

# Théorie des champs

Responsable de publication : Sophie Chouaf

En application du Code de la Propriété Intellectuelle et notamment de ses articles L. 122.4, L. 122-5 et L. 335-2, toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite. Une telle représentation ou reproduction constituerait un délit de contrefaçon, puni de deux ans d'emprisonnement et de 300 000 euros d'amende.

Ne sont autorisés que les copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective, ainsi que les analyses et courtes citations, sous réserve que soient indiqués clairement le nom de l'auteur et la source.

© *Les Presses de l'ENSTA,*

*Imprimé en France*

ISBN 978-2-7225-0955-9

*[www.ensta-paristech.fr](http://www.ensta-paristech.fr)*

---

Jérôme Perez

# Théorie des champs

Les équations de la physique

Mécanique analytique

Relativité restreinte et générale

Mécanique quantique

PALAISEAU

LES PRESSES DE L'ENSTA

828, Boulevard des maréchaux, 91120 Palaiseau



---

# Table des matières

---

## Partie I Mécanique analytique

---

<b>1</b>	<b>L'incroyable legs de Joseph-Louis Lagrange</b> . . . . .	3
1.1	L'origine . . . . .	3
1.2	La méthode de variation des constantes. . . . .	4
1.2.1	Ordre 1 : l'idée . . . . .	4
1.2.2	Ordres supérieurs : la généralisation . . . . .	5
1.3	Application au mouvement de la Lune . . . . .	7
1.3.1	L'idée de la méthode . . . . .	7
1.3.2	La mise en œuvre . . . . .	8
1.4	Généralisation à tous les problèmes de mécanique conservative	12
1.4.1	La construction des équations de Lagrange . . . . .	12
1.4.2	La fin de son œuvre. . . . .	15
1.5	Compléments historiques sur la vie, l'œuvre de Joseph Louis Lagrange et son contexte.. . . . .	17
<b>2</b>	<b>Formulation lagrangienne</b> . . . . .	25
2.1	Coordonnées généralisées . . . . .	25
2.2	Principe de moindre action . . . . .	26
2.2.1	Quelles équations pour $\mathcal{L}$ ? . . . . .	27
2.2.2	Propriétés du lagrangien . . . . .	28
2.3	Principe de relativité . . . . .	30
2.4	Détermination de la fonction de Lagrange. . . . .	31
2.4.1	Particule libre . . . . .	31
2.4.2	Systèmes conservatifs . . . . .	32
2.4.3	Systèmes non conservatifs . . . . .	35
2.4.4	Lagrangien et mouvement d'un solide . . . . .	38

2.4.5	Contraintes . . . . .	41
2.4.6	Lagrangien d'une particule dans un champ électromagnétique. . . . .	47
2.5	Lagrangien, symétries et lois de conservation . . . . .	50
2.5.1	Symétries . . . . .	50
2.5.2	Théorème de Noether . . . . .	51
2.5.3	Trois exemples fondamentaux. . . . .	52
<b>3</b>	<b>Formulation hamiltonienne</b> . . . . .	<b>57</b>
3.1	Equations de Hamilton. . . . .	57
3.2	Nature de la fonction de Hamilton . . . . .	59
3.3	Interprétation des équations de Hamilton . . . . .	60
3.4	Systèmes dynamiques hamiltoniens . . . . .	63
3.4.1	Crochets de Poisson . . . . .	63
3.4.2	Le choix des coordonnées . . . . .	71
3.4.3	Intégrabilité d'un système dynamique hamiltonien . . . . .	77
3.4.4	Transformations canoniques et crochets de Poisson. . . . .	86

---

## Partie II Relativité restreinte et électromagnétisme

---

<b>4</b>	<b>Relativité restreinte</b> . . . . .	<b>97</b>
4.1	Insuffisances de la mécanique classique . . . . .	97
4.2	Transformation de Lorentz . . . . .	99
4.2.1	Nature de la transformation de Lorentz . . . . .	99
4.2.2	Loi de composition des vitesses relativistes . . . . .	102
4.3	Notations quadridimensionnelles. . . . .	103
4.3.1	Le 4–vecteur position d'un évènement. . . . .	104
4.3.2	Un peu de géométrie . . . . .	109
4.3.3	Les 4–vecteurs de la physique . . . . .	115
4.3.4	Les Tenseurs . . . . .	122
4.3.5	Calcul vectoriel différentiel en relativité restreinte . . . . .	128
<b>5</b>	<b>Principe de moindre action et relativité</b> . . . . .	<b>133</b>
5.1	Construction de l'action . . . . .	133
5.1.1	Particule libre . . . . .	133
5.1.2	Particule soumise à des forces. . . . .	134
5.2	Equation de la dynamique relativiste. . . . .	135
5.2.1	Action toujours varie, bien fol est qui s'y fie . . . . .	135

5.2.2	Propriétés du tenseur champ . . . . .	137
5.2.3	Premier groupe d'équations de Maxwell . . . . .	139
5.2.4	Force de Lorentz . . . . .	141
5.3	Sources du champ : 2 <sup>eme</sup> groupe d'équations de Maxwell . . . . .	142
5.3.1	Action d'interaction entre un champ et une assemblée de charges en mouvement . . . . .	142
5.3.2	Action d'auto-interaction du champ . . . . .	143
5.3.3	Invariance de jauge . . . . .	147
5.4	Les équations de Lagrange en relativité restreinte . . . . .	148
5.4.1	Equation de la dynamique d'une particule relativiste . . . . .	148
5.4.2	Les équations du champ . . . . .	150
5.5	Le théorème de Noether en théorie des champs . . . . .	152
5.5.1	Le théorème pour un champ scalaire . . . . .	152
5.5.2	Un exemple pour champ de 4-vecteurs . . . . .	152

---

## Partie III Théorie du champ de gravitation : Relativité générale

---

<b>6</b>	<b>Principe d'équivalence</b> . . . . .	157
<b>7</b>	<b>Application du principe d'équivalence</b> . . . . .	161
7.1	Mouvement géodésique . . . . .	161
7.2	Connexion affine et tenseur métrique . . . . .	162
7.3	La connexion affine est-elle un tenseur ? . . . . .	164
7.4	Dérivée covariante . . . . .	165
7.4.1	Nécessité d'une nouvelle dérivée . . . . .	165
7.4.2	Construction . . . . .	166
7.4.3	Propriétés . . . . .	168
7.4.4	Dérivée covariante le long d'une courbe . . . . .	169
7.5	Déviation géodésique : Courbure . . . . .	171
7.6	Propriétés de la courbure de Riemann-Christoffel . . . . .	172
7.6.1	Symétrie . . . . .	173
7.6.2	Courbure complètement covariante . . . . .	173
7.6.3	Contractions : Tenseur de Ricci et courbure scalaire . . . . .	174
7.6.4	Une remarque remarquable . . . . .	175
<b>8</b>	<b>Equations d'Einstein</b> . . . . .	177
8.1	Remarques préliminaires . . . . .	177
8.2	L'action de courbure et sa variation . . . . .	178

8.3	L'action de matière et sa variation . . . . .	182
8.3.1	Le tenseur énergie-impulsion . . . . .	182
8.3.2	Un petit exemple... . . . . .	185
8.4	Equations du champ gravitationnel . . . . .	186
8.4.1	Courbure = matière. . . . .	186
8.4.2	Choix de la constante . . . . .	188

---

**Partie IV Mécanique quantique et autres crochets**

---

<b>9</b>	<b>Fondements de la mécanique quantique.</b> . . . . .	195
9.1	Postulats de la mécanique quantique . . . . .	195
9.1.1	Espace des états de la mécanique quantique . . . . .	195
9.1.2	Grandeurs physiques . . . . .	198
9.1.3	Évolution temporelle . . . . .	203
9.1.4	Représentation spatiale . . . . .	207
9.2	Formulation de Dirac de la mécanique quantique. . . . .	211
9.2.1	Définitions, propriétés . . . . .	211
9.2.2	Relation de fermeture et applications . . . . .	213
9.2.3	Opérateurs. . . . .	214
9.2.4	Observables . . . . .	215
9.2.5	Fonction d'onde . . . . .	219
9.2.6	Commutation, compatibilité et indétermination.. . . .	222
9.2.7	Évolution temporelle des observables . . . . .	225
9.3	La mécanique quantique par le théorème de Noether . . . . .	229
9.3.1	Le lagrangien quantique . . . . .	229
9.3.2	La symétrie du lagrangien quantique. . . . .	232
<b>10</b>	<b>La physique non-dissipative en quelques crochets</b> . . . . .	235
10.1	Mécanique classique . . . . .	235
10.2	Électromagnétisme . . . . .	237
10.3	Relativité générale. . . . .	239
10.4	Physique statistique et mécanique des fluides . . . . .	240

---

**Partie V De la théorie à la pratique**

---

<b>11</b>	<b>Exercices</b> . . . . .	247
-----------	----------------------------	-----

11.1 Mécanique classique . . . . .	247
11.1.1 Juste pour vérifier ... . . . .	247
11.1.2 Pour se faire plaisir . . . . .	247
11.1.3 La chaîne d'oscillateurs couplés . . . . .	248
11.1.4 Le brachystochrone . . . . .	248
11.1.5 Avec contraintes . . . . .	249
11.1.6 Noether très simple . . . . .	249
11.1.7 De l'harmonique à Kepler . . . . .	250
11.1.8 Lagrange, Hamilton et le champ magnétique . . . . .	250
11.2 Relativités . . . . .	251
11.2.1 Quelques manipulations . . . . .	251
11.2.2 Electromagnétisme tensoriel . . . . .	251
11.2.3 Énergie-impulsion d'un fluide parfait. . . . .	252
11.2.4 Un point commun entre l'électromagnétisme et la relativité générale . . . . .	253
11.2.5 Le champ de Schwarzschild. . . . .	254
11.2.6 Une approche lagrangienne du problème des 2 corps : classique et relativiste . . . . .	255
11.3 Autour du théorème de Noether en théorie des champs . . . . .	258
11.3.1 Le courant de Noether. . . . .	258
11.3.2 Symétrie de jauge en électromagnétisme . . . . .	259
11.3.3 Symétrie infinitésimale. . . . .	259
11.3.4 Invariance par translation . . . . .	260
11.3.5 Tenseur de moment angulaire. . . . .	260
11.3.6 Non unicité du tenseur énergie-impulsion. . . . .	261
11.3.7 Tenseur énergie impulsion du champ scalaire . . . . .	261
11.3.8 Tenseur énergie impulsion du champ de Maxwell . . . . .	261
11.3.9 Généralisation du tenseur énergie impulsion . . . . .	262
<b>Bibliographie</b> . . . . .	263
<b>Index</b> . . . . .	265



---

# Objectifs et notations

## Objectifs

Le présent ouvrage est la concrétisation matérielle d'un cours donné par l'auteur à l'Ecole Nationale Supérieure de Techniques Avancées depuis de nombreuses années dans le cadre de l'enseignement thématique *Vision géométrique de la physique*. L'objectif de ce cours est simple à formuler, agréable à mettre en œuvre, mais nécessite un public de qualité.

C'est par une prise en compte des résultats généraux de la physique de Newton que la formulation analytique de la mécanique apparaît avec Lagrange dès la fin du XVIII<sup>ème</sup> siècle. Cette théorie prend alors le temps de mûrir dans l'esprit des physiciens pour permettre au début du XX<sup>ème</sup> siècle une formulation variationnelle de l'électromagnétisme, sous les traits de la relativité restreinte, puis de la gravitation dans le cadre de la relativité générale.

L'objectif de ce cours est donc simple : il s'agit dans un premier temps de mettre en place la formulation variationnelle de la mécanique classique afin d'en tirer par la suite, le meilleur parti dans le contexte des nouveaux principes relativistes. Cet objectif n'a rien d'original, on trouvera d'ailleurs dans les diverses références citées tous les éléments nécessaires à sa concrétisation. L'intérêt du présent ouvrage réside ainsi pour beaucoup dans son unité. Cet objectif est agréable à mettre en œuvre dans la mesure où il s'adresse à des étudiants en début de second cycle universitaire normalement érudits dans divers domaines de la physique mais n'ayant pas encore goûté aux beautés de formulations unificatrices. Ces étudiants – ou simples amateurs – à qui l'on a demandé dans le meilleur des cas de patienter pour en savoir plus, trouveront ici deux issues possibles à leurs investigations en philosophie de la nature : soit un point final unificateur à l'enseignement

scientifique qu'ils ont reçu depuis de longues années, soit l'ouverture qui leur permettra de s'enivrer de physique moderne.

## Notations

Etant donné que l'un des objectifs de cet ouvrage est d'apprendre à lire et à écrire la physique moderne, il est important d'adopter des notations précises et les plus représentatives possible de celles utilisées par les professionnels de la physique théorique.

Dans ce livre nous manipulerons plusieurs type d'objets mathématiques dans différents ensembles afin de représenter les choses de la Nature. Une grande partie de ces objets sont des tenseurs de différents ordres affectés par différents types opérateurs différentiels.

La notion de tenseur n'est pas si compliquée que l'on prétend. Sa présentation précise nécessite simplement d'avoir reçu, ou subit, un cours de base d'algèbre linéaire. Devant la carence en ce domaine dans de nombreux livres de physique, nous avons décidé d'y consacrer plusieurs pages. Mais nous ne le ferons qu'après être entré en contact avec suffisamment de cas, c'est-à-dire vers le milieu du livre. Nous pensons ainsi que le lecteur fidèle, au lieu d'être effrayé par un peu de mathématiques, y trouvera du réconfort et des explications. Les lecteurs plus érudits ou pressés peuvent néanmoins se rendre directement à la page 122 pour lever le voile sur ce point.

Pour être précis en physique comme en mathématiques, il faut préciser la référence que l'on utilise. Cette référence est souvent une base ou un référentiel (une base complétée par un point origine). Sans précision locale, cette base est la base canonique de l'espace considéré.

Les quantités dont la valeur ne dépend pas du référentiel d'étude sont appelées des scalaires. Ces scalaires peuvent avoir une valeur globale ou bien varier en fonction de la position, de la vitesse, du temps, etc., on parle alors de champ de scalaire ou bien de scalaire «tout court».

Ces scalaires ou bien ces champs scalaires seront notés en lettres italiques non grasses : la vitesse de la lumière  $c$  et toutes les constantes fondamentales de la nature sont des scalaires globaux, alors que la densité volumique  $\rho$  de masse ou de charge ont des valeurs qui dépendent de la position, de la vitesse, du temps, mais pas du référentiel : ce sont des champs scalaires tout comme l'énergie cinétique  $T$  ou l'énergie potentielle  $U$ . Généralement ces champs scalaires sont «fabriqués» à partir de produits scalaires définis dans les espaces considérés.

Les scalaires sont des tenseurs d'ordre 0.

Les quantités dont la valeur dépend du référentiel d'étude sont généralement des tenseurs. Cette ou ces valeurs sont leurs composantes dans le référentiel ou la base choisie. Quelque soit l'ordre de ces tenseurs, ces quantités seront généralement notées en gras italique :  $\mathbf{x}$ ,  $\mathbf{X}$ ,  $\boldsymbol{\eta}$ ,  $\mathbf{g}$ , etc.

Une exception notable à cette règle dans ce livre sont les vecteurs de  $\mathbb{R}^3$  pour lesquels nous utiliserons une notation avec une flèche à laquelle les étudiants sont habitués. Ainsi, la position ou la vitesse d'une particule ponctuelle sera repérée dans l'espace à un instant  $t$  par la notation  $\vec{r}$  ou  $\vec{v}$ . Par contre la coordonnée généralisée d'un système possédant  $s$  degrés de liberté qui regroupe toutes les informations concernant la position dans l'espace d'un ensemble de  $N$  particules et qui possède donc  $s \leq 3N$  composantes sera notée  $\mathbf{q}$ , même si le cas  $N = 1$  pourrait imposer des flèches.

Cette notation qui permet un passage en douceur de la notion de vecteur à celle de tenseur, permet en outre de simplifier quelques écritures de composantes. Par exemple, dans l'espace-temps – que nous noterons  $\mathbb{M}^4$  car il diffère de  $\mathbb{R}^4$  par ses propriétés métriques – on pourra ainsi exprimer facilement la 4-position  $\mathbf{x}$  d'un évènement en fonction de  $t$  et de  $\vec{r}$ .

Les champs de vecteurs ou de tenseurs seront donc également notés de cette manière grasse ou fléchée en fonction de la dimension de l'espace dans lequel ils vivent ou travaillent. Un champ de force dans l'espace sera donc noté  $\vec{F}(\vec{r})$ , il donne le vecteur force de  $\mathbb{R}^3$  qui s'exerce en chaque point de cet espace sur une particule repérée par sa position  $\vec{r}$ . Par contre le champ de vecteur permettant de changer de référentiel galiléen dans  $\mathbb{M}^4$  s'écrira  $\mathbf{L}(\mathbf{u}) = \mathbf{u}'$ .

Enfin, les composantes de ces vecteurs ou tenseurs, qui sont des nombres réels ou complexes, seront notées en lettres italique non grasses :  $r_i$ ,  $x^\mu$ ,  $L^\mu_\nu$ ,  $\eta^{\mu\nu}$ , etc. La position de ces différents indices sera précisée dans la section consacrée aux tenseurs (voir page 122).

Un dernier point concerne les notations concernant la représentation des dérivées de ces différents types d'objets. Ce point est fondamental car la physique est gourmande de dérivées !

La dérivée d'une fonction  $f$  par rapport à son unique variable réelle  $x$  sera indiquée par les notations traditionnelles. Si  $f : \mathbb{R} \rightarrow \mathbb{R}$ , on écrira quand tout se passe bien

$$f'(x) = \frac{df}{dx} : \begin{array}{c} \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto \lim_{y \rightarrow 0} \frac{f(x+y) - f(x)}{y} \end{array}$$

C'est la notation introduite par Joseph-Louis Lagrange.

Dans le cas particulier d'une fonction de la variable temporelle, le ' peut être remplacé par un ` , c'est la notation introduite par Isaac Newton. Dans les deux cas, on parle en physique de dérivée *totale*.

La dérivée d'un champ scalaire  $g$  d'une variable vectorielle  $\mathbf{x} = (x_1, \dots, x_n)$  par rapport à une seule de ses variable  $x_i$  est donc une dérivée *partielle*. Elle est traditionnellement notée  $\frac{\partial g}{\partial x_i}$ , sa signification est simplement la généralisation du cas monodimensionnel pour la seule variable  $x_i$  de la fonction  $g(\mathbf{x})$ .

Le vecteur composé de toutes les dérivées de  $g$  par rapport à toutes ses variables est le gradient de  $g$  dans la direction du vecteur  $\mathbf{q}$ . La notation officielle de ce vecteur est la suivante :

$$\text{Si } g : \mathbb{R}^n \rightarrow \mathbb{R} \quad \mathbf{x} \mapsto g(\mathbf{x}) \quad \text{on notera } \text{grad}_{\mathbf{x}}(g) := \frac{\partial g}{\partial \mathbf{x}} = \left[ \frac{\partial g}{\partial x_1}, \dots, \frac{\partial g}{\partial x_n} \right]^T$$

La première formulation fait appel à l'opérateur grad. Elle est classique mais prend de la place, d'autant plus qu'il faut préciser dans notre contexte la direction dans laquelle on prend ce gradient, ici  $\mathbf{x}$ . C'est pourquoi on lui préférera souvent la seconde qui utilise la notation d'une dérivée partielle par rapport à un vecteur.

Nous ne nous priverons pas d'étendre cette notion à des fonctions de plusieurs variables vectorielles. Par exemple, nous verrons qu'un lagrangien  $\mathcal{L}$  est un champ scalaire des coordonnées généralisées  $\mathbf{q} = (q_1, \dots, q_s) \in \mathbb{R}^s$  et des vitesses généralisées  $\dot{\mathbf{q}} = (\dot{q}_1, \dots, \dot{q}_s) \in \mathbb{R}^s$ . De façon très pratique, nous écrivons :

$$\text{Si } \mathcal{L} = \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}), \text{ on écrira } \left\{ \begin{array}{l} \frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \left[ \frac{\partial \mathcal{L}}{\partial q_1}, \dots, \frac{\partial \mathcal{L}}{\partial q_s} \right]^T \in \mathbb{R}^s \\ \text{ou bien} \\ \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = \left[ \frac{\partial \mathcal{L}}{\partial \dot{q}_1}, \dots, \frac{\partial \mathcal{L}}{\partial \dot{q}_s} \right]^T \in \mathbb{R}^s \end{array} \right.$$

D'une manière générale le produit scalaire dans un espace vectoriel euclidien est noté par un point central entre les deux vecteurs impliqués. Par exemple, si deux vecteurs de  $\mathbb{R}^n$ ,  $\mathbf{x}$  et  $\mathbf{y}$ , ont pour composantes respectives dans sa base canonique  $[x_1, \dots, x_n]^T$  et  $[y_1, \dots, y_n]^T$ , leur produit scalaire s'écrira

$$\forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n \quad \mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i \in \mathbb{R}$$

Dans l'espace de Minkowski  $\mathbb{M}^4$  de la relativité restreinte, la métrique n'est plus euclidienne. Le produit scalaire s'en trouve affecté et devient dans les mêmes conditions

$$\forall(\mathbf{x}, \mathbf{y}) \in \mathbb{M}^4 \times \mathbb{M}^4 \quad \mathbf{x} \cdot \mathbf{y} = x_0 y_0 - \sum_{i=1}^3 x_i y_i \in \mathbb{R}$$

On notera que la première composante est indicée par 0, c'est une autre spécificité des éléments de  $\mathbb{M}^4$ . Les raisons de ces spécificités se feront jour le moment venu.

Le choix de la signature  $(+, -, -, -)$  est conventionnel et ne change aucun résultat physique. Ce choix est généralement celui de la théorie des champs orientée vers la physique des particules, dans les ouvrages spécifiques de relativité générale on utilise plutôt la convention opposée.

## Remerciements

L'auteur remercie toutes les personnes qui, par leurs encouragements et les corrections qu'ils ont apportés au présent ouvrage, ont contribué à sa réalisation. Généralement les étudiants apprécient ce cours qui leur permet de faire le point avec leur physique, je les remercie globalement pour leur enthousiasme et leurs questions qui ont enrichi ce document. Enfin, je remercie plus particulièrement Marc Lenoir, Michel Garcia et Luc André dont les relectures initiales furent d'une précision redoutable, Franz Glessgen qui m'a aidé en apprenant lors de la relecture de la seconde édition. C'est le confinement du printemps 2020 qui m'a donné l'occasion d'unifier les notations et de clarifier certains chapitres, qu'il en soit remercié!



Mécanique analytique



# L'incroyable legs de Joseph-Louis Lagrange

Ce chapitre est issu d'un long travail historique mené par l'auteur et visant à clarifier l'origine de la théorie des champs classique : comment Lagrange a-t-il obtenu ses résultats et que lui doit-on réellement ?

## 1.1 L'origine

L'initiateur de Lagrange est sans conteste le grand Newton. Publiée en 1687 mais entreprise intellectuellement en 1666, la «*Philosophiae naturalis principia mathematica*» d'Isaac Newton donne pour la première fois dans l'histoire des sciences du grain à moudre à qui veut comprendre la philosophie naturelle. En fait, la grande avancée de ce texte difficile est la définition de la notion de force : une force est une quantité physique qui mesure, par unité de masse, la variation temporelle de la vitesse d'un corps. Tout part de là : un corps dont la vitesse ne varie pas n'est soumis à aucune force nette; la pomme et la lune ont des vitesses qui varient et la force universelle qui les anime est mise en place par des arguments comparatifs; tous les cas simples sont étudiés comme les forces proportionnelles à la position ou à la vitesse; finalement le cas miraculeux du mouvement gravitationnel de deux corps est résolu et les lois de Kepler retrouvées.

Même si son contemporain et peut-être inspirateur Hooke a incontestablement compris une partie des travaux de Newton, tant de nouveautés techniques vont mettre près d'un siècle pour être complètement assimilées par ses pairs. Ce n'est qu'en 1742 que l'on pourra voir Mac Laurin expliquer qu'il faut projeter la relation de Newton sur trois axes d'espace pour obtenir des équations différentielles dont les solutions correspondent au mouvement... C'est justement sur ce genre de problème technique que se penche le jeune Joseph-Louis Lagrange : les méthodes de résolution des équations différen-

tielles linéaires. Il semblerait que Lagrange ait compris très tôt (pendant sa période turinoise i.e. avant 1766) la portée de la méthode de la variation des constantes qu'il ne publiera pourtant que tardivement vers 1808. Je propose donc de retracer un cheminement didactique des idées qui permettent de comprendre comment l'on peut construire ce que l'on appelle désormais le formalisme lagrangien à partir du problème qui occupait les physiciens depuis la mort de Newton : le mouvement de la Lune, i.e. le problème des deux corps perturbé.

Nous ne tenons aucune preuve que le cheminement intellectuel du génial Lagrange s'est effectivement produit de la façon que nous allons présenter; mais tout est bien présent dans l'œuvre du Maître, pas forcément dans le bon ordre ou tout aussi explicitement dit.

## 1.2 La méthode de variation des constantes

### 1.2.1 Ordre 1 : l'idée

Depuis Lagrange, tout le monde sait que pour résoudre l'équation différentielle

$$y' + a(x)y = f(x) \quad \text{où } y' = \frac{dy}{dx} \quad (1.1)$$

Il faut tout d'abord résoudre l'équation amputée du second membre  $f(x)$ ,  $h' + a(x)h = 0$  improprement<sup>1</sup> appelée équation homogène associée à (1.1). En suivant Lagrange, on procède pour cela par séparation des variables et, lorsque tout ceci à un sens, on obtient  $h(x) = \lambda \exp(-\int a(x) dx)$  ou  $\lambda$  est une constante arbitraire, ici réelle. Pour trouver l'ensemble des solutions de l'équation (1.1), il ne suffit alors plus que d'en trouver une solution particulière  $p(x)$ . C'est encore Lagrange qui, utilisant la structure linéaire de l'équation, propose de faire varier la constante  $\lambda$  en essayant  $p(x) = \lambda(x) \exp(-\int a(x) dx)$ .

Pour qu'un tel  $p(x)$  soit bien solution de (1.1) il suffit que

$$\frac{d\lambda}{dx} = f(x) \exp\left(+\int a(x) dx\right)$$

La fonction  $\lambda(x)$  est ainsi parfaitement définie comme la primitive d'une certaine fonction, elle permet donc de trouver la solution particulière cherchée. La solution de l'équation différentielle linéaire (1.1) est donc complète,

---

<sup>1</sup> Une équation homogène est une équation invariante par changement d'échelle. Ainsi, une

c'est le premier pas effectué par des méthodes obscures et en utilisant des notations malhabiles par notre piémontais dès 1765<sup>2</sup>.

### 1.2.2 Ordres supérieurs : la généralisation

Lorsque l'équation différentielle reste linéaire mais voit son ordre dépasser l'unité, les problèmes arrivent. La solution de l'équation

$$y'' + a(x)y' + b(x)y = f(x) \quad \text{où } y'' = \frac{d^2y}{dx^2} \quad (1.2)$$

s'écrit toujours comme la somme de la solution générale de l'équation sans second membre  $h(x)$  et d'une solution particulière de l'équation complète  $p(x)$ , i.e.  $y(x) = h(x) + p(x)$ ; mais la méthode de séparation des variables ne nous permet plus de trouver l'expression générale de  $h(x)$ . En utilisant des concepts modernes, apparus seulement vers la fin du XIX<sup>e</sup> siècle, nous savons cependant que  $h(x)$  est élément d'un espace vectoriel de dimension 2. Il s'écrit donc dans la base  $[h_1, h_2]$  sous la forme de la combinaison linéaire  $h(x) = \lambda h_1(x) + \mu h_2(x)$  où  $\lambda$  et  $\mu$  sont deux constantes toujours réelles; le problème réside dans le fait que nous ne disposons plus d'une méthode générale permettant de trouver les fonctions  $h_1$  et  $h_2$ .

Sans disposer des concepts, Lagrange avait saisi la nature de la solution et avait pleine conscience du problème. L'équation étant toujours linéaire, il propose de chercher la solution générale sous la forme  $p(x) = \lambda(x)h_1(x) + \mu(x)h_2(x)$ ; il s'agit comme précédemment de trouver une seule fonction permettant de construire ce qui s'avérera être l'espace affine des solutions de (1.2).

L'idée est excellente mais les choses se compliquent : en utilisant la forme précise de  $p(x)$  et le fait que  $h_1$  et  $h_2$  sont solutions de l'équation sans second membre, on obtient assez vite une relation de la forme

$$\lambda''h_1 + \lambda'(2h_1' + ah_1) + \mu''h_2 + \mu'(2h_2' + ah_2) = f. \quad (1.3)$$

La résolution de cette équation dont les inconnues sont les fonctions  $\lambda$  et  $\mu$  conduit au système

$$\begin{cases} u'h_1 + u(2h_1' + ah_1) = f \\ v'h_2 + v(2h_2' + ah_2) = 0 \end{cases} \quad \text{avec } u = \lambda' \text{ et } v = \mu'. \quad (1.4)$$

<sup>1</sup> équation linéaire sans second membre est homogène mais la réciproque est fausse.

<sup>2</sup> Voir les *Solution de différents problèmes de calcul intégral*, Miscellanea Taurinensia, t. III, 1762-1765, Tome 1 des Œuvres complètes p.471-668

Ce système est découplé et l'on peut donc, toujours en utilisant la méthode de séparation des variables, trouver les expressions de  $u$  et  $v$  qui sont en fait les primitives des fonctions  $\lambda$  et  $\mu$  recherchées. Sans être explicite, car on ne connaît généralement pas les expressions des fonctions  $h_1(x)$  et  $h_2(x)$ , la méthode fonctionne toujours !

Dans son génie, Lagrange propose alors un court-circuit qui va se révéler d'une fécondité extraordinaire.

Au lieu de considérer l'expression la plus générale  $p' = \lambda h_1' + \mu h_2' + \lambda' h_1 + \mu' h_2$  et compte-tenu du fait que l'on ne cherche en fait qu'une seule solution particulière, Lagrange propose de tuer les deux derniers termes de cette dérivée qui sont à l'origine dans le calcul de  $p''$  de complications inutiles... Il impose donc aux deux fonctions inconnues  $\lambda$  et  $\mu$  de vérifier la relation

$$\lambda' h_1 + \mu' h_2 = 0. \tag{1.5}$$

Ce court-circuit est salutaire dans les calculs qui donnent alors directement

$$\lambda' h_1' + \mu' h_2' = f$$

Adjointe à la contrainte (1.5) cette équation forme un système linéaire qui n'est plus différentiel et que l'on peut généralement inverser si le déterminant de Wronski<sup>3</sup>  $w = h_1 h_2' - h_1' h_2$  n'est pas identiquement nul. On trouve alors

$$\lambda' = -\frac{f h_2}{w} \text{ et } \mu' = +\frac{f h_1}{w}$$

On voit ici le pragmatisme de Lagrange, mais qu'en est-il de la généralité de l'approche ? Si l'on reste conscient du fait que l'on ne cherche qu'une solution particulière  $p$  pour compléter l'espace vectoriel des solutions de l'équation sans second membre, on est convaincu par la simplification pour peu que les primitives soient acceptables. Si l'on y regarde d'un peu plus près, on se rend même compte que la simplification est amplement espérée par le fait que les solutions  $\lambda$  et  $\mu$  du système (1.4) dépendent de 4 constantes d'intégration là où 2 seulement étaient attendues compte-tenu de l'ordre de l'équation initiale.

Remarquons enfin que la contrainte (1.5) n'est pas unique. En physique théorique moderne, elle a pris le nom de jauge et va se révéler tout à fait fondamentale dans la compréhension des symétries des équations de la physique

---

<sup>3</sup> Josef Hoëné-Wronski est un contemporain du vieux Lagrange avec lequel il a travaillé et s'est même fâché en 1812... L'appellation de Wronskien pour ce genre de déterminant est due à

que Lagrange est ici en train d'ébaucher. A la mémoire de ce grand initiateur, nous proposons le nom de *jauge de Lagrange* pour la contrainte (1.5).

Une fois ce travail compris, assimilé et généralisé aux équations linéaires d'un ordre  $n$  quelconque, Lagrange passe à l'application de cette théorie au cas particulier qui occupait les mécaniciens du ciel de cette époque : le mouvement de la Lune.

## 1.3 Application au mouvement de la Lune

Le grand succès de Newton avait été, entre autres, de résoudre le problème des deux corps et de retrouver, par le calcul, les lois obtenues empiriquement par Kepler. Son grand souci, qu'il ne manquait pas de reconnaître, était de s'être avoué vaincu lorsque l'on sortait du cadre académique du problème qu'il avait résolu : soit que les corps en question n'avaient pas la symétrie sphérique, soit qu'ils étaient plus que deux, soit qu'une autre force non gravitationnelle intervenait dans le problème.

Dans le cas du mouvement de la Lune, Newton lui-même s'était rendu compte du fait que sa théorie n'était applicable qu'à l'ordre le plus bas. Et si l'on veut suivre notre satellite sur quelques périodes orbitales, on s'écarte très vite de la solution académique. Les principaux astronomes du début du XVIII<sup>e</sup> siècle comme Clairaut, d'Alembert, Bernoulli ou Euler se cassèrent d'ailleurs les dents sur ce problème. Celui qui vint apporter la première véritable théorie viable du mouvement de la Lune est bel est bien Joseph-Louis Lagrange stimulé par son concurrent et collègue Pierre-Simon Laplace.

### 1.3.1 L'idée de la méthode

Pour ce problème, Lagrange, illustre mathématicien, va user de méthodes approximatives dont il sait pertinemment qu'elles ne sont pas forcément licites dans le contexte général du problème traité. Il le dit explicitement à de nombreuses reprises au début de ses nombreux articles sur le sujet, mais il les utilise quand même, pour chercher l'inspiration, en considérant que les perturbations sont «petites».

De quoi s'agit-il ?

Lagrange remarque que les équations générales du mouvement de la Lune, incluant des perturbations, sont trois équations différentielles d'ordre deux. Dans le problème non perturbé de Newton, six constantes du mouvement

---

Thomas Muir à la toute fin du XIX<sup>e</sup> siècle.

apparaissent (voir figure 1.1): deux pour chacune des trois équations. Dans le cas d'un problème lié (i.e. d'énergie négative), la solution de ce problème est une ellipse définie dans l'espace par six paramètres : la géométrie de l'ellipse avec son demi-grand axe  $a$  et son excentricité  $e$ ; la position du plan orbital fixe par rapport à un système de référence avec l'inclinaison  $i$  et l'argument de la ligne des nœuds  $\Omega$ ; enfin la position des points, l'un étant au foyer de l'ellipse et l'autre décalé d'un angle  $f(t)$  du périastre, lui-même décalé d'un angle constant  $\omega$  par rapport au plan de référence. L'angle  $f(t)$  dépend du temps car le point  $B$  bouge sur l'ellipse par contre la direction du périastre  $P$  est fixe dans ce problème. Pour rajouter une dernière constante, on introduit un temps  $\tau$  correspondant à l'instant du dernier passage de  $B$  au périastre  $P$ .

Les six constantes que nous venons d'introduire sont rassemblées dans un «vecteur<sup>4</sup>»  $\mathbf{C} = (a, e, i, \omega, \Omega, \tau) \in \mathbb{R}^6$  caractérisant la solution  $\vec{r} \in \mathbb{R}^3$  du problème des deux corps à chaque instant  $t$  :  $\vec{r} = \vec{r}(t, \mathbf{C})$ . Dans le cas du mouvement elliptique le vecteur  $\mathbf{C}$  est constitué des éléments elliptiques, on peut évidemment considérer d'autres représentations du mouvement. L'idée de Lagrange consiste à dire qu'en extrapolant la technique de la variation des constantes, la solution du problème à deux corps perturbé sera de la forme  $\vec{r} = \vec{r}(t, \mathbf{C}(t))$ . Il est bien conscient du fait que le problème n'est pas linéaire, mais si la perturbation est petite, la déformation de l'ellipse devrait faire apparaître des «éléments osculateurs elliptiques»  $\mathbf{C}(t)$  dépendant maintenant du temps.

### 1.3.2 La mise en œuvre

Dans le référentiel où l'un des deux corps perturbé est fixe, en notant  $\vec{r}$  le vecteur position du second, le problème à deux corps perturbé s'écrit

$$\ddot{\vec{r}} = -\mu \frac{\vec{r}}{|\vec{r}|^3} + \vec{F} \quad \text{avec } \dot{\phantom{x}} = \frac{d}{dt}$$

où  $\mu = G(M + m)$  et  $\vec{F}$  est le vecteur décrivant la force perturbant le problème à deux corps de masses  $m$  et  $M$ . Comme on le fait pour la méthode de variation des constantes, on remplace les trois inconnues que sont les coordonnées de  $\vec{r}$  par les six que représentent celles du vecteur  $\mathbf{C}(t)$  tel que  $\vec{r} = \vec{r}(t, \mathbf{C}(t))$ , on a donc

---

<sup>4</sup> Ce n'en est pas vraiment un car si tel était le cas, il faudrait préciser la base dans laquelle on travaille. Ce n'est donc qu'une notation bien pratique...

