et c'est pour cela que nous avons toujours placé les indices en bas dans ce cas. Par contre, pour une base générale, la distinction entre les deux types de composantes est essentielle. Elle est aussi importante dans la description des quadrivecteurs en relativité, car l'espace-temps de Minkowski n'est pas euclidien, mais pseudo-euclidien.

Tenseurs

Si $\{\mathbf{e}_i\}$ est une base de l'espace vectoriel V , alors les produits $\{\mathbf{e}_i \otimes \mathbf{e}_j\}$ forment une base de l'espace produit tensoriel $V \otimes V$. on peut aussi utiliser la base duale, ou deux bases mixtes : $\{\mathbf{e}^i \otimes \mathbf{e}^j\}$, $\{\mathbf{e}_i \otimes \mathbf{e}^j\}$ et

$\{\mathbf{e}^i \otimes \mathbf{e}_j\}$. Dans cet espace produit, on définit des tenseurs de rang 2, caractérisés par des composantes à deux indices :

$$\mathbf{T} = T^{ij} \mathbf{e}_i \otimes \mathbf{e}_j = T_{ij} \mathbf{e}^i \otimes \mathbf{e}^j = T^i_j \mathbf{e}_i \otimes \mathbf{e}^j = T_i^j \mathbf{e}^i \otimes \mathbf{e}_j \quad (7.68)$$

On constate immédiatement que, lors d'un changement de base, ces composantes se transforment de la manière suivante :

$$\begin{aligned} T'^{ij} &= S^i_k S^j_l T^{kl} \\ T'_{ij} &= (S^{-1})^k_i (S^{-1})^l_j T_{kl} \\ T'^i_j &= S^i_k (S^{-1})^l_j T^k_l \\ T_i^j &= (S^{-1})^k_i S^j_l T_k^l \end{aligned}$$

Par abus de langage, on donne le nom de tenseur aux composantes elles-mêmes, et on appelle T_{ij} un tenseur covariant, T^{ij} un tenseur contravariant, alors que les ensembles T_i^j et T^i_j sont qualifiés de *tenseurs mixtes*.

On définit généralement des tenseurs de rang N, sur l'espace produit tensoriel $V \otimes V \otimes \dots \otimes V$ (N facteurs), qui comportent N indices. Un tenseur de rang un est un vecteur. Un tenseur de rang zéro est un scalaire (une quantité invariante, c'est-à-dire qui est indépendante de la base utilisée).

Tenseur métrique

La base $\{\mathbf{e}_i\}$ peut bien sûr être exprimée dans la base duale :

$$\mathbf{e}_i = g_{ij} \mathbf{e}^j \quad (7.69)$$

Les coefficients g_{ij} sont également les produits scalaires des vecteurs de base :

$$\mathbf{e}_i \cdot \mathbf{e}_j = g_{ik} \mathbf{e}^k \cdot \mathbf{e}_j = g_{ik} \delta_j^k = g_{ij} \quad (7.70)$$

La relation inverse s'exprime ainsi :

$$\mathbf{e}^i = g^{ij} \mathbf{e}_j \quad \text{où} \quad g^{ik} g_{kj} = \delta_j^i \quad \text{et} \quad \mathbf{e}^i \cdot \mathbf{e}^j = g^{ij} \quad (7.71)$$

Comme les vecteurs de base \mathbf{e}_i et \mathbf{e}^i sont respectivement covariants et contravariants, les quantités g_{ij} et g^{ij} se transforment bien comme les composantes covariantes et contravariantes d'un tenseur de rang 2. On l'appelle le *tenseur métrique*.

Le produit scalaire $\mathbf{A} \cdot \mathbf{B}$ s'exprime comme suit en fonction du tenseur métrique :

$$\mathbf{A} \cdot \mathbf{B} = A^i B^j \mathbf{e}_i \cdot \mathbf{e}_j = A^i B^j g_{ij} \quad (7.72)$$

Le tenseur métrique peut être utilisé pour passer d'une composante covariante à contravariante et vice-versa. Ceci est vrai (par définition) pour les vecteurs de base, mais aussi pour les composantes. Explicitement :

$$\begin{aligned} \mathbf{A} = A^i \mathbf{e}_i = A^i g_{ij} \mathbf{e}^j &\implies A_j = g_{ij} A^i \\ \mathbf{A} = A_i \mathbf{e}^i = A_i g^{ij} \mathbf{e}_j &\implies A^j = g^{ij} A_i \end{aligned}$$

Plus généralement, le tenseur métrique peut être appliqué de cette manière à tout indice, covariant ou contravariant :

$$T_{ij} = g_{ik} g_{jl} T^{kl} \quad , \quad T_i^j = g^{jk} T_{ik} \quad , \quad \text{etc.} \quad (7.73)$$

Contraction

On appelle contraction d'un indice covariant et d'un indice contravariant la somme sur les valeurs égales de ces indices, comme dans un produit scalaire. Ainsi, dans l'expression $A^i e_i$, on dit que les deux indices sont contractés. Autre exemple : le produit scalaire de deux vecteurs \mathbf{A} et \mathbf{B} s'exprime ainsi :

$$\mathbf{A} \cdot \mathbf{B} = (A^i \mathbf{e}_i) \cdot (B_j \mathbf{e}^j) = A^i B_j \delta_i^j = A^i B_i \quad (7.74)$$

Les indices d'un tenseur de rang deux peuvent aussi être contractés : T^i_i .

En général, la contraction de deux indices mène soit à une quantité invariante, soit à une quantité de rang inférieur de deux à la quantité originale. Par exemple, d'un tenseur de rang 4 R_{iklm} (comme le tenseur de courbure géométrie riemannienne), on peut former un tenseur de rang deux par contraction : $R_{ik} = R^l_{ilk} = g^{ml} R_{mil k}$.