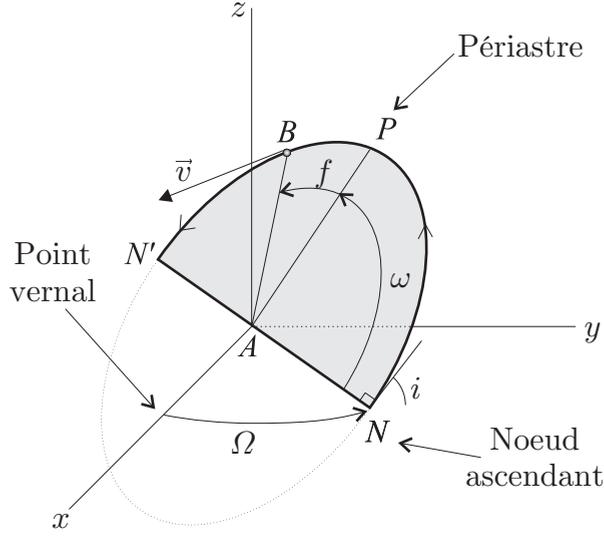


Fig. 1.1. Éléments elliptiques pour une orbite képlérienne

$$\dot{\vec{r}} = \frac{\partial \vec{r}}{\partial t} + \sum_{i=1}^6 \frac{\partial \vec{r}}{\partial C_i} \dot{C}_i$$

Si l'on se place dans la jauge de Lagrange, ce qui est vital ici pour continuer à y voir clair, conformément aux indications de la section 1.2.2 (eq. 1.5), il convient d'imposer la condition

$$\sum_{i=1}^6 \frac{\partial \vec{r}}{\partial C_i} \dot{C}_i = 0 \quad (1.6)$$

Ce qui correspond à trois conditions sur chacune des coordonnées de  $\vec{r}$ . On peut alors calculer dans cette jauge la dérivée temporelle seconde du vecteur position, il vient

$$\ddot{\vec{r}} = \frac{\partial^2 \vec{r}}{\partial t^2} + \sum_{i=1}^6 \frac{\partial \dot{\vec{r}}}{\partial C_i} \dot{C}_i$$

L'équation du mouvement du problème à deux corps non perturbé s'écrit  $\ddot{\vec{r}} = \frac{\partial^2 \vec{r}}{\partial t^2} = -\mu \frac{\vec{r}}{|\vec{r}|^3}$ , il est donc naturel d'identifier la totalité de la perturbation dans

$$\vec{F} = \sum_{i=1}^6 \frac{\partial \dot{\vec{r}}}{\partial C_i} \dot{C}_i \quad (1.7)$$

Le génie de Lagrange fait la suite : pour combiner entre-elles les relations (1.6) et (1.7), il faut faire le produit scalaire de l'une par  $\frac{\partial \dot{\vec{r}}}{\partial C_j}$  et de l'autre

par  $\frac{\partial \vec{r}}{\partial C_j}$  pour un  $j$  fixé. En retranchant ces deux produits scalaires on obtient alors

$$\sum_{i=1}^6 \left( \frac{\partial \vec{r}}{\partial C_j} \cdot \frac{\partial \dot{\vec{r}}}{\partial C_i} - \frac{\partial \dot{\vec{r}}}{\partial C_j} \cdot \frac{\partial \vec{r}}{\partial C_i} \right) \dot{C}_i = \frac{\partial \vec{r}}{\partial C_j} \cdot \vec{F} \quad (1.8)$$

On introduit les parenthèses de Lagrange

$$(C_j, C_i) = \frac{\partial \vec{r}}{\partial C_j} \cdot \frac{\partial \dot{\vec{r}}}{\partial C_i} - \frac{\partial \dot{\vec{r}}}{\partial C_j} \cdot \frac{\partial \vec{r}}{\partial C_i}$$

puis l'on restreint le problème à des perturbations qui dérivent d'une «fonction perturbatrice» notée  $R(\vec{r})$  et telle que

$$\vec{F} = + \text{grad}_{\vec{r}} R = \frac{\partial R}{\partial \vec{r}}$$

Ici encore cette hypothèse restreint la généralité du problème mais elle est pleine de pragmatisme : sans elle, on ne peut rien faire avec la relation (1.8), avec elle, Lagrange invente ni plus ni moins que la notion d'énergie potentielle dont bien sûr il comprendra l'importance et se resservira dans sa mécanique analytique. Sous cette hypothèse et avec ces notations, on peut écrire la relation (1.8) sous forme matricielle

$$\mathbf{L} \dot{\mathbf{C}} = \text{grad}_{\mathbf{C}} R \quad \text{avec } L_{ij} = (C_j, C_i)$$

La matrice  $\mathbf{L}$ , que j'appellerai matrice de Lagrange, contient donc 6 lignes et 6 colonnes, elle est antisymétrique et ses éléments sont des parenthèses de Lagrange. Aidé de son élève Poisson, Lagrange montrera au soir de sa vie que ces parenthèses ne dépendent pas explicitement du temps, ce sont donc des intégrales premières du mouvement. Pour les calculer on peut ainsi se placer en un point particulier de la trajectoire, le périhélie par exemple, qui permet de simplifier les calculs. Sans l'algèbre linéaire, mais assez efficacement tout de même, Lagrange inverse le système et l'écrit explicitement sous une forme qui dans une écriture moderne deviendrait

$$\dot{\mathbf{C}} = \mathbf{L}^{-1} \text{grad}_{\mathbf{C}} (R) \quad (1.9)$$

Ces équations sont connues en mécanique céleste comme les «équations planétaires de Lagrange». Leur utilisation dans le problème du mouvement de la Lune autour de la Terre perturbée par le Soleil (voir [20]) a permis au Maître d'obtenir l'explication de certains mouvement du périhélie et de la ligne des nœuds lunaire observés depuis l'antiquité et sans explication jusque



## 1.4 Généralisation à tous les problèmes de mécanique conservative

### 1.4.1 La construction des équations de Lagrange

Avant Lagrange la seule énergie connue était l'énergie cinétique. Étudiée et définie sous diverses formes par Leibniz et Bernoulli, c'est Huygens qui comprendra son rôle effectif dans le mouvement des corps. Ce dernier point est rappelé par Lagrange dans l'introduction historique de sa mécanique analytique dans laquelle il précise l'état de l'art avant de tout reformuler...

Lagrange entreprend donc de réécrire les équations du mouvement (de Newton) en utilisant l'énergie cinétique et en introduisant une coordonnée généralisée  $\mathbf{q} \in \mathbb{R}^s$  dont les composantes  $q_i$  généralisent les  $C_i$  du problème à deux corps perturbé. La dimension de  $\mathbf{q}$  est fixée par le nombre  $s$  de degré de liberté du système considéré, qu'une étude préalable devra déterminer. Le nombre  $s$  est optimal ainsi les variables  $q_i$  sont indépendantes

$$\forall i, j \in [1, s] \quad \frac{\partial q_i}{\partial q_j} = \delta_{ij}$$

Tout comme les  $C_i$  ne sont pas nécessairement homogènes à des longueurs, les  $q_i$  ne sont que  $s$  quantités indépendantes dont la connaissance de l'évolution permet de déterminer entièrement le système considéré. Ces coordonnées généralisées permettent de définir des vitesses généralisées par la relation  $\dot{\mathbf{q}} = \frac{d\mathbf{q}}{dt}$ . Les  $s$  quantités ainsi définies ne sont elles aussi pas forcément homogènes à des vitesses, mais elles sont par construction indépendantes entre elles

$$\forall i, j \in [1, s] \quad \frac{\partial \dot{q}_i}{\partial \dot{q}_j} = \delta_{ij}$$

et strictement indépendantes des coordonnées généralisées

$$\forall i, j \in [1, s] \quad \frac{\partial q_i}{\partial \dot{q}_j} = 0$$

Le contexte étant posé, il est temps de passer à l'action : si cela est possible, pour généraliser ses équations planétaires Lagrange doit combiner entre elles certaines dérivées d'une énergie par rapport à des coordonnées généralisées pour faire apparaître des vitesses généralisées. La seule énergie dont il dispose est comme nous l'avons dit, l'énergie cinétique qui, pour  $N$  particules de masse  $m_{k=1, \dots, N}$  repérées par les vecteurs  $\vec{r}_{k=1, \dots, N}$ , s'écrit

$$T = \sum_{k=1}^N \frac{1}{2} m_k \dot{\vec{r}}_k^2$$

Par définition, les vecteurs position  $\vec{r}_k$  ne dépendent que des coordonnées généralisées :  $\vec{r}_k = \vec{r}_k(t, \mathbf{q})$ ; ainsi l'énergie cinétique dépend à la fois des coordonnées et des vitesses généralisées. Pour généraliser ses calculs perturbatifs du problème à deux corps, Lagrange doit donc se livrer au calcul des deux dérivées :

$$\frac{\partial T}{\partial q_\ell} \quad \text{et} \quad \frac{\partial T}{\partial \dot{q}_\ell}$$

pour chaque composante  $\ell = 1, \dots, s$  du vecteur  $\mathbf{q}$ . Ce calcul est un exploit pour l'époque où les notations et le formalisme ne sont pas ce qu'ils sont aujourd'hui. Songez que l'invention du calcul différentiel ne date à cette époque que d'une dizaine d'années...

Après de laborieux efforts que l'on peut suivre dans sa mécanique analytique, ou bien dans un style plus lisible dans le chapitre suivant de cet ouvrage (voir section 2.4.2), il arrive au nœud gordien de l'affaire en remarquant que

$$\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_\ell} \right) - \frac{\partial T}{\partial q_\ell} = \sum_{k=1}^N m_k \ddot{\vec{r}}_k \cdot \frac{\partial \vec{r}_k}{\partial q_\ell} \quad (1.10)$$

En suivant Newton, il reconnaît que  $m_k \ddot{\vec{r}}_k = \vec{F}_k$  représente la force appliquée par tous les autres corps sur le corps  $k$ . Arrivé à ce stade et sans autre hypothèse, tout comme dans le problème du mouvement perturbé de la Lune, il ne peut simplifier davantage cette expression... Mais justement, l'expérience acquise sur ce problème lui ouvre alors la voie : tout se simplifie s'il suppose que les forces dérivent d'un champ scalaire  $U$  qui généralise la fonction perturbatrice  $R$  du vénérable problème.

$$\text{Si } \exists U(\vec{r}_1, \dots, \vec{r}_N) \text{ tel que } \vec{F}_k = -\text{grad}_{\vec{r}_k} U$$

alors l'équation (1.10) se simplifie en

$$\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_\ell} \right) - \frac{\partial T}{\partial q_\ell} = -\frac{\partial U}{\partial q_\ell}$$

Quelques remarques s'imposent!

Le signe  $-$  qui apparaît devant le champ scalaire  $U$  vient de Lagrange qui dans son texte utilise la lettre  $V$ ... Il est bien expliqué<sup>5</sup> que les composantes

<sup>5</sup> page 305 du tome 1 de la seconde édition de la mécanique analytique...

du gradient (notées  $P$ ,  $Q$  et  $R\dots$ ) duquel dérivent les forces tendent à «*diminuer ces lignes*», il faut donc comprendre que Lagrange considère l'opposé des forces! Dans les livres de mécanique moderne la quantité  $V$  n'est autre que l'énergie potentielle.

Dans son texte<sup>6</sup> Lagrange note  $Z = T - V$ , ce n'est que bien plus tard, dans le premier quart du XX<sup>e</sup> siècle<sup>7</sup> que cette quantité est devenue le lagrangien tel que nous l'appelons aujourd'hui.

Les  $s$  équations que représentent la notation

$$\frac{d}{dt} (\text{grad}_{\dot{\mathbf{q}}} T) = \text{grad}_{\mathbf{q}} (T - U)$$

sont maintenant connues sous le nom d'équation de Lagrange de la mécanique analytique. En introduisant le lagrangien conservatif

$$\mathcal{L} = T - U$$

et comme  $U$  ne dépend pas des vitesses généralisées, on peut les réécrire sous une forme plus conventionnelle

$$\forall \ell = 1, \dots, s \quad \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_\ell} - \frac{\partial \mathcal{L}}{\partial q_\ell} = 0 \quad (1.11)$$

C'est le grand Euler qui, sur un cas particulier initial que lui soumis le jeune Lagrange, remarqua que cette famille d'équations correspondait à l'extremum d'une quantité

$$S = \int_{t_1}^{t_2} \mathcal{L} dt$$

adjointe de bonnes conditions aux limites. C'est pourquoi l'on donne quelquefois, et fort injustement, le nom de système d'Euler-Lagrange aux équations (1.11).

Les choses ayant été clarifiées, dans les prochains chapitres, afin d'adopter une présentation plus déductive et moins historique, nous partirons de ce «principe de moindre action» pour reconstruire la mécanique tout d'abord classique et enfin relativiste.

Mais avant de faire tourner la machine lagrangienne, nous allons voir que Lagrange ne s'est pas arrêté là ! Comment aurait-il pu le faire sachant

<sup>6</sup> voir par exemple page 325 du tome 1 de la seconde édition de sa mécanique analytique

<sup>7</sup> Sans doute sous l'impulsion de Paul Dirac comprenant l'importance de la mécanique analytique

qu'il poursuivait le but de retrouver l'équivalent de ses équations planétaires d'ordre 1 en temps alors les équation de Lagrange sont d'ordre 2... Et enfin, que sont devenues les parenthèses de Lagrange ?

Les équations de Lagrange n'étaient donc qu'une étape pour leur inventeur !

### 1.4.2 La fin de son œuvre.

Les équations de Hamilton ont été écrites par Lagrange !

Il suffit pour le vérifier de lire la seconde édition (1811) du tome 1 de la mécanique analytique au milieu de la page 336. Tout n'est pas très explicite mais Lagrange part bien de son lagrangien qu'il note  $Z$  dépendant des variables  $(\xi, \psi, \varphi, \xi', \psi', \varphi')$  vérifiant des équations du second ordre pour obtenir une nouvelle fonction qu'il note  $\Omega$  dépendant de nouvelles variables appelées  $(\alpha, \beta, \gamma, \lambda, \mu, \nu)$  et solutions d'équations du premier ordre. Tout est expliqué entre les pages 329 et 336 de ce merveilleux livre dans lequel l'introduction précise qu'«*on ne trouvera point de Figures dans cet Ouvrage*». Notons enfin que la quantité  $\mathcal{H} = T + V$  est elle aussi de Lagrange, on trouve cette notation à de nombreuses reprises dans son livre<sup>8</sup>. La lettre  $\mathcal{H}$  est introduite sans autre référence qu'une précision qui revient comme un leitmotiv « *$\mathcal{H}$  est une constante arbitraire qui renferme la conservation des forces vives*», or on sait, puisqu'il le rappelle longuement dans son introduction que c'est Huygens qui proposa ce principe... Si l'on avait lu Lagrange, c'est donc naturellement *huygensien* que l'on devrait dire en lieu et place du traditionnel hamiltonien. C'est sans doute Dirac le fautif, mais nul ne peut lui en vouloir !

Traduisons en des termes modernes ce passage de l'ordre 2 à l'ordre 1 dans les équations du mouvement.

La méthode est maintenant largement utilisée : le lagrangien  $\mathcal{L}$  est une fonction des variables  $\mathbf{q}$  et  $\dot{\mathbf{q}}$ , en posant  $\mathbf{p} = \text{grad}_{\dot{\mathbf{q}}} \mathcal{L}$  et en effectuant ce que l'on appelle désormais une transformation de Legendre<sup>9</sup>, on introduit le champ scalaire  $\mathcal{H} = \mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{L}$  pour lequel on montre que si tout va bien

$$\mathcal{H} = \mathcal{H}(\mathbf{q}, \mathbf{p}) \quad \text{et} \quad \begin{cases} \dot{\mathbf{q}} = + \text{grad}_{\mathbf{p}} \mathcal{H} \\ \dot{\mathbf{p}} = - \text{grad}_{\mathbf{q}} \mathcal{H} \end{cases} \quad (1.12)$$

dans la structure des équations de la toute nouvelle mécanique quantique.

<sup>8</sup> voir par exemple en haut de la page 4 du second tome de la seconde édition de la mécanique analytique.

<sup>9</sup> Adrien-Marie Legendre remplace en 1812 Joseph-Louis Lagrange au Bureau des longitudes...

Nous reviendrons au chapitre 3 sur les détails d'une telle transformation mais restons pour le moment dans les délices de l'histoire. Remarquons tout de suite qu'en posant

$$\mathbf{Z} = \begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} \in \mathbb{R}^{2s} \text{ et } \mathbf{J} = \begin{bmatrix} 0 & +\text{Id}_{\mathbb{R}^s} \\ -\text{Id}_{\mathbb{R}^s} & 0 \end{bmatrix}$$

Les équations (1.12) malencontreusement dénommées «de Hamilton», deviennent

$$\dot{\mathbf{Z}} = \mathbf{J} \text{grad}_{\mathbf{Z}} \mathcal{H} \quad (1.13)$$

où l'on ne peut que reconnaître la forme générale du premier ordre généralisant à tout problème de mécanique les équations planétaires de Lagrange (1.9). Tout y est : la matrice antisymétrique  $\mathbf{L}^{-1}$  est devenue  $\mathbf{J}$ , la fonction perturbatrice  $R$  que Lagrange note  $\Omega$  devient le malheureux hamiltonien  $\mathcal{H}$ , et  $2s$  correspond bien au double des trois degrés de liberté de la Lune dans l'espace. Lagrange a donc bien mené son idée jusqu'au bout et généralisé l'idée de la variation des constantes à un problème quelconque de mécanique.

La dernière innovation de la variation des constantes lunaires non encore revisitée restait les fameuses parenthèses de Lagrange. Elles aussi ont été abordées par le vieux maître en collaboration avec son jeune élève Poisson. Stimulé par un article de Poisson publié en octobre 1808<sup>10</sup>, Lagrange publie en 1809<sup>11</sup> un remarquable mémoire dans lequel il donne la touche finale à son grand-œuvre. Il ne cherche ici absolument pas à s'approprier les travaux de son jeune élève, il tient cependant à clarifier ses «*calculs longs et délicats*» qui constituent une «*savante analyse qui est comme l'inverse de la mienne*». On notera l'utilisation de l'adjectif inverse chez Lagrange qui résume en un mot des développements compliqués pour l'époque : il avait tout compris !

Le formalisme des «crochets de Poisson» constitue le point final de la mécanique analytique classique, il est également l'ingrédient essentiel qui permet la généralisation vers la théorie des champs classique puis la physique théorique moderne, rien que cela !

La matrice de Lagrange  $\mathbf{L}$  fait apparaître des parenthèses antisymétriques dont les ingrédients sont des  $\frac{\partial \vec{r}}{\partial C_j}$  et  $\frac{\partial \vec{r}}{\partial C_i}$ , mais comme nous l'avons vu ce n'est pas cette matrice qui donne les équations du mouvement mais son inverse  $\mathbf{L}^{-1}$ . La prise en compte des nouvelles variables du formalisme d'ordre 1 dit

<sup>10</sup> 15<sup>e</sup> Cahier du Journal de l'Ecole Polytechnique, page 266.

<sup>11</sup> Voir par exemple les œuvres complètes de Lagrange, Tome 6, *Second mémoire sur la théorie de la variation des constantes arbitraires dans les problèmes de mécanique, dans lequel on simplifie l'application des formules générales à ces problèmes*, pages 809-816

d'Hamilton qui veut que  $\mathbf{C}$  devient  $\mathcal{H}$ , les  $\vec{r}_k$  deviennent  $\mathbf{q}$ , les  $\vec{r}_k$  deviennent  $\mathbf{p}$  nous fait entrevoir que la généralisation de  $\mathbf{L}^{-1}$  fera intervenir des ingrédients de la forme  $\text{grad}_{\mathbf{q}} \mathcal{H}$  et  $\text{grad}_{\mathbf{p}} \mathcal{H}$ ... C'est exactement le programme proposé par les deux articles de l'élève et du maître que l'on peut résumer en quelques lignes.

Si l'on considère une fonction quelconque  $f = f(\mathbf{q}, \mathbf{p})$ , on a immédiatement

$$\dot{f} = \dot{\mathbf{q}} \cdot \text{grad}_{\mathbf{q}} f + \dot{\mathbf{p}} \cdot \text{grad}_{\mathbf{p}} f$$

en utilisant les équations de Lagrange du premier ordre (1.12) – dites de Hamilton – il vient

$$\begin{aligned} \dot{f} &= \text{grad}_{\mathbf{p}} \mathcal{H} \cdot \text{grad}_{\mathbf{q}} f - \text{grad}_{\mathbf{q}} \mathcal{H} \cdot \text{grad}_{\mathbf{p}} f \\ &= \{f, H\} \end{aligned}$$

Dans les articles de 1808-1809 à la place du crochet  $\{, \}$  moderne on utilise des  $[, ]$  mais peu importe tout est là et sous la forme annoncée ! Il ne restera plus qu'à comprendre la structure d'algèbre de Lie sous-jacente et à Poincaré d'apporter la théorie des systèmes dynamiques pour tirer tout le parti d'un tel formalisme. Dans tous les cas on ne peut que remercier Joseph-Louis Lagrange pour avoir à ce point transformé la physique en essayant de comprendre le mouvement de notre Lune.

## 1.5 Compléments historiques sur la vie, l'œuvre de Joseph Louis Lagrange et son contexte.

Joseph-Louis Lagrange est né à Turin, dans ce que l'on appelait à l'époque le royaume de Piémont-Sardaigne, le 25 janvier 1736. Ses études locales le conduisirent à un poste de professeur de mathématiques à l'âge de seize ans à l'école royale d'artillerie de cette même ville. En se rassemblant avec ses meilleurs élèves, il fonde alors ce qui allait devenir l'académie de Turin. Le premier travail important de Lagrange concerne la propagation du son. Suite aux travaux de d'Alembert, qui propose en 1747 une équation pour modéliser les vibrations d'une corde, (Daniel) Bernoulli propose que la solution de cette équation doit pouvoir s'obtenir comme la somme de fonctions trigonométriques de fréquences multiples, sans autre justification que son intuition. Ce genre de démarche déplaît au jeune Lagrange qui entreprend de démontrer rigoureusement une telle affirmation. Il publie entre 1759 et 1761 une série de quatre articles sur la nature et la propagation du son

dans laquelle il pose les fondements de ce qui deviendra un demi-siècle plus tard ni plus ni moins que l'analyse de Fourier. Il publie<sup>12</sup> également entre 1762 et 1765 la méthode de la variation des constantes pour les équations différentielles linéaires que nous avons exposée plus haut. Ces travaux et un autre sur la recherche d'extremum de fonctions parviennent à Euler qui règne en maître sur les mathématiques de l'époque et qui entreprend une relation épistolaire suivie avec son jeune collègue. Au XVIII<sup>e</sup> siècle, on faisait des mathématiques par lettre. On peut retrouver dans une de ces lettres, datée du 2 octobre 1759, la grande estime que lui porte très tôt Euler : *« Votre solution du problème des isopérimètres ne laisse rien à désirer, et je me réjouis que ce sujet, dont je m'étais presque seul occupé depuis les premières tentatives, ait été porté par vous au plus haut degré de perfection. L'importance de la matière m'a excité à en tracer, à l'aide de vos lumières, une solution analytique à laquelle je ne donnerai aucune publicité jusqu'à ce que vous même ayez publié la suite de vos recherches pour ne vous enlever aucune partie de la gloire qui vous est due. »* Lagrange n'avait pas vingt quatre ans et Euler qui *« tenait le sceptre des mathématiques »* sait reconnaître son successeur. Le jeune homme est lancé et ses travaux se diversifient : théorie des nombres, analyse, mécanique des fluides, théorie des fonctions et bien sûr mécanique céleste pour laquelle ses travaux sur les libérations de la Lune sont récompensés par le Grand Prix de l'Académie des Sciences de Paris en 1764. Turin devient alors trop petite... Euler qui cherche à quitter Berlin pour rejoindre Saint-Petersbourg et d'Alembert qui a les faveurs du Roi de Prusse proposent Lagrange pour le poste de directeur de l'académie de Prusse pour les sciences physico-mathématiques. Il y prend fonction le 6 novembre 1766, il vient d'avoir 30 ans. Commence alors une période de production scientifique d'une intensité et d'une variété rare. Il travaille sans relâche tous les jours pendant douze heures sans interruption, il ne reçoit personne tout en occupant un poste d'une grande responsabilité. On le laissait sans doute tout à son art sans l'encombrer de tâches administratives. C'est pendant la période berlinoise qu'il bâtit sa mécanique analytique, en mélangeant toutes ses idées sur la variation des constantes, la théorie de la Lune et le mouvement des corps en mécanique. Louis XVI le veut à Paris et charge Mirabeau de lui en faire la proposition : une position égale, un salaire équivalent et la publication de son livre *« La mécanique analitique »* aux frais du Roi. Lagrange se renseigne sur la stabilité et la vie dans ce pays que son bisaïeul, capitaine de cavalerie au service du Roi de France,

---

<sup>12</sup> Solution de différents problèmes de calcul intégral, p.471-668, Miscellanea Taurinensia, t. III, 1762-1765

avait servi et accepte la proposition. L'accueil parisien fut à la hauteur du personnage, on peut s'en faire une idée en lisant l'éloge de Delambre : « *Ce fut en 1787 que M. Lagrange vint à Paris, siéger à l'Académie des Sciences, dont il était depuis quinze ans associé étranger. Pour lui donner droit au suffrage on changea ce titre en celui de pensionnaire vétérane. Ses nouveaux confrères se montrèrent à l'envi heureux et glorieux de le posséder; la Reine l'accueillit avec bienveillance; elle le considérait comme un allemand; il lui avait été recommandé de Vienne. On lui donna un logement au Louvre; il y vécut heureux jusqu'à la révolution.* »

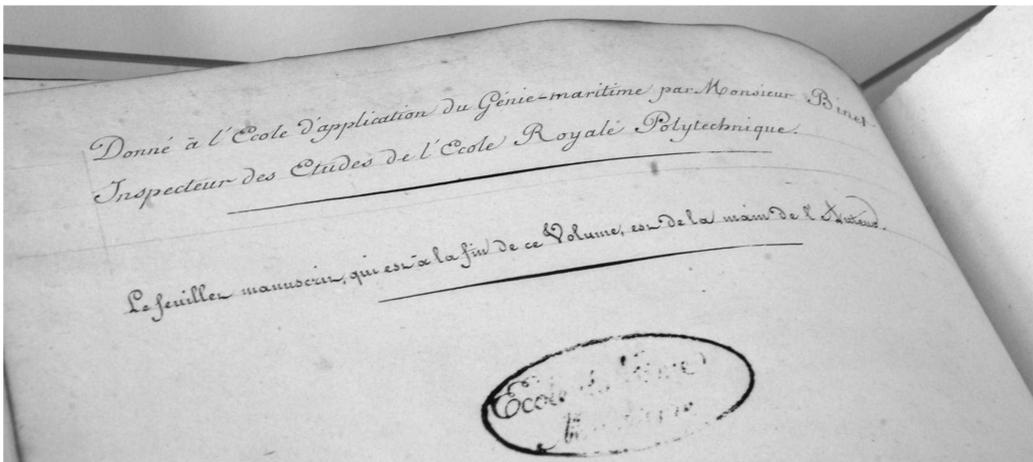
Bien entendu la révolution vint tout changer, et l'on aurait pu s'inquiéter du sort qu'elle aurait pu réserver à un étranger, protégé du Roi et habitant au Louvre... Mais il ne fut jamais inquiété sans doute protégé par son aura. Lagrange lui-même aurait pu fuir pour trouver asile ailleurs en Europe où sa notoriété lui aurait assuré une place éminente. Il choisit de rester car il fut impressionné par « *ce peuple qui prend en main les rennes de son destin* » et il décida de s'investir dans la construction d'un nouvel ordre : il devint révolutionnaire à cinquante ans passés ! Sa contribution est fondamentale et nombreux sont ceux qui l'ont oublié. Elle se situe dans deux domaines : l'enseignement et les unités de mesure. Sur ce dernier point, on doit reconnaître en Lagrange l'un des principaux promoteur et constructeur du système métrique décimal. A la suppression des académies et aux moments les plus troubles de la révolution, il fut ainsi nommé Président de la commission chargée de l'élaboration du nouveau système des poids et mesures quand Lavoisier, Coulomb, Borda, Laplace et bien d'autres furent évincés ou pire. Il mena sa tâche à terme et c'est peu de dire que le monde lui doit beaucoup tant le système métrique décimal révolutionnaire non seulement la métrologie mais la science toute entière.

Sur le plan pédagogique son œuvre est encore plus grande. C'est tout simplement sur ses conseils que l'on créera l'Ecole Polytechnique pour doter la république de savants et d'ingénieurs et l'Ecole Normale pour assurer la formation de professeurs susceptibles d'éduquer le peuple. Il s'investi même dans le fonctionnement de ces deux écoles en y devenant le professeur de mathématiques dès leurs créations en 1794. Face à cette charge d'enseignant, nouvelle pour lui et qu'il n'avait qu'effleuré à Turin, il composa une série de cours rassemblés dans sa *Théorie des fonctions analytiques* et ses *Leçons sur le calcul des fonctions* qui initient une méthode nouvelle de l'enseignement des mathématiques. En les analysant en détail, on peut même aller jusqu'à dire qu'il invente une nouvelle façon d'enseigner les mathématiques qui pose

les fondations de ce qui allait devenir la «tradition de l'École Française de Mathématique» toujours vivace deux siècles plus tard...

Vers la fin de sa vie, stimulé par l'un de ses meilleurs élèves, Siméon Denis Poisson, il entreprendra d'écrire une nouvelle édition de sa mécanique analytique «*revue et augmentée par l'auteur*». Le premier tome fut terminé et publié en 1811; mais comme on peut le lire dans son avertissement «*La publication de ce deuxième volume de la Mécanique Analytique a éprouvé un retard dont nous allons exposer les principaux motifs. M. Lagrange en avait déjà fait imprimer les premières feuilles lorsque la mort l'enleva aux sciences.*» le 8 avril 1813. C'est à l'un de ses élèves de l'École Polytechnique, Jacques-Philippe Binet, que l'on confia la lourde tâche de terminer la publication de ce second tome à partir des manuscrits du Maître. C'est par cette anecdote qu'indirectement j'ai pu prendre connaissance et conscience de la grandeur de l'Œuvre de Lagrange et de son importance dans l'histoire des sciences toute entière. Je terminerai ce chapitre par quelques explications à ce sujet.

C'est en fouillant dans les archives de l'École Nationale Supérieure de Techniques Avancées au tournant du second millénaire que j'ai découvert un très vieux livre d'un auteur dont je connaissais le nom, Lagrange, mais pas encore l'ampleur du legs. En l'ouvrant, je tombais sur une note manuscrite portée sur la seconde de couverture et dont le fac-similé est présenté sur la figure 1.3.



**Fig. 1.3.** Seconde de couverture de l'exemplaire de la seconde édition de la mécanique analytique du fond ancien de l'Ensta-Paris

Me reportant à la fin de l'ouvrage comme l'indique la note, je tombais effectivement sur un manuscrit de la main de l'auteur inséré à la reliure dans l'ouvrage, comme on peut s'en rendre compte sur la figure 1.5.

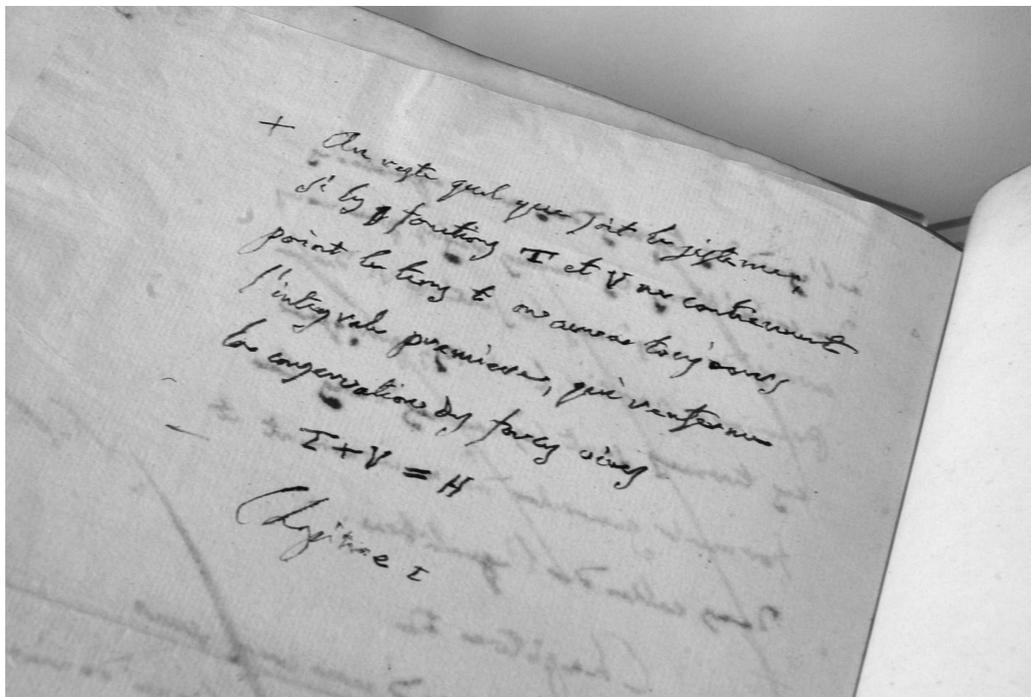


Fig. 1.4. Détail du manuscrit original inséré dans la couverture de l'ouvrage

C'est en lisant ce manuscrit (voir figure 1.4) que je compris que Lagrange avait écrit  $H = T + V$  alors qu'Hamilton n'avait que cinq ans et que j'enseignais ce «hamiltonien» depuis quelques années déjà dans mes différents cours ! Je voulus en savoir d'avantage sur Joseph-Louis Lagrange et ce que j'ai appris est reporté partiellement dans ce chapitre. Cette histoire m'amena également au Panthéon<sup>13</sup> de Paris pour commémorer le bicentenaire de sa disparition en 2013 et me fit commissaire d'une grande exposition qui anima le grand hall de l'Ensta pendant trois mois au printemps 2014.

Comment ce livre était-il arrivé à l'Ensta ? Les détours de l'histoire sont parfois simples si l'on ne cherche pas trop de complications. Alors que Jacques Philippe Binet terminait la rédaction du second tome de la *Mécanique Analytique* sur la base des manuscrits de son maître, Napoléon Bona-

<sup>13</sup> Voir [https://www.youtube.com/watch?v=fHCbdl\\_SI-M](https://www.youtube.com/watch?v=fHCbdl_SI-M)

parte périlait et la restauration s'installait en France. Dans ses premières mesures, le nouveau roi Louis XVIII redistribua le pouvoir, l'Ecole Impériale Polytechnique devint l'Ecole Royale Polytechnique et son nouveau directeur des études fut Jacques Philippe Binet... L'une des nouvelles missions<sup>14</sup> de cette Ecole était d'accueillir dans le giron de ses écoles d'applications l'Ecole du Génie Maritime, qui deviendra l'Ensta en 1970. L'Ecole du Génie Maritime héritière de l'Ecole des vaisseaux royaux avait en effet été ostracisée à la révolution pour ses liens trop étroits avec le Roi et c'est donc un retour en grâce imposé par Louis XVIII à la restauration. L'interprétation la plus simple est alors évidente. Jacques Philippe Binet terminant la nouvelle édition de la *Mécanique Analytique* tint à faire un cadeau de marque au directeur de cette ancienne école de retour en grâce. Il fit réaliser un exemplaire de prestige<sup>15</sup> contenant dans la reliure l'un des manuscrits les plus représentatifs du génie de son maître en sa possession. Le cadeau fut sans doute remis au directeur de l'Ecole du Génie Maritime entre 1817 et 1820 par Jacques Philippe Binet dont c'est bien l'écriture tout comme celle de Lagrange dans le manuscrit. Ce directeur s'empressa d'enfermer le livre dans les archives de son Ecole là où j'ai eu le bonheur de l'y retrouver presque deux siècles plus tard<sup>16</sup>...

La dernière question que l'on peut se poser concerne l'oubli dans lequel est tombé Joseph-Louis Lagrange. Que ce soit parmi les mathématiciens ou les physiciens bien peu se rendent compte de l'importance de ce personnage dans notre histoire scientifique et même sociale comme nous l'avons évoqué. Et dans le grand public, où il est largement confondu avec *Léo* Lagrange, un ministre des sports du front populaire...

---

<sup>14</sup> Voir l'almanach royal pour l'année 1817 page 615 sur la nouvelle organisation de l'Ecole Royale Polytechnique.

<sup>15</sup> Nous n'en connaissons pas d'autre !

<sup>16</sup> Pour voir ce fameux livre, on peut également retrouver un petit reportage sur «*H* ou le trésor de l'ENSTA» sur le site : <https://www.youtube.com/watch?v=T1MwsZsS1mg>

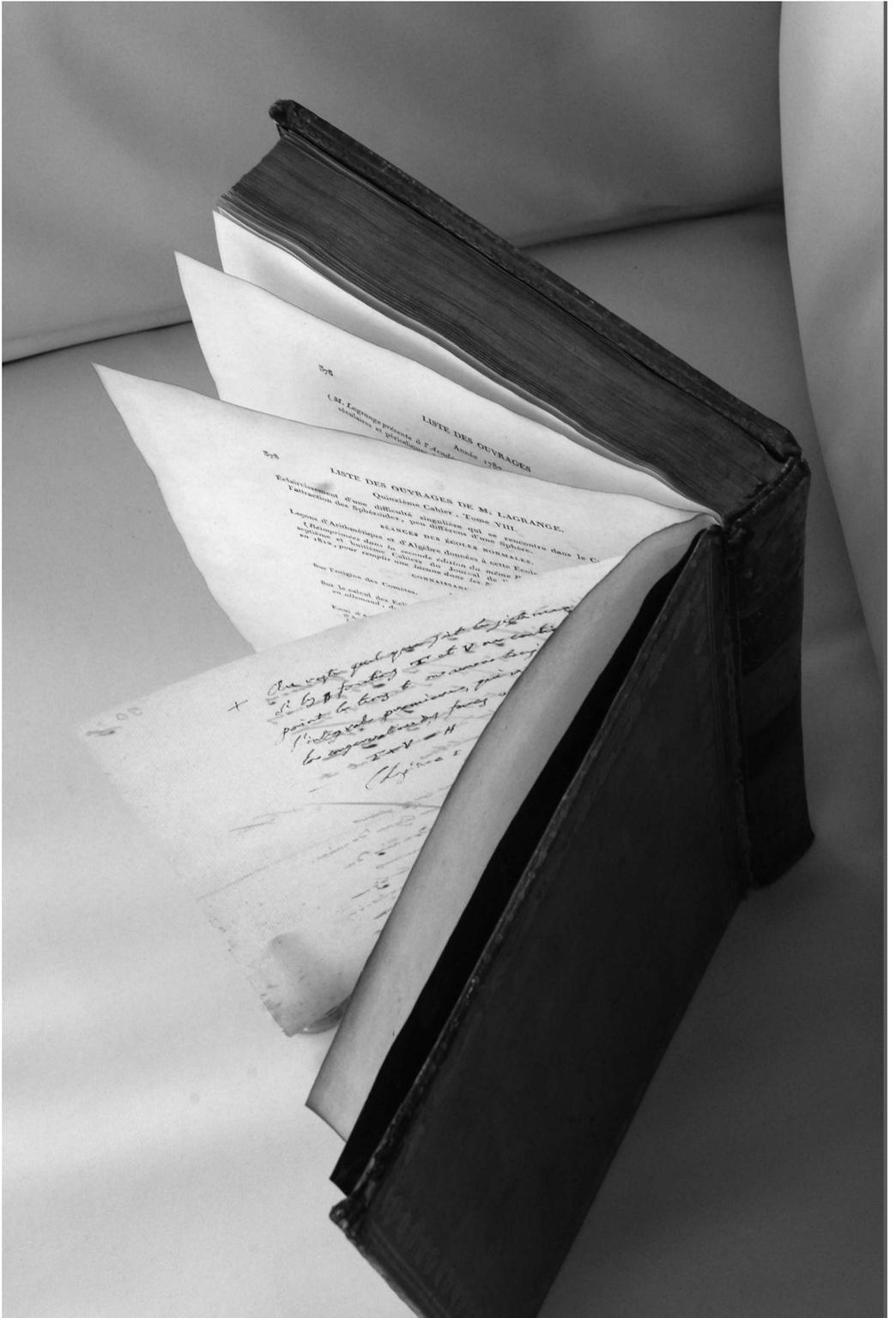


Fig. 1.5. Le manuscrit original inséré dans la couverture de l'ouvrage



## Formulation lagrangienne

### 2.1 Coordonnées généralisées

En physique classique, un point matériel est repéré dans l'espace par un vecteur position  $\vec{r}$  de  $\mathbb{R}^3$ . Ce vecteur dépend généralement du temps  $t$ . Le choix d'une base de  $\mathbb{R}^3$  permet d'attribuer des coordonnées, ou composantes, au vecteur position. Si cette base est indépendante du temps, on parlera alors de repère et de coordonnées cartésiennes.<sup>1</sup>

La dérivée totale du vecteur position par rapport au temps définit, à chaque instant  $t$ , le vecteur vitesse instantanée

$$\vec{v} := \frac{d\vec{r}}{dt} \quad (2.1)$$

La dérivée seconde totale du vecteur position par rapport au temps définit, à chaque instant  $t$ , le vecteur accélération instantanée

$$\vec{a} := \frac{d^2\vec{r}}{dt^2} \quad (2.2)$$

Un système de  $N$  points matériels est repéré dans  $\mathbb{R}^3$  par  $N$  vecteurs positions  $\vec{r}_{k=1,\dots,N}$ , c'est-à-dire un ensemble de  $3N$  coordonnées cartésiennes. Le nombre  $s$  de grandeurs indépendantes nécessaires à la connaissance complète de la position du système, est en général inférieur ou égal à  $3N$  : en effet, si le système contient des liaisons ou des contraintes internes ou bien si certaines symétries conduisent à un système de coordonnées autre que cartésien on a en général

$$s = 3N - f \quad (2.3)$$

---

<sup>1</sup> En fait les véritables coordonnées cartésiennes sont moins générales que cela, mais cette ex-

$f$  étant le nombre de ces contraintes ou liaisons *indépendantes* entre les coordonnées. Le nombre  $s$  est appelé nombre de degrés de liberté du système. On définit alors des coordonnées généralisées  $q_i$ , pour chaque  $i = 1, \dots, s$ . Pour avoir des formules plus compactes, on utilisera souvent le vecteur  $\mathbf{q} \in \mathbb{R}^s$  et ses dérivées. Les coordonnées généralisées sont par définition indépendantes

$$\forall i \neq j = 1, \dots, s \quad \frac{\partial q_i}{\partial q_j} = 0 \quad (2.4)$$

Les  $N$  vecteurs positions sont des fonctions bijectives des coordonnées généralisées

$$\forall k = 1, \dots, N \quad \vec{r}_k = \vec{r}_k(q_1, \dots, q_s) = \vec{r}_k(\mathbf{q}) \quad (2.5)$$

Les dérivées totales par rapport au temps des coordonnées généralisées définissent les vitesses généralisées, notées  $\dot{q}_i$

$$\forall i = 1, \dots, s \quad \dot{q}_i = \frac{dq_i}{dt} . \quad (2.6)$$

L'expérience montre que la donnée simultanée des coordonnées et des vitesses détermine complètement l'état du système et permet de prévoir son mouvement futur. Ainsi la donnée d'une équation différentielle du second ordre de la forme

$$F_i(\ddot{\mathbf{q}}, \dot{\mathbf{q}}, \mathbf{q}, t) = 0 \quad \text{avec } \mathbf{q} \in \mathbb{R}^s \quad (2.7)$$

pour chaque degré de liberté  $1 \leq i \leq s$  assortie d'un jeu de conditions initiales, doit s'avérer suffisante pour décrire le mouvement dans sa totalité. Une telle relation est appelée équation du mouvement.

## 2.2 Principe de moindre action

Lors du premier chapitre historico-technique nous avons vu comment les équations de Lagrange et le principe de moindre action trouvaient leurs origines. De façon moins déductive et en l'état actuel de la situation nous pouvons toujours émettre le vœu pieux suivant :

*Tout système mécanique est caractérisé par une fonction scalaire  $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$ . Si à 2 instants  $t_1$  et  $t_2$  le système occupe 2 positions déterminées par 2 ensembles de valeurs de coordonnées généralisées  $\{\mathbf{q}_{(1)}\}$  et  $\{\mathbf{q}_{(2)}\}$ , entre ces deux positions le système se meut de telle façon que l'intégrale*

$$S := \int_{t_1}^{t_2} \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) dt \quad (2.8)$$

ait la plus petite valeur possible.

La fonction  $\mathcal{L}$  est appelée fonction de Lagrange, ou lagrangien, la quantité  $S$  est appelée action du système.

Outre le temps, le lagrangien ne dépend que des  $q_i$  et des  $\dot{q}_i$  et non des dérivées temporelles d'ordre supérieur, seules ces deux premières quantités définissent l'état dynamique du système.

### 2.2.1 Quelles équations pour $\mathcal{L}$ ?

Le lagrangien est une fonction, un champ scalaire en fait, de la forme  $\mathcal{L} = \mathcal{L}(\dot{\mathbf{q}}, \mathbf{q}, t)$ . Supposons qu'il existe un vecteur  $\tilde{\mathbf{q}} = \tilde{\mathbf{q}}(t)$  qui minimise l'action  $S$ . Tout choix de  $\mathbf{q}(t) = \tilde{\mathbf{q}}(t) + \delta\mathbf{q}(t)$  implique un accroissement de  $S$ . Aux instants  $t_1$  et  $t_2$  le vecteur  $\mathbf{q}$  est connu et a des valeurs  $\mathbf{q}(t_1)$  et  $\mathbf{q}(t_2)$ , tous les choix de  $\mathbf{q}$  doivent conduire à ces valeurs, ainsi

$$\delta\mathbf{q}(t_1) = \delta\mathbf{q}(t_2) = \mathbf{0} \quad (2.9)$$

La variation  $\mathbf{q} \rightarrow \mathbf{q} + \delta\mathbf{q}$  se traduit par la variation

$$\delta S = \int_{t_1}^{t_2} \mathcal{L}(\mathbf{q} + \delta\mathbf{q}, \dot{\mathbf{q}} + \delta\dot{\mathbf{q}}, t) dt - \int_{t_1}^{t_2} \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) dt \quad (2.10)$$

qui s'écrit au premier ordre

$$\delta S = \int_{t_1}^{t_2} [\text{grad}_{\mathbf{q}}(\mathcal{L}) \cdot \delta\mathbf{q} + \text{grad}_{\dot{\mathbf{q}}}(\mathcal{L}) \cdot \delta\dot{\mathbf{q}}] dt \quad (2.11)$$

il est toujours possible de permuter les variations  $d$  et  $\delta$  ainsi  $\delta\dot{\mathbf{q}} = \frac{d}{dt}(\delta\mathbf{q})$  et il vient

$$\delta S = [\text{grad}_{\dot{\mathbf{q}}}(\mathcal{L}) \cdot \delta\mathbf{q}]_{t_1}^{t_2} + \int_{t_1}^{t_2} \left[ \text{grad}_{\mathbf{q}}(\mathcal{L}) - \frac{d}{dt}(\text{grad}_{\dot{\mathbf{q}}}(\mathcal{L})) \right] \cdot \delta\mathbf{q} dt \quad (2.12)$$

en vertu des conditions aux limites (2.9), le terme tout intégré disparaît et la condition nécessaire de minimum pour  $S$  s'écrit

$$\delta S = \int_{t_1}^{t_2} \left[ \text{grad}_{\mathbf{q}} (\mathcal{L}) - \frac{d}{dt} (\text{grad}_{\dot{\mathbf{q}}} (\mathcal{L})) \right] \cdot \delta \mathbf{q} dt = 0 \quad (2.13)$$

Cette variation de l'action s'écrit comme un produit scalaire entre deux vecteurs dont l'un est la variation  $\delta \mathbf{q}$ . Cette dernière est quelconque ce qui implique la nullité de toutes les composantes de l'autre terme. En conséquence, si le vecteur  $\tilde{\mathbf{q}}(t)$  rend l'action minimale, ses composantes  $\tilde{q}_i(t)$  pour  $i = 1, \dots, s$  doivent vérifier

$$\forall i = 1, \dots, s \quad \frac{\partial \mathcal{L}}{\partial \tilde{q}_i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \left( \frac{d\tilde{q}_i}{dt} \right)} = 0 \quad (2.14)$$

Réciproquement, pour trouver  $\tilde{\mathbf{q}}$  connaissant  $\mathcal{L}$ , il «suffit» de résoudre l'équation (2.14), qui est bien, comme annoncé, une équation différentielle du second ordre.

Les  $s$  équations du mouvement d'un système possédant  $s$  degrés de liberté définissant chacun une coordonnée généralisée  $q_i$  pour  $i = 1, \dots, s$  sont de la forme

$$\forall i = 1, \dots, s \quad \frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = 0 \quad (2.15)$$

Elles sont appelées *équations de Lagrange*.

Les solutions de ces équations du mouvement contiendront chacune deux constantes d'intégration qu'il conviendra de mettre en correspondance avec le jeu de conditions initiales décrivant le système.

Les équations de Lagrange imposent que le mouvement corresponde à un extremum de l'action. Le fait que celui-ci soit un minimum reste à vérifier et imposera le signe de certaines constantes.

## 2.2.2 Propriétés du lagrangien

### Relativité

Pour des raisons évidentes de simplicité et d'efficacité, la fonction de Lagrange doit être la même dans tous les référentiels.

### Additivité

Soit un système mécanique composé de deux parties  $A$  et  $B$  isolées. Chacune des parties d'un système n'interagissant pas avec l'autre, les équations du

mouvement de la partie  $A$  (resp.  $B$ ) ne peuvent pas contenir des grandeurs se rapportant à la partie  $B$  (resp.  $A$ ). La fonction de Lagrange globale doit donc vérifier  $\mathcal{L} = \mathcal{L}_A + \mathcal{L}_B$ .

### Multiplication par une constante

Il est clair que la relation  $\delta S = 0$  laisse la possibilité de multiplier  $\mathcal{L}$  par une constante arbitraire sans changer les équations du mouvement.

Il n'en résulte cependant pas d'indétermination.

La contrainte d'additivité permet de voir que le choix de la constante multiplicative doit être global et correspond donc à un choix des unités de mesure des grandeurs physiques intervenant dans le problème.

### Choix de jauge

La forme de l'action  $S$  impose la remarque suivante : considérons deux fonctions de Lagrange  $\mathcal{L}(\dot{\mathbf{q}}, \mathbf{q}, t)$  et  $\mathcal{L}'(\dot{\mathbf{q}}, \mathbf{q}, t)$  ne différant l'une de l'autre que par la dérivée totale par rapport au temps d'un champ scalaire quelconque  $\varphi(\mathbf{q}, t)$

$$\mathcal{L}' = \mathcal{L} + \frac{d\varphi}{dt} \quad (2.16)$$

nous avons donc

$$\begin{aligned} S' &= \int_{t_1}^{t_2} \mathcal{L}'(\mathbf{q}, \dot{\mathbf{q}}, t) dt = \int_{t_1}^{t_2} \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) dt + \int_{t_1}^{t_2} \frac{d\varphi}{dt} dt \\ &= S + [\varphi]_{t_1}^{t_2} = S - \varphi(\mathbf{q}(t_1), t_1) + \varphi(\mathbf{q}(t_2), t_2) \\ &= S + cste \end{aligned} \quad (2.17)$$

ainsi la condition  $\delta S' = 0$  est équivalente à la condition  $\delta S = 0$ . Les équations de Lagrange fournies par  $\mathcal{L}$  sont donc équivalentes à celles fournies par  $\mathcal{L}'$ .

La fonction de Lagrange d'un système est donc définie à l'addition près d'une dérivée totale par rapport au temps d'un champ scalaire arbitraire ne dépendant que des coordonnées généralisées  $\mathbf{q}$  et du temps. Cette liberté est appelée liberté de jauge, elle joue un rôle fondamental dans de nombreuses théories.

## 2.3 Principe de relativité

Pour étudier les propriétés d'un système, il convient de choisir un système de référence.

Par rapport à certains systèmes de référence, l'espace est inhomogène et anisotrope, le temps lui-même peut ne pas être uniforme : des instants différents peuvent ne pas être équivalents.

L'hypothèse de Galilée est de supposer l'existence d'un système de référence dans lequel l'espace est homogène et isotrope et le temps s'écoule uniformément. En particulier, dans un tel référentiel dit galiléen, un corps libre et au repos à un instant donné restera au repos pendant un temps illimité.

Quelle est la forme de la fonction de Lagrange d'un corps libre au repos dans un tel référentiel galiléen ?

- L'homogénéité de l'espace et l'uniformité du temps impose au lagrangien de ne pas dépendre de  $\vec{r}$  ni de  $t$ , on doit donc avoir  $\mathcal{L} = \mathcal{L}(\vec{v})$ .
- L'isotropie de l'espace impose au lagrangien d'être fonction d'une variable isotrope, ainsi tenant compte de la précédente propriété on doit avoir  $\mathcal{L} = \mathcal{L}(|\vec{v}|) = \mathcal{L}(\vec{v}^2)$

La prise en compte de ces contraintes pour le lagrangien permet d'écrire les équations de Lagrange en coordonnées cartésiennes pour

$$\forall i = 1, 2, 3 \quad \frac{\partial \mathcal{L}}{\partial r_i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial v_i} = - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial v_i} = 0 \quad (2.18)$$

ainsi

$$\forall i = 1, 2, 3 \quad \frac{\partial \mathcal{L}}{\partial v_i} = cste \quad (2.19)$$

compte-tenu du fait que le lagrangien ne dépend que du module de la vitesse, on a donc

$$\vec{v} = \overrightarrow{cste} \quad (2.20)$$

Dans un référentiel galiléen tout mouvement libre s'effectue à vitesse constante en grandeur et en direction. C'est le principe d'inertie qui n'est donc plus un axiome, mais une conséquence de notre vœu pieux et des propriétés de symétrie de l'espace et du temps évoquées plus haut.

Conséquence supplémentaire, tout référentiel  $\mathcal{R}$  en translation uniforme à vitesse constante  $\vec{V}$  par rapport à un référentiel galiléen  $\mathcal{R}'$  est galiléen.

---

tension profite à notre exposé sans lui nuire d'aucune façon.

Il s'agit là du principe d'inertie de Galilée, associé à la transformation de Galilée

$$\begin{cases} \vec{r} = \vec{r}' + \vec{V}t \\ t = t' \end{cases} \quad (2.21)$$

## 2.4 Détermination de la fonction de Lagrange

### 2.4.1 Particule libre

Soient deux référentiels galiléens  $\mathcal{R}$  et  $\mathcal{R}'$ , le principe de relativité impose que  $\mathcal{R}'$  est en translation uniforme à vitesse constante  $\vec{\varepsilon}$  par rapport à  $\mathcal{R}$ .

Dans le cas d'une particule libre, nous avons vu que le lagrangien devait être de la forme  $\mathcal{L} = f(\vec{v}^2)$  dans  $\mathcal{R}$  ou  $\mathcal{L}' = f(\vec{v}'^2)$  dans  $\mathcal{R}'$ , la fonction  $f$  devant être la même dans tous les référentiels. Les équations du mouvement doivent être équivalentes dans  $\mathcal{R}$  et  $\mathcal{R}'$ , ainsi en vertu de la liberté de jauge  $\mathcal{L}$  et  $\mathcal{L}'$  ne peuvent différer que par la dérivée totale par rapport au temps d'une fonction de  $\varphi(\vec{r}, t)$ . Mais

$$\mathcal{L}' = f(\vec{v}'^2) = f([\vec{v} + \vec{\varepsilon}]^2) = f(\vec{v}^2 + 2\vec{v} \cdot \vec{\varepsilon} + \vec{\varepsilon}^2) \quad (2.22)$$

En considérant dans un premier temps que  $|\vec{\varepsilon}|$  est petit devant  $|\vec{v}|$ , un développement au premier ordre en  $\vec{\varepsilon}$  donne

$$\mathcal{L}' = \mathcal{L} + 2 \frac{df}{d\vec{v}^2} \vec{v} \cdot \vec{\varepsilon} + O(\vec{\varepsilon}^2) = \mathcal{L} + 2 \frac{df}{d\vec{v}^2} \frac{d\vec{r}}{dt} \cdot \vec{\varepsilon} + O(\vec{\varepsilon}^2) \quad (2.23)$$

pour entrer dans la liberté de jauge, la solution la plus simple<sup>2</sup> est de choisir  $f$  linéaire en  $\vec{v}^2$ , c'est-à-dire  $f = \mu \vec{v}^2$ . Le vecteur  $\vec{\varepsilon}$  étant constant, le choix  $\varphi(\vec{r}, t) = 2\mu \vec{r} \cdot \vec{\varepsilon}$  permet d'écrire  $\mathcal{L}' = \mathcal{L} + d\varphi/dt$ . Le coefficient  $\mu$  doit être positif afin que l'action soit minimale pour le mouvement, nous poserons  $\mu = m/2$  où le coefficient positif  $m$  est appelé masse de la particule.

Le lagrangien d'une assemblée de  $N$  particules libres est donc son énergie cinétique totale  $T$

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^N m \vec{v}_i^2 = T \quad (2.24)$$

A posteriori on peut vérifier que si  $|\vec{\varepsilon}|$  n'est plus petit devant  $|\vec{v}|$  le résultat demeure : en effet, si  $\mathcal{L} = \mu \vec{v}^2$  alors

<sup>2</sup> Il y a d'autres possibilités mais elles sont beaucoup plus complexes, essayons dans un premier

$$\mathcal{L}' = \mu(\vec{v} + \vec{\varepsilon}')^2 = \mu\vec{v}^2 + 2\mu\vec{v} \cdot \vec{\varepsilon}' + \mu\vec{\varepsilon}'^2 \quad (2.25)$$

il suffit alors de prendre  $\varphi(\vec{r}, t) = 2\mu\vec{r} \cdot \vec{\varepsilon}' + \mu t\vec{\varepsilon}'^2$  pour que  $\mathcal{L}$  et  $\mathcal{L}'$  ne diffèrent que d'une jauge.

Ce raisonnement est assez intuitif, il repose sur un «principe» de simplicité. Il illustre très bien la démarche à suivre pour la recherche d'un lagrangien : respecter les grandes propriétés de symétrie, de relativité correspondant au contexte, et en cas d'ambiguïté choisir la solution la plus simple. Nous citerons à ce sujet Albert Einstein qui disait «*Dieu a écrit la physique dans le langage complexe des mathématiques, mais il n'est pas méchant !*»

## 2.4.2 Systèmes conservatifs

Considérons une assemblée de  $N$  particules en interaction repérées dans un référentiel galiléen par  $k = 1, \dots, N$  vecteurs positions  $\vec{r}_k$ .

Sans information supplémentaire, le nombre de degrés de liberté de ce système est  $s = 3N$  car chaque composante de chaque  $\vec{r}_k$  est un degré de liberté<sup>3</sup>.

Si l'on considère une base cartésienne  $C = (\vec{e}_x, \vec{e}_y, \vec{e}_z)$  pour représenter les vecteurs  $\vec{r}_k$ , les  $s$  coordonnées généralisées seront alors  $q_1 = \vec{r}_1 \cdot \vec{e}_x$ ,  $q_2 = \vec{r}_1 \cdot \vec{e}_y$ ,  $q_3 = \vec{r}_1 \cdot \vec{e}_z$ ,  $q_4 = \vec{r}_2 \cdot \vec{e}_x$  et ainsi de suite jusqu'à  $q_{s=3N} = \vec{r}_N \cdot \vec{e}_z$ . Mais ce n'est qu'un exemple et le choix de la base est laissé à l'utilisateur qui devra donc avoir une certaine expérience en mécanique pour choisir la représentation la mieux adaptée à son système d'étude.

Nous avons vu qu'un système comportant  $s$  degrés de liberté et dont les particules sont libres de toute interaction admet pour lagrangien son énergie cinétique totale  $T$ . Les équations du mouvement pour un tel système sont donc

$$\forall \ell, 1 \leq \ell \leq s \quad \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_\ell} \right) - \frac{\partial T}{\partial q_\ell} = 0 \quad (2.26)$$

Calculons formellement chacun des deux termes présents dans chacune de ces équations. Par définition, l'énergie cinétique totale  $T$  d'un système est donnée par la relation

$$T = \frac{1}{2} \sum_{k=1}^N m_k \left( \frac{d\vec{r}_k}{dt} \right)^2 \quad (2.27)$$

temps de tester les solutions simples !

<sup>3</sup> L'entier  $s = 3N$  est un maximum car toute contrainte entre ces degrés de liberté comme des liaisons rigides, ou bien toutes les propriétés des interactions qui peuvent contraindre le mouvement, diminuent le nombre de degré de liberté.

Commençons par les vitesses généralisées, pour tout  $\ell$ , nous avons

$$\frac{\partial T}{\partial \dot{q}_\ell} = \sum_{k=1}^N m_k \frac{d\vec{r}_k}{dt} \cdot \frac{\partial}{\partial \dot{q}_\ell} \frac{d\vec{r}_k}{dt} \quad (2.28)$$

Une propriété fondamentale de la position est de ne dépendre que des coordonnées généralisées et pas des vitesses généralisées, ainsi pour les  $k = 1, \dots, N$  particules nous aurons  $\vec{r}_k = \vec{r}_k(\mathbf{q}, t)$ . Cette dépendance générale est d'ailleurs souvent très modeste : si par exemple les particules sont libres, en coordonnées cartésiennes nous aurons  $\vec{r}_1 = \vec{r}_1(q_1, q_2, q_3)$ ,  $\vec{r}_2 = \vec{r}_2(q_4, q_5, q_6)$ , etc. Néanmoins, qui peut le plus peut le moins et l'on peut écrire dans le cas général

$$\forall k = 1, \dots, N \quad \frac{d\vec{r}_k}{dt} = \sum_{i=1}^s \frac{\partial \vec{r}_k}{\partial q_i} \dot{q}_i$$

on peut à présent expliciter la dérivée par rapport à  $\dot{q}_\ell$  dans l'expression (2.28), il vient

$$\frac{\partial T}{\partial \dot{q}_\ell} = \sum_{k=1}^N m_k \frac{d\vec{r}_k}{dt} \cdot \sum_{i=1}^s \left( \dot{q}_i \frac{\partial^2 \vec{r}_k}{\partial \dot{q}_\ell \partial q_i} + \frac{\partial \vec{r}_k}{\partial q_i} \frac{\partial \dot{q}_i}{\partial \dot{q}_\ell} \right) \quad (2.29)$$

par hypothèse le vecteur  $\vec{r}_k$  ne dépend que des coordonnées généralisées, ainsi pour tout  $\ell$  nous aurons

$$\frac{\partial \vec{r}_k}{\partial \dot{q}_\ell} = 0 \quad \implies \quad \frac{\partial^2 \vec{r}_k}{\partial \dot{q}_\ell \partial q_i} = 0, \quad (2.30)$$

nous avons introduit les coordonnées et les vitesses généralisées de manière à rendre les quantités dynamiques indépendantes, nous avons donc

$$\frac{\partial \dot{q}_i}{\partial \dot{q}_\ell} = \delta_{i\ell} \quad , \quad (2.31)$$

la relation (2.29) se réduit finalement à

$$\frac{\partial T}{\partial \dot{q}_\ell} = \sum_{k=1}^N m_k \frac{d\vec{r}_k}{dt} \cdot \frac{\partial \vec{r}_k}{\partial q_\ell}$$

et le premier terme des équations de Lagrange pour la particule libre s'écrit

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\ell} = \sum_{k=1}^N m_k \frac{d^2 \vec{r}_k}{dt^2} \cdot \frac{\partial \vec{r}_k}{\partial q_\ell} + \sum_{k=1}^N m_k \frac{d\vec{r}_k}{dt} \cdot \frac{d}{dt} \frac{\partial \vec{r}_k}{\partial q_\ell} \quad (2.32)$$

Pour des particules libres de toute force, le premier terme du second membre de cette égalité est identiquement nul. Par contre, si chaque particule repérée par son indice  $k$  est soumise à une force conservative  $\vec{F}_k$ , par linéarité du gradient, ces forces dérivent toutes du même potentiel scalaire  $U = U(\vec{r}_1, \dots, \vec{r}_N)$  et nous aurons

$$\forall k = 1, \dots, N \quad m_k \frac{d^2 \vec{r}_k}{dt^2} = \vec{F}_k = -\text{grad}_{\vec{r}_k}(U) \quad (2.33)$$

Par ailleurs, un petit tour de passe-passe sur la dérivée temporelle et celle par rapport à  $q_\ell$  permet de voir que

$$\frac{\partial}{\partial q_\ell} \frac{d\vec{r}_k}{dt} = \frac{\partial}{\partial q_\ell} \sum_{i=1}^s \frac{\partial \vec{r}_k}{\partial q_i} \dot{q}_i = \sum_{i=1}^s \frac{\partial^2 \vec{r}_k}{\partial q_\ell \partial q_i} \dot{q}_i + \sum_{i=1}^s \frac{\partial \vec{r}_k}{\partial q_i} \frac{\partial \dot{q}_i}{\partial q_\ell}$$

la dernière somme de cette égalité est nulle car les coordonnées généralisées sont indépendantes des vitesses généralisées, ainsi

$$\frac{\partial}{\partial q_\ell} \frac{d\vec{r}_k}{dt} = \sum_{i=1}^s \frac{\partial^2 \vec{r}_k}{\partial q_\ell \partial q_i} \dot{q}_i = \sum_{i=1}^s \dot{q}_i \frac{\partial}{\partial q_i} \frac{\partial \vec{r}_k}{\partial q_\ell} = \frac{d}{dt} \frac{\partial \vec{r}_k}{\partial q_\ell} \quad (2.34)$$

En reportant les résultats (2.33) et (2.34) dans (2.32) on trouve finalement

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\ell} = - \sum_{k=1}^N \text{grad}_{\vec{r}_k}(U) \cdot \frac{\partial \vec{r}_k}{\partial q_\ell} + \sum_{k=1}^N m_k \frac{d\vec{r}_k}{dt} \cdot \frac{\partial}{\partial q_\ell} \frac{d\vec{r}_k}{dt} \quad (2.35)$$

Il est trivial de remarquer que la dernière somme de cette égalité est tout simplement  $\frac{\partial T}{\partial q_\ell}$ . Les équations de Lagrange pour un système conservatif s'écrivent donc

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\ell} - \frac{\partial T}{\partial q_\ell} = - \sum_{k=1}^N \text{grad}_{\vec{r}_k}(U) \cdot \frac{\partial \vec{r}_k}{\partial q_\ell} = \frac{\partial U}{\partial q_\ell} \quad (2.36)$$

La dernière égalité repose entièrement sur le fait que l'énergie potentielle d'un système conservatif est un champ scalaire ne dépendant que des positions des particules qui y sont soumises, comme nous l'avons dit plus haut  $U = U(\vec{r}_1, \dots, \vec{r}_N)$ . Si tel n'est pas le cas, le système n'est que faussement conservatif et le présent calcul devra être adapté. On étudiera ces cas pathologiques en mécanique classique avec la force de Lorentz (voir 2.4.6) ou

bien la rotation qui peut, dans les cas simples, introduire des forces d'inertie que l'on peut tout de même récupérer dans notre histoire (voir 2.4.4).

Si le potentiel ne dépend pas des vitesses, il ne dépend pas non plus des vitesses généralisées, i.e.  $\frac{\partial U}{\partial \dot{q}_\ell} = 0$ , ainsi dans ce cas particulier, les équations (2.36) deviennent

$$\forall \ell, 1 \leq \ell \leq s \quad \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_\ell} - \frac{\partial \mathcal{L}}{\partial q_\ell} = 0 \quad \text{avec } \mathcal{L} = T - U \quad (2.37)$$

souvent appelées équations de Lagrange de première espèce.

Dans le cas de forces conservatives, le lagrangien d'un système est donc la différence entre l'énergie cinétique totale et l'énergie potentielle totale contenue dans ce système.

La condition de moindre action est dans ce contexte automatiquement vérifiée. En effet, si  $\mathcal{L} = T(\dot{\mathbf{q}}) - U(\mathbf{q}) = \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$  la condition d'extremum de l'action impose (via une intégration par parties)

$$\delta S = \int_{t_1}^{t_2} \left[ \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \right) - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} \right] \cdot \delta \mathbf{q} dt = 0$$

qui se réalise bien pour un mouvement décrit par les équations (2.37).

La déduction de la forme du lagrangien conservatif est ici basée sur l'écriture préalable des équations de Lagrange (dans le cas libre) et du principe de moindre action. Tout ceci peut paraître assez magique, et l'on peut légitimement être en droit de se demander d'où vient ce principe. Seule l'histoire racontée au chapitre 1 peut apporter une réponse claire à cette question. En y regardant de plus près, on s'aperçoit qu'un problème de mécanique céleste se cache derrière toute cette aventure et l'on se reportera sans tarder vers le premier chapitre !

### 2.4.3 Systèmes non conservatifs

Les forces conservatives s'écrivent comme le gradient par rapport aux variables d'espace d'un potentiel  $U$ . Dans ce cas nous avons vu au paragraphe précédent que le lagrangien du système était la simple différence entre énergies cinétique et potentielle. De façon très générale on peut toujours écrire les forces sous la forme

$$\vec{f}_k = - \frac{\partial U(\vec{r}_1, \dots, \vec{r}_N)}{\partial \vec{r}_k} + \vec{Z}_k \quad (2.38)$$

où le terme  $\vec{Z}_k$  ne peut pas être introduit dans le gradient, bien souvent parce qu'il fait intervenir la vitesse. On qualifie souvent ce terme de non conservatif, par analogie avec le cas le plus simple où il est linéaire en vitesse et que cette force modélise des frottements visqueux introduisant une non conservation de l'énergie. Le traitement de telles forces phénoménologiques pourrait être écarté d'un revers de main tant ces dernières ont peu de sens physique et ne doivent leur existence qu'à leur grande efficacité pour la description de certains mouvements dans des conditions particulières. Mais qu'à cela ne tienne, nous allons voir qu'il est souvent possible de les incorporer dans le formalisme lagrangien.

Il suffit de reprendre la déduction du paragraphe précédent en s'arrêtant à la relation (2.36)

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\ell} - \frac{\partial T}{\partial q_\ell} = \sum_{k=1}^N \vec{f}_k \cdot \frac{\partial \vec{r}_k}{\partial q_\ell} \quad (2.39)$$

on introduit alors la force la plus générale (2.38), et en conservant la définition conservative  $\mathcal{L} = T - U$  on obtient pour chacun des degrés de liberté  $\ell = 1, \dots, s$

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_\ell} - \frac{\partial \mathcal{L}}{\partial q_\ell} = \sum_{k=1}^N \vec{Z}_k \cdot \frac{\partial \vec{r}_k}{\partial q_\ell} \quad (2.40)$$

c'est alors que l'on introduit la force généralisée<sup>4</sup>

$$Q_\ell = \sum_{k=1}^N \vec{Z}_k \cdot \frac{\partial \vec{r}_k}{\partial q_\ell} \quad (2.41)$$

pour obtenir finalement les équations de Lagrange de seconde espèce

$$\forall \ell, 1 \leq \ell \leq s \quad \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_\ell} - \frac{\partial \mathcal{L}}{\partial q_\ell} = Q_\ell \quad (2.42)$$

Il va de soi que ces équations sont équivalentes aux équations du mouvement en présence de forces non conservatives. Qu'en est-il de la validité du principe variationnel, vœu pieux que nous avons posé initialement?

L'action demeure toujours

$$S := \int_{t_1}^{t_2} \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) dt \quad (2.43)$$

<sup>4</sup> Comme le lecteur pourra le vérifier cette force généralisée est très particulière car il ne s'agit

sa variation donne

$$\delta S = \left[ \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \cdot \delta \mathbf{q} \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \left[ \frac{\partial \mathcal{L}}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \right] \cdot \delta \mathbf{q} dt \quad (2.44)$$

le terme tout intégré s'annule toujours mais à cause des forces non conservatives et en vertu de l'équation (2.42) cette relation devient

$$\delta S = - \int_{t_1}^{t_2} \sum_{\ell=1}^s \sum_{k=1}^N \vec{Z}_k \cdot \frac{\partial \vec{r}_k}{\partial q_\ell} \delta q_\ell dt \quad (2.45)$$

soit

$$\delta S = - \sum_{k=1}^N \int_{t_1}^{t_2} \vec{Z}_k \cdot \delta \vec{r}_k dt \quad (2.46)$$

Les équations du mouvement, équations de Lagrange de seconde espèce correspondent donc à un extremum de l'action si l'on a

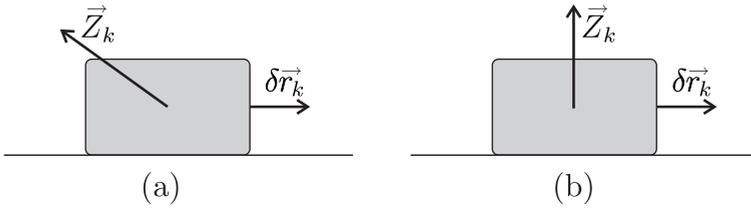
$$\sum_{k=1}^N \int_{t_1}^{t_2} \vec{Z}_k \cdot \delta \vec{r}_k dt = 0 \quad (2.47)$$

La quantité qui doit s'annuler correspond physiquement à une somme de travaux, c'est pourquoi elle est souvent appelée condition des travaux virtuels, ou principe de d'Alembert. Elle est aussi équivalente au principe de l'action et de la réaction.

L'exemple typique, qui donne son nom aux forces non conservatives est celui des frottements sec et/ou visqueux. Ces forces phénoménologiques proportionnelles à une puissance de la vitesse apparaissent comme une représentation effective moyenne d'interactions microscopiques conservatives, ce caractère étant perdu à l'échelle macroscopique. Les deux régimes de frottements sec (lois de Coulomb) correspondent à une force proportionnelle à une puissance nulle de la vitesse, ils ne la font donc pas apparaître explicitement dans l'expression de la force. L'introduction d'un cône de frottement permet alors de statuer sur l'existence ou non d'une vitesse. Dans le cas dynamique, deux types de frottements sont généralement considérés : les frottements parfaits (figure 2.1-(b)) et les frottements réels (figure 2.1-(a)).

---

pas d'un vecteur et de plus, dans certains cas, elle n'a pas la dimension d'une force.



**Fig. 2.1.** Frottement sec parfait (b), réel (a).

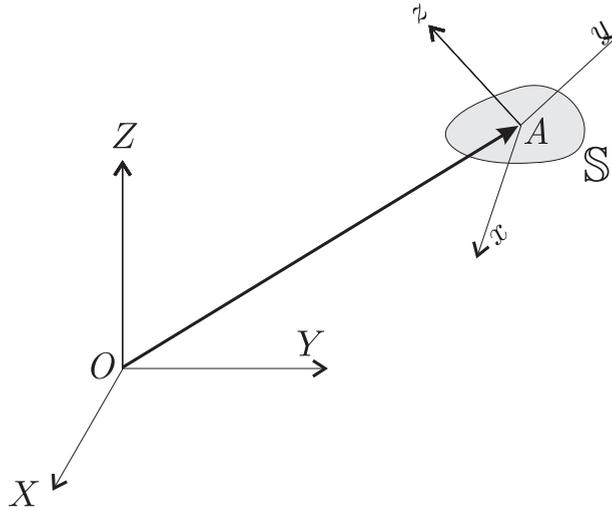
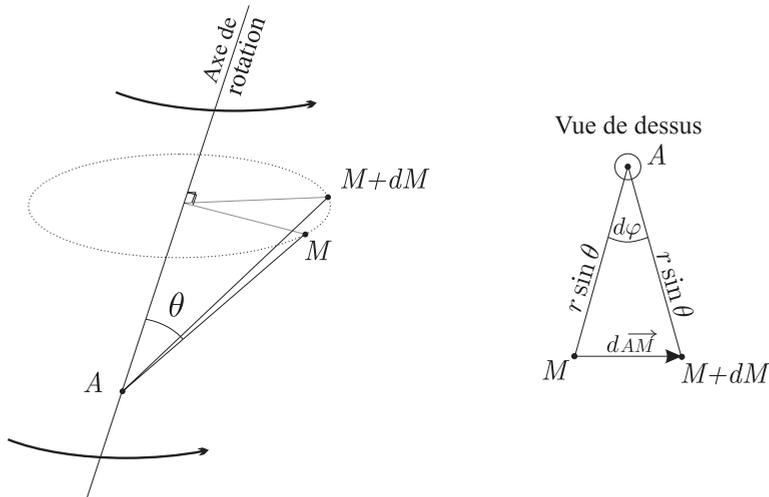
Dans le cas parfait, la somme des travaux est nulle car la force est constamment orthogonale à la variation infinitésimale de position. Les équations de Lagrange de seconde espèce sont bien équivalentes à un principe variationnel. Dans le cas le plus général des frottements non parfaits, le formalisme lagrangien semble s'affaiblir et perdre une partie de sa cohérence. En fait le problème provient surtout de notre mauvaise représentation des frottements. Une description de ce phénomène à l'échelle atomique entre tout à fait dans le formalisme lagrangien, c'est donc le passage au macroscopique qui pose des problèmes... mais ceci est une autre histoire !

#### 2.4.4 Lagrangien et mouvement d'un solide

Jusqu'à présent nous n'avons considéré que le mouvement d'une assemblée de points libres ou en interaction : il s'agit d'un problème académique et très théorique. Un cas moins académique est celui du mouvement d'un solide, c'est-à-dire d'un ensemble de points dont les distances relatives sont invariables<sup>5</sup>. Pour traiter un tel système, il nous faut considérer deux référentiels : l'un  $(O_{XYZ})$  fixe et galiléen, l'autre  $(A_{xyz})$  dont la seule propriété est d'être en liaison rigide avec un point  $A$  du solide, ces deux référentiels sont représentés sur la figure 2.2. Un solide peut avoir un mouvement de rotation propre, c'est ce raffinement qui le distingue du point matériel. Chaque solide possède donc 6 degrés de liberté : 3 de translation et 3 de rotation. Nous considérerons ici que l'axe de rotation du solide est fixe. C'est une restriction importante qui nous permettra de mener à bien l'ensemble des calculs, nous laisserons les lecteurs férus de mécanique plus générale, et donc plus appliquée, consulter des ouvrages plus spécialisés.

Considérons un point  $M$  quelconque du solide, nous avons  $\overrightarrow{OM} = \overrightarrow{OA} + \overrightarrow{AM}$ . Le mouvement infinitésimal du point  $M$  induit par un petit mouvement

<sup>5</sup> Un passage à la limite continue est aussi envisageable, il ne pose pas de problème dans les cas simples.


**Fig. 2.2.** Référentiel fixe et référentiel attaché au solide

**Fig. 2.3.** Rotation d'un solide autour d'un axe fixe

du solide se décompose en deux parties :  $d\vec{OM} = d\vec{OA} + d\vec{AM}$ . En posant  $r = |\vec{AM}|$ , comme on peut le voir sur la figure 2.3, on a

$$\sin \frac{d\varphi}{2} = \frac{|d\vec{r}|}{2|\vec{r}|\sin \theta} . \quad (2.48)$$

Sous l'hypothèse d'une rotation infinitésimale, nous avons  $d\varphi \ll 1$  et par

conséquent, à l'ordre 1,  $|d\vec{r}| = |\vec{r}| \sin \theta d\varphi$ . En notant  $d\vec{\varphi}$  le vecteur dont la longueur est l'angle de rotation et dont la direction coïncide avec l'axe de rotation, nous avons finalement

$$d\overrightarrow{AM} = d\vec{\varphi} \wedge \overrightarrow{AM} \quad (2.49)$$

nous obtenons donc finalement la vitesse instantanée du solide en un point quelconque de celui-ci

$$\frac{d\vec{r}}{dt} = \frac{d\overrightarrow{OA}}{dt} + \vec{\Omega} \wedge \overrightarrow{AM} \quad (2.50)$$

où le vecteur  $\vec{\Omega} := d\vec{\varphi}/dt$  est indépendant du point considéré si l'axe de rotation est fixe, il est communément appelé vitesse angulaire.

Munis de ces précisions nous pouvons calculer l'énergie cinétique totale du solide. Par définition, cette quantité est la somme de toutes les énergies cinétiques de tous les points constituant le solide.

$$\begin{aligned} T &= \sum_{i \in \mathbb{S}} \frac{m_i}{2} \left( \frac{d\vec{r}_i}{dt} \right)^2 = \sum_{i \in \mathbb{S}} \frac{m_i}{2} \left( \frac{d\overrightarrow{OA}}{dt} + \vec{\Omega} \wedge \overrightarrow{AM}_i \right)^2 \\ &= \sum_{i \in \mathbb{S}} \frac{m_i}{2} \left( \frac{d\overrightarrow{OA}}{dt} \right)^2 + \left( \frac{d\overrightarrow{OA}}{dt} \wedge \vec{\Omega} \right) \cdot \sum_{i \in \mathbb{S}} m_i \overrightarrow{AM}_i \\ &\quad + \sum_{i \in \mathbb{S}} \frac{m_i}{2} \left[ \vec{\Omega}^2 \overrightarrow{AM}_i^2 - \left( \vec{\Omega} \cdot \overrightarrow{AM}_i \right)^2 \right] \end{aligned}$$

Si l'on fait l'hypothèse que  $A$  est le centre d'inertie de  $\mathbb{S}$ , le terme central de cette dernière somme est nul. En se plaçant dans la base fixe on peut doter les vecteurs de coordonnées, soit  $\vec{\Omega} = [\Omega_1, \Omega_2, \Omega_3]$  et  $\overrightarrow{AM}_i = [x_i, y_i, z_i]$ , il vient

$$T = \sum_{i \in \mathbb{S}} \frac{m_i}{2} \left( \frac{d\overrightarrow{OA}}{dt} \right)^2 + \sum_{\alpha, \beta=1}^3 \frac{1}{2} \Omega_\alpha \Omega_\beta I^{\alpha\beta} \quad (2.51)$$

où nous avons introduit le tenseur d'inertie du solide  $\mathbb{S}$ , dont on peut fournir toutes les composantes dans un tableau : si  $\alpha$  est la ligne et  $\beta$  la colonne de ce tableau on a

$$I^{\alpha\beta} = \sum_{i \in \mathbb{S}} m_i \begin{bmatrix} y_i^2 + z_i^2 & x_i y_i & x_i z_i \\ y_i x_i & x_i^2 + z_i^2 & y_i z_i \\ z_i x_i & z_i y_i & x_i^2 + y_i^2 \end{bmatrix} \quad (2.52)$$

La relation (2.51) montre à l'évidence que l'énergie cinétique d'un solide en rotation autour d'un axe fixe se décompose en la somme d'une énergie cinétique de translation et d'une énergie cinétique de rotation. Le terme de translation ne fait intervenir que la vitesse du centre d'inertie du solide, le terme de rotation fait intervenir un élément essentiel en mécanique des solides, la matrice d'inertie, couplée à la vitesse angulaire.

Si des forces conservatives sont appliquées au solide, son lagrangien sera simplement  $\mathcal{L} = T - U$ , où nous avons noté  $U$  la somme des énergies potentielles de tous les points du solide. Si les forces sont plus complexes, de fortes et variées théories abondent dans les plus merveilleux livres de mécanique, mais il ne s'agit plus alors de physique théorique ...

### 2.4.5 Contraintes

Les contraintes ou liaisons doivent absolument être prises en compte dans le formalisme lagrangien. En effet, la variation de l'action s'écrit en phase terminale

$$\delta S = \int_{t_1}^{t_2} \left[ \frac{\partial \mathcal{L}}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \right] \cdot \delta \mathbf{q} dt \quad (2.53)$$

la condition  $\delta S = 0$  est donc équivalente aux simples équations de Lagrange si et seulement si les  $s$  composantes du vecteur  $\delta \mathbf{q}$  sont indépendantes. Si ce n'est pas le cas, deux situations sont à envisager.

#### Contraintes holonomes

On parle de contraintes ou liaisons holonomes (lois entières) dans le cas où la dépendance des  $\delta q_\ell$  est intégrable en  $\Lambda$  relations de la forme

$$\forall \lambda = 1, \dots, \Lambda \quad f_\lambda(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t) = 0 \quad (2.54)$$

Ces contraintes sont indépendantes, c'est-à-dire que pour tous les vecteurs positions  $\vec{r}_1, \dots, \vec{r}_N$  et à chaque instant  $t$  la matrice

$$\begin{bmatrix} \frac{\partial f_1}{\partial \vec{r}_1} & \cdots & \frac{\partial f_1}{\partial \vec{r}_N} \\ \vdots & & \vdots \\ \frac{\partial f_\Lambda}{\partial \vec{r}_1} & \cdots & \frac{\partial f_\Lambda}{\partial \vec{r}_N} \end{bmatrix} \quad \Lambda \leq N \quad (2.55)$$

est de rang plein (=  $\Lambda$ ).

La prise en compte de contraintes holonomes dans le formalisme lagrangien est tout à fait sympathique, il suffit en effet simplement de diminuer le nombre de coordonnées généralisées du nombre de contraintes holonomes.

L'exemple du pendule double planaire est dans ce contexte très parlant.

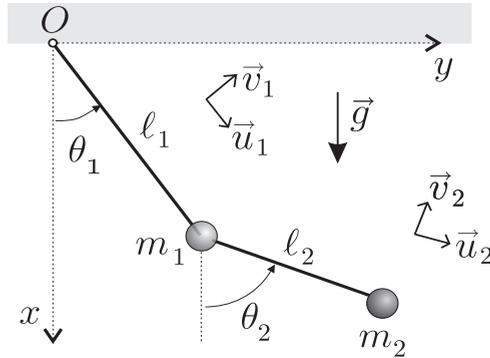


Fig. 2.4. Le pendule double planaire

Les deux billes ponctuelles de masse  $m_1$  et  $m_2$  évoluent dans un plan au bout de deux tiges rigides de longueurs respectives  $\ell_1$  et  $\ell_2$  (voir figure 2.4) et dont les masses sont négligeables devant celles des billes. Ce système possède en théorie 4 degrés de liberté : les deux coordonnées cartésiennes de chaque bille pourraient être considérées comme les 4 coordonnées généralisées. La rigidité des tiges impose cependant deux contraintes holonomes

$$\sqrt{x_1^2 + y_1^2} = \ell_1 \quad \text{et} \quad \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} = \ell_2 \quad (2.56)$$

les composantes du vecteur  $\mathbf{q} \in \mathbb{R}^4$  dont les composantes seraient  $(x_1, y_1, x_2, y_2)$  ne sont pas indépendantes... Un peu d'expérience en mécanique, permet cependant de voir immédiatement qu'en se plaçant en coordonnées polaires dans le plan du mouvement on libère automatiquement les variables angulaires. On peut s'en rendre compte en déterminant le lagrangien de ce système. Nous avons trivialement

$$\begin{aligned} \vec{r}_1 &= \ell_1 \vec{u}_1 & \text{et} & \quad \frac{d\vec{r}_1}{dt} = \ell_1 \dot{\theta}_1 \vec{v}_1 \\ \vec{r}_2 &= \ell_1 \vec{u}_1 + \ell_2 \vec{u}_2 & \text{et} & \quad \frac{d\vec{r}_2}{dt} = \ell_1 \dot{\theta}_1 \vec{v}_1 + \ell_2 \dot{\theta}_2 \vec{v}_2 \end{aligned} \quad (2.57)$$

on en déduit directement l'énergie cinétique

$$T = \frac{1}{2} \left[ m_1 \left( \frac{d\vec{r}_1}{dt} \right)^2 + m_2 \left( \frac{d\vec{r}_2}{dt} \right)^2 \right] \quad (2.58)$$

$$= \frac{1}{2} (m_1 + m_2) \ell_1^2 \dot{\theta}_1^2 + \frac{1}{2} m_2 \ell_2^2 \dot{\theta}_2^2 + m_2 \ell_1 \ell_2 \dot{\theta}_1 \dot{\theta}_2 \cos(\theta_1 - \theta_2) .$$

L'énergie potentielle est purement d'origine gravitationnelle, il vient

$$U = -(m_1 + m_2)g\ell_1 \cos(\theta_1) - m_2g\ell_2 \cos(\theta_2) \quad (2.59)$$

en introduisant les coordonnées généralisées  $q_1 = \theta_1$  et  $q_2 = \theta_2$  composantes du vecteur  $\mathbf{q} \in \mathbb{R}^2$ , on constate que le lagrangien habituel  $\mathcal{L} = T - U$  possède la forme requise  $\mathcal{L} = \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$ . Ce sont donc bien  $4 - 2 = 2$  degrés de liberté que compte ce système. Il n'y a que deux équations du mouvement qui sont les deux composantes de la relation vectorielle

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \mathbf{0} \quad (2.60)$$

Nous ne les écrivons pas car la résolution des problèmes de la physique est une activité différente de celle qui consiste à comprendre ses équations...

Les contraintes holonomes semblent donc être une aubaine en diminuant uniquement le nombre de degrés de liberté sans ajouter de complication supplémentaire. Certains cas sont cependant moins triviaux que celui que nous avons traité!

### Contraintes non holonomes

On parle de contraintes non holonomes, lorsque les relations de dépendance entre les  $\delta q_\ell$  ne peuvent pas être intégrées. Ces  $\Lambda$  contraintes prennent donc la forme générale

$$1 \leq \lambda \leq \Lambda \quad \sum_{k=1}^N \omega_k^\lambda(\vec{r}_1, \dots, \vec{r}_N, t) d\vec{r}_k = 0 \quad (2.61)$$

Cette situation est beaucoup plus complexe que le cas holonome et se traite au cas par cas en fonction de la forme des fonctions  $\omega_k^\lambda$ .

Un seul cas supporte une théorie générale, c'est celui de contraintes linéaires en vitesse :

$$1 \leq \lambda \leq \Lambda \quad \sum_{k=1}^N a_{\lambda k} \frac{d\vec{r}_k}{dt} = 0 \quad (2.62)$$

ce type de contraintes entre bien dans le cadre des liaisons non holonomes les fonctions  $\omega_k^\lambda$  se réduisant ici à de simples «inverses de  $dt$ ».

Puisque nous sommes dans ce cas particulier, on dit d'une contrainte qu'elle est rhéonome si elle dépend du temps et scléronome si elle n'en dépend pas.

Dans le cas (2.62) que nous évoquions de contrainte à la fois rhéonome et non holonome inverse en  $dt$ , la condition d'extremum pour l'action

$$\delta S = \int_{t_1}^{t_2} \sum_{i=1}^s \left[ \frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right] \delta q_i dt = 0 \quad (2.63)$$

n'est pas remplie complètement en imposant les seules équations de Lagrange, les  $\delta q_i$  n'étant plus indépendants. On peut toutefois «sauver» la situation en introduisant  $\Lambda$  multiplicateurs de Lagrange  $\beta_\lambda$  (inventés pour l'occasion), et en résolvant le système

$$\begin{cases} \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = \sum_{\lambda=1}^{\Lambda} a_{\lambda i} \beta_\lambda \\ \sum_{i=1}^s a_{\lambda i} \dot{q}_i = 0 \end{cases} \quad (2.64)$$

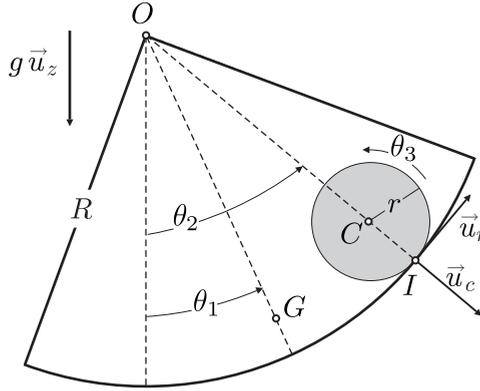
ce système est complet pour les variables  $\beta_{\lambda(1 \leq \lambda \leq \Lambda)}$  et  $q_{i(1 \leq i \leq s)}$ ,  $\Lambda$  est le nombre de contraintes (et donc de multiplicateurs de Lagrange à introduire) et  $s$  le nombre de degrés de liberté du système étudié.

Un cas fondamental de ce type de contraintes est celui du roulement sans glissement.

Traitons pour l'exemple un problème célèbre car il occupa les candidats du concours commun Mines-Ponts en 2008 qui devaient le traiter sans le formalisme lagrangien...

Il s'agit d'un rail suspendu en forme de portion de cercle de masse  $M$  à l'intérieur duquel roule sans glisser une sphère de masse  $m$ . Le problème est considéré plan et sans frottement. A un instant initial, on écarte le rail de sa position d'équilibre et on lâche la sphère depuis une certaine position. Des oscillations couplées apparaissent alors, c'est très joli!

Contrairement à ce qu'ont pensé les candidats à ce concours la difficulté de ce problème n'est pas l'obtention des équations du mouvement, mais l'étude des propriétés de leurs solutions. La mécanique analytique permet sans coup



**Fig. 2.5.** Un rail, une sphère et du bonheur !

férier de les écrire en utilisant les quelques principes énoncés plus haut. En ce qui concerne leur propriétés, on peut y avoir accès en utilisant certaines des méthodes présentées dans les prochains chapitres.

Plaçons-nous dans le référentiel galiléen centré sur  $O$  et utilisons les vecteurs unitaires présentés sur la figure 2.5. Les forces appliquées à ce système sont uniquement les poids du rail et de la sphère, le lagrangien est donc simplement la différence entre les énergies cinétique et potentielle. Détaillons un minimum car sinon la résolution ne prendrait que trois lignes...

S'agissant de solides leur énergie cinétique se décompose en une partie due à la translation de leur centre d'inertie et une due à leur rotation :

- Le centre  $C$  de la sphère possède une vitesse  $\vec{v} = \frac{d}{dt} [(R - r) \vec{u}_c] = (R - r) \dot{\theta}_2 \vec{u}_r$ . Si l'on note  $J_s$  son moment d'inertie par rapport à son axe de rotation, l'utilisation de la relation (2.51) dans un cas très simple donne son énergie cinétique

$$T_s = \frac{1}{2} m \vec{v}^2 + \frac{1}{2} J_s \dot{\theta}_3^2 = \frac{1}{2} m (R - r)^2 \dot{\theta}_2^2 + \frac{1}{2} J_s \dot{\theta}_3^2$$

- Le centre d'inertie  $G$  du rail n'est qu'en rotation autour de  $O$ , son énergie cinétique est entièrement due à cette dernière. En notant  $J_r$  son moment d'inertie par rapport à l'axe normal au plan du mouvement passant par  $O$ , on a donc

$$T_r = \frac{1}{2} J_r \dot{\theta}_1^2$$

Pour l'énergie potentielle, seules les variations d'altitude des deux centres d'inertie ont une contribution. En notant  $\ell$  la distance constante entre  $O$  et  $G$ , il vient

$$U = -Mgl \cos \theta_1 - mg(R - r) \cos \theta_2 + cste$$

La constante dépendant du choix de l'origine des altitudes, prenons-la nulle!

Un dernier point reste à régler, il s'agit de la condition de roulement sans glissement. Sous ce vocable on entend en fait que la vitesse du point  $I$  considéré comme un point de la sphère est la même que celle du point  $I$  considéré comme un point du rail. Cette condition s'écrit donc comme la contrainte

$$(R - r) \dot{\theta}_2 + r \dot{\theta}_3 = R \dot{q}_1$$

Le lagrangien du système est donc prêt, il s'écrit

$$\begin{aligned} \mathcal{L} &= T_s + T_r - U \\ &= \frac{1}{2} J_r \dot{\theta}_1^2 + \frac{1}{2} m (R - r)^2 \dot{\theta}_2^2 + \frac{1}{2} J_s \dot{\theta}_3^2 + Mgl \cos \theta_1 + mg(R - r) \cos \theta_2 \\ &= \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) \quad \text{avec } \mathbf{q} = (\theta_1, \theta_2, \theta_3) \in \mathbb{R}^3 \end{aligned}$$

Notons qu'une fois n'est pas coutume nous avons noté un vecteur de  $\mathbb{R}^3$  sans son antique flèche, pour conserver un peu d'inertie dans les notations compte-tenu du caractère accidentel de cette situation. Le système possède donc 3 degrés de liberté, il est soumis à une unique contrainte ( $\Lambda = 1$ ) qui ne fait apparaître qu'un seul vecteur  $\mathbf{a}_1$  décrivant la liaison entre les vitesses sous la forme

$$\mathbf{a}_1 \cdot \dot{\mathbf{q}} = 0 \quad \text{avec } \mathbf{a}_1 = (-R, R - r, r)$$

Comme  $\Lambda = 1$ , un seul multiplicateur de Lagrange est nécessaire. En notant  $\beta_1$  ce dernier, les 3 équations du mouvement sont donc les 3 composantes de l'équation vectorielle ci-dessous, complétées de la contrainte qui permet de déterminer  $\beta_1$ .

$$\begin{cases} \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \beta_1 \mathbf{a}_1 \\ \mathbf{a}_1 \cdot \dot{\mathbf{q}} = 0 \end{cases}$$

Que l'on ne détaille même pas car là n'est plus la difficulté! Le problème est d'ailleurs tellement simple que l'on peut ici utiliser la contrainte pour faire disparaître un degré de liberté... Pour bien enfoncer le clou, notons qu'il ne serait pas rendu beaucoup plus compliqué par l'introduction d'un ressort rectiligne entre le rail et le point de suspension  $O$ , il suffirait d'introduire un quatrième degré de liberté qui pourrait être l'allongement du ressort.

Essayez de traiter un tel problème dans le formalisme de Newton... ceux qui y étaient parvenus avaient réussi leur écrit, s'ils avaient connu Lagrange ils auraient été moins bien triés!

### 2.4.6 Lagrangien d'une particule dans un champ électromagnétique

Le cas d'une particule dans un champ électromagnétique est particulièrement important pour plusieurs raisons : tout d'abord, si les forces d'origines électriques sont bien conservatives, en ce sens qu'elles sont le gradient spatial d'un potentiel spatial, il n'en va pas de même des forces magnétiques qui font intervenir la vitesse de la charge. Une charge en mouvement dans un champ électromagnétique subit donc une force non conservative. Enfin, l'importance physique de l'interaction électromagnétique ne nous autorise pas à la laisser sur le bord du chemin variationnel que nous avons entrepris depuis le début de ce chapitre.

Considérons donc une assemblée de  $k = 1, \dots, N$  charges  $e_k$  repérées par  $N$  vecteurs positions  $\vec{r}_k$  et placées dans un champ électromagnétique extérieur  $(\vec{E}, \vec{B})$ . Chaque charge  $e_k$  perçoit donc des champs vectoriels  $\vec{E}_k = \vec{E}(\vec{r}_k, t)$  et  $\vec{B}_k = \vec{B}(\vec{r}_k, t)$  associés aux potentiel vecteur  $\vec{A}_k = \vec{A}(\vec{r}_k, t)$  et scalaire  $V_k = V(\vec{r}_k, t)$  tels que

$$\vec{E} = -\frac{\partial \vec{A}}{\partial t} - \vec{\nabla} V \quad \text{et} \quad \vec{B} = \vec{\nabla} \wedge \vec{A} \quad (2.65)$$

Pour la première fois dans ce livre, nous voyons apparaître le «vecteur nabla», il est bien utile :

$$\vec{\nabla} = \frac{\partial}{\partial \vec{r}}$$

En mécanique classique, on constate expérimentalement que la charge  $e_k$  subit une force, dite de Lorentz<sup>6</sup>, et que l'équation donnant son mouvement s'écrit

$$m_k \ddot{\vec{r}}_k = e_k \left( \vec{E}_k + \dot{\vec{r}}_k \wedge \vec{B}_k \right) = \vec{F}_k \quad (2.66)$$

Le but du jeu, car il ne s'agit pas d'autre chose pour le moment, consiste à trouver un lagrangien tel que l'équation de Lagrange associée redonne l'équation (2.66).

Pour ce faire, commençons par remplacer dans l'expression de la force de Lorentz, les champs par leurs potentiels, il vient

$$\vec{F}_k = e_k \left[ -\frac{\partial \vec{A}_k}{\partial t} - \vec{\nabla} V_k + \dot{\vec{r}}_k \wedge \left( \vec{\nabla} \wedge \vec{A}_k \right) \right] \quad (2.67)$$

<sup>6</sup> Nous reviendrons en détail sur cette force dans le chapitre consacré à la relativité restreinte.

en remarquant que

$$\begin{aligned} \frac{d\vec{A}_k}{dt} &= \frac{\partial \vec{A}_k}{\partial t} + \frac{d\vec{r}_k}{dt} \frac{\partial \vec{A}_k}{\partial \vec{r}_a} \\ &= \frac{\partial \vec{A}_k}{\partial t} + \left( \dot{\vec{r}}_k \cdot \vec{\nabla} \right) \vec{A}_k \end{aligned} \quad (2.68)$$

cette force se réécrit

$$\vec{F}_k = e_k \left[ -\frac{d\vec{A}_k}{dt} - \vec{\nabla} V_k + \left( \dot{\vec{r}}_k \cdot \vec{\nabla} \right) \vec{A}_k + \dot{\vec{r}}_k \wedge \left( \vec{\nabla} \wedge \vec{A}_k \right) \right] \quad (2.69)$$

en se souvenant d'une formule d'analyse vectorielle fort utile

$$\vec{\nabla} (\vec{u} \cdot \vec{v}) = \vec{u} \wedge (\vec{\nabla} \wedge \vec{v}) + \vec{v} \wedge (\vec{\nabla} \wedge \vec{u}) + (\vec{u} \cdot \vec{\nabla}) \vec{v} + (\vec{v} \cdot \vec{\nabla}) \vec{u} \quad (2.70)$$

et en notant que l'opérateur spatial  $\vec{\nabla}$  a une action nulle sur la vitesse  $\dot{\vec{r}}_k$ , la force de Lorentz peut finalement se mettre sous la forme

$$\vec{F}_k = e_k \left[ -\frac{d\vec{A}_k}{dt} + \vec{\nabla} \left( \dot{\vec{r}}_k \cdot \vec{A}_k - V_k \right) \right] \quad (2.71)$$

l'équation du mouvement (2.66) s'écrit donc

$$\frac{d}{dt} \left[ m_k \dot{\vec{r}}_k + e_k \vec{A}_k \right] = \vec{\nabla} \left( e_k \dot{\vec{r}}_k \cdot \vec{A}_k - e_k V_k \right) \quad (2.72)$$

pour retrouver la forme si particulière de l'équation de Lagrange il faut faire apparaître une dérivée par rapport à la vitesse à l'intérieur du terme de dérivée temporelle totale, c'est possible en écrivant

$$\frac{d}{dt} \left[ m_k \dot{\vec{r}}_k + e_k \vec{A}_k \right] = \frac{d}{dt} \left[ \frac{d}{d\dot{\vec{r}}_k} \left( \frac{1}{2} m_k \dot{\vec{r}}_k^2 + e_k \dot{\vec{r}}_k \cdot \vec{A}_k \right) \right] \quad (2.73)$$

l'équation (2.72) donne donc

$$\frac{d}{dt} \left[ \frac{d}{d\dot{\vec{r}}_k} \left( \frac{1}{2} m_k \dot{\vec{r}}_k^2 + e_k \dot{\vec{r}}_k \cdot \vec{A}_k \right) \right] = \frac{d}{d\dot{\vec{r}}_k} \left( e_k \dot{\vec{r}}_k \cdot \vec{A}_k - e_k V_k \right) \quad (2.74)$$

à l'intérieur de la dérivée par rapport à la vitesse (terme de gauche), on peut rajouter le terme  $-e_k V_k$  qui ne dépend que de la position. De même, à l'intérieur de la dérivée par rapport à la position (terme de droite), on

peut rajouter le terme  $m_k \dot{\vec{r}}_k^2/2$  qui ne dépend que de la vitesse, ainsi en introduisant le lagrangien

$$\mathcal{L} = \sum_{k=1}^N \frac{1}{2} m_k \dot{\vec{r}}_k^2 + e_k \dot{\vec{r}}_k \cdot \vec{A}_k - e_k V_k \quad (2.75)$$

$$= \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) \quad \text{avec } \mathbf{q} = (\vec{r}_1, \dots, \vec{r}_N) \in \mathbb{R}^{3N} \quad (2.76)$$

l'équation du mouvement est une équation de Lagrange et le tour est joué !

Ce lagrangien ressemble à une différence entre un terme cinétique et un terme potentiel, mais ici l'énergie potentielle

$$U = \sum_{k=1}^N e_k V(\vec{r}_k) - e_k \dot{\vec{r}}_k \cdot \vec{A}(\vec{r}_k) = U(\mathbf{q}, \dot{\mathbf{q}}) \quad (2.77)$$

dépend surnoisement de la vitesse, ainsi la force (ou plus précisément sa partie magnétique) ne dérive pas d'un vrai potentiel, i.e. ne dépendant que des positions. On retrouve là un résultat bien connu qui dit que les forces électromagnétiques ne sont pas conservatives. Sauf, bien entendu, si le terme  $e_k \dot{\vec{r}}_k \cdot \vec{A}_k$  est la dérivée totale par rapport au temps d'une fonction, auquel cas il entre dans la liberté de jauge et peut être oublié ! C'est le cas par exemple d'un potentiel vecteur uniforme et indépendant du temps.

Il est enfin bien connu qu'une transformation (dite de jauge) des champs de la forme

$$\begin{cases} \vec{A} \rightarrow \vec{A}' = \vec{A} + \frac{d\phi}{d\vec{r}} \\ V \rightarrow V' = V - \frac{\partial\phi}{\partial t} \end{cases} \quad (2.78)$$

où  $\phi = \phi(\vec{r}, t)$  est un champ scalaire quelconque, laisse inchangées les équations du mouvement. Dans le contexte du formalisme lagrangien on a

$$\mathcal{L} \rightarrow \mathcal{L}' = \mathcal{L} + \sum_{k=1}^N e_k \left( \frac{\partial\phi}{\partial t} + \frac{d\vec{r}}{dt} \frac{d\phi}{d\vec{r}} \right) = \mathcal{L} + \frac{d\phi}{dt} \sum_{k=1}^N e_k \quad (2.79)$$

Si la charge totale  $Q := \sum_{k=1}^N e_k$  est constante, lors d'une transformation de jauge des champs, on ajoute au lagrangien la dérivée totale par rapport au temps d'une fonction de  $\vec{r}$  et de  $t$ , et tout va pour le mieux dans le meilleur des mondes. Si la charge totale n'est pas conservée, alors la symétrie lagrangienne est brisée, mais c'est bien la moindre des choses car il faut expliquer pourquoi la charge varie ...

## 2.5 Lagrangien, symétries et lois de conservation

### 2.5.1 Symétries

Afin de parler précisément de choses précises, nous préférons commencer cette section par un ensemble de définitions qui nous seront bien utiles.

Commençons par la notion de symétrie.

Soit  $g$  une application de l'espace  $E$  des configurations<sup>7</sup> dans lui-même

$$\begin{aligned} E &\rightarrow E \\ \mathbf{q} = (q_1, \dots, q_s) &\mapsto \mathbf{g}(\mathbf{q}) = (g^1(\mathbf{q}), \dots, g^s(\mathbf{q})) \end{aligned} \quad (2.80)$$

et  $\mathbf{g}^*$  l'application induite dans l'espace  $\Gamma$  des phases<sup>8</sup>

$$\begin{aligned} \Gamma &\rightarrow \Gamma \\ (\mathbf{q}, \dot{\mathbf{q}}) &\mapsto \mathbf{g}^*(\mathbf{q}, \dot{\mathbf{q}}) = \left( \mathbf{g}(\mathbf{q}), \frac{d\mathbf{g}(\mathbf{q})}{dt} \right) \end{aligned} \quad (2.81)$$

Munis de ces précisions, nous pouvons définir la notion de symétrie.

**Définition 2.1.** *On appelle symétrie du système décrit par le lagrangien  $\mathcal{L}$ , toute fonction  $g$  de l'espace de configuration telle que son application induite  $g^*$  laisse  $\mathcal{L}$  inchangé*

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = \mathcal{L}(\mathbf{g}^*(\mathbf{q}, \dot{\mathbf{q}})) \quad (2.82)$$

Le cas malheureux où le lagrangien dépendrait explicitement du temps est lui aussi envisageable. Il suffit alors d'incorporer  $t$  comme une nouvelle coordonnée généralisée et de définir un lagrangien étendu (voir relation (2.102) et suivantes). Une conséquence immédiate de cette définition est la suivante : si  $g$  est une symétrie d'un système, l'image par  $g$  d'une solution des équations du mouvement  $\mathbf{q}(t)$  est aussi solution de ces équations. La démonstration de ce petit théorème est immédiate car le lagrangien étant invariant, l'action l'est aussi, si elle est extrémale pour  $\mathbf{q}(t)$  elle l'est donc aussi pour  $\mathbf{g}(\mathbf{q})$ .

On dit que le système est invariant par rapport à  $g$  ou bien que  $g^*$  commute avec l'évolution.

**Définition 2.2.** *On dit que  $\mathbb{G} = \{g_\sigma, \sigma \in \mathbb{R}\}$  est un groupe de symétrie à un paramètre réel  $\sigma$  si et seulement si*

$$\begin{cases} g_{\sigma_1} \circ g_{\sigma_2} = g_{\sigma_1 + \sigma_2} \\ g_{\sigma=0} = \text{id}_{\mathbb{R}} \end{cases} \quad (2.83)$$

<sup>7</sup> L'espace des configurations est l'ensemble de toutes les positions atteignables par le système pendant un temps déterminé.

<sup>8</sup> L'espace des phases est l'espace des configurations complété de l'ensemble des vitesses acces-

Munis de ces quelques définitions nous pouvons en venir au fait.

### 2.5.2 Théorème de Noether

Le théorème suivant est l'œuvre d'Emmy Noether, mathématicienne allemande, démontré pour la première fois dans un article de 1918, il symbolise la suprématie du formalisme lagrangien pour le traitement des problèmes de mécanique théorique. Ses prolongements dans la physique théorique moderne sont aussi nombreux que fondamentaux. En substance ce théorème énonce qu'une grandeur est conservée dès qu'une symétrie apparaît dans un système.

**Théorème 2.1.** *Si  $\mathbb{G}$  est un groupe de symétrie à un paramètre réel  $\sigma$  du système décrit par le lagrangien  $\mathcal{L}$  alors la quantité*

$$I(\mathbf{q}, \dot{\mathbf{q}}) := \sum_{\ell=1}^s \frac{\partial \mathcal{L}}{\partial \dot{q}_\ell} \frac{dg_\sigma^\ell(\mathbf{q})}{d\sigma} \Big|_{\sigma=0} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \cdot \frac{d\mathbf{g}_\sigma(\mathbf{q})}{d\sigma} \Big|_{\sigma=0} \quad (2.84)$$

est une constante du mouvement.

#### Démonstration

Par hypothèse  $\mathbb{G}$  est un groupe de symétrie pour  $\mathcal{L}$  ainsi

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = \mathcal{L}(\mathbf{g}_\sigma^*(\mathbf{q}, \dot{\mathbf{q}})) \quad (2.85)$$

de cette égalité nous pouvons déduire que

$$\frac{d}{d\sigma} \mathcal{L}(\mathbf{g}_\sigma^*(\mathbf{q}, \dot{\mathbf{q}})) = \frac{d}{d\sigma} \mathcal{L}\left(\mathbf{g}_\sigma(\mathbf{q}), \frac{d\mathbf{g}_\sigma(\mathbf{q})}{dt}\right) = 0 \quad (2.86)$$

soit

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}} \cdot \frac{d\mathbf{g}_\sigma(\mathbf{q})}{d\sigma} + \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \cdot \frac{d}{d\sigma} \left( \frac{d\mathbf{g}_\sigma(\mathbf{q})}{dt} \right) = 0 \quad (2.87)$$

Si  $\mathbf{q}(t)$  est solution des équations du mouvement, les équations de Lagrange donnent

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \right) \quad (2.88)$$

en incorporant ce résultat dans le premier terme de l'équation (2.87) et en permutant<sup>9</sup>, dans le deuxième terme de cette même équation, la dérivée par rapport à  $t$  et celle par rapport à  $\sigma$ , il vient

sibles.

<sup>9</sup> Ce qui est toujours possible au moins par hypothèse...

$$\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \right) \cdot \frac{d\mathbf{g}_\sigma(\mathbf{q})}{d\sigma} + \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \cdot \frac{d}{dt} \left( \frac{d\mathbf{g}_\sigma(\mathbf{q})}{d\sigma} \right) = 0 \quad (2.89)$$

On reconnaît finalement la dérivée totale par rapport au temps d'un produit

$$\frac{dI(\mathbf{q}, \dot{\mathbf{q}})}{dt} = 0 \quad \text{avec} \quad I(\mathbf{q}, \dot{\mathbf{q}}) = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \cdot \frac{d\mathbf{g}_\sigma(\mathbf{q})}{d\sigma} \quad (2.90)$$

ainsi  $I(\mathbf{q}, \dot{\mathbf{q}})$  est bien une constante du mouvement. Le fait de prendre  $\sigma = 0$  est un cas particulier qui permet d'éliminer ce paramètre, mais le théorème reste vrai dans le cas général.  $\square$

### 2.5.3 Trois exemples fondamentaux

Pour les trois exemples qui suivent, nous considérons des systèmes conservatifs pour lesquels le lagrangien s'écrit donc  $\mathcal{L} = T - U$ . C'est par exemple le cas d'un système de  $N$  particules sans contraintes repérées par les  $k = 1, \dots, N$  vecteurs position  $\vec{r}_k \in \mathbb{R}^3$ . Dans ce cas le nombre de degrés de liberté est  $s = 3N$ , et les coordonnées généralisées sont  $\mathbf{q} = (\vec{r}_1, \dots, \vec{r}_N)$ .

Afin de pouvoir écrire des choses lisibles on attribue des masses à chaque coordonnée généralisée : si pour chaque  $k = 1, \dots, N$  l'on note  $m(k)$  la masse de la particule  $k$  et par groupe de 3 :

$$\begin{aligned} m_1 &= m_2 = m_3 = m(1), \\ m_4 &= m_5 = m_6 = m(2), \\ &\dots, \\ m_{s-2} &= m_{s-1} = m_s = m(N). \end{aligned}$$

Nous aurons

$$T = \sum_{\ell=1}^s \frac{1}{2} m_\ell \dot{q}_\ell^2 = T(\dot{\mathbf{q}}) \quad \text{et} \quad U = U(\mathbf{q}) \quad (2.91)$$

### Systèmes invariants par translation

Soit  $\mathbf{g}_\sigma$  la translation de vecteur  $\mathbf{n} = (n_1, \dots, n_s)$  dans l'espace de configuration telle que

$$\mathbf{g}_\sigma : \mathbf{q} \mapsto \mathbf{q} + \sigma \mathbf{n} \quad (2.92)$$

c'est-à-dire

$$g_\sigma^\ell(\mathbf{q}) := q_\ell + \sigma n_\ell$$

il est bien évident que  $\mathbb{T} = \{\mathbf{g}_\sigma, \sigma \in \mathbb{R}\}$  forme un groupe de symétrie. Pour que  $\mathcal{L}$  soit invariant sous ce groupe de symétrie, il suffit que  $U$  le soit, c'est-à-dire

$$\forall \sigma \in \mathbb{R}, U(\mathbf{q} + \sigma \mathbf{n}) = U(\mathbf{q}) \quad (2.93)$$

l'espace est alors dit homogène.

Si tel est le cas, le théorème de Noether affirme que

$$I(\mathbf{q}, \dot{\mathbf{q}}; t) = \sum_{\ell=1}^s \frac{\partial \mathcal{L}}{\partial \dot{q}_\ell} \frac{dg_\sigma^\ell(\mathbf{q})}{d\sigma} \Bigg|_{\sigma=0} = cste \quad (2.94)$$

de façon triviale

$$\frac{dg_\sigma^\ell(\mathbf{q})}{d\sigma} \Bigg|_{\sigma=0} = n_\ell \quad \text{et} \quad \frac{\partial \mathcal{L}}{\partial \dot{q}_\ell} = m_\ell \dot{q}_\ell \quad (2.95)$$

Si l'espace est homogène nous avons donc

$$\sum_{i=1}^s m_i \dot{q}_i n_i = cste \quad (2.96)$$

Cette constante est clairement la projection sur  $\mathbf{n}$  du vecteur quantité de mouvement étendu sur  $\mathbb{R}^{3N}$  : en posant  $M\mathbf{v} = (m(1)\dot{\mathbf{r}}_1, \dots, m(N)\dot{\mathbf{r}}_N)$ , nous pouvons écrire

$$M\mathbf{v} \cdot \mathbf{n} = cste$$

qui se résume symboliquement par

$$\boxed{\text{Espace homogène}} \Rightarrow \boxed{\text{Conservation de la quantité de mouvement}} \quad (2.97)$$

## Systèmes invariants par rotation

Considérons à présent un système composé d'une particule de masse  $m$  évoluant dans un potentiel présentant une symétrie de rotation autour d'un axe fixe. Pour fixer les idées supposons que l'espace de configuration de cette particule soit  $\mathbb{R}^3$  repéré par la base cartésienne  $(\vec{e}_x, \vec{e}_y, \vec{e}_z)$  et que l'axe de la rotation soit  $(O, \vec{e}_z)$ . Les trois coordonnées généralisées sont donc les trois coordonnées cartésiennes de la particule  $\mathbf{q} = \vec{r} = (x, y, z)$  et la symétrie est telle que

$$\mathbf{g}_\sigma(\mathbf{q}) = \mathbf{R}_\sigma \mathbf{q} \quad \text{où} \quad \mathbf{R}_\sigma = \begin{bmatrix} +\cos \sigma & +\sin \sigma & 0 \\ -\sin \sigma & +\cos \sigma & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.98)$$

est la matrice de rotation autour de  $(O, \vec{e}_z)$ . Le potentiel est invariant par rotation autour de cet axe, ce qui signifie que  $U(\mathbf{q}) = U(\mathbf{R}_\sigma \mathbf{q})$ . L'énergie cinétique est ici simplement donnée par  $T(\dot{\mathbf{q}}) = \frac{1}{2} m \dot{\mathbf{q}} \cdot \dot{\mathbf{q}}$ . Le lagrangien est donc

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = T(\dot{\mathbf{q}}) - U(\mathbf{q}).$$

Il est supposé invariant par l'application induite par  $\mathbf{g}_\sigma$  dans l'espace des phases soit  $\mathbf{g}_\sigma^* : (\mathbf{q}, \dot{\mathbf{q}}) \mapsto (\mathbf{R}_\sigma \mathbf{q}, \mathbf{R}_\sigma \dot{\mathbf{q}})$ . En vertu du théorème de Noether, l'invariance du lagrangien sous cette symétrie se traduit par la conservation de

$$I(\mathbf{q}, \dot{\mathbf{q}}) = \left. \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \cdot \frac{d\mathbf{g}_\sigma(\mathbf{q})}{d\sigma} \right|_{\sigma=0} = \left. \frac{\partial T}{\partial \dot{\mathbf{q}}} \cdot \frac{d(\mathbf{R}_\sigma \mathbf{q})}{d\sigma} \right|_{\sigma=0} \quad (2.99)$$

les deux termes de ce produit scalaire sont facilement calculables, tout d'abord dans ce cas très simple  $\frac{\partial T}{\partial \dot{\mathbf{q}}} = m\dot{\mathbf{q}}$ , puis

$$\left. \frac{d(\mathbf{R}_\sigma \mathbf{q})}{d\sigma} \right|_{\sigma=0} = \begin{bmatrix} -\sin \sigma & +\cos \sigma & 0 \\ -\cos \sigma & -\sin \sigma & 0 \\ 0 & 0 & 0 \end{bmatrix}_{\sigma=0} \quad \mathbf{q} = \mathbf{q} \wedge \vec{u}_z$$

Soit finalement

$$I(\mathbf{q}, \dot{\mathbf{q}}) = m\dot{\mathbf{q}} \cdot (\mathbf{q} \wedge \vec{u}_z) = \vec{u}_z \cdot (m\dot{\mathbf{q}} \wedge \mathbf{q}) = -L_z. \quad (2.100)$$

la composante  $z$  du moment cinétique de  $m$  est conservée.

Si le lagrangien est invariant par rotation autour des deux autres axes, le système est alors qualifié d'isotrope et l'on a conservation du moment cinétique total. La partie cinétique du lagrangien est toujours isotrope, un système sera isotrope si son énergie potentielle l'est : un potentiel radial  $U = U(|\mathbf{q}|)$  en est un exemple classique.

$$\boxed{\text{Système isotrope}} \Rightarrow \boxed{\begin{array}{c} \text{Conservation} \\ \text{du} \\ \text{moment cinétique} \end{array}} \quad (2.101)$$

## Systèmes invariants par translation dans le temps

Dans le cas général, le lagrangien  $\mathcal{L}$  dépend du temps, au prix d'un petit élargissement de l'espace des phases, le temps peut devenir une coordonnée généralisée supplémentaire.

En posant  $q_0 = t$  et

$$\forall \mu = 0, \dots, s \quad \dot{q}_\mu = \frac{dq_\mu}{d\lambda} \quad (2.102)$$

$\lambda$  étant un paramètre décrivant l'évolution du système, on peut introduire un lagrangien étendu

$$\mathcal{L}^{(e)}(q_0, \dots, q_s, \dot{q}_0, \dots, \dot{q}_s) := \mathcal{L}\left(q_1, \dots, q_s, \frac{\dot{q}_1}{\dot{q}_0}, \dots, \frac{\dot{q}_s}{\dot{q}_0}, q_0\right) \dot{q}_0 \quad (2.103)$$

Ce nouveau lagrangien ne dépend alors plus du paramètre d'évolution.

Le mouvement s'effectue maintenant de telle manière que

$$\forall \mu = 0, \dots, s \quad \frac{d}{d\lambda} \left( \frac{\partial \mathcal{L}^{(e)}}{\partial \dot{q}_\mu} \right) - \frac{\partial \mathcal{L}^{(e)}}{\partial q_\mu} = 0 \quad (2.104)$$

en effet, la condition d'extremum pour l'action s'écrit

$$\delta S = \int_{t_1}^{t_2} \mathcal{L} dt = 0 \Leftrightarrow \int_{\lambda_1}^{\lambda_2} \mathcal{L} \frac{dt}{d\lambda} d\lambda = 0 \Leftrightarrow \int_{\lambda_1}^{\lambda_2} \mathcal{L}^{(e)} d\lambda = 0 \quad (2.105)$$

qui est bien équivalente aux équations (2.104) dans le cas de notre système conservatif.

Introduisons à présent le groupe des translations temporelles défini par le générateur

$$g_\sigma : (q_0, \dots, q_s) \mapsto (q_0 + \sigma, \dots, q_s) \quad (2.106)$$

c'est-à-dire

$$g_\sigma^0(\mathbf{q}) = q_0 + \sigma \quad \text{et} \quad \forall 1 \leq i \leq s, \quad g_\sigma^i(\mathbf{q}) = q_i$$

Si  $g_\sigma$  est une symétrie du système, c'est-à-dire qu'en fait le lagrangien est indépendant du temps, alors le théorème de Noether (étendu) indique la conservation de la quantité

$$I := \sum_{\mu=0}^s \frac{\partial \mathcal{L}^{(e)}}{\partial \dot{q}_\mu} \frac{dg_\sigma^\mu(\mathbf{q})}{d\sigma} \Bigg|_{\sigma=0} \quad (2.107)$$

la forme particulièrement simple du générateur des translations temporelles ne laisse persister dans cette relation que le premier terme de la somme

$$I = \frac{\partial \mathcal{L}^{(e)}}{\partial \dot{q}_0} = \mathcal{L} + \dot{q}_0 \sum_{i=1}^s \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{d}{d\dot{q}_0} \left( \frac{\dot{q}_i}{\dot{q}_0} \right) \quad (2.108)$$

soit

$$I = \mathcal{L} - \sum_{i=1}^s \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{\dot{q}_i}{\dot{q}_0} = \mathcal{L} - \sum_{i=1}^s \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{dq_i}{dt} \quad (2.109)$$

dans cette dernière relation  $\dot{q}_i$  peut s'entendre comme  $dq_i/dt$ , le lagrangien  $\mathcal{L}$  ne dépend en effet pas du temps dans le cadre de notre symétrie, on peut donc choisir  $\lambda = t$  dans  $\mathcal{L}$ .

Dans le cas conservatif, nous avons  $\mathcal{L} = T - U$  ainsi (en coordonnées cartésiennes)

$$\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = m_i \dot{q}_i \quad (2.110)$$

et finalement

$$I = \sum_{i=1}^s \frac{m_i}{2} \dot{q}_i^2 - U - \sum_{i=1}^s m_i \dot{q}_i^2 = - \sum_{i=1}^s \frac{m_i}{2} \dot{q}_i^2 - U := -H \quad (2.111)$$

un lagrangien conservatif indépendant du temps assure donc la conservation de l'énergie mécanique totale du système.

$$\boxed{\text{lagrangien conservatif}} \Rightarrow \boxed{\begin{array}{c} \text{Conservation} \\ \text{de} \\ \text{l'énergie totale} \end{array}} \quad (2.112)$$

Les trois résultats que nous venons d'obtenir sont bien connus des mécaniciens et depuis fort longtemps<sup>10</sup>, mais ils se révèlent ici intimement liés à un seul et même résultat plus fondamental. Lorsque la mécanique analytique se révélera être le fondement de la physique moderne, le théorème de Noether s'affirmera être un puissant outil générateur de concepts.

<sup>10</sup> Christiaan Huygens au XVII<sup>e</sup> siècle évoquait déjà la conservation de l'intégrale des forces vives (énergie) pour les systèmes isolés.

## Formulation hamiltonienne

Comme nous l'avons vu dans le chapitre 1 historique et introductif, les équations de Hamilton ont été écrites par Lagrange au début du XIX<sup>e</sup> siècle. Elles correspondent à un formalisme d'ordre 1 en temps là où le formalisme lagrangien est d'ordre 2. Les équations de la physique sous leur forme la plus générale sont écrites à partir de ce formalisme d'ordre 1 : c'est une partie de son intérêt.

### 3.1 Equations de Hamilton

On considère un système lagrangien, c'est-à-dire un système pour lequel il existe un champ scalaire  $\mathcal{L} = \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$  tel que la solution des équations de Lagrange

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} = 0 \quad (3.1)$$

décrit le mouvement du système au cours du temps :  $\{\mathbf{q}(t), t \in \mathbb{R}\}$ .

Si le lagrangien dépend du temps et que l'on peut choisir un autre paramètre  $\lambda$  pour décrire l'évolution de  $\mathbf{q}$ , on peut toujours étendre le lagrangien et se ramener au problème initial dans un espace de configuration étendu (voir (2.103) et (2.104)).

Les équations de Lagrange représentent  $s$  équations différentielles du second ordre en caractérisant les  $s$  composantes du vecteur de coordonnées généralisées  $\mathbf{q}$ . L'origine du formalisme hamiltonien naît dans la volonté de transformer ce système en  $2s$  équations du premier ordre, caractérisant toujours les  $s$  composantes de  $\mathbf{q}$  mais aussi celles d'un nouveau vecteur  $\mathbf{p}$  indépendant que  $\mathbf{q}$ , appelé impulsion généralisée.

Dans le formalisme de Lagrange les variables  $\mathbf{q}$  et  $\dot{\mathbf{q}}$  sont indépendantes. C'est cette propriété qui va être mise à profit pour pratiquer une transformation de Legendre du Lagrangien.

Posons

$$\mathbf{p} := \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \in \mathbb{R}^s \quad (3.2)$$

pour préserver l'indépendance entre les  $\mathbf{q}$  et  $\dot{\mathbf{q}}$ , la définition des  $\mathbf{p}$  doit se traduire par les conditions

$$1 \leq \alpha, \beta \leq s \quad \frac{\partial p_\alpha}{\partial q_\beta} = 0 \quad \text{et} \quad \frac{\partial p_\alpha}{\partial p_\beta} = \delta_\alpha^\beta \quad (3.3)$$

Nous pouvons alors procéder au changement de variable via les formes différentielles.

Puisque  $\mathcal{L} = \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$  nous avons

$$d\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \mathbf{q}} \cdot d\mathbf{q} + \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \cdot d\dot{\mathbf{q}} \quad (3.4)$$

En utilisant les nouvelles variables  $\mathbf{p}$  dans les équations de Lagrange, ces dernières s'écrivent

$$\dot{\mathbf{p}} = \frac{\partial \mathcal{L}}{\partial \mathbf{q}} \quad (3.5)$$

on peut alors introduire ce résultat dans la relation (3.4) pour obtenir

$$\begin{aligned} d\mathcal{L} &= \dot{\mathbf{p}} \cdot d\mathbf{q} + \mathbf{p} \cdot d\dot{\mathbf{q}} \\ &= \dot{\mathbf{p}} \cdot d\mathbf{q} + d(\mathbf{p} \cdot \dot{\mathbf{q}}) - \dot{\mathbf{q}} \cdot d\mathbf{p} \end{aligned} \quad (3.6)$$

que l'on peut toujours écrire sous la forme

$$d(\mathcal{L} - \mathbf{p} \cdot \dot{\mathbf{q}}) = \dot{\mathbf{p}} \cdot d\mathbf{q} - \dot{\mathbf{q}} \cdot d\mathbf{p} \quad (3.7)$$

Cette dernière relation montre alors d'une part que le champ scalaire <sup>1</sup>

$$\mathcal{H} = \mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{L} \quad (3.8)$$

ne dépend que des variables indépendantes  $\mathbf{p}$  et  $\mathbf{q}$  et que d'autre part

$$\frac{\partial \mathcal{H}}{\partial \mathbf{q}} = -\dot{\mathbf{p}} \quad \text{et} \quad \frac{\partial \mathcal{H}}{\partial \mathbf{p}} = +\dot{\mathbf{q}} \quad (3.9)$$

<sup>1</sup> Attention au changement de signe introduit pour des raisons de commodité future.

La fonction  $\mathcal{H} = \mathcal{H}(\mathbf{q}, \mathbf{p})$  est appelée fonction de Hamilton ou hamiltonien, les  $2s$  équations différentielles du premier ordre (3.9) sont appelées équations canoniques de Hamilton. Les variables  $\mathbf{q}$  et  $\mathbf{p}$  sont dites canoniquement conjuguées par les relations (3.3).

Remarques :

- La transformation (3.8) élimine les variables  $\dot{\mathbf{q}}$  au profit des variables  $\mathbf{q}$ . Il s'agit d'une transformation de Legendre :

$$\mathcal{H}(\mathbf{q}, \mathbf{p}) = \dot{\mathbf{q}} \cdot \mathbf{p} - \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$$

- Si le lagrangien dépend explicitement du temps  $t$ , ces relations s'étendent !
- Comme nous l'avons vu dans le chapitre 1 section 1.4.2, historiquement cette fonction a été introduite par Lagrange au tout début du XIX<sup>e</sup> siècle, avant même la naissance de Hamilton. La notation  $\mathcal{H}$  de Lagrange était un hommage à Huygens, la postérité en a décidé autrement...

## 3.2 Nature de la fonction de Hamilton

Dans le contexte, simple mais assez général en mécanique théorique, où nous nous sommes placés pour introduire la fonction de Hamilton, le lagrangien s'écrit

$$\mathcal{L} = \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = T(\dot{\mathbf{q}}) - U(\mathbf{q}) \quad (3.10)$$

Pour un système de particules de  $k = 1, \dots, N$  particules de masses  $m(k)$  sans liaison, les coordonnées cartésiennes permettent d'écrire simplement l'énergie cinétique en reprenant les notations définies en (2.91) pour les masses

$$T(\dot{\mathbf{q}}) = \sum_{\alpha=1}^{s=3N} \frac{1}{2} m_{\alpha} \dot{q}_{\alpha}^2 \quad (3.11)$$

Sous ces conditions les impulsions généralisées deviennent

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = \frac{\partial T}{\partial \dot{\mathbf{q}}} \quad \text{soit } p_{\alpha} = m_{\alpha} \dot{q}_{\alpha} \quad (3.12)$$

nous constatons donc que dans ce cas théorique très idéalisé, impulsions généralisées et quantités de mouvements se confondent, *ce n'est toutefois pas une règle générale* comme nous le verrons dans certaines situations.

La définition générale du hamiltonien  $\mathcal{H} = \mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{L}$  s'écrit donc en coordonnées cartésiennes

$$\mathcal{H} = \sum_{\alpha=1}^s m_{\alpha} \dot{q}_{\alpha}^2 - \sum_{\alpha=1}^s \frac{1}{2} m_{\alpha} \dot{q}_{\alpha}^2 + U(\mathbf{q}) \quad (3.13)$$

soit enfin

$$\mathcal{H} = \sum_{\alpha=1}^s \frac{1}{2} m_{\alpha} \dot{q}_{\alpha}^2 + U(\mathbf{q}) = T + U \quad (3.14)$$

Dans le cas d'un système conservatif sans liaison interne non holonome, le hamiltonien s'interprète donc comme l'énergie mécanique totale du système. Il ne faut pas chercher plus loin la raison de l'hommage de Lagrange à Huygens qui avait constaté la conservation de l'énergie dans un système isolé, propriété que nous avons retrouvée grâce au théorème de Noether.

### 3.3 Interprétation des équations de Hamilton

Les équations de Lagrange sont équivalentes aux équations du mouvement et permettent de les interpréter «géométriquement» via le principe de moindre action.

Les équations de Hamilton en font nécessairement autant ! Dans le cadre très simple des coordonnées cartésiennes et sans contrainte d'aucune sorte, nous venons de constater que

$$\mathcal{H} = T + U = \sum_{\alpha=1}^s \frac{p_{\alpha}^2}{2m_{\alpha}} + U(\mathbf{q}) \quad (3.15)$$

La première équation de Hamilton donne

$$\dot{\mathbf{q}} = + \frac{\partial \mathcal{H}}{\partial \mathbf{p}} \quad \text{soit} \quad \dot{q}_{\alpha} = \frac{p_{\alpha}}{m_{\alpha}} \quad (3.16)$$

c'est-à-dire la définition de l'impulsion généralisée  $p_{\alpha}$ . La seconde équation s'écrit quant à elle

$$\dot{\mathbf{p}} = - \frac{\partial \mathcal{H}}{\partial \mathbf{q}} \quad \text{soit} \quad \dot{p}_{\alpha} = - \frac{\partial U}{\partial q_{\alpha}} \quad (3.17)$$

dans le cas conservatif, la somme des forces appliquées sur une particule s'écrit comme l'opposé du gradient de la somme des énergies potentielles correspondantes. Cette deuxième équation de Hamilton est donc équivalente

au principe fondamental de la dynamique, et donc aux équations du mouvement. Tout ceci n'est qu'une histoire d'écriture et donc de formalisme, mais tout est là!

Une simple vérification permet en outre de se convaincre que les équations de Hamilton forment, tout comme les équations de Lagrange, une condition nécessaire d'extremum pour l'action.

Le formalisme hamiltonien lorsqu'il existe, est donc complètement équivalent au formalisme lagrangien. Son intérêt n'est cependant pas à négliger, il apparaîtra clairement plus loin.

La nature et la mise œuvre du formalisme Hamiltonien que nous venons de mettre en évidence dans un contexte très simplifié se généralise à chaque cas d'étude. Mais il ne faut sauter aucune étape et procéder avec ordre et rigueur :

- on écrit dans un premier temps le lagrangien du système en fonction des variables  $(\mathbf{q}, \dot{\mathbf{q}})$ ;
- on procède alors à la transformation de Legendre qui se matérialise par le changement de variable  $\mathbf{p} := \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}$
- on écrit ensuite le hamiltonien  $\mathcal{H} = \mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{L}$  en terme des variables  $(\mathbf{q}, \mathbf{p})$ .

Ecrire un Hamiltonien dans d'autres variables, ou en laissant trainer des vitesses généralisées dans son expression n'a donc pas de sens!

Considérons l'exemple des particules chargées en mouvement dans un champ électromagnétique. Nous avons construit lagrangien d'un tel système (voir 2.75 et 2.76), en reprenant les mêmes notations nous avons  $\mathbf{q} = (\vec{r}_1, \dots, \vec{r}_N) \in \mathbb{R}^{3N}$

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = \sum_{k=1}^N \frac{1}{2} m_k \dot{r}_k^2 + e_k \dot{r}_k \cdot \vec{A}_k - e_k V_k$$

le terme en  $e_k \dot{r}_k \cdot \vec{A}_k$  associé aux forces magnétiques vient ici «perturber» la situation dans laquelle les forces dérivent d'un potentiel ne dépendant que des positions. Pour «passer» dans le formalisme hamiltonien, il faut tout d'abord déterminer les impulsions généralisées. Leur définition se matérialise ici par la relation

$$\vec{p}_i := \frac{\partial \mathcal{L}}{\partial \dot{r}_i} = m_i \dot{r}_i + e_i \vec{A}_i \quad (3.18)$$

un terme supplémentaire qui vient donc s'ajouter à la quantité de mouvement dans la définition de l'impulsion. Pour obtenir le hamiltonien de ce système on écrit donc sa définition

$$\begin{aligned}
 \mathcal{H} &= \sum_{k=1}^N \vec{p}_k \cdot \dot{\vec{r}}_k - \mathcal{L} \\
 &= \sum_{k=1}^N \left[ \left( m_k \dot{\vec{r}}_k + e_k \vec{A}_k \right) \cdot \dot{\vec{r}}_k - \frac{1}{2} m_k \dot{\vec{r}}_k^2 - e_k \dot{\vec{r}}_k \cdot \vec{A}_k + e_k V_k \right] \\
 &= \sum_{k=1}^N \left[ \frac{1}{2} m_k \dot{\vec{r}}_k^2 + e_k V_k \right]
 \end{aligned} \tag{3.19}$$

Mais ce n'est surtout pas terminé car cette expression ne fait pas intervenir les bonnes variables! On doit éliminer la variable  $\dot{\mathbf{q}}$  matérialisée par la présence des  $\dot{\vec{r}}_k$  au profit de l'impulsion généralisée  $\mathbf{p} = (\vec{p}_1, \dots, \vec{p}_N) \in \mathbb{R}^{3N}$ . On utilise pour cela l'expression (3.18) pour écrire la bonne expression du hamiltonien, dans les bonnes variables

$$\mathcal{H} = \mathcal{H}(\mathbf{q}, \mathbf{p}) = \sum_{k=1}^N \frac{\left[ \vec{p}_k - e_k \vec{A}(\vec{r}_k) \right]^2}{2m_k} + e_k V(\vec{r}_k) \tag{3.20}$$

la première équation de Hamilton est alors comme dans le cas conservatif une «définition» de la quantité de mouvement, et la deuxième redonne les équations du mouvement. En utilisant la relation classique

$$\vec{\nabla}(\vec{u} \cdot \vec{u}) = 2 \left[ \vec{u} \wedge (\vec{\nabla} \wedge \vec{u}) + (\vec{u} \cdot \vec{\nabla}) \vec{u} \right] \tag{3.21}$$

quelques lignes de calcul et un retour aux champs permettent en effet de se rendre compte que

$$\dot{\mathbf{p}} = - \frac{\partial \mathcal{H}}{\partial \mathbf{q}} \implies m_k \frac{d^2 \vec{r}_k}{dt^2} = e_k \left[ \vec{E}(\vec{r}_k) + \dot{\vec{r}}_k \wedge \vec{B}(\vec{r}_k) \right] \tag{3.22}$$

bien connue des spécialistes !

Comme nous venons de le voir, le formalisme hamiltonien est d'ordre 1 en temps : le hamiltonien ne contient pas de dérivées temporelles de ses variables et les équations de Hamilton ne font intervenir que des dérivées premières de variables canoniques  $(\mathbf{q}, \mathbf{p})$ .

Le formalisme lagrangien était lui d'ordre 2, mais il ne faut pas s'y méprendre : les deux formalismes lagrangien et hamiltonien ne sont pour le moment que deux méthodes subtiles permettant d'obtenir efficacement les équations du mouvement dans de très nombreux contextes là où l'équation de Newton a bien du mal!

Le principe de moindre action est aussi l'un des avantages associés à ces formalismes, mais ce n'est que plus tard que nous en comprendrons la portée car pour le moment, nous n'avons traité que les équations de la mécanique classique, mais patience!

Il est temps maintenant de changer de domaine en nous intéressant non plus à l'obtention des équations du mouvement mais aux propriétés de leurs solutions. Cette activité est devenue en physique théorique et même en mathématique un domaine à part entière : les systèmes dynamiques hamiltoniens. Initiés par Lagrange en personne lorsqu'il inventa ses parenthèses, puis aidé de son élève avec les crochets de Poisson, instaurés plus tard en domaine de recherche par Poincaré, les systèmes dynamiques sont devenus depuis l'objet d'études merveilleuses.

Nous ne les aborderons que très modestement et en restant dans le cadre des systèmes hamiltoniens.

## 3.4 Systèmes dynamiques hamiltoniens

L'écriture des équations du mouvement sous forme hamiltonienne met en évidence un système d'équations différentielles qu'il est intéressant d'étudier. Non pas que l'on sache le résoudre, sauf dans des cas exceptionnels, mais parce que l'on peut très souvent obtenir de très nombreuses informations sur la dynamique locale ou globale. Il faut pour cela utiliser des outils très fins introduits par Henri Poincaré à la fin du XIX<sup>e</sup> siècle sur la base des travaux de Lagrange, Hamilton ou Jacobi. Ces travaux fondateurs initièrent un domaine des mathématiques et de la physique théorique, la théorie des systèmes dynamiques. Nous ne présenterons ici que quelques éléments très introductifs du vocabulaire et des méthodes de cette théorie. On pourra trouver de très nombreux compléments mathématiques dans les livres d'Arnold ([2],[1]) ou physiques dans ceux de Gignoux et Silvestre-Brac ([7],[8]).

### 3.4.1 Crochets de Poisson

#### Construction

Comme nous l'avons vu dans le premier chapitre, les crochets de Poisson correspondent à l'étape ultime du travail de Lagrange, qui mit un point d'honneur à généraliser ses parenthèses, ou plutôt leur inverse...

Comment faire apparaître des opérateurs antisymétriques à partir des équations de Hamilton? Rien de plus simple : il suffit de rendre symétriques

ces dernières... Partant des équations de Hamilton (3.9), on peut toujours écrire

$$1 \leq \alpha \leq s \quad \left\{ \begin{array}{l} \dot{p}_\alpha = \sum_{\beta=1}^s - \frac{\partial \mathcal{H}}{\partial q_\beta} \times \delta_{\alpha\beta} \\ \dot{q}_\alpha = \sum_{\beta=1}^s + \frac{\partial \mathcal{H}}{\partial p_\beta} \times \delta_{\alpha\beta} \end{array} \right.$$

Les propriétés imposées aux coordonnées hamiltoniennes font que

$$\forall 1 \leq \alpha, \beta \leq s, \quad \frac{\partial q_\alpha}{\partial q_\beta} = \frac{\partial p_\alpha}{\partial p_\beta} = \delta_{\alpha\beta} \text{ et } \frac{\partial q_\alpha}{\partial p_\beta} = \frac{\partial p_\alpha}{\partial q_\beta} = 0$$

on peut donc astucieusement écrire

$$\begin{aligned} \dot{p}_\alpha &= \sum_{\beta=1}^s \left( 0 \times \frac{\partial \mathcal{H}}{\partial p_\beta} - \delta_{\alpha\beta} \times \frac{\partial \mathcal{H}}{\partial q_\beta} \right) = \sum_{\beta=1}^s \left( \frac{\partial p_\alpha}{\partial q_\beta} \frac{\partial \mathcal{H}}{\partial p_\beta} - \frac{\partial p_\alpha}{\partial p_\beta} \frac{\partial \mathcal{H}}{\partial q_\beta} \right) \\ \dot{q}_\alpha &= \sum_{\beta=1}^s \left( \delta_{\alpha\beta} \times \frac{\partial \mathcal{H}}{\partial p_\beta} - 0 \times \frac{\partial \mathcal{H}}{\partial q_\beta} \right) = \sum_{\beta=1}^s \left( \frac{\partial q_\alpha}{\partial q_\beta} \frac{\partial \mathcal{H}}{\partial p_\beta} - \frac{\partial q_\alpha}{\partial p_\beta} \frac{\partial \mathcal{H}}{\partial q_\beta} \right) \end{aligned}$$

Pour une paire de champ scalaires dérivables  $f(\mathbf{q}, \mathbf{p})$  et  $g(\mathbf{q}, \mathbf{p})$ , on peut alors définir le crochet de Poisson

$$\{f, g\} = \frac{\partial f}{\partial \mathbf{q}} \cdot \frac{\partial g}{\partial \mathbf{p}} - \frac{\partial f}{\partial \mathbf{p}} \cdot \frac{\partial g}{\partial \mathbf{q}} = \sum_{\beta=1}^s \left( \frac{\partial f}{\partial q_\beta} \frac{\partial g}{\partial p_\beta} - \frac{\partial f}{\partial p_\beta} \frac{\partial g}{\partial q_\beta} \right)$$

et les équations du mouvement deviennent alors belles et symétriques sous la forme

$$\dot{\mathbf{q}} = \{\mathbf{q}, \mathcal{H}\} \text{ et } \dot{\mathbf{p}} = \{\mathbf{p}, \mathcal{H}\}$$

Cette symétrie implique en fait qu'il n'y a pas deux équations du mouvement, mais une seule considérée pour une fonction quelconque. Soit en effet  $\varphi = \varphi(\mathbf{q}, \mathbf{p}, t)$ , une fonction de classe  $C^1$  des impulsions, des coordonnées généralisées et du temps. Nous avons évidemment

$$\frac{d\varphi}{dt} = \frac{\partial \varphi}{\partial t} + \frac{\partial \varphi}{\partial \mathbf{q}} \cdot \dot{\mathbf{q}} + \frac{\partial \varphi}{\partial \mathbf{p}} \cdot \dot{\mathbf{p}} \tag{3.23}$$

en utilisant les équations de Hamilton nous aurons

$$\frac{d\varphi}{dt} = \frac{\partial\varphi}{\partial t} + \frac{\partial\varphi}{\partial\mathbf{q}} \cdot \frac{\partial\mathcal{H}}{\partial\mathbf{p}} + \frac{\partial\varphi}{\partial\mathbf{p}} \cdot \left( -\frac{\partial\mathcal{H}}{\partial\mathbf{q}} \right) \quad (3.24)$$

et donc finalement

$$\frac{d\varphi}{dt} = \frac{\partial\varphi}{\partial t} + \{\varphi, \mathcal{H}\} \quad (3.25)$$

C'est cette équation qui traduit à elle seule le mouvement.

Elle qui constitue ce que l'on pourrait appeler l'*équation fondamentale de la mécanique classique*. La dénomination fondamentale n'est pas usurpée! Nous verrons plus loin comment cette équation devient une sorte de prototype des équations de la physique, à commencer dans le chapitre 9 section 9.1.3, comment elle peut faire apparaître l'évolution temporelle d'un état en mécanique quantique! Puis, dans le chapitre 10, comment elle se généralise à de nombreux domaines de la physique non dissipative.

### Propriétés du crochet de Poisson

Commençons par de l'algèbre. L'application  $\{, \}$  de  $C^1 \times C^1 \rightarrow C^1$  jouit de nombreuses propriétés intéressantes que l'on vérifiera.

Si  $f$ ,  $g$  et  $h$  sont des champs scalaires alors :

- Le crochet de Poisson anticommute :

$$\{f, g\} = -\{g, f\}$$

- Le crochet de Poisson est une forme bilinéaire :

$$\{f + g, h\} = \{f, h\} + \{g, h\} \quad (3.26)$$

$$\{f, g + h\} = \{f, g\} + \{f, h\} \quad (3.27)$$

Pour tout réel  $\lambda$ , on a  $\{\lambda f, g\} = \{f, \lambda g\} = \lambda \{f, g\}$ . Le crochet est bien une forme car son résultat est un produit scalaire.

- Le crochet de Poisson satisfait l'identité de Jacobi :

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

- Le crochet de Poisson jouit de la propriété des dérivées, que l'on appelle aussi «règle de Leibniz» :

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

- Le crochet de poisson peut servir de gradient :

$$\{\mathbf{q}, f\} = \text{grad}_{\mathbf{p}}(f) = \frac{\partial f}{\partial \mathbf{p}} \quad (3.28)$$

$$\{g, \mathbf{p}\} = \text{grad}_{\mathbf{q}}(g) = \frac{\partial g}{\partial \mathbf{q}} \quad (3.29)$$

- Les propriétés d'indépendance des  $(\mathbf{q}, \mathbf{p})$  définissent les relations de commutation canonique :

$$1 \leq \alpha, \beta \leq s, \quad \begin{cases} \{q_\alpha, q_\beta\} = \{p_\alpha, p_\beta\} = 0 \\ \{q_\alpha, p_\beta\} = \delta_{\alpha\beta} \end{cases} \quad (3.30)$$

Pour les lecteurs ayant oublié certaines définitions de base, il est bon à présent de faire un petit rappel.

**Définition 3.1.** *Une algèbre sur un corps  $\mathbb{K}$ , (ou  $\mathbb{K}$ -algèbre) est un espace vectoriel  $E$  sur  $\mathbb{K}$  muni d'une application bilinéaire.*

Plus explicitement, il s'agit d'un espace vectoriel  $E(\mathbb{K}, +)$  muni d'une loi de composition  $\star$  interne de  $E \times E \rightarrow E$  telle que  $\forall(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in E^3$  et  $\forall(\lambda, \mu) \in \mathbb{K}^2$

$$\begin{aligned} (\mathbf{x} + \mathbf{y}) \star \mathbf{z} &= \mathbf{x} \star \mathbf{z} + \mathbf{y} \star \mathbf{z} \\ \mathbf{x} \star (\mathbf{y} + \mathbf{z}) &= \mathbf{x} \star \mathbf{y} + \mathbf{x} \star \mathbf{z} \\ (\lambda \mathbf{x}) \star (\mu \mathbf{y}) &= \lambda \mu (\mathbf{x} \star \mathbf{y}) \end{aligned}$$

**Définition 3.2.** *On appelle algèbre de Lie toute algèbre dont l'application bilinéaire est anticommutative et vérifie l'identité de Jacobi.*

Ces deux définitions appellent le théorème suivant dont la démonstration vient en vérifiant les axiomes.

**Théorème 3.1.** *L'espace des champs scalaires  $C^1$  des variables  $(\mathbf{p}, \mathbf{q})$  et du temps  $t$  muni de la loi crochet de Poisson est une algèbre de Lie.*

Nous voici à présent dans un contexte favorable !

Mais nous ne poursuivrons pas plus loin dans cette direction laissant la main aux mathématiciens (on pourra consulter avec bénéfice l'ouvrage [22], très complet à ce sujet).

Terminons plutôt par une propriété fort utile aux sages.

**Théorème 3.2.** *Si  $f(\mathbf{p}, \mathbf{q})$  et  $g(\mathbf{p}, \mathbf{q})$  sont deux champs scalaires constants lors de l'évolution temporelle, alors leur crochet de Poisson est aussi constant lors de cette évolution.*

On dit que  $f$ ,  $g$  et  $h = \{f, g\}$  sont des constantes ou des intégrales premières du mouvement

**Démonstration**

Par hypothèse

$$\frac{df}{dt} = \frac{dg}{dt} = 0 \quad (3.31)$$

en appliquant la relation (3.25) on a directement

$$\frac{d\{f, g\}}{dt} = \frac{\partial \{f, g\}}{\partial t} + \{\{f, g\}, \mathcal{H}\} \quad (3.32)$$

La définition du crochet de Poisson permet de voir instantanément que la dérivée partielle agit sur lui comme sur un produit, on a donc

$$\frac{d\{f, g\}}{dt} = \left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\} - \{\mathcal{H}, \{f, g\}\} \quad (3.33)$$

une application de l'identité de Jacobi sur le dernier terme permet alors d'obtenir

$$\begin{aligned} \frac{d\{f, g\}}{dt} &= \left\{ \frac{\partial f}{\partial t}, g \right\} - \{\{\mathcal{H}, f\}, g\} + \{f, \{g, \mathcal{H}\}\} + \left\{ f, \frac{\partial g}{\partial t} \right\} \\ &= \left\{ \frac{\partial f}{\partial t} + \{f, \mathcal{H}\}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} + \{g, \mathcal{H}\} \right\} \end{aligned} \quad (3.34)$$

c'est-à-dire

$$\frac{d\{f, g\}}{dt} = \left\{ \frac{df}{dt}, g \right\} + \left\{ f, \frac{dg}{dt} \right\} \quad (3.35)$$

Ces deux derniers crochets sont bien nuls par hypothèse.  $\square$

<sup>2</sup> Si elle en dépend, ce n'est peut être pas bien grave. On peut toujours essayer d'augmenter la dimension de l'espace de configuration d'une unité comme nous l'avons fait avec le lagrangien étendu au moment du théorème de Noether...

## L'exemple qui tue!

Une application essentielle des crochets de Poisson permet de faire apparaître ce qu'en mécanique quantique on appelle un propagateur.

Considérons une fonction  $\varphi = \varphi(\mathbf{p}, \mathbf{q})$  qui ne dépend pas explicitement du temps<sup>2</sup>, l'équation fondamentale de la mécanique classique nous dit alors que

$$\dot{\varphi} = \{\varphi, \mathcal{H}\} \quad (3.36)$$

En considérant le crochet de Poisson comme un opérateur on peut toujours écrire

$$\dot{\varphi} = -\{\mathcal{H}, \cdot\} \varphi = -\sum_{\alpha=1}^s \left[ \frac{\partial \mathcal{H}}{\partial q_{\alpha}} \frac{\partial}{\partial p_{\alpha}} - \frac{\partial \mathcal{H}}{\partial p_{\alpha}} \frac{\partial}{\partial q_{\alpha}} \right] \varphi := \mathcal{A} \varphi \quad (3.37)$$

Nous avons ici une équation différentielle linéaire. Si  $\mathcal{H}$  ne dépend pas du temps, l'opérateur  $\mathcal{A}$  n'en dépend pas non plus. En fait pour tous les temps  $t_1$  et  $t_2$ , l'opérateur  $\mathcal{A}$  évalué en  $t = t_1$  commute avec ce même opérateur évalué en  $t = t_2$  :

$$[\mathcal{A}_{t_1}, \mathcal{A}_{t_2}] = \mathcal{A}_{t_1} \mathcal{A}_{t_2} - \mathcal{A}_{t_2} \mathcal{A}_{t_1} = 0$$

La théorie des équations différentielles nous indique alors que la résolvante de l'équation (3.37) s'écrit

$$R(t_1, t_0) = e^{-\{\mathcal{H}, \cdot\}(t_1 - t_0)} \quad (3.38)$$

Ainsi, la solution générale de cette même équation sera

$$\varphi(\mathbf{p}, \mathbf{q})|_t = e^{-(t-t_0)\{\mathcal{H}, \cdot\}} \varphi(\mathbf{p}, \mathbf{q})|_{t=t_0} \quad (3.39)$$

Cette «nouvelle» façon de voir les problèmes de mécanique possède certains avantages :

- En exhibant la résolvante d'un système, on peut étudier certaines propriétés globales de ce système sans expliciter les solutions des équations du mouvement : solutions bornées, périodiques, symétriques ...
- L'opérateur  $R(t, t_0) = e^{-(t-t_0)\{\mathcal{H}, \cdot\}}$  résolvante du système s'interprète comme un opérateur qui déplace le système de l'instant  $t_0$  à l'instant  $t$ , il s'agit d'un opérateur d'évolution temporelle ou, comme on le dit en mécanique quantique, un propagateur. Ceci permet de «voir» la mécanique sous un aspect plus géométrique.

- Dans les systèmes que nous considérons,  $\mathcal{H}$  est l'énergie totale du système, c'est donc cette dernière qui intervient dans l'opérateur d'évolution temporelle. Lorsque nous avons étudié le théorème de Noether, nous avons déjà remarqué le lien étroit qui existe entre le temps et l'énergie, nous avons même établi l'existence de deux autres couples : translations-quantité de mouvement et rotations-moment cinétique. On peut alors montrer dans ce contexte que l'équation (3.36) se généralise à chacun des couples  $(\lambda, A)$  intervenant dans le théorème de Noether de la façon suivante

$$\forall \varphi(\mathbf{p}, \mathbf{q}) \quad \frac{d\varphi}{d\lambda} = \{\varphi, A\} \quad (3.40)$$

si  $A$  est l'énergie,  $\lambda$  est le temps, si  $A$  est le moment cinétique,  $\lambda$  est un angle de rotation, etc.

Comment utiliser de telles notions : un exemple s'impose !

Considérons par exemple l'oscillateur harmonique à une dimension de masse  $m$  et de pulsation  $\omega$ . En variables canoniques, le hamiltonien s'écrit  $\mathcal{H} = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2q^2$ . En prenant  $\varphi(p, q) = q$ ,  $t_0 = 0$  et  $t_1 = t$  on a donc

$$\begin{aligned} q|_t = e^{-\{\mathcal{H}, \cdot\}t} q|_{t=0} &= \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!} t^n [(\{\mathcal{H}, \cdot\})^n q]_{t=0} \\ &= q_0 - t \{\mathcal{H}, q\}|_{t=0} + \frac{1}{2}t^2 \{\mathcal{H}, \{\mathcal{H}, q\}\}|_{t=0} + \dots \end{aligned}$$

Où l'on a noté  $q_0 = q(t=0)$ . Un rapide calcul, que l'on se régale de faire sous forme algébrique en utilisant les propriétés du crochet de Poisson, permet de voir que

$$\begin{aligned} \{\mathcal{H}, q\} &= \left\{ \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2q^2, q \right\} \\ &= \frac{1}{2m} \{p^2, q\} + \frac{1}{2}m\omega^2 \{q^2, q\} = \frac{p}{m} \{p, q\} + m\omega^2q \{q, q\} \end{aligned}$$

les relations de commutation canonique achèvent l'affaire en

$$\{\mathcal{H}, q\} = -\frac{p}{m}$$

Ceci étant fait, on peut calculer

$$\{\mathcal{H}, \{\mathcal{H}, q\}\} = -\frac{1}{m} \{\mathcal{H}, p\} = -\frac{1}{m^2}p \{p, p\} - \omega^2q \{q, p\} = -\omega^2q$$

soit

$$\{\mathcal{H}, p\} = m\omega^2 q \quad (3.41)$$

et ainsi de suite; avec des notations évidentes  $\{\mathcal{H}, q\}_{n+1} = \{\mathcal{H}, \{\mathcal{H}, q\}_n\}$  on obtient

$$\begin{aligned} \{\mathcal{H}, q\}_3 &= \{\mathcal{H}, \{\mathcal{H}, q\}_2\} = -\omega^2 \{\mathcal{H}, q\}_1 \\ \{\mathcal{H}, q\}_4 &= \{\mathcal{H}, \{\mathcal{H}, q\}_3\} = -\omega^2 \{\mathcal{H}, \{\mathcal{H}, q\}_1\} = -\omega^2 \{\mathcal{H}, q\}_2 \end{aligned}$$

soit par récurrence

$$\forall k \in \mathbb{N}^* \quad \begin{cases} \{\mathcal{H}, q\}_{2k+1} = \frac{1}{m} (-1)^{k+1} \omega^{2k} p \\ \{\mathcal{H}, q\}_{2k} = (-1)^k \omega^{2k} q \end{cases}$$

On peut donc expliciter les termes de la série donnant  $q$  pour avoir

$$\begin{aligned} q &= q_0 \left( 1 - \frac{t^2}{2} \omega^2 + \frac{t^4}{4!} \omega^4 - \frac{t^6}{6!} \omega^6 + \dots \right) + \frac{p_0}{m} \left( t - \frac{t^3}{3!} \omega^2 + \frac{t^5}{5!} \omega^4 + \dots \right) \\ &= q_0 \cos \omega t + \frac{1}{\omega} \frac{p_0}{m} \sin \omega t \end{aligned}$$

On retrouve ainsi un résultat bien connu que d'autres méthodes permettraient d'obtenir directement... Il est simplement réconfortant et il nous encourage à utiliser cette méthode dans des contextes où il n'y en a vraisemblablement pas d'autres !

Si l'on veut par exemple obtenir une solution approchée à un ordre donné, disons  $o(t^n)$ , pour un problème plus complexe dont les équations du mouvement sont abjectes, il suffira avec cette méthode de calculer seulement  $n$  crochets de Poisson !

Essayez par exemple de trouver l'expression à l'ordre  $o(t^2)$  d'une fonction  $\phi$  suffisamment dérivable de  $\mathbf{q} = (\theta_1, \theta_2, \theta_3)$  pour le problème décrit sur la figure 2.5 page 45 qui terrorisa les candidats au concours Mines-Ponts en 2008 : sans cette méthode, vous allez écrire les équations du mouvement – en formalisme de Newton tant qu'on y est – puis tenter de bricoler celles-ci afin d'espérer une expression des 3 solutions  $\theta_\alpha(t)$ . Puis espérer que vous aller pouvoir en déduire ce que vous cherchez... Alors qu'en fait il suffit de déterminer le hamiltonien du système et de calculer 2 crochets !

Pour bien enfoncer le clou précisons que  $t$  n'est que l'un des paramètres possible pour exprimer  $\varphi(\mathbf{q}, \mathbf{p})$  grâce à  $\mathcal{H}$ , on peut utiliser en fait n'importe quel couple de Noether  $(\lambda, \Lambda)$  adapté à la situation cf. (3.40). La méthode continue à fonctionner : rappelons la maxime «*Les bons outils font les bons ouvriers*».

### 3.4.2 Le choix des coordonnées

Dans certains cas très particuliers les solutions des équations de Hamilton sont très simples. Par exemple, si le Hamiltonien ne dépend pas de l'une des variables canoniques – on dit alors qu'elle est cyclique – sa coordonnée conjuguée est une intégrale première du mouvement : elle est conservée dans l'évolution temporelle. Par ailleurs, rien n'impose le choix d'un système de coordonnées pour peu qu'il permette de représenter convenablement le système. L'exemple du plan est ici instructif. Les coordonnées généralisées peuvent être cartésiennes  $\mathbf{q} = (q_1, q_2) = (x, y)$  ou polaires  $\mathbf{q} = (r, \theta)$  ces deux systèmes sont strictement équivalents, pourtant dans le cas d'un système soumis à des forces centrales le hamiltonien ne dépend que de  $r$ ; la coordonnée  $\theta$  est donc cyclique. Il semble donc clair, au moins sur cet exemple, que l'existence de coordonnées cycliques dépend intimement du choix du système de représentation.

La question que l'on peut alors légitimement se poser est la suivante : pour un système donné, existe-t-il un jeu de coordonnées généralisées qui soient toutes cycliques et si oui comment faire pour trouver le changement de variable qu'il faut s'empresse d'effectuer pour rendre la dynamique triviale ?

### Notion de transformation canonique

Considérons un système possédant  $s$  degrés de liberté, décrit par des coordonnées généralisées  $\mathbf{q} \in \mathbb{R}^s$  et des impulsions  $\mathbf{p} \in \mathbb{R}^s$  dont l'évolution dynamique est fixée par le hamiltonien  $H(\mathbf{q}, \mathbf{p})$ . Ce système peut éventuellement être non autonome, auquel cas le temps est l'un de ces degré de liberté et c'est d'un hamiltonien étendu qu'il s'agit (voir (2.103) et (2.104)).

Les transformations les plus générales acceptables que l'on puisse effectuer pour passer des variables  $(\mathbf{q}, \mathbf{p})$  aux variables  $(\mathbf{Q}, \mathbf{P})$  requièrent l'existence de deux difféomorphismes

$$\mathbf{Q} : \begin{cases} \mathbb{R}^s \times \mathbb{R}^s \rightarrow \mathbb{R}^s \\ (\mathbf{q}, \mathbf{p}) \mapsto \mathbf{Q}(\mathbf{q}, \mathbf{p}) \end{cases} \quad \text{et} \quad \mathbf{P} : \begin{cases} \mathbb{R}^s \times \mathbb{R}^s \rightarrow \mathbb{R}^s \\ (\mathbf{q}, \mathbf{p}) \mapsto \mathbf{P}(\mathbf{q}, \mathbf{p}) \end{cases} \quad (3.42)$$

La propriété de difféomorphisme est capitale pour préserver les relations d'indépendance des nouvelles variables. Ces nouvelles coordonnées seront acceptables si elles permettent de décrire le mouvement du même système. Pour que ces nouvelles variables soient physiquement intéressantes il faut qu'elle jouissent d'une propriété supplémentaire, qui leur permette d'être toujours hamiltonienne. Nous entendons par là, le fait qu'il existe un champ scalaire  $\mathcal{K}(\mathbf{Q}, \mathbf{P})$  tel que

$$\dot{\mathbf{Q}} = +\frac{\partial \mathcal{K}}{\partial \mathbf{P}} \quad \text{et} \quad \dot{\mathbf{P}} = -\frac{\partial \mathcal{K}}{\partial \mathbf{Q}} \quad (3.43)$$

En tant que solution des équations du mouvement, les anciennes coordonnées correspondaient à un extremum de l'action. La définition (3.8) du hamiltonien implique donc la relation

$$\delta S = \int_{t_1}^{t_2} \delta (\mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{H}) dt = 0 \quad (3.44)$$

Vérifiant les nouvelles équations de Hamilton (3.43), les nouvelles coordonnées doivent quant à elles vérifier la relation

$$\int_{t_1}^{t_2} \delta (\mathbf{P} \cdot \dot{\mathbf{Q}} - \mathcal{K}) dt = 0 \quad (3.45)$$

Si tel est bien le cas, on dira que la transformation

$$(\mathbf{q}, \mathbf{p}) \mapsto (\mathbf{Q}, \mathbf{P}) \quad (3.46)$$

est *canonique*. La réalisation simultanée des équations (3.44) et (3.45) ne signifie pas que les intégrands soient égaux, mais stipule toutefois qu'ils ne peuvent différer au plus que par une dérivée totale par rapport au temps d'un champ scalaire arbitraire  $F$ . Ce champ permet, comme nous allons le voir, de définir complètement la transformation canonique. Ce champ est traditionnellement appelé fonction génératrice de la transformation canonique.

## Les fonctions génératrices

### *Classification générale*

En plus du temps la fonction  $F$  peut dépendre des anciennes et des nouvelles coordonnées, c'est-à-dire en tout de  $4s + 1$  variables pour un système possédant  $s$  degrés de liberté. Le caractère bijectif de la transformation permet néanmoins de restreindre l'étude du cas général à l'un des 4 cas particuliers suivants :

$$F_1(\mathbf{q}, \mathbf{Q}, t), F_2(\mathbf{q}, \mathbf{P}, t), F_3(\mathbf{p}, \mathbf{Q}, t), \text{ ou } F_4(\mathbf{p}, \mathbf{P}, t) \quad (3.47)$$

les circonstances du problème fixant le bon choix !

Examinons pour commencer le premier cas.

Les intégrands des équations (3.45) et (3.44) sont reliés par une condition du type

$$\mathbf{p}\dot{\mathbf{q}} - \mathcal{H} = \mathbf{P}\dot{\mathbf{Q}} - \mathcal{K} + \frac{dF_1}{dt} \quad (3.48)$$

Le fait que  $F_1 = F_1(\mathbf{q}, \mathbf{Q}, t)$  permet d'expliciter sa dérivée totale :

$$\frac{dF_1}{dt} = \frac{\partial F_1}{\partial t} + \frac{\partial F_1}{\partial \mathbf{q}} \cdot \dot{\mathbf{q}} + \frac{\partial F_1}{\partial \mathbf{Q}} \cdot \dot{\mathbf{Q}} \quad (3.49)$$

la relation (3.48) devient donc

$$\dot{\mathbf{q}} \cdot \left( \mathbf{p} - \frac{\partial F_1}{\partial \mathbf{q}} \right) - \dot{\mathbf{Q}} \cdot \left( \mathbf{P} + \frac{\partial F_1}{\partial \mathbf{Q}} \right) - \mathcal{H} + \mathcal{K} - \frac{\partial F_1}{\partial t} = 0 \quad (3.50)$$

les coordonnées  $\mathbf{q}$ , tout comme les  $\mathbf{Q}$ , étant indépendantes entre elles, on obtient la relation (3.50) en imposant

$$\begin{cases} \mathbf{p} = \frac{\partial F_1}{\partial \mathbf{q}} \\ \mathbf{P} = -\frac{\partial F_1}{\partial \mathbf{Q}} \end{cases} \quad \text{et} \quad \mathcal{K} = \mathcal{H} + \frac{\partial F_1}{\partial t} \quad (3.51)$$

En se donnant une fonction génératrice de type 1, i.e.  $F_1(\mathbf{q}, \mathbf{Q}, t)$ , les deux équations vectorielles du haut et du bas de (3.51) sont deux systèmes de chacun  $s$  équations que l'on peut résoudre pour obtenir les  $s$  composantes des vecteurs  $\mathbf{Q} = \mathbf{Q}(\mathbf{q}, \mathbf{p})$  et  $\mathbf{P} = \mathbf{P}(\mathbf{q}, \mathbf{p})$ . Le nouvel hamiltonien  $\mathcal{K}$  s'obtiendra simplement en écrivant l'équation du milieu de (3.51) en prenant soin de ne faire apparaître dans son expression, uniquement les variables  $(\mathbf{Q}, \mathbf{P})$ .

La connaissance de  $F_1$  augmentée des équations (3.51) permet donc de définir complètement le changement de variable à effectuer. Cette certitude est assurée par le caractère bijectif de la transformation.

Que faire à présent si les symétries du problème, ou l'intuition de l'utilisateur, indiquent que les arguments indépendants de la fonction génératrice doivent être les vecteurs  $\mathbf{q}$  et les  $\mathbf{P}$ , c'est-à-dire de type  $F_2$  ?

Comme pour la transformation du lagrangien au hamiltonien qui consistait à passer de  $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$  à  $\mathcal{H}(\mathbf{q}, \mathbf{p})$  en posant

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \quad \text{et} \quad \mathcal{H} = \mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{L}, \quad (3.52)$$

la structure hamiltonienne, que nous avons préservée, permet de procéder de façon analogue en ajustant les signes : on passe de  $F_1(\mathbf{q}, \mathbf{Q})$  à  $F_2(\mathbf{q}, \mathbf{P})$ , en définissant

$$F_2(\mathbf{q}, \mathbf{P}) = F_1(\mathbf{q}, \mathbf{Q}) + \mathbf{P} \cdot \mathbf{Q} \quad (3.53)$$

avec les conditions

$$\left\{ \begin{array}{l} \mathbf{p} = \frac{\partial F_2}{\partial \mathbf{q}} \\ \mathbf{Q} = \frac{\partial F_2}{\partial \mathbf{P}} \end{array} \right. \quad \text{et} \quad \mathcal{K} = \mathcal{H} + \frac{\partial F_2}{\partial t} \quad (3.54)$$

pour les variables et le hamiltonien.

De la même façon, les fonctions  $F_3$  et  $F_4$  sont aussi obtenues par des transformations de Legendre appropriées sur la fonction  $F_1$ , on trouve

$$F_3(\mathbf{p}, \mathbf{Q}) = F_1(\mathbf{q}, \mathbf{Q}) - \mathbf{p} \cdot \mathbf{q} \quad (3.55)$$

associée à

$$\left\{ \begin{array}{l} \mathbf{q} = -\frac{\partial F_3}{\partial \mathbf{p}} \\ \mathbf{P} = -\frac{\partial F_3}{\partial \mathbf{Q}} \end{array} \right. \quad \text{et} \quad \mathcal{K} = \mathcal{H} + \frac{\partial F_3}{\partial t} \quad (3.56)$$

pour  $F_4$  il faut appliquer une double transformation sur  $F_1$

$$F_4(\mathbf{p}, \mathbf{P}) = F_1(\mathbf{q}, \mathbf{Q}) + \mathbf{P} \cdot \mathbf{Q} - \mathbf{p} \cdot \mathbf{q} \quad (3.57)$$

associée à

$$\left\{ \begin{array}{l} \mathbf{q} = -\frac{\partial F_4}{\partial \mathbf{p}} \\ \mathbf{Q} = \frac{\partial F_4}{\partial \mathbf{P}} \end{array} \right. \quad \text{et} \quad \mathcal{K} = \mathcal{H} + \frac{\partial F_4}{\partial t} \quad (3.58)$$

La méthode proposée ici pour obtenir des fonctions génératrices est assez abstraite, elle permet de traiter tous les cas en peu de lignes et de façon formelle. Dans la pratique, en fonction du problème à traiter, on se donne souvent une fonction génératrice que l'on construit à partir de celui-ci. On en déduit alors les nouvelles variables. Il faut alors généralement vérifier que les équations obtenues sont bien canoniques car le caractère bijectif du chagement de variable n'est pas forcément garanti, étant donné qu'on ne l'impose nulle part. Donnons quelques exemples ci-dessous.

*Exemples*

Soit la fonction génératrice

$$F_2(\mathbf{q}, \mathbf{P}) = \mathbf{q} \cdot \mathbf{P} \quad (3.59)$$

comme  $F_2$  est explicitement indépendante du temps, le nouvel hamiltonien est égal à l'ancien  $\mathcal{K} = \mathcal{H}$ . De plus, les équations de la transformation de type 2 fournissent

$$\mathbf{p} = \frac{\partial F_2}{\partial \mathbf{q}} = \mathbf{P} \quad \text{et} \quad \mathbf{Q} = \frac{\partial F_2}{\partial \mathbf{P}} = \mathbf{q} \quad (3.60)$$

La fonction génératrice (3.59) engendre donc l'identité!

Il faut donc se méfier de l'apparente complexité d'une fonction génératrice...

La fonction génératrice mélangeait des positions et des impulsions généralisées, essayons maintenant de ne mélanger que des positions généralisées, soit

$$F_1(\mathbf{q}, \mathbf{Q}) = \mathbf{q} \cdot \mathbf{Q} \quad (3.61)$$

Le hamiltonien est toujours inchangé mais les variables sont quant à elles échangées :

$$\mathbf{p} = \frac{\partial F_1}{\partial \mathbf{q}} = \mathbf{Q} \quad \text{et} \quad \mathbf{P} = -\frac{\partial F_1}{\partial \mathbf{Q}} = -\mathbf{q} \quad (3.62)$$

À un signe près cette fonction  $F_1$  engendre donc l'échange entre les  $\mathbf{p}$  et les  $\mathbf{q}$ . Les coordonnées conjuguées de Hamilton sont donc quasiment interchangeables, elles ont en tout cas un rôle très symétrique, la seule chose qui importe est que l'on puisse définir deux familles distinctes et canoniquement conjuguées.

On voit sur cet exemple que la notion de grandeur physique perd un peu de sa splendeur : en fait peu importe qui est la position ou qui représente l'impulsion, ce qui est important c'est qu'il existe deux familles de variables indépendantes. Le fait que ces deux familles correspondent à ce que nous sommes habitués à interpréter comme la position et la vitesse peut ne pas s'avérer très utile pour comprendre la dynamique du système considéré.

Considérons un dernier exemple à la fois miraculeux et fondateur.

Soit un oscillateur harmonique à une dimension déjà étudié plus haut, dont le hamiltonien s'écrit ( $\omega > 0$ )

$$\mathcal{H}(p, q) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 \quad (3.63)$$

Appliquons-lui la transformation canonique engendrée par la fonction de type 1

$$F(q, Q) = \frac{1}{2}m\omega q^2 \cot Q \quad (3.64)$$

Celle-ci n'est bien sûr pas choisie au hasard et nous tenterons de fournir plus loin quelques explications sur la ou les méthodes que l'on peut mettre en œuvre pour la construire. En appliquant les formules (3.51), les nouvelles variables s'écrivent

$$p = \frac{\partial F}{\partial q} = m\omega q \cot Q \quad (3.65)$$

et

$$P = -\frac{\partial F}{\partial Q} = \frac{m\omega q^2}{2 \sin^2 Q} \Rightarrow q = \sqrt{\frac{2P}{m\omega}} \sin Q \quad (3.66)$$

en injectant cette dernière relation dans l'avant-dernière, il vient

$$p = \sqrt{2m\omega P} \cos Q \quad (3.67)$$

en terme des nouvelles variables le hamiltonien s'écrit donc

$$\mathcal{H}(P, Q) = \omega P \quad (3.68)$$

la variable  $Q$  apparaît donc cyclique : la dynamique en variables  $(Q, P)$  est donc triviale

$$P = \frac{\mathcal{H}}{\omega} = cste \quad (3.69)$$

$$Q = \omega t + \varphi$$

Même si cette dynamique est particulièrement simple, celle en variable  $(q, p)$  n'était déjà pas très complexe et l'on pourrait penser qu'un tel cas soit exceptionnel. Ce cas particulier est en fait l'archétype d'un système dit intégrable.

Ce qu'a montré la théorie des systèmes dynamiques au xx<sup>e</sup> siècle est qu'effectivement ces systèmes intégrables sont assez exceptionnels, mais, que la plupart des systèmes hamiltoniens étaient en fait toujours très proches d'un système intégrable. C'est le fameux théorème KAM pour Kolmogorov, Arnold et Moser.

<sup>3</sup> Un hamiltonien est dit autonome s'il ne dépend pas explicitement du temps.

### 3.4.3 Intégrabilité d'un système dynamique hamiltonien

Considérons un système possédant  $s$  degrés de liberté et décrit par un hamiltonien  $\mathcal{H} = \mathcal{H}(\mathbf{q}, \mathbf{p})$ . Nous considérons le cas autonome<sup>3</sup> ainsi  $\mathbf{q}$  et  $\mathbf{p}$  sont des vecteurs de  $\mathbb{R}^s$  et toute transformation canonique conserve le hamiltonien.

Ce système est dit intégrable si et seulement s'il existe une transformation canonique  $F$  permettant de passer des variables  $(\mathbf{q}, \mathbf{p})$  aux variables  $(\mathbf{Q}, \mathbf{P})$  pour lesquelles le nouvel hamiltonien ne dépend plus que d'une seule famille de variables :  $\mathcal{H} = \mathcal{H}(\mathbf{Q})$  ou  $\mathcal{H} = \mathcal{H}(\mathbf{P})$ . La transformation  $(\mathbf{q}, \mathbf{p}) \rightarrow (\mathbf{Q}, \mathbf{P})$  est bien bijective car canonique, mais l'expression du nouvel hamiltonien en fonction des nouvelles variables ne fait apparaître qu'une des deux familles de coordonnées.

La dynamique d'un tel système est alors triviale.

Le fait que la transformation soit canonique assure que les équations de la dynamique seront toujours hamiltoniennes en terme des nouvelles variables : c'est tout l'intérêt de la chose ! En choisissant par exemple le cas  $\mathcal{H} = \mathcal{H}(\mathbf{P})$ , elles s'écrivent pour la variable  $P$

$$\dot{\mathbf{P}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{Q}} = 0 \implies \mathbf{P} = \boldsymbol{\nu} = \mathbf{cste}$$

Les  $s$  composantes du vecteur impulsion  $\mathbf{P}$  de ce système intégrable sont donc  $\ell = 1, \dots, s$  constantes réelles que nous avons notées  $\nu_\ell$ .

Pour les coordonnées généralisées ce n'est guère plus compliqué :

$$\dot{\mathbf{Q}} = +\frac{\partial \mathcal{H}}{\partial \mathbf{P}} = \boldsymbol{\varphi}(\boldsymbol{\nu}) = \boldsymbol{\alpha} = \mathbf{cste}$$

En effet, puisque  $\mathcal{H}$  est un champ scalaire dépendant uniquement de  $\mathbf{P}$ , son gradient sera un champ de vecteur  $\boldsymbol{\varphi}$  de cette même variable qui ne dépend pas du temps comme nous l'avons vu. Ainsi,

$$\mathbf{Q} = \boldsymbol{\alpha}t + \mathbf{Q}(t=0)$$

Les systèmes dynamiques hamiltonniens se répartissent en deux catégories : les intégrables pour lequel la dynamique est triviale et les non intégrables pour lesquels la dynamique est plutôt compliquée voire peut-être même chaotique. Pour être plus précis, on dira en fait simplement que la dynamique d'un système intégrable n'est pas chaotique. Pour la réciproque, c'est une vaste question associée à un vaste problème...

Commençons par une vaste famille de systèmes intégrables.

## Systèmes isolés à un degré de liberté

Le hamiltonien d'une particule isolée ne possédant qu'un seul degré de liberté dans un potentiel  $U(q)$  est simplement son énergie mécanique

$$H = \frac{p^2}{2m} + U(q) \quad (3.70)$$

Cette énergie est conservée car le système est isolé et donc invariant par translation dans le temps. Ce système est intégrable. Pour le montrer il suffit de choisir un nouveau système de coordonnées  $(P, Q)$  tel que  $\lambda P = H$  où  $\lambda$  est une constante permettant de fixer les détails. Il ne reste plus qu'à trouver la transformation canonique, ce que l'on peut faire ici de façon un peu formelle mais efficace. L'une des nouvelles coordonnées est proportionnelle au hamiltonien, pour construire l'autre on peut par exemple écrire que

$$p = \pm \sqrt{2m(\lambda P - U(q))} \quad (3.71)$$

L'impulsion  $p$  n'est fonction que de  $P$  et  $q$ , pour faire simple on peut ici utiliser une fonction génératrice de type  $F_2 = F_2(q, P)$ , sur des cas plus particuliers on peut adapter la démarche. En combinant (3.54) et (3.71), on obtient

$$p = \frac{\partial F_2}{\partial q} = \pm \sqrt{2m(\lambda P - U(q))} \text{ soit } F_2(q, P) = \pm \int \sqrt{2m(\lambda P - U(q))} dq$$

et ainsi

$$Q = \pm \frac{\partial}{\partial P} \int \sqrt{2m(\lambda P - U(q))} dq$$

En prenant par exemple  $\lambda = 1/m$ , le système décrit en coordonnées  $(p, q)$  par le hamiltonien (3.70) est donc décrit en variables  $(P, Q)$  par le hamiltonien  $H = mP$  avec

$$P = \frac{p^2}{2} + mU(q) \text{ et } Q = \pm \frac{\partial}{\partial P} \int \sqrt{2P - 2mU(q)} dq$$

La dynamique pour ces nouvelles variables est triviale

$$\begin{aligned} \dot{P} &= -\frac{\partial H}{\partial Q} = 0 \implies P(t) = \lambda H = \omega \\ \dot{Q} &= +\frac{\partial H}{\partial P} = \lambda \implies Q(t) = \lambda t + Q_0 = \frac{\omega t}{H} + Q_0 \end{aligned}$$

Dans cet exemple la variable cyclique est  $Q$  car c'est d'elle dont ne dépend pas le nouvel hamiltonien.

Si la coordonnée généralisée cyclique est un angle<sup>4</sup>, on vérifie que sa coordonnée conjuguée est homogène à une énergie multipliée par un temps : ce qui correspond à une action. C'est pour cette raison que l'on parle de façon générale de variables angle-action, même si dans le cas le plus général  $Q$  n'est pas forcément un angle et donc  $P$  une action.

Quelques exemples pour rendre tout ceci transparent :

*Exemple 3.1.* Toujours l'oscillateur harmonique...

L'oscillateur harmonique est toujours un bon exemple dans ce contexte. Son hamiltonien s'écrit  $\mathcal{H} = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2$ . Il est constant et peut donc jouer le rôle de l'action : on pose donc  $P = \lambda\mathcal{H}$  et l'on cherche la bonne transformation canonique. La symétrie du hamiltonien permet d'en envisager directement au moins deux, écrivons pour changer

$$q = \pm \sqrt{\frac{2\mathcal{H}}{m\omega^2} - \frac{p^2}{m^2\omega^2}} = \pm \sqrt{\frac{2P}{\lambda m\omega^2} - \frac{p^2}{m^2\omega^2}} = q(P, p) \quad (3.72)$$

Le signe  $\pm$  n'a pas grande importance et indique simplement un sens de parcours de la trajectoire, gardons le cas négatif pour une raison que la phrase suivante va permettre de comprendre. On peut ici utiliser une fonction génératrice de type 4 pour laquelle l'une des relations (3.58) donne

$$q = -\frac{\partial F(p, P)}{\partial p}$$

soit

$$\begin{aligned} F(p, P) &= \int \sqrt{\frac{2P}{\lambda m\omega^2} - \frac{p^2}{m^2\omega^2}} dp \\ &= \frac{2P}{\lambda\omega} \int \sqrt{1 - u^2} du \quad \text{avec } u = \sqrt{\frac{\lambda p^2}{2mP}} \end{aligned}$$

Le calcul de la primitive ne pose pas de problème on trouve

$$F(p, P) = \frac{1}{\lambda\omega} \left[ \sqrt{aP} \sqrt{1 - \frac{a}{P}} + P \arcsin \sqrt{\frac{a}{P}} \right] \quad \text{avec } a = \frac{\lambda p^2}{2m} \quad (3.73)$$

Pour obtenir l'expression de  $Q$  on utilise l'autre relation contenue dans (3.58), il vient

<sup>4</sup> ce qui est souvent le cas pour une coordonnée généralisée

$$Q = \frac{\partial F}{\partial P} = \frac{1}{\lambda\omega} \arcsin \left( \sqrt{\frac{\lambda p^2}{2mP}} \right) \text{ soit } p^2 = \frac{2mP}{\lambda} \sin^2(\lambda\omega Q) \quad (3.74)$$

en revenant à la définition (3.72) on trouve

$$q^2 = \frac{2P}{\lambda m\omega^2} \cos^2(\lambda\omega Q) \quad (3.75)$$

Les deux relations (3.74) et (3.75) permettent bien d'exprimer  $(p, q)$  en fonction de  $(P, Q)$  grâce à la transformation canonique (3.73). En utilisant les coordonnées  $(P, Q)$  le nouvel hamiltonien s'écrit bien  $\mathcal{H} = P/\lambda$ , les variables  $(P, Q)$  sont des variables angle-action de l'oscillateur harmonique. En prenant  $\lambda = \omega^{-1}$  on retrouve<sup>5</sup> les effets de la transformation (3.64) qui semblait sortir de nulle part !

*Exemple 3.2.* Orbite d'une particule de masse  $m$  dans un potentiel radial

Ce type de problème correspond par exemple au mouvement d'une étoile dans un amas globulaire sous une hypothèse de champ moyen et à symétrie sphérique. La force subie par une étoile supposée ponctuelle, de masse  $m$  et repérée dans l'espace par le vecteur  $\vec{r}$ , est radiale. En notant  $q = |\vec{r}|$ , cette force dérive de l'énergie potentielle  $m\psi(q)$  créée par l'ensemble des étoiles de cet amas considéré comme un continuum. Les propriétés physiques de cet amas confère à cette énergie potentielle certaines propriétés comme le fait que  $\psi(q)$  soit négative et croissante,  $\lim_{q \rightarrow +\infty} \psi(q) = 0^-$  car le système est supposé isolé et  $\forall \alpha > 1$ ,  $\lim_{q \rightarrow 0} q^\alpha \psi(q) = 0^-$  pour que la masse totale de l'amas soit finie. La force radiale que subit l'étoile confine sa trajectoire dans un plan. Le système possède donc a priori 2 degrés de liberté, mais la conservation du moment cinétique  $\vec{L} = m\vec{r} \wedge \dot{\vec{r}}$  permet en fait de n'en considérer qu'un seul : la distance  $q$  de l'étoile au centre de l'amas. En considérant le module  $L$  du moment cinétique comme un paramètre, le hamiltonien du système s'écrit

$$\mathcal{H}(q, p) = \frac{p^2}{2m} + \frac{L^2}{2q^2} + m\psi(q) < 0$$

Il est conservé au cours du mouvement sa valeur est négative car l'étoile est supposée liée au système. Posons

$$P = P(q, p) = \mathcal{H}(p, q) = \frac{p^2}{2m} + \frac{L^2}{2q^2} + m\psi(q) \quad (3.76)$$

<sup>5</sup> Pour retrouver exactement la transformation (3.64), il aurait fallu prendre une fonction de

on a donc formellement

$$p = \pm \sqrt{2m [P - m\psi(q)] - \frac{mL^2}{q^2}}$$

la fonction  $p$  dépend ainsi des variables  $P$  et  $q$ , on peut l'engendrer avec une fonction génératrice de type  $F_2 = F_2(q, P)$  pour laquelle

$$p = \frac{\partial F_2}{\partial q} = \pm \sqrt{2m [P - m\psi(q)] - \frac{mL^2}{q^2}}$$

ainsi

$$F_2(q, P) = \pm \int \sqrt{2m [P - m\psi(q)] - \frac{mL^2}{q^2}} dq$$

On peut alors utiliser l'autre propriété de cette fonction génératrice pour obtenir

$$Q = \frac{\partial F_2}{\partial P} = \sqrt{m} \int \frac{dq}{\sqrt{2 [P - m\psi(q)] - \frac{L^2}{q^2}}}$$

C'est ce qui fait dire à certains qu'un système hamiltonien isolé à un degré de liberté est intégrable par quadrature. Cette primitive rarement explicitable est a priori telle que  $Q = Q(P, q)$  mais une fois le calcul fait on peut remplacer l'expression (3.76) de  $P$  pour obtenir finalement  $Q = Q(p, q)$ . On peut assez facilement interpréter certaines expressions tirées de cette variable  $Q$ .

En effet, pour  $P < 0$  fixé, la trajectoire de l'étoile est périodique et confinée entre deux valeurs extrêmes appelées apocentre  $q_{\max} = \max_q Q(q, P)$  et péricentre  $q_{\min} = \min_q Q(q, P)$ . On vérifie assez facilement<sup>6</sup> que la quantité

$$\tau = 2 [Q(q_{\max}, P) - Q(q_{\min}, P)] = \sqrt{m} \int_{q_{\min}}^{q_{\max}} \frac{dq}{\sqrt{2 [P - m\psi(q)] - \frac{L^2}{q^2}}}$$

est la période temporelle de l'oscillation radiale d'une étoile d'énergie  $P$  dans le plan considéré. La différence entre ces deux valeurs de  $Q$  donne un temps, ce n'est pas très étonnant car nous savions déjà que

$$\dot{Q} = \frac{\partial \mathcal{H}}{\partial P} = 1 \implies Q(t) = t + t_0$$

type 1. Mais «abondance de biens ne nuit pas» et le lecteur pourra s'y essayer !

<sup>6</sup> Il suffit de remarquer qu'en écrivant l'une des équations de Hamilton  $\dot{q} = \frac{dq}{dt} = p/m$  puis

On peut également rapprocher toutes ces notions de celle souvent mieux connue en mécanique de potentiel effectif. Cette quantité correspond ici à la fonction

$$\psi_{\text{eff}}(q) = \mathcal{H}(q, p = 0) = \frac{L^2}{2q^2} + m\psi(q)$$

Les propriétés de  $\psi(q)$  permettent rapidement de montrer que

$$\lim_{q \rightarrow 0} \psi_{\text{eff}}(q) = +\infty \text{ et } \lim_{q \rightarrow +\infty} \psi_{\text{eff}}(q) = 0^-$$

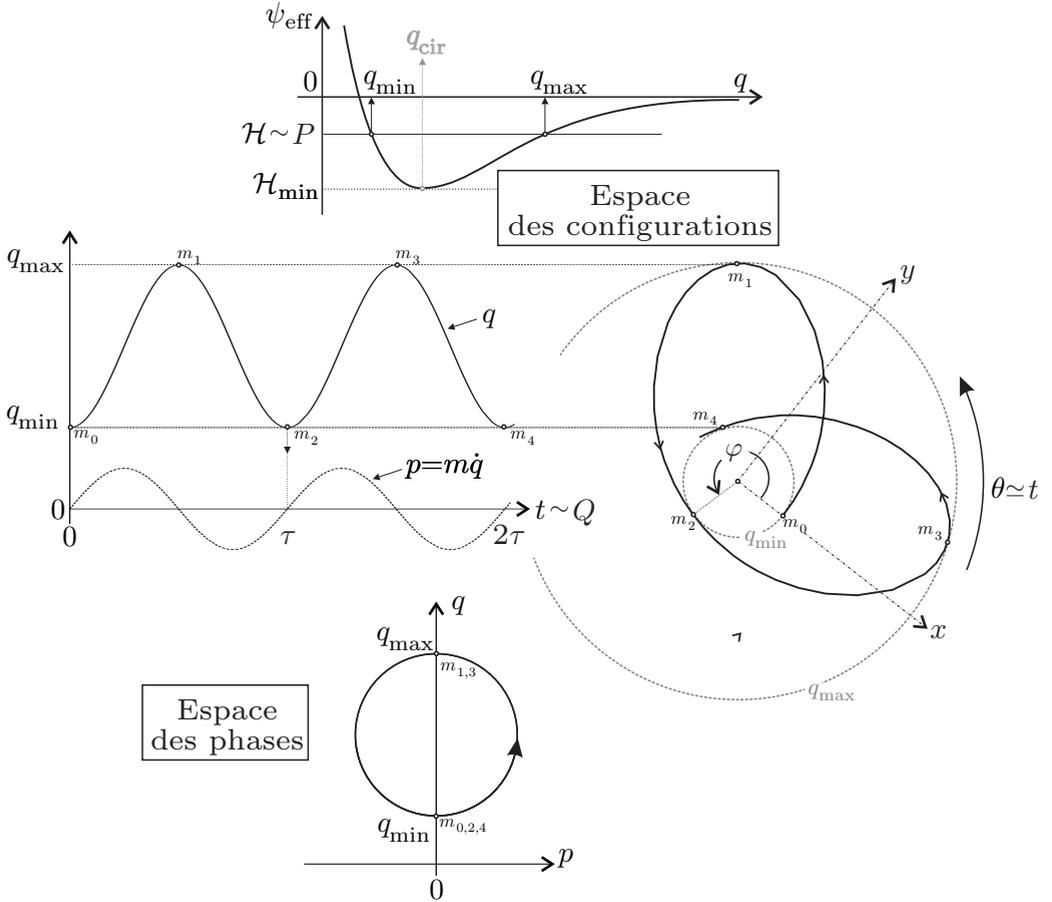
Les orbites liées sont celles d'énergie  $\mathcal{H}$  négative, les intersections entre la courbe  $\psi_{\text{eff}}(q)$  et la constante  $\mathcal{H} < 0$  correspondent aux deux valeurs extrêmes  $q_{\text{min}}$  et  $q_{\text{max}}$ . Lorsque ces deux valeurs sont égales, la trajectoire est circulaire et l'énergie extrême.

En variable  $(Q, P)$  la dynamique est triviale :  $Q$  est le temps et  $P$  l'énergie, ces deux grandeurs sont donc des variables angle-action. Les variables généralisées  $(q, p)$  sont des fonctions périodiques : par exemple  $q$  peut être la distance au centre qui, si l'énergie est négative et le potentiel croissant, oscille entre le périastre  $q_{\text{min}}$  et l'apoastre  $q_{\text{max}}$ ; la variable conjuguée  $p$  sera alors proportionnelle à la vitesse radiale. L'orbite n'est fermée que si le déphasage  $\varphi$  sur une période  $\tau$  est un multiple fractionnaire de  $2\pi$ . Si on utilise les variables cartésiennes  $(x, y, p_x, p_y)$  on a rien compris car en fait le système ne possède qu'un seul degré de liberté... On peut voir les différents liens entre toutes ses variables sur la figure 3.1.

Si le potentiel effectif n'est plus en forme de cuvette ou si l'énergie est positive, l'orbite n'est plus nécessairement périodique mais le système reste intégrable.

## Représentation et stabilité des systèmes intégrables

Comme nous venons de le voir, chaque système hamiltonien isolé ne possédant qu'un seul degré de liberté est intégrable. Pour chaque jeu de conditions initiales  $(q_0, p_0)$ , la dynamique de ce système est entièrement déterminée par la donnée d'une constante  $P$  – proportionnelle à son énergie qui est une fonction de  $(q_0, p_0)$  – et d'une fonction affine du temps  $Q(t) = \lambda t + Q_0$ . Si nécessaire, on peut alors en déduire, plus ou moins explicitement, les expressions de  $q$  et  $p$  pour n'importe quel temps  $t$ . Lorsque l'on représente les variables associées aux degrés de liberté du système en fonction de leur paramètre d'évolution, on représente l'espace des configurations du système.

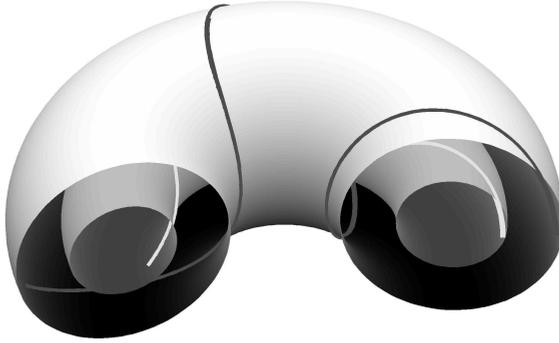


**Fig. 3.1.** Différentes variables pour une particule de masse  $m$  dans un potentiel radial

Dans cet espace, on représente par exemple  $q(t)$  ou  $p(t)$ . On peut également tenter de dessiner directement l'évolution de  $q(p)$  ou de  $p(q)$ , on représente alors l'espace des phases du système.

L'espace des phases d'un système intégrable et confiné ne possédant qu'un seul de degré de liberté est un cercle dès lors que l'on a choisi une condition initiale  $(q_0, p_0)$ . On peut voir sa représentation sur la figure 3.1. Partant de la condition initiale, l'évolution dynamique enroule la trajectoire sur le cercle dont le rayon est fixé par  $(q_0, p_0)$ . En faisant varier les conditions initiales, si les trajectoires restent bornées, on obtient une collection de cercles concentriques. Qu'en est-il si maintenant on augmente le nombre de degré de libertés ? Si le système est toujours intégrable et si les trajectoires

d'écrire que dans le cas considéré  $\tau = 2 \int_{q_{\min}}^{q_{\max}} dt$ .



**Fig. 3.2.** Partie de l'espace des phases d'un système intégrable possédant deux degrés de liberté et dont les trajectoires sont bornées. Pour un seul degré de liberté, l'espace des phases est un cercle ( $S_1$ ). Avec deux degrés de liberté, l'espace des phases devient  $S_1 \times S_1$ , il est de dimension 4. Afin de pouvoir le visualiser, nous n'en avons représenté qu'une partie :  $p_1 \times (q_0, p_0)$  pour deux conditions initiales différentes.

restent toujours bornées, on montre (voir [2],[1]) que ces cercles deviennent des tores sur lesquels les trajectoires s'enroulent. On a représenté sur la figure 3.2, l'espace des phases d'un système intégrable possédant deux degrés de liberté et dont les trajectoires sont bornées. L'espace des phases est en fait de dimension 4, il est entièrement décrit par la donnée de deux angles et de deux constantes. L'une de ces constantes n'a pas été représentée sur la figure 3.2, mais on y voit deux trajectoires issues de deux conditions initiales différentes.

Dans ce contexte l'un des grands achèvements du  $xx^e$  siècle fut la démonstration du théorème KAM sur la stabilité de ce type de système soumis à des perturbations. Nous allons en évoquer quelques éléments sans entrer dans les détails ce qui nous mènerait trop loin.

Considérons un système hamiltonien intégrable possédant  $n$  degrés de liberté. Il existe donc un système de coordonnées  $(\mathbf{Q}, \mathbf{P}) \in \mathbb{R}^n \times \mathbb{R}^n$  dans lequel son hamiltonien s'écrit  $\mathcal{H}_0 = \mathcal{H}_0(\mathbf{P})$ . Pour rester dans le cadre de ce que nous avons évoqué jusqu'à présent nous supposons que les variations de  $\mathbf{Q}$  sont bornées. Nous admettrons que cela permet de les assimiler à des angles aux variations périodiques, généralisant ainsi les tores de la figure 3.2 au cas  $n$ -dimensionnel. Les équations de Hamilton fournissent

$$\mathbf{P} = cste \quad \text{et} \quad \dot{\mathbf{Q}} = \frac{\partial \mathcal{H}_0}{\partial \mathbf{P}} = \boldsymbol{\omega}(\mathbf{P}) = cste$$

Le vecteur  $\boldsymbol{\omega} = \text{grad}_{\mathbf{P}}(\mathcal{H}_0)$  est appelé vecteur des fréquences, il est constant lors de l'évolution temporelle du système intégrable.

Depuis Laplace, Lagrange et Gauss on avait pu montrer que certains systèmes intégrables – comme le problème de  $N$  corps en interaction gravitationnelle avec un  $N + 1$  ème sans considérer les interactions entre les  $N$  premiers – restaient généralement stables à l'ordre  $\varepsilon^2$  si l'on les perturbait à l'ordre  $\varepsilon$  – par exemple en considérant que les interactions entre les  $N$  premiers corps étaient de cet ordre. Dans le cas du système solaire, si l'on ne considère que le Soleil comme source du mouvement des planètes, chacune d'elle est en orbite képlérienne bornée autour de notre étoile : chacune des fréquences  $\omega_i$  est donné par la 3<sup>è</sup> loi de Kepler en fonction du demi-grand axe de chaque ellipse. Comme nous l'avons vu, on peut prendre pour  $\mathbf{P}$  le vecteur dont les composantes sont les énergies, conservées et négatives, de chacune des planètes. Si l'on prend en compte les interactions gravitationnelles mutuelles des planètes comme une perturbation globale d'ordre  $\varepsilon$  par rapport à l'interaction avec le Soleil<sup>7</sup>. Le hamiltonien du système perturbé s'écrira sous la forme standard suivante

$$\mathcal{H}_\varepsilon(\mathbf{P}, \mathbf{Q}) = \mathcal{H}_0(\mathbf{P}) + \varepsilon \mathcal{H}_1(\mathbf{P}, \mathbf{Q}) \quad (3.77)$$

Le résultat de ces 3 illustres mathématiciens indiquait la stabilité sur des temps longs (d'ordre  $1/\varepsilon^2$ ) du système perturbé en l'absence de résonance du système non perturbé. Par résonance on entend classiquement la non commensurabilité des fréquences, c'est-à-dire le fait que

$$\forall \mathbf{k} \in \mathbb{Z}^{n,*}, \quad \mathbf{k} \cdot \boldsymbol{\omega} \neq 0$$

Pour ne pas avoir de résonance, il faut ne pas pouvoir annuler des combinaisons linéaires des fréquences avec des coefficients entiers.

En 1954, Kolmogorov donne la première lettre du théorème KAM, en montrant que la stabilité reste acquise pour la plupart des systèmes hamiltoniens de la forme (3.77) sous une condition dite diophantienne

$$\forall \mathbf{k} \in \mathbb{Z}^{n,*}, \quad \forall \kappa, \exists \gamma > 0, \quad |\mathbf{k} \cdot \boldsymbol{\omega}| \geq \frac{\kappa}{|\mathbf{k}|^\gamma}$$

La notion de stabilité est même quelque peu précisée : les orbites – bornées et dessinées sur un tore  $n$  dimensionnel – engendrée par le hamiltonien  $\mathcal{H}_0$  et issues d'une condition initiale donnée, sont simplement déformées lorsque la dynamique est pilotée par  $\mathcal{H}_\varepsilon$  : les tores se déforment mais conservent leurs

<sup>7</sup> Les masses respectives du soleil et des planètes permettent ce genre d'estimation.

propriétés topologiques. Dans le jargon des systèmes dynamiques ces tores déformés sont qualifiés de persistants. Si par contre certaines fréquences rendent la condition diophantienne impossible, les tores correspondant explosent dans la dynamique gouvernée par  $\mathcal{H}_\varepsilon$ . L'espace des phases n'est plus folié en  $n$  tores  $n$  dimensionnels mais contient des zones denses confinées entre deux tores persistants. Les deux autres lettres du théorème KAM sont associées à Arnold et Moser pour des compléments de la démonstration de ce résultat, sur la généralisation des systèmes auquel il s'applique et la clarification ou la simplification de l'hypothèse diophantine. Ces différents résultats se sont étalés sur toute la seconde moitié du XX<sup>e</sup> siècle.

Outre le flot de mathématiques qu'il a et qu'il continue d'engendrer, le théorème KAM est particulièrement intéressant d'un point de vue physique. Dans une certaine réalité, on peut en effet toujours faire l'hypothèse que l'on peut écrire le hamiltonien d'un système physique sous la forme (3.77) requise par le théorème KAM : un intervalle de temps suffisamment faible, le confinement suffisamment proche d'une solution connue et intégrable, etc. On peut ainsi en conclure que pendant cet intervalle de temps ou bien dans ce voisinage, la dynamique du système physique correspondant sera simple. En physique cette simplicité se caractérise par l'absence de chaos, ce mot restant à définir. Sur des temps plus longs ou bien hors de ces voisinages, c'est plutôt l'inverse qui est garanti. Comprenne qui pourra!

### 3.4.4 Transformations canoniques et crochets de Poisson

#### Un théorème fondamental

Nous avons vu comment effectuer des changements de variables permettant de préserver la structure canonique des équations du mouvement. Une question essentielle que l'on peut se poser est la suivante : tout changement de variable constitue-t-il une transformation canonique ?

C'est en effet souvent dans ce sens que se pose la question. Pour y répondre, nous disposons d'un élégant théorème reposant une fois de plus sur les crochets de Poisson.

**Théorème 3.3.** *Une transformation est canonique si et seulement si elle laisse invariant le crochet de Poisson.*

$$(\mathbf{p}, \mathbf{q}) \mapsto (\mathbf{P}, \mathbf{Q}) \iff \{f, g\}_{\mathbf{p}, \mathbf{q}} = \{f, g\}_{\mathbf{P}, \mathbf{Q}} \quad (3.78)$$

#### Démonstration

L'une des implications est triviale : si une transformation de coordonnées préserve le crochet de Poisson, les équations du mouvement (3.4.1) seront inchangées et donc les nouvelles variables décriront le mouvement du même système ce qui est la définition d'une transformation canonique.

La réciproque peut être prouvée de deux manières, en suivant le discours hautement physique de Landau, ou en faisant un calcul. Nous choisirons ici cette deuxième voie peut-être plus convaincante ?

Considérons le crochet de Poisson

$$\{f, g\}_{\mathbf{p}, \mathbf{q}} = \frac{\partial f}{\partial \mathbf{q}} \cdot \frac{\partial g}{\partial \mathbf{p}} - \frac{\partial g}{\partial \mathbf{q}} \cdot \frac{\partial f}{\partial \mathbf{p}} \quad (3.79)$$

dans lequel  $f = f(\mathbf{q}, \mathbf{p})$  et  $g = g(\mathbf{q}, \mathbf{p})$  sont deux champs scalaires quelconques. On effectue alors le changement de variable  $\mathbf{Q} = \mathbf{Q}(\mathbf{q}, \mathbf{p})$  et  $\mathbf{P} = \mathbf{P}(\mathbf{q}, \mathbf{p})$  après quelques lignes de calcul qui se simplifient facilement, le crochet (3.79) devient quelque chose qu'il est préférable d'écrire en composante plutôt qu'en faisant apparaître des produits scalaires :

$$\{f, g\}_{\mathbf{p}, \mathbf{q}} = \sum_{\alpha, \beta=1}^s \left[ \frac{\partial f}{\partial Q_\beta} \frac{\partial g}{\partial P_\beta} - \frac{\partial g}{\partial Q_\beta} \frac{\partial f}{\partial P_\beta} \right] \left[ \frac{\partial Q_\beta}{\partial q_\alpha} \frac{\partial P_\beta}{\partial p_\alpha} - \frac{\partial Q_\beta}{\partial p_\alpha} \frac{\partial P_\beta}{\partial q_\alpha} \right] \quad (3.80)$$

C'est le moment de faire l'hypothèse que la transformation est canonique. Choisissons-là par exemple de type 1, ainsi

$$\begin{cases} \mathbf{p} = \frac{\partial F_1}{\partial \mathbf{q}} \\ \mathbf{P} = -\frac{\partial F_1}{\partial \mathbf{Q}} \end{cases} \quad \text{et} \quad F_1 = F_1(\mathbf{q}, \mathbf{Q}) \quad (3.81)$$

Sous cette hypothèse, le second terme de la multiplication de l'équation (3.80) se réécrit

$$\frac{\partial Q_\beta}{\partial q_\alpha} \frac{\partial P_\beta}{\partial p_\alpha} - \frac{\partial Q_\beta}{\partial p_\alpha} \frac{\partial P_\beta}{\partial q_\alpha} = \sum_{\alpha=1}^s \left[ -\frac{\partial Q_\beta}{\partial q_\alpha} \frac{\partial^2 F_1}{\partial p_\alpha \partial Q_\beta} + \frac{\frac{\partial}{\partial q_\alpha} \frac{\partial F_1}{\partial Q_\beta}}{\frac{\partial}{\partial Q_\beta} \frac{\partial F_1}{\partial q_\alpha}} \right] \quad (3.82)$$

Le premier terme de cette dernière somme est nul car  $F_1$  ne dépend pas de  $\mathbf{p}$ , de plus l'indépendance des  $\mathbf{q}$  et des  $\mathbf{Q}$  montre que le deuxième terme est un symbole de Kronecker  $\delta_{\alpha\beta}$ , la relation (3.80) devient donc

$$\{f, g\}_{\mathbf{p}, \mathbf{q}} = \sum_{\alpha, \beta=1}^s \left[ \frac{\partial f}{\partial Q_\beta} \frac{\partial g}{\partial P_\beta} - \frac{\partial g}{\partial Q_\beta} \frac{\partial f}{\partial P_\beta} \right] \delta_{\alpha\beta} = \{f, g\}_{\mathbf{P}, \mathbf{Q}} \quad (3.83)$$

La vérification de cette même propriété kroneckerienne dans les cas de transformations de type 2, 3 ou 4 complète la démonstration.  $\square$

Ce résultat est fondamental pour les systèmes dynamiques, il permet de nombreuses investigations. L'exemple de la section suivante en est l'illustration.

### Le mouvement comme transformation canonique

Considérons un système de particules possédant  $s$  degrés de liberté. Les variables canoniquement conjuguées sont  $\mathbf{q}$  et  $\mathbf{p}$  tous les deux dans  $\mathbb{R}^s$ . Soit  $\tau$  un instant quelconque, la transformation

$$[\mathbf{q}(t), \mathbf{p}(t)] \mapsto [\mathbf{Q}(t), \mathbf{P}(t)] = [\mathbf{q}(t + \tau), \mathbf{p}(t + \tau)] \quad (3.84)$$

est évidemment canonique car si  $(\mathbf{q}, \mathbf{p})$  vérifie les équations du mouvement, alors  $(\mathbf{Q}, \mathbf{P})$  les vérifie aussi !

La fonction génératrice de cette transformation est l'action. En effet,

$$S = \int_{t_1}^{t_2} \mathcal{L} dt = \int_{t_1}^{t_2} (\mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{H}) dt = \int_{t_1}^{t_2} (\mathbf{p} \cdot d\mathbf{q} - \mathcal{H} dt) \quad (3.85)$$

prenons  $t_1 = t$  et  $t_2 = t + \tau$  nous avons

$$dS = [\mathbf{p} \cdot d\mathbf{q} - \mathcal{H} dt]_t^{t+\tau} \quad (3.86)$$

supposons  $\mathcal{H}$  indépendant du temps (ceci est bien compatible avec le fait que la fonction génératrice est l'action qui est elle-même indépendante du temps), en explicitant (3.86) il vient

$$\begin{aligned} dS &= \mathbf{p}(t + \tau) \cdot d\mathbf{q}(t + \tau) - \mathbf{p}(t) \cdot d\mathbf{q}(t) \\ &= \mathbf{P} \cdot d\mathbf{Q} - \mathbf{p} \cdot d\mathbf{q} \end{aligned} \quad (3.87)$$

relation qui assure que  $S = S(\mathbf{q}, \mathbf{Q})$  telle que

$$\mathbf{P} = \frac{\partial S}{\partial \mathbf{Q}} \quad \text{et} \quad \mathbf{p} = -\frac{\partial S}{\partial \mathbf{q}} \quad (3.88)$$

c'est-à-dire une fonction génératrice de type 1 (au signe près, que l'on peut rattraper en prenant  $F_2 = -S$ ).

## Le théorème de Liouville et ses conséquences

Maintenant que nous sommes certains de la nature de la transformation associée au mouvement, considérons la quantité

$$V(t) = \int dqd\mathbf{p} \quad (3.89)$$

qui représente le volume d'espace des phases occupé par le système à l'instant  $t$ . Un petit calcul complémentaire, que nous laissons à la perspicacité du lecteur, montre que le symbole de Kronecker apparaissant dans la fin de la démonstration du théorème 3.3 implique que le jacobien de toute transformation canonique est l'unité. Ainsi en conservant les notations du paragraphe précédent

$$V(t) = \int dqd\mathbf{p} = \int d\mathbf{Q}d\mathbf{P} = V(t + \tau) \quad (3.90)$$

Ce résultat valant pour chaque valeur de  $\tau$ , il assure la conservation de  $V$  au cours du temps. Par conséquent

$$\frac{dV}{dt} = 0 \quad (3.91)$$

Ce résultat constitue le théorème de Liouville. La démonstration que nous venons de proposer est basée sur les transformations canoniques et la propriété fondamentale des crochets de Poisson.

En fait tout ceci repose sur l'antisymétrie déjà évoquée des équations de la physique. En mathématiques, en effet, le théorème est bien connu de ceux qui étudient les équations différentielles autonomes qui s'écrivent comme chacun sait

$$\mathbf{z} \in \mathbb{R}^n, \quad \dot{\mathbf{z}} = \mathbf{F}(\mathbf{z})$$

La fonction  $\mathbf{F}$  de  $\mathbb{R}^n \rightarrow \mathbb{R}^n$  est alors appelée un champ de vecteurs. Le transport de la solution d'un instant à l'autre est assurée par le flot  $(t, \mathbf{z}) \mapsto \varphi(t, \mathbf{z})$  solution du problème

$$\frac{\partial \varphi}{\partial t} = \mathbf{F}(\varphi) \quad \text{avec} \quad \varphi(0, \mathbf{z}) = \mathbf{z}$$

Le flot  $\varphi$  est donc associé au champ de vecteur  $\mathbf{F}$ .

L'application partielle  $\varphi_t : \mathbf{z} \mapsto \varphi(t, \mathbf{z})$  transporte bien la condition initiale  $\mathbf{z}(t=0)$  en la solution  $\mathbf{z}(t)$  correspondante au temps  $t$  selon la relation  $\varphi_t(\mathbf{z}(0)) = \mathbf{z}(t)$ . Pour des solutions dites maximales et en supposant la solution unique pour une condition initiale donnée, cette relation ne dépend pas de l'instant de départ, on montre en effet que  $\varphi_{t+\tau}(\mathbf{z}) = \varphi_t \circ \varphi_\tau(\mathbf{z})$ .

Le théorème de Liouville s'appuie alors sur trois remarques :

- La divergence d'un champ de vecteur est la trace de la *différentielle* de ce champ de vecteurs :

$$\operatorname{div} \mathbf{F}(\mathbf{z}) = \sum_{i=1}^n \frac{\partial F_i(\mathbf{z})}{\partial z_i} = \operatorname{tr} [D\mathbf{F}(\mathbf{z})]$$

- Si l'on considère à l'instant  $t = 0$  une partie  $A$  mesurable de  $\mathbb{R}^n$  pour la mesure  $d\mathbf{z}$  son volume sera

$$V[A] = \int_A d\mathbf{z}$$

l'application  $\varphi_t$  envoie chaque point de  $A$  à l'instant 0 vers le point correspondant à l'instant  $t$ , le volume correspondant sera

$$V[A(t)] = \int_{\varphi_t(A)} d\mathbf{z}$$

la formule du changement de variable pour une intégrale multiple donne alors

$$V[A(t)] = \int_A |\det D\varphi_t(\mathbf{z})| d\mathbf{z}$$

- La dernière remarque est en fait une propriété fondamentale du flot d'un champ de vecteur qui montre que sa différentielle est la résolvante  $R(t, 0)$  de son linéarisé :  $D\varphi_t = R(t, 0)$ . Cette résolvante est donc la solution de l'équation matricielle

$$\begin{cases} \frac{dR(t, 0)}{dt} = D\mathbf{F}(\mathbf{z}) R(t, 0) \\ R(0, 0) = Id \end{cases}$$

La différentielle du déterminant d'une matrice inversible est bien connue,  $D(\det A) H = \det(A) \operatorname{tr}[A^{-1}H]$ , elle permet de calculer

$$\begin{aligned} \frac{d}{dt} \det[R(t, 0)] &= \det[R(t, 0)] \operatorname{tr} \left[ R(t, 0)^{-1} \frac{dR(t, 0)}{dt} \right] \\ &= \det[R(t, 0)] \operatorname{tr} \left[ R(t, 0)^{-1} D\mathbf{F}(\mathbf{z}) R(t, 0) \right] \\ &= \det[R(t, 0)] \operatorname{tr} [D\mathbf{F}(\mathbf{z})] \end{aligned}$$

La conjonction de ces trois remarques permet d'affirmer le théorème de Liouville des équations différentielles : le volume occupé par un champ de vecteur à divergence nulle est constant, donc conservé !

Le fait que tout cela s'applique à la physique est également le résultat d'un constat, plus calculatoire celui-ci : nous savons depuis Lagrange (voir l'équation (1.9) page 10, ou bien l'équation 1.13 page 16) que les équations d'évolution temporelle d'un système hamiltonien s'écrivent sous la forme

$$\dot{\mathbf{z}} = \mathbf{F}(\mathbf{z}) \quad \text{avec} \quad \begin{cases} \mathbf{z} = (\mathbf{q}, \mathbf{p}) \\ \mathbf{F} = J \nabla_{\mathbf{z}} \mathcal{H}(\cdot) \quad \text{et} \quad J = \begin{bmatrix} 0 & \text{Id}_n \\ -\text{Id}_n & 0 \end{bmatrix} \end{cases}$$

Le théorème de Liouville des physiciens est enfin prêt. En effet, la matrice  $J$  qui intervient dans cette équation fondamentale de la physique est anti-symétrique mais elle est également telle que  $J^{-1} = J^{\top} = -J$  on peut donc vérifier que<sup>8</sup>

$$\begin{aligned} \text{div } \mathbf{F} &= \text{tr} [D\mathbf{F}] = \text{tr} [J \nabla_{\mathbf{z}}^2 \mathcal{H}] = -\text{tr} [J^{-1} \nabla_{\mathbf{z}}^2 \mathcal{H}] = -\text{tr} [JJ^{-1} \nabla_{\mathbf{z}}^2 \mathcal{H} J^{-1}] \\ &= -\text{tr} [\nabla_{\mathbf{z}}^2 \mathcal{H} J^{-1}] = -\text{tr} [\nabla_{\mathbf{z}}^2 \mathcal{H} J^{\top}] = -\text{tr} [(J \nabla_{\mathbf{z}}^2 \mathcal{H})^{\top}] \end{aligned}$$

Ainsi  $\text{div } \mathbf{F} = -\text{div } \mathbf{F}$  et la divergence d'un champ de vecteur hamiltonien est donc nulle : la conservation du volume est assurée lors de l'évolution temporelle !

Une remarque importante doit être mise en avant : en présence de dissipation, le système n'est plus hamiltonien et le théorème de Liouville ne s'applique plus. La transformation n'est plus générée par l'action car le hamiltonien dépend du temps, elle n'est plus canonique. Le volume de l'espace des phases occupé à un instant par le système ne se conserve pas.

Les conséquences du théorème du Liouville sont tout simplement gigantesques : c'est lui qui permet d'établir les fondements de la physique statistique...

Considérons en effet un système composé de  $N$  particules et possédant  $s$  degrés de liberté. Si  $N$  (ou plutôt  $s$ ) est très grand, il est vain de chercher à représenter le système par le gigantesque couple  $(\mathbf{q}, \mathbf{p})$ ; dans le contexte de la physique statistique, on préfère introduire la probabilité de présence dans l'espace des phases, un nouveau champ scalaire... Cette fonction notée ici  $w(\mathbf{q}, \mathbf{p}, t)$  donne la probabilité de trouver le système dans l'état  $(\mathbf{q}, \mathbf{p})$  à l'instant  $t$ . Cette probabilité est normalisée et vérifie donc à chaque instant

$$\int w \, d\mathbf{q}d\mathbf{p} = 1 \tag{3.92}$$

<sup>8</sup> Pour les étourdis on rappelle que le Hessien  $\nabla_{\mathbf{z}}^2 \mathcal{H}$  du Hamiltonien est symétrique et que la

une application plus ou moins directe du théorème de Liouville permet alors de montrer que

$$\frac{dw}{dt} = 0$$

soit en utilisant les crochets de Poisson

$$\frac{\partial w}{\partial t} + \{w, \mathcal{H}\} = 0 \tag{3.93}$$

ou encore plus explicitement

$$\frac{\partial w}{\partial t} = \frac{\partial \mathcal{H}}{\partial \mathbf{q}} \cdot \frac{\partial w}{\partial \mathbf{p}} - \frac{\partial \mathcal{H}}{\partial \mathbf{p}} \cdot \frac{\partial w}{\partial \mathbf{q}} \tag{3.94}$$

Cette équation appelée équation de Liouville est tout simplement la pierre de voûte d'une partie de la physique statistique, de toute la théorie cinétique, et par conséquent de la physique des plasmas ou de la mécanique des fluides (voir 10.4). Mentionnons quelques pistes à suivre, ou pas, dans ce répertoire.

La fonction  $w$  dépend outre du temps de la position et de l'impulsion de chacune des particules du système considéré  $w = w(t, \mathbf{q}, \mathbf{p})$ . On parle de fonction de distribution à  $N$  particules. Si ce n'est son caractère probabiliste, l'équation de Liouville est exacte dans le sens où elle ne fait aucune hypothèse sur les particules elle-mêmes. On peut également s'intéresser comme le fit Boltzmann à l'équation vérifiée par une fonction plus simple décrivant une particule moyenne décrite par une loi marginale à une seule particule décrite<sup>9</sup> par une position  $\vec{r} \in \mathbb{R}^3$  et une impulsion  $\vec{p} \in \mathbb{R}^3$

$$f(t, \vec{r}, \vec{p}) = \int w(t, \underbrace{\vec{r}, \vec{q}_2, \dots, \vec{q}_N}_{\mathbf{q}}, \underbrace{\vec{p}, \vec{p}_2, \dots, \vec{p}_N}_{\mathbf{p}}) \prod_{i=2}^N d\vec{q}_i d\vec{p}_i$$

Si toutes les particules possèdent la même loi marginale de probabilité  $f$  (les particules sont alors indiscernables); si le hamiltonien se sépare en la somme d'une énergie cinétique  $T = T(\vec{p})$  et d'une énergie potentielle  $U = U(\vec{r})$  on montre alors que

$$\frac{\partial f}{\partial t} + \vec{p} \cdot \frac{\partial f}{\partial \vec{r}} = Q(f, f)$$

Le second membre  $Q(f, f)$  est généralement appelé opérateur de collision, il est l'objet de nombreux débats depuis sa première formulation par

---

trace d'une matrice est invariante lors d'un changement de base.

<sup>9</sup> On se place ici dans le cas où l'espace des configurations du système est de dimension 3.

Boltzmann dès 1870 dans sa fameuse équation. Ces débats ont pour enjeu l'irréversibilité des phénomènes macroscopiques ou plus explicitement les sources de la dissipation en physique. L'interaction entre les particules peut être modélisée selon plusieurs schémas. Dans les théories de champ moyen, on considère que l'interaction entre une particule et l'ensemble de toutes les autres est résumée par des quantités moyennes. On peut également prendre en compte des interactions locales comme des passages proches dans un champ électrostatiques ou gravitationnels qui vont modifier les trajectoires des particules concernées : on parle généralement de collisions sans que pourtant il n'y ait de choc, simplement une interaction locale entre deux, trois ou plus de particules. Les propriétés générales de l'opérateur  $Q(f, f)$  permettent d'établir dans de nombreux cas le théorème  $H$  démontrant l'augmentation de l'entropie  $S = -k_B \int f \ln f d\vec{r}d\vec{p}$  au cours du temps. Ce fut l'une des œuvres de jeunesse de Boltzmann.

Nous mentionnerons deux exemples fondamentaux pour  $Q(f, f)$ : le cas du champ moyen non collisionnel (équation de Vlasov) et le champ moyen dans l'approximation linéaire (équation de Boltzmann).

### 1. L'équation de Vlasov

Si le système est conservatif toutes les forces subies par les particules dérivent d'un potentiel. En ce qui concerne la particule moyenne que nous considérons, elle subit une force

$$\vec{F}(f) = -\frac{\partial U}{\partial \vec{r}}$$

dérivant du potentiel de champ moyen

$$U(\vec{r}) = g(\vec{r}) *_{\vec{r}} \int f d\vec{p}$$

Cette équation généralise l'équation de Poisson de la gravitation ou de l'électrostatique : elle décrit via la convolution l'interaction non forcément locale entre une source (densité spatiale de particules) et une fonction de Green  $g$  (réponse impulsionnelle pour les physiciens) pour l'interaction considérée. L'opérateur de collision est alors réduit à sa plus simple expression

$$Q(f, f) = \vec{F}(f) \cdot \frac{\partial f}{\partial \vec{p}}$$

L'équation de Vlasov permet d'obtenir les équations de la physique statistique non dissipative : équation de continuité, équation d'Euler de la mécanique des fluides, etc. On calcule pour cela les moments d'ordre 0, 1, etc. de cette équation sur l'espace des impulsions.

## 2. L'équation de Boltzmann

$$Q(f, f) = \vec{F}(f) \cdot \frac{\partial f}{\partial \vec{p}} + \frac{f - f_0}{\tau}$$

Le temps  $\tau$  représente le temps de vol entre deux interactions dans le système considéré. La fonction  $f_0$  est une solution stationnaire de ce système. C'est ce type d'équation, dans cette approximation du temps de vol, qui permet par exemple d'obtenir toutes les relations linéaires entre les courants et les différences de potentiel dans divers cas de transport de particules : loi d'Ohm pour les charges électriques, loi de Fourier pour les photons, loi de Fick pour les neutrons, etc.

Relativité restreinte et électromagnétisme



## Relativité restreinte

### 4.1 Insuffisances de la mécanique classique

Il est tout à fait naturel de penser que comme tout flatteur, le mouvement d'un point vit aux dépens de celui qui l'observe : c'est la fable classique du voyageur et du chef de gare<sup>1</sup>.

*Au passage devant une station,  
Un voyageur assis dans le train,  
Au contrôleur, pose une question.*

...

Pour le chef de gare le contrôleur se déplace à la même vitesse que le train, pour le voyageur le contrôleur est immobile. L'analyse de ces observations augmentée de l'isotropie de l'espace et de l'uniformité du temps conduit à la transformation de Galilée, ou loi de composition classique des vitesses. Avec la découverte en 1873 des équations de l'électromagnétisme par Maxwell de gros problèmes surgissent. Selon ces équations, il existe en effet des ondes associées au champ électromagnétique, et celles-ci se propagent à une vitesse  $c$  dans un référentiel qu'il convient de choisir. Ce dernier devient alors le seul dans lequel les équations de Maxwell peuvent être écrites. Bien pire encore : comme on le pense à l'époque et comme le son se propage dans l'air, toute onde se propage dans un milieu. Il est alors toujours possible d'imaginer que l'univers entier baigne dans un milieu, l'éther, support de la propagation des ondes électromagnétiques. Cela pourrait peut-être aussi régler le problème de l'interaction gravitationnelle de Monsieur Newton qui agit à distance et sans support apparent. Les propriétés de l'éther sont toutefois extraordinaires : omniprésent, infiniment rigide (pour propager les ondes

---

<sup>1</sup> Nous sommes ici en présence d'un nouveau type de fable d'environ neuf pieds.

transverses à grande vitesse), impondérable, infiniment perméable aux objets matériels puisqu'il n'affecte pas le mouvement de la Terre depuis des siècles. L'histoire du Calorique (fluide hypothétique imaginé au XVIII<sup>e</sup> siècle pour assurer la propagation de la chaleur) ou du Phlogistique (milieu créé par les anciens chimistes pour expliquer la combustion, et dont Lavoisier montra l'inexistence ...) doit alors nous faire raison garder ! D'autant plus que philosophiquement, l'intervention de l'éther est anti-démocratique : en le créant, on crée un référentiel despote qui force l'écriture de certaines équations sous son empire. Expérimentalement, si la vitesse de ces ondes obéit à la transformation de Galilée, on doit pouvoir mesurer des variations de cette vitesse pour des mouvements assez rapides par rapport à l'éther. Le mouvement de la Terre sur son orbite autour du Soleil à 30 km/s est suffisamment rapide pour mettre en évidence de telles variations par des expériences d'interférométrie optique. Pourtant la célèbre expérience de Michelson (1881) - qu'il reconduisit en 1887 avec Morley - n'a rien mis en évidence du tout !

On peut malgré tout essayer de réparer la théorie de l'éther, en disant que ce dernier est entraîné par la Terre dans son voisinage, et patati et patata ... On peut aussi comme le fit Albert Einstein en 1905, écarter l'éther d'un revers de main et reformuler la cinématique en prenant en compte le caractère nouveau imposé par la propagation de la lumière.

Pour cela, adoptons avec lui un principe qui nous semble bien plus acceptable que la tyrannie étherifique : « *Les équations de la physique s'écrivent de la même façon dans tous les référentiels en translations uniformes les uns par rapports aux autres (référentiels galiléens).* » Ce principe, dit de relativité restreinte, a une conséquence très claire : les équations de Maxwell s'écrivant de la même façon dans tous les référentiels galiléens, la vitesse de la lumière  $y$  est donc la même et ne dépend donc pas du référentiel choisi. C'est sur cette invariance que nous allons construire la nouvelle cinématique et en déduire la nouvelle dynamique.

Comme dans toutes les parties de ce livre, nous n'obtiendrons que les équations de relativité restreinte et ne chercherons pas à les résoudre dans diverses situations. De nombreux ouvrages présentent la relativité restreinte de façon bien plus approfondie et avec des applications. Nous citerons par exemple le superbe ouvrage d'Eric Gourgouhlon [11] et les notes de cours de Jean-Michel Raymond [21] dont nous nous sommes inspirés.

## 4.2 Transformation de Lorentz

### 4.2.1 Nature de la transformation de Lorentz

Soit un point  $M$ , de coordonnées  $(x, y, z)$  à l'instant  $t$  dans un repère  $\mathcal{R}$  et  $(x', y', z')$  à l'instant  $t'$  dans un repère  $\mathcal{R}'$ . Nous nous plaçons d'emblée dans le cadre d'une transformation de Lorentz spéciale (TLS), c'est-à-dire que nous considérons que  $\mathcal{R}'$  est animé d'un mouvement de translation rectiligne uniforme de vitesse  $\vec{v}$  dirigée selon  $Ox$  (voir figure 4.1).

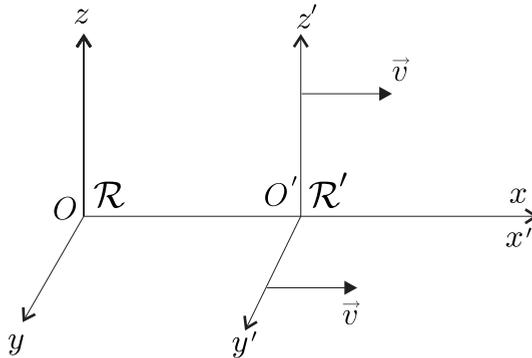


Fig. 4.1. Transformation de Lorentz spéciale

En mécanique classique, la loi de composition des vitesses (transformation de Galilée) s'écrit simplement (en posant  $v = \|\vec{v}\|$ )

$$\begin{cases} x' = x - vt \\ y' = y \\ z' = z \\ t' = t \end{cases} \quad (4.1)$$

Pour satisfaire le Principe de Relativité Restreinte, nous devons donc chercher une relation plus compliquée de la forme

$$\begin{cases} x' = F(x, y, z, t) \\ y' = G(x, y, z, t) \\ z' = H(x, y, z, t) \\ t' = K(x, y, z, t) \end{cases} \quad (4.2)$$

L'application de ce principe impose que :

1. Un mouvement rectiligne uniforme dans  $\mathcal{R}$ , doit être uniforme dans  $\mathcal{R}'$ , les fonctions  $F, G, H$  et  $K$  sont donc linéaires.
2. Les coefficients de la transformation (4.2) ne peuvent être que des fonctions de la vitesse d'entraînement de  $\mathcal{R}$  par rapport à  $\mathcal{R}'$ , donc de  $\vec{v}$ .
3. Les relations donnant  $x, y, z$  et  $t$  en fonction de  $x', y', z'$  et  $t'$  doivent être les mêmes que celles donnant  $x', y', z'$  et  $t'$  en fonction de  $x, y, z$  et  $t$  en changeant  $\vec{v}$  en  $-\vec{v}$ .
4. Si un évènement se produit le long de  $Ox$  (ou  $Ox'$ ), il est tel que  $y = y' = z = z' = 0$ , ceci pour toute valeur de  $x$  (ou  $x'$ ). La coordonnée  $x'$  ne dépend donc que de  $x$  et  $t$ , et les coordonnées  $y'$  et  $z'$ , ne dépendent ni de  $x$ , ni de  $t$ .

On a donc forcément

$$\begin{cases} x' = ax + bt \\ y' = y \\ z' = z \\ t' = dx + et \end{cases} \quad (4.3)$$

Pour plus d'homogénéité dans les formules, considérons la variable  $ct$  en lieu et place du temps  $t$ ,  $c$  étant bien sûr la vitesse de la lumière qui est donc la même dans les deux référentiels<sup>2</sup>. Dans ces conditions la nouvelle transformation s'écrit

$$\begin{bmatrix} x' \\ ct' \end{bmatrix} = \begin{bmatrix} A & B \\ D & E \end{bmatrix} \begin{bmatrix} x \\ ct \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} y' \\ z' \end{bmatrix} = Id_{\mathbb{R}^2} \begin{bmatrix} y \\ z \end{bmatrix} \quad (4.4)$$

Pour déterminer les réels  $A, B, D$  et  $E$  deux remarques s'imposent :

1. Le point  $O(x = 0)$  a une vitesse  $-\vec{v}$  dans  $\mathcal{R}'$ , pour ce point nous avons donc  $x' = -vt'$  alors que

$$\begin{cases} x' = Bct \\ ct' = Ect \end{cases} \quad (4.5)$$

Nous en tirons

$$\frac{B}{E} = -\frac{v}{c} := -\beta \quad (4.6)$$

2. En un temps  $t = 0$ , une source située en  $O$ , émet un signal lumineux se propageant à la vitesse (de la lumière)  $c$ , dans toutes les directions de l'espace. En un temps  $t$  quelconque, un observateur placé en  $x, y, z$  dans  $\mathcal{R}$ , reçoit le signal, on a donc  $x^2 + y^2 + z^2 = c^2t^2$ . Dans le repère  $\mathcal{R}'$ , et

<sup>2</sup> On pourrait envisager n'importe quelle autre vitesse choisie constante, on simplifie les notations en considérant celle de la lumière.

puisque  $c$  est invariant selon le principe de relativité restreinte, on doit donc avoir  $x'^2 + y'^2 + z'^2 = c^2 t'^2$ . En regroupant ces deux équations il vient donc

$$x^2 - c^2 t^2 = x'^2 - c^2 t'^2 \quad (4.7)$$

En incorporant la transformation (4.4) dans la relation (4.7), il vient

$$x^2 - c^2 t^2 = (Ax + Bct)^2 - (Dx + Ect)^2 \quad (4.8)$$

$$= (A^2 - D^2)x^2 + 2xct(AB - ED) - c^2 t^2(E^2 - B^2) \quad (4.9)$$

Cette relation est vraie pour tout  $x$  et à chaque instant  $t$ , ainsi, on doit avoir simultanément

$$A^2 - D^2 = 1 \quad (4.10)$$

$$E^2 - B^2 = 1 \quad (4.11)$$

$$AB - ED = 0 \quad (4.12)$$

Les deux relations (4.10) et (4.11) impliquent

$$\exists(\psi_1, \psi_2) \in \mathbb{R}^2 \text{ tel que } \begin{cases} A = \cosh \psi_1 & , & E = \cosh \psi_2 \\ D = \sinh \psi_1 & , & B = \sinh \psi_2 \end{cases} \quad (4.13)$$

avec ces nouvelles variables, la relation (4.12) devient

$$0 = \cosh(\psi_1) \sinh(\psi_2) - \cosh(\psi_2) \sinh(\psi_1) = \sinh(\psi_2 - \psi_1) \quad (4.14)$$

la fonction sinus hyperbolique étant bijective nous en déduisons que

$$\psi_2 = \psi_1 := \psi \quad (4.15)$$

La grandeur sans dimension  $\psi$  est appelée rapidité. C'est le moment de «ressortir» la relation (4.6) pour constater que

$$\tanh \psi = -\beta \Rightarrow \psi = -\arg \tanh \beta = -\frac{1}{2} \ln \left( \frac{1 + \beta}{1 - \beta} \right) \quad (4.16)$$

quelques lignes de trigonométrie hyperbolique donnent alors

$$A = E = \cosh(-\arg \tanh \beta) = +\frac{1}{\sqrt{1 - \beta^2}} := \gamma \quad (4.17)$$

$$D = B = \sinh(-\arg \tanh \beta) = -\frac{\beta}{\sqrt{1 - \beta^2}} = -\beta\gamma \quad (4.18)$$

La transformation recherchée s'écrit donc finalement

$$\begin{bmatrix} x' \\ ct' \end{bmatrix} = \begin{bmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{bmatrix} \begin{bmatrix} x \\ ct \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} y' \\ z' \end{bmatrix} = Id_{\mathbb{R}^2} \begin{bmatrix} y \\ z \end{bmatrix} \quad (4.19)$$

Il s'agit d'une transformation de Lorentz spéciale (TLS). Cette transformation a été obtenue en imposant la conservation de la forme quadratique de la relation (4.7) et l'invariance de la vitesse de la lumière. Il semble clair que cette forme quadratique reste invariante sous une réflexion du temps ou une quelconque isométrie spatiale. L'axe  $Ox$  que nous avons considéré est arbitraire, on montre donc assez directement le résultat fondamental suivant

**Théorème 4.1.** *Toute transformation linéaire des coordonnées  $(ct, x, y, z)$  qui conserve l'intervalle  $\ell^2 = (ct)^2 - x^2 - y^2 - z^2$  est la composée d'une TLS avec des isométries spatiales et des réflexions du temps.*

L'ensemble de ces transformations forme un groupe  $SO(1, 3)$  pour 1 signe  $+$  et 3 signes  $-$ . Il est appelé groupe de Lorentz. Dans la limite classique  $\gamma = 1$ , le groupe de Lorentz se réduit à un sous-groupe dit de Galilée, qui décrit la loi de composition des vitesses de la mécanique classique. Si l'on rajoute les translations d'espace-temps au groupe de Lorentz, on obtient le cas le plus général : le groupe de Poincaré.

## 4.2.2 Loi de composition des vitesses relativistes

Considérons à présent le changement successif de 2 référentiels :  $\mathcal{R} \rightarrow \mathcal{R}' \rightarrow \mathcal{R}''$  avec des vitesses de translations respectives  $\vec{v}$  et  $\vec{v}'$ . Nous nous plaçons ici encore une fois dans le cadre de transformations de Lorentz spéciales selon le même axe, la généralisation ne pose pas de problèmes autres que calculatoires. Les composantes d'un évènement  $E$  sont repérées par les quadruplets  $(ct, x, y, z)^\top$  dans  $\mathcal{R}$ ,  $(ct', x', y', z')^\top$  dans  $\mathcal{R}'$  et  $(ct'', x'', y'', z'')^\top$  dans  $\mathcal{R}''$ .

En posant,  $\psi = -\arg \tanh \beta$  avec  $\beta = v/c$  et  $\psi' = -\arg \tanh \beta'$  avec  $\beta' = v'/c$ , les résultats du paragraphe précédent nous permettent d'écrire successivement

$$\begin{bmatrix} x' \\ ct' \end{bmatrix} = \begin{bmatrix} \cosh \psi & \sinh \psi \\ \sinh \psi & \cosh \psi \end{bmatrix} \begin{bmatrix} x \\ ct \end{bmatrix} \quad \text{avec} \quad \begin{cases} \cosh \psi = \gamma \\ \sinh \psi = -\beta\gamma \end{cases} \quad (4.20)$$

puis

$$\begin{bmatrix} x'' \\ ct'' \end{bmatrix} = \begin{bmatrix} \cosh \psi' & \sinh \psi' \\ \sinh \psi' & \cosh \psi' \end{bmatrix} \begin{bmatrix} x' \\ ct' \end{bmatrix} \quad \text{avec} \quad \begin{cases} \cosh \psi' = \gamma' \\ \sinh \psi' = -\beta'\gamma' \end{cases} \quad (4.21)$$

en mélangeant ces deux transformations, et en utilisant quelques règles simples de trigonométrie hyperbolique, il vient

$$\begin{bmatrix} x'' \\ ct'' \end{bmatrix} = \begin{bmatrix} \cosh \psi'' & \sinh \psi'' \\ \sinh \psi'' & \cosh \psi'' \end{bmatrix} \begin{bmatrix} x \\ ct \end{bmatrix} \quad (4.22)$$

où l'on a posé  $\psi'' = \psi + \psi'$ . En appelant  $\vec{v}''$  la vitesse qui permet de passer directement de  $\mathcal{R}$  à  $\mathcal{R}''$ , nous aurons automatiquement  $\cosh \psi'' = \gamma''$  et  $\sinh \psi'' = -\beta''\gamma''$ . En explicitant  $\psi''$  dans ces deux dernières relations, il vient

$$\begin{aligned} \gamma'' &= \cosh(\psi + \psi') = \cosh \psi \cosh \psi' + \sinh \psi \sinh \psi' = \gamma\gamma'(1 + \beta\beta') \\ -\beta''\gamma'' &= \sinh(\psi + \psi') = \cosh \psi \sinh \psi' + \sinh \psi \cosh \psi' = -\gamma\gamma'(\beta + \beta') \end{aligned}$$

le rapport de ces deux dernières équations fournit gracieusement

$$\beta'' = \frac{\beta + \beta'}{1 + \beta\beta'} \quad (4.23)$$

qui n'est autre que la formule de composition des vitesses relativistes. Cette relation généralise la loi d'addition des vitesses de la mécanique classique à laquelle elle se réduit lorsque  $\beta$  et  $\beta'$  sont très inférieurs à 1. Une remarque pour finir, la représentation de la transformation de Lorentz par la rapidité est additive : une composition des transformations revient à une addition des paramètres. La rapidité est donc le bon paramètre de la cinématique relativiste, adieu la vitesse concept essentiellement classique.

### 4.3 Notations quadridimensionnelles

Il est temps de prendre de bonnes habitudes !

Comme dans beaucoup de domaines de la physique ou des mathématiques, les notations sont fondamentales en relativité. Pour écrire les choses simplement et de façon compréhensible, il faut adopter des notations efficaces. Nous allons les présenter dans un premier temps de façon progressive et intuitive afin de pouvoir partager un langage commun dont nous comprendrons l'origine. Une fois muni de ce peu d'expérience, nous reviendrons en détail (section 4.3.4) sur ces notations et sur la notion de tenseur qui sera définie dans un contexte plus général que celui dont nous avons besoin en relativité. Il nous a semblé préférable et pédagogique de procéder de cette façon constructive, le lecteur pointilleux ou pressé pourra tout à fait lire ou se référer à cette section quand il le souhaite.

### 4.3.1 Le 4–vecteur position d'un évènement

#### Composante contravariante

Dans un référentiel donné, un évènement est parfaitement déterminé par la donnée de 4 nombres : l'instant  $t$  où se produit cet évènement et les 3 coordonnées spatiales que nous avons appelées  $x$ ,  $y$  et  $z$  dans la base d'usage de  $\mathbb{R}^3$ , ces 3 coordonnées sont donc celles du vecteur position  $\vec{r}$  repérant l'évènement dans notre espace euclidien habituel.

Soit  $\mathbb{M}_{4,\mathbb{R}}$ , le  $\mathbb{R}$ –espace vectoriel de dimension 4 muni d'une base  $B = \{\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ . Tout évènement  $\mathcal{E}$  peut être associé à un vecteur  $\mathbf{x} \in \mathbb{M}_{4,\mathbb{R}}$  par le biais de ses composantes, notées  $x^\mu$ , dans cette base :

$$\mathbf{x} = \sum_{\mu=0}^3 x^\mu \mathbf{e}_\mu \quad (4.24)$$

On dira que l'évènement  $\mathcal{E}$  est associé à la 4–position  $\mathbf{x}$  dont les composantes dans la base  $B$  de  $\mathbb{M}_{4,\mathbb{R}}$  sont  $x^0 = ct$ ,  $x^1 = x$ ,  $x^2 = y$  et  $x^3 = z$ .

Afin de simplifier les notations en éliminant les signes  $\sum$ , nous adopterons la convention d'Einstein qui nous autorisera à identifier la présence d'un même indice en «haut» et en «bas» dans un produit de termes avec la sommation sur cet indice.

Nous adopterons aussi la convention habituelle de sommer sur l'ensemble des valeurs possibles pour les indices grecs minuscules ( $\mu = 0, 1, 2, 3$ ) et sur la partie spatiale uniquement pour les indices latins ( $i = 1, 2, 3$ ), ainsi, sans aucune équivoque, nous écrirons la relation (4.24) sous la forme simplifiée

$$\mathbf{x} = x^\mu \mathbf{e}_\mu = x^0 \mathbf{e}_0 + x^i \mathbf{e}_i \quad (4.25)$$

Les quatre nombres  $x^\mu$  sont appelés composantes contravariantes de  $\mathbf{x}$ .

Il est toujours possible de munir  $\mathbb{M}_{4,\mathbb{R}}$  d'un produit scalaire. Si nous considérons deux évènements  $\mathbf{x} = x^\mu \mathbf{e}_\mu$  et  $\mathbf{y} = y^\nu \mathbf{e}_\nu$ , nous aurons

$$\mathbf{x} \cdot \mathbf{y} := x^\mu y^\nu \mathbf{e}_\mu \cdot \mathbf{e}_\nu \quad (4.26)$$

Il ne reste plus qu'à définir le produit  $\mathbf{e}_\mu \cdot \mathbf{e}_\nu$  : c'est ce qui permet généralement de définir des longueurs qui comme nous le savons bien ne dépendent pas de la base considérée. Si nous étions par exemple dans la base canonique de  $\mathbb{R}^4$ , nous aurions simplement

$$\mathbf{e}_\mu \cdot \mathbf{e}_\nu = \delta_{\mu\nu} \quad \text{et} \quad \forall \mathbf{x} \in \mathbb{R}^4, \quad \mathbf{x} \cdot \mathbf{x} = (ct)^2 + x^2 + y^2 + z^2$$

Mais cela ne nous servirait pas beaucoup car les longueurs ne sont que des intervalles d'espace !

Avec la transformation de Lorentz, nous venons de voir que la nature impose de rajouter le temps à cette histoire. Les intervalles invariants que nous pouvons mesurer dans des expériences doivent être considérés dans l'espace *et* le temps : c'est-à-dire dans l'*espacetemps*<sup>3</sup>

En conséquence, si nous souhaitons relier le produit scalaire  $\mathbf{x} \cdot \mathbf{y}$  à la nouvelle notion d'intervalle invariant sous les transformations de Lorentz, et donc identifier  $\mathbb{M}_{4,\mathbb{R}}$  à l'espacetemps de la relativité restreinte, nous sommes amenés à poser

$$\forall \mathbf{x} \in \mathbb{M}_{4,\mathbb{R}}, \quad \mathbf{x} \cdot \mathbf{x} = c^2 t^2 - x^2 - y^2 - z^2 \quad (4.27)$$

C'est tout à fait possible à condition de définir pour les vecteurs de la base canonique de  $\mathbb{M}_{4,\mathbb{R}}$  les quantités notées

$$\eta_{\mu\nu} := \mathbf{e}_\mu \cdot \mathbf{e}_\nu = \begin{cases} +1 & \text{si } \mu = \nu = 0 \\ -1 & \text{si } \mu = \nu = 1, 2 \text{ ou } 3 \\ 0 & \text{sinon} \end{cases} \quad (4.28)$$

qui, en utilisant la convention d'Einstein, permettent d'écrire

$$\forall (\mathbf{x}, \mathbf{y}) \in \mathbb{M}_{4,\mathbb{R}} \times \mathbb{M}_{4,\mathbb{R}} \quad \mathbf{x} \cdot \mathbf{y} = \eta_{\mu\nu} x^\mu y^\nu \quad (4.29)$$

Les 16 valeurs de  $\eta_{\mu\nu}$  permettent de doter l'espace vectoriel  $\mathbb{M}_{4,\mathbb{R}}$  de la notion d'intervalle si utile en relativité. Cette notion n'est pas définie positive, il ne faudrait donc pas parler de produit scalaire, mais il n'y a pas d'ambiguïté si l'on ne l'applique qu'à des vecteurs de  $\mathbb{M}_{4,\mathbb{R}}$ . Afin de les particulariser nous les appellerons des 4-vecteurs.

Il existe un vocabulaire précis pour parler de toutes ces choses. L'espace vectoriel  $\mathbb{R}^n$ , muni du produit scalaire canonique (4.3.1) est un espace euclidien. L'espace vectoriel  $\mathbb{M}_{4,\mathbb{R}}$  est quant à lui un espace de Minkowski en hommage à l'inventeur de ce formalisme et de ces notations. Les composantes  $\eta_{\mu\nu}$  sont celle d'un objet que l'on notera  $\boldsymbol{\eta}$  et qui permet d'obtenir un scalaire lorsqu'on lui fournit 2 vecteurs : nous verrons qu'il s'agit donc d'un tenseur d'ordre 2, le tenseur métrique de  $\mathbb{M}_{4,\mathbb{R}}$ . La notation complète  $\mathbb{M}_{4,\mathbb{R}}$  pour espace de Minkowski de dimension 4 sur le corps des réels sera noté  $\mathbb{M}^4$  s'il n'y a pas d'ambiguïté. Les gens pointilleux parlent de pseudo-vecteurs et de pseudo-produit scalaire pour qualifier les atours de  $\mathbb{M}^4$ , mais laissons les pseudos aux romantiques !

<sup>3</sup> Ce mot ne fait malheureusement pas partie de la langue française. Il est néanmoins tellement

### Composante covariante

De manière formelle posons

$$y_\mu = \eta_{\mu\nu} y^\nu \tag{4.30}$$

dans cette relation, il y a bien sûr sommation sur l'indice  $\nu$  ce qui explique son absence dans le terme de gauche de l'équation. Dans ces conditions, pour tous les éléments de  $\mathbb{M}^4$  le produit scalaire s'écrit

$$\mathbf{x} \cdot \mathbf{y} = x^\mu y_\mu \tag{4.31}$$

expression bien évidemment sommée sur l'indice  $\mu$  : c'est la dernière fois que nous le précisons, mais il faut y aller doucement au début ! Le terme  $y_\mu$  est appelé composante covariante de l'évènement  $\mathbf{y}$ . Elle permet à elle seule, lorsqu'on l'applique à la composante contravariante de  $\mathbf{x}$  d'obtenir un scalaire. Il s'agit donc d'une forme qui vit dans le dual de  $\mathbb{M}^4$ . Nous avons donc le schéma de principe suivant pour l'affectation des composantes d'un vecteur

$$\begin{array}{ccccc}
 \mathbb{M}^4 & & \text{Base} & \rightarrow & \text{composante} \\
 & & & & \text{contravariante} & x^\mu \\
 & \nearrow & & & & \searrow \\
 \text{vecteur : } \mathbf{x} & & & & & x_\nu = \eta_{\mu\nu} x^\mu \\
 & \searrow & & & & \nearrow \\
 \mathbb{M}^{4*} & & \text{Base} & \rightarrow & \text{composante} \\
 & & \text{duale} & & \text{covariante} & x_\nu
 \end{array} \tag{4.32}$$

En notant  $\eta^{\mu\nu}$ , l'inverse de  $\eta_{\mu\nu}$ , c'est-à-dire avec notre convention

$$\eta^{\mu\nu} = \begin{cases} +1 & \text{si } \mu = \nu = 0 \\ -1 & \text{si } \mu = \nu = 1, 2 \text{ ou } 3 \\ 0 & \text{sinon} \end{cases} \tag{4.33}$$

on vérifie bien que  $\eta^{\mu\alpha}\eta_{\alpha\nu} = \delta^\mu_\nu$ , représentation de l'identité de  $\mathbb{M}^4$  dans la base  $\mathbf{e}_\mu$  (souvent appelé symbole de Kronecker  $\delta^\mu_\nu = 1$  si  $\mu = \nu$  et 0 sinon). Nous avons alors

$$x^\mu = \delta^\mu_\nu x^\nu = \eta^{\mu\alpha}\eta_{\alpha\nu} x^\nu = \eta^{\mu\alpha} x_\alpha \tag{4.34}$$

Ainsi  $\eta_{\mu\nu}$  [resp.  $\eta^{\mu\nu}$ ] permet de «transformer» la composante contravariante [ resp. covariante ] en composante covariante [ resp. contravariante ]. D'un

---

dans la mouvance relativiste, dans le lien inextricable qu'il crée entre temps et espace, que nous l'adopterons dorénavant.

point de vue purement formel, les notations introduites permettent de constater que les symboles  $\eta_{\mu\nu}$  et  $\eta^{\mu\nu}$  s'utilisent comme des objets permettant de faire «monter» ou «descendre» les indices des composantes d'un vecteur. Nous verrons un peu plus loin que ce constat continue de fonctionner pour les composantes d'un tenseur.

A titre d'exercice d'application, montrons que le produit scalaire que nous avons défini est bien symétrique, i.e.  $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ , il s'agit d'un simple jeu d'écriture

$$\mathbf{x} \cdot \mathbf{y} = x^\mu y_\mu = \eta^{\mu\alpha} x_\alpha \eta_{\beta\mu} y^\beta = \eta^{\mu\alpha} \eta_{\beta\mu} x_\alpha y^\beta = \delta^\alpha_\beta x_\alpha y^\beta = y^\alpha x_\alpha = \mathbf{y} \cdot \mathbf{x} \quad (4.35)$$

### Changement de base

Nous avons vu au chapitre précédent que lors d'un changement de référentiel, ce que nous appelons maintenant la composante contravariante de l'évènement associé à  $\mathbf{x}$ , subissait une transformation de Lorentz. La forme particulièrement simple de cette transformation se réduit à une simple multiplication matricielle

$$x'^\mu = L^\mu_\nu x^\nu \quad (4.36)$$

si l'on appelle  $x^\nu$  et  $x'^\mu$  les composantes contravariantes respectives de l'évènement dans  $\mathcal{R}$  et  $\mathcal{R}'$  et si l'on note<sup>4</sup>

$$L^\mu_\nu = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.37)$$

Le reste n'est quasiment plus qu'une histoire de manipulation !

On peut en effet toujours écrire

$$x'^\mu = L^\mu_\nu x^\nu \Leftrightarrow \eta^{\mu\alpha} x'_\alpha = L^\mu_\nu \eta^{\nu\beta} x_\beta \quad (4.38)$$

ainsi

$$x'_\alpha = \eta_{\mu\alpha} L^\mu_\nu \eta^{\nu\beta} x_\beta \quad (4.39)$$

dans cette dernière relation le terme  $\eta_{\mu\alpha} L^\mu_\nu \eta^{\nu\beta}$  est sommé sur les indices  $\mu$  et  $\nu$ , on peut donc l'écrire

<sup>4</sup> Nous explicitons ici la représentation d'une transformation de Lorentz spéciale, comme nous l'avons déjà mentionné plus haut, le cas général fait appel à une matrice moins creuse et plus

$$\eta_{\mu\alpha} L^\mu{}_\nu \eta^{\nu\beta} = L_\alpha{}^\beta \quad (4.40)$$

la relation (4.39) est donc une transformation de Lorentz pour les composantes covariantes, elle s'écrit <sup>5</sup>

$$x'_\alpha = L_\alpha{}^\beta x_\beta \quad (4.41)$$

Une question brûle alors les lèvres de notre lecteur : quel lien existe-t-il entre  $L^\mu{}_\nu$  et  $L_\mu{}^\nu$  ? La réponse nous vient du principe de relativité restreinte qui stipule l'invariance de l'intervalle d'espace-temps sous les changements de référentiels. Grâce à notre définition du produit scalaire, cela se traduit en composantes par la relation

$$x'^\mu y'_\mu = x^\mu y_\mu \quad (4.42)$$

en explicitant les termes primés à travers les opérateurs de Lorentz, cette égalité devient

$$L^\mu{}_\alpha L_\mu{}^\beta x^\alpha y_\beta = x^\mu y_\mu \quad (4.43)$$

on doit donc avoir

$$L^\mu{}_\alpha L_\mu{}^\beta = \delta^\beta{}_\alpha \quad (4.44)$$

L'œil désormais averti aura remarqué que cette dernière relation n'est pas un produit matriciel, mais impose pour le moins à  $L^\mu{}_\nu$  d'être l'inverse de  $L_\mu{}^\nu$ .

Par simple jeu, vérifions que l'abaissement et l'élévation des indices «fonctionne» toujours aussi bien en utilisant le tenseur métrique sur les opérateurs de Lorentz, i.e.

$$L^\alpha{}_\beta = \eta_{\mu\beta} \eta^{\alpha\rho} L_\rho{}^\mu \quad (4.45)$$

c'est très simple<sup>6</sup> ...

Partons de l'invariance du produit scalaire :

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &= x^\mu y_\mu = x^\mu \eta_{\mu\alpha} y^\alpha = x'^\mu \eta_{\mu\alpha} y'^\alpha \\ &= L^\mu{}_\nu x^\nu \eta_{\mu\alpha} L^\alpha{}_\rho y^\rho \\ &= x^\mu L^\sigma{}_\mu \eta_{\sigma\beta} L^\beta{}_\alpha y^\alpha \end{aligned} \quad (4.46)$$

ainsi

---

complexe mais tout fonctionne de la même façon.

<sup>5</sup> Pour ne pas se tromper dans les écritures lors des changements de repère, il faut toujours que les indices sommés soient en «contact» ainsi c'est bien  $L_\alpha{}^\beta x_\beta$  et non pas  $L^\beta{}_\alpha x_\beta$  qui désigne la TL adéquate.

<sup>6</sup> Nous détaillons ici volontairement les calculs et manipulations d'indices, nous nous en excusons

$$\eta_{\mu\alpha} = L^\sigma{}_\mu \eta_{\sigma\beta} L^\beta{}_\alpha \quad (4.47)$$

en multipliant par  $\eta^{\nu\alpha}$  et en écrivant que  $\delta^\nu{}_\mu = L^\sigma{}_\mu L_{\sigma\nu}$ , il vient alors

$$L^\sigma{}_\mu L_{\sigma\nu} = L^\sigma{}_\mu \eta^{\nu\alpha} \eta_{\sigma\beta} L^\beta{}_\alpha \quad (4.48)$$

d'où

$$L_{\sigma\nu} = \eta^{\nu\alpha} \eta_{\sigma\beta} L^\beta{}_\alpha \quad \square \quad (4.49)$$

Nous pouvons alors «sauter» sur l'occasion pour définir deux nouvelles quantités :

$$L^{\alpha\beta} = \eta^{\alpha\mu} L_\mu{}^\beta \quad \text{et} \quad L_{\alpha\beta} = \eta_{\alpha\mu} L_\beta{}^\mu \quad (4.50)$$

Ces deux nouveaux «tenseurs lorentziens» permettent le passage simultané d'un référentiel à un autre et d'un type de composante à l'autre ! Nous avons en effet

$$x'^\mu = L^\mu{}_\nu x^\nu = L^\mu{}_\nu \eta^{\rho\nu} x_\rho = L^{\mu\rho} x_\rho \quad (4.51)$$

et de même

$$x'_\mu = L_{\mu\rho} x^\rho \quad (4.52)$$

### 4.3.2 Un peu de géométrie

#### Géométrie algébrique

En renonçant partiellement à la géométrie euclidienne, i.e. en rejetant l'hypothèse d'un temps universel, nous sommes donc conduits à décrire les phénomènes physiques dans un espace à 4 dimensions, l'espacetemps ou Univers :  $\mathbb{M}^4$

La structure géométrique de cet espace s'exprime à travers son produit scalaire introduit au paragraphe précédent : pour tout couple de vecteurs ou évènements  $(\mathbf{x}, \mathbf{y})$  de  $\mathbb{M}^4 \times \mathbb{M}^4$ , on note  $x^\mu$  et  $y^\mu$  les composantes contravariantes dans la base  $\mathbf{e}_\mu$  ou  $x_\nu$  et  $y_\nu$  les composantes covariantes dans la base duale. Nous avons défini un produit scalaire tel que

$$\mathbf{x} \cdot \mathbf{y} = \eta_{\mu\nu} x^\mu y^\nu = \eta^{\mu\nu} x_\mu y_\nu = x_\mu y^\mu = x^\nu y_\nu \quad (4.53)$$

Comme nous l'avons évoqué plus haut, ce produit permet de faire le lien entre la notion d'intervalle spatio-temporel et les structures habituelles de l'algèbre linéaire. Il n'en demeure pas moins que la norme induite par ce produit scalaire n'en est pas une. Il existe en effet des vecteurs non nuls de longueur nulle. C'est d'ailleurs ce petit souci qui va doter  $\mathbb{M}^4$ , et donc l'espacetemps dans lequel nous vivons, de propriétés géométriques rocambolesques.

On partitionne généralement  $\mathbb{M}^4$  en trois régions :

- Le cône chronologique  $\mathfrak{C}$  qui rassemble tous les vecteurs de genre temps

$$\mathfrak{C} = \{ \mathbf{x} \in \mathbb{M}^4, \mathbf{x} \cdot \mathbf{x} > 0 \} \quad (4.54)$$

- Le cône lumineux  $\mathfrak{L}$  qui rassemble tous les vecteurs de genre lumière

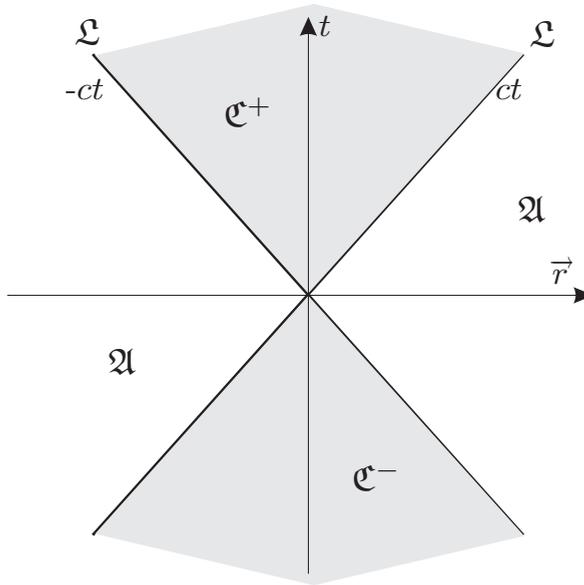
$$\mathfrak{L} = \{ \mathbf{x} \in \mathbb{M}^4, \mathbf{x} \cdot \mathbf{x} = 0 \} \quad (4.55)$$

ce sont justement les vecteurs de «longueur» nulle.

- L'ailleurs  $\mathfrak{A}$  qui rassemble tous les vecteurs de genre espace

$$\mathfrak{A} = \{ \mathbf{x} \in \mathbb{M}^4, \mathbf{x} \cdot \mathbf{x} < 0 \} \quad (4.56)$$

Ces régions sont représentées dans un diagramme d'espace-temps bidimensionnel sur la figure 4.2. L'ensemble  $\bar{\mathfrak{C}} = \mathfrak{C} \cup \mathfrak{L}$  des vecteurs qui ne sont pas



**Fig. 4.2.** Partition de l'espace-temps

de genre espace rassemble tous les vecteurs causaux. On dit que  $\bar{\mathfrak{C}}$  est le cône causal. De façon triviale, on montre que

$$\bar{\mathfrak{C}} \cup \mathfrak{A} = \mathbb{M}^4 \text{ et } \bar{\mathfrak{C}} \cap \mathfrak{A} = \emptyset \quad (4.57)$$

Le cône chronologique est lui-même séparé en deux régions distinctes :

- Le demi-cône du futur  $\mathfrak{C}^+$  qui contient tous les vecteurs de genre temps qui ont une composante temporelle positive

$$\mathfrak{C}^+ = \{\mathbf{x} \in \mathfrak{C}, x^0 > 0\} \quad (4.58)$$

- Le demi-cône du passé  $\mathfrak{C}^-$  qui contient tous les vecteurs de genre temps qui ont une composante temporelle négative

$$\mathfrak{C}^- = \{\mathbf{x} \in \mathfrak{C}, x^0 < 0\} \quad (4.59)$$

On montre alors que

$$\mathfrak{C}^\pm \cap \mathfrak{L} = \emptyset \quad (4.60)$$

Plus généralement, on peut vérifier directement que le cône chronologique ne contient aucun vecteur dont la composante temporelle est nulle.

$\mathbb{M}^4$ , l'espace-temps de la relativité restreinte, est plat<sup>7</sup>, les courbes de longueur minimale reliant deux points (géodésiques) sont des segments de droites. Avec cette topologie, on montre directement que les deux demi-cônes du futur et du passé sont deux parties convexes disjointes de  $\mathbb{M}^4$ . Cette propriété est tout à fait importante d'un point de vue physique : 2 évènements de notre futur strict ne sont reliés que par des évènements de notre futur strict, et idem pour notre passé...

C'est le scénariste de *Retour vers le futur* qui en prend un coup !

En notant  $\perp$ , l'orthogonal au sens du produit scalaire de  $\mathbb{M}^4$ , il est un jeu d'enfant de prouver les affirmations suivantes

$$\mathfrak{C}^\perp \subset \mathfrak{A}, \quad \mathfrak{A}^\perp \subset \mathfrak{C}, \quad \mathfrak{L}^\perp = \mathfrak{L} \quad (4.61)$$

Finalement, méditons sur les propriétés suivantes que nos lecteurs courageux démontreront sans peine : soient  $\mathbf{x}$  et  $\mathbf{y}$  deux évènements de  $\mathbb{M}^4$

- Si  $(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) > 0$ , on dit que  $\mathbf{x}$  et  $\mathbf{y}$  sont séparés par un intervalle de genre temps. Il existe alors un référentiel galiléen dans lequel les composantes spatiales de  $\mathbf{x}$  et de  $\mathbf{y}$  sont identiques<sup>8</sup>. Ils occupent dans ce référentiel la même position spatiale. Par contre, ils ne sont simultanés pour aucun observateur.
- Si  $(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) < 0$ , on dit que  $\mathbf{x}$  et  $\mathbf{y}$  sont séparés par un intervalle de genre espace. Ces deux évènements ne peuvent occuper la même position

auprès des lecteurs déjà habitués à ces petits tours de magie ...

<sup>7</sup> Nous pourrions le vérifier plus tard en relativité générale

<sup>8</sup> Cela signifie que l'on peut toujours trouver une transformation de Lorentz qui réalise cet

spatiale dans aucun référentiel galiléen. Par contre, certains observateurs fixes dans des référentiels bien choisis peuvent les déclarer simultanés. C'est-à-dire constater qu'ils ont la même composante temporelle.

- Si  $\mathbf{x} \cdot \mathbf{y} = 0$ , ces deux évènements peuvent être reliés par un rayon lumineux, et s'ils sont eux-mêmes lumineux ils sont proportionnels !

Nous nous limiterons ici à ces quelques propriétés de l'espace-temps en renvoyant le lecteur appâté vers des ouvrages plus avancés comme celui de D. Giulini [9] ou bien directement la bible d'Eric Gourgoulhon [11].

## Géométrie *dessinatoire*

Il est temps de prendre quelques exemples et de faire quelques dessins afin de bien comprendre ce que nous avons fait !

On peut représenter les transformations de Lorentz sous forme de diagrammes d'espace-temps. Nous nous limiterons ici au cas traité plus haut qui est celui d'une transformation de Lorentz spéciale, le référentiel  $\mathcal{R}'$  étant, par rapport à  $\mathcal{R}$ , en translation rectiligne uniforme à la vitesse  $v$  constante et portée par l'axe  $(O, \mathbf{e}_1)$  et dans son sens.

Oublions les coordonnées  $y$  et  $z$ , par exemple en les considérant nulles. Celà revient à se placer dans le plan  $\mathcal{P} = \text{vect}(\mathbf{e}_0, \mathbf{e}_1)$  de  $\mathbb{M}^4$ . Tout est comme en 3-physique mis à part que  $\mathbf{e}_1$  est un vecteur de genre espace car  $\mathbf{e}_1 \cdot \mathbf{e}_1 = -1$ , alors que  $\mathbf{e}_0$  est de genre temps car  $\mathbf{e}_0 \cdot \mathbf{e}_0 = 1$ . Ce petit détail change beaucoup de choses car ces deux vecteurs sont ceux qui permettent de repérer les évènements que nous considérons et qui s'écrivent  $\mathbf{m} = ct\mathbf{e}_0 + x\mathbf{e}_1$  dans  $\mathcal{R}$  ou bien  $\mathbf{m} = ct'\mathbf{e}'_0 + x'\mathbf{e}'_1$  dans  $\mathcal{R}'$ . Les vecteurs  $\mathbf{e}'_0$  et  $\mathbf{e}'_1$  étant les images respectives des vecteurs  $\mathbf{e}_0$  et  $\mathbf{e}_1$  par la transformation de Lorentz que nous considérons.

En terme de composantes, avec beaucoup d'abus de langage<sup>9</sup> on dira que la composante contravariante de  $\mathbf{m}$  est  $m^\mu = (ct, x, 0, 0)^\top$  dans  $\mathcal{R}$ ,  $m'^\mu = (ct', x', 0, 0)^\top$  dans  $\mathcal{R}'$  et que l'on a

$$m'^\nu = L^\nu{}_\mu m^\mu \quad \text{i.e.} \quad \begin{cases} x' = \gamma x - \beta \gamma ct \\ ct' = -\beta \gamma x + \gamma ct \end{cases} \quad (4.62)$$

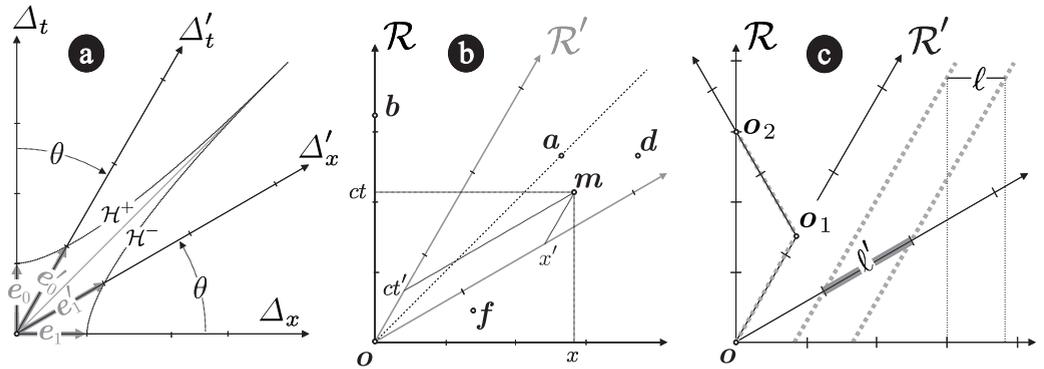
L'axe spatial de  $\mathcal{R}$  que l'on notera  $\Delta_x$  est ici la droite  $(O, \mathbf{e}_1)$ . De même, l'axe temporel de  $\mathcal{R}$  que l'on notera  $\Delta_t$  est la droite  $(O, \mathbf{e}_0)$ . Dans  $\mathcal{R}'$ , c'est

exploit.

<sup>9</sup> Ce qui est très courant en algèbre linéaire où l'on ne précise jamais que l'on se place dans les bases canoniques des espaces vectoriels que l'on considère...

la même chose avec des primes... Cependant, l'axe  $\Delta'_x$  est aussi le lieu des points tels que  $t' = 0$ . Le système (4.62) permet donc de trouver son équation dans le référentiel  $\mathcal{R}$ , elle s'écrit  $ct = \beta x$ . De même, l'axe  $\Delta'_t$  est le lieu des points tels que  $x' = 0$ , le même système permet toujours de trouver son équation dans le référentiel  $\mathcal{R}$ , elle s'écrit maintenant  $ct = \beta^{-1}x$ .

Géométriquement, si l'on définit l'angle  $\theta$  tel que  $\tan \theta = \beta$ , sans qu'il y ait de problème de définition car  $\beta \leq 1$ , on peut donc voir que d'une part les 4 axes  $\Delta_x$ ,  $\Delta'_x$ ,  $\Delta_t$  et  $\Delta'_t$  se coupent en un point  $O$ , et que d'autre part si  $\rho_{O,\alpha}$  est la rotation de centre  $O$  et d'angle  $\alpha$  alors  $\Delta'_x = \rho_{O,\theta}(\Delta_x)$  et  $\Delta'_t = \rho_{O,-\theta}(\Delta_t)$ . Ces axes sont représentés sur la figure 4.3-a.



**Fig. 4.3.** Géométrie de la transformation de Lorentz sur un diagramme d'espace-temps

Supposons que l'on choisisse des unités de mesure identiques pour les expériences que l'on peut faire dans les référentiels  $\mathcal{R}$  et  $\mathcal{R}'$  : on peut mesurer par exemple les temps en seconde et les distances en mètre. En mécanique classique, les résultats d'une mesure exprimés dans ces unités seront les mêmes dans les deux référentiels. Ce n'est plus le cas en relativité où ces résultats sont relatifs aux référentiels : c'est l'origine du nom de la théorie !

Seuls les scalaires ont la même valeur dans tous les référentiels : par exemple  $\mathbf{e}_1 \cdot \mathbf{e}_1 = \mathbf{e}'_1 \cdot \mathbf{e}'_1 = -1$ . Mais de manière évidente  $\mathbf{e}_1$  compte les unités d'espace dans  $\mathcal{R}$  alors que  $\mathbf{e}'_1$  fait la même chose dans  $\mathcal{R}'$ . Pour comparer les résultats d'une mesure, il faut donc se placer dans le même référentiel pour les effectuer.

Les vecteurs  $\mathbf{e}_1$  et  $\mathbf{e}'_1$  sont dans l'ensemble  $\mathcal{H}^- = \{\mathbf{m} \in \mathbb{M}^4, \mathbf{m} \cdot \mathbf{m} = -1\}$ . Dans le référentiel  $\mathcal{R}$ , compte-tenu de notre restriction à  $\mathcal{P}$ , cet ensemble est caractérisé par l'équation  $c^2t^2 - x^2 = -1$  qui est celle d'une hyperbole équilatère. Appartenant à la fois à  $\Delta'_x$  et à  $\mathcal{H}^-$ , le vecteur  $\mathbf{e}'_1$

est donc simplement défini par l'intersection entre cette droite et cette hyperbole qui est unique pour  $x > 0$ . Un raisonnement parfaitement symétrique montre que le vecteur  $\mathbf{e}'_0$  est défini par l'intersection entre  $\Delta'_t$  et  $\mathcal{H}^+ = \{\mathbf{m} \in \mathbb{M}_{4,\mathbb{R}}, \mathbf{m} \cdot \mathbf{m} = +1\}$ . Les vecteurs  $\mathbf{e}'_0$  et  $\mathbf{e}'_1$  peuvent donc être représentés sur la figure 4.3-a.

On n'oubliera pas que les coordonnées d'un vecteur dans un repère s'obtiennent par projection sur un axe parallèlement aux autres comme on peut le voir sur la figure 4.3-b. Sur cette même figure, nous pouvons maintenant visualiser un certain nombre de choses :

- Les évènements  $\mathbf{o}$  et  $\mathbf{b}$  se produisent au même endroit dans  $\mathcal{R}$ , en  $x = 0$ , mais pas en même temps car  $\mathbf{b}$  est visiblement dans le futur de  $\mathbf{o}$ . Par contre, dans  $\mathcal{R}'$ , il ne sont ni simultanés ni ne se produisent au même endroit ! On constate que le vecteur  $\mathbf{b} - \mathbf{o}$  est proportionnel à  $\mathbf{e}_0$ , il est donc de genre temps : nous sommes dans un cas particulier d'une propriété énoncée un peu plus haut.
- Les évènements  $\mathbf{a}$  et  $\mathbf{d}$  sont simultanés dans  $\mathcal{R}$  mais pas dans  $\mathcal{R}'$ . Il est clair que  $\mathbf{d} - \mathbf{a}$  est de genre espace car il est proportionnel à  $\mathbf{e}_1$ . Nous l'avons déjà dit : deux évènements distincts simultanés dans un référentiel sont forcément séparés par un intervalle de genre espace.
- Les évènements  $\mathbf{o}$  et  $\mathbf{a}$  sont tous les deux sur la bissectrice des repères : les vecteurs  $\mathbf{o}$ ,  $\mathbf{a}$  et  $\mathbf{a} - \mathbf{o}$  sont tous des vecteurs de lumière !
- On peut vérifier à la règle et au compas que les évènements  $\mathbf{m}$  et  $\mathbf{d}$  sont simultanés dans  $\mathcal{R}'$ , ainsi le vecteur  $\mathbf{d} - \mathbf{m}$  est de genre espace. Notons qu'ils ne sont clairement pas simultanés dans  $\mathcal{R}$ .
- Les évènements  $\mathbf{f}$  et  $\mathbf{a}$  se produisent au même endroit dans  $\mathcal{R}'$ , l'évènement  $\mathbf{a}$  ayant lieu un peu plus tard que l'évènement  $\mathbf{f}$ . Le vecteur  $\mathbf{f} - \mathbf{a}$  est de genre temps.

Toujours en faisant des dessins, et en observant le dernier diagramme (figure 4.3-c), on peut déjà apercevoir deux propriétés importantes des transformations de Lorentz :

- Une règle de longueur propre  $\ell'$  unité est immobile dans  $\mathcal{R}'$ . Elle se déplace donc à vitesse constante  $v$  par rapport à  $\mathcal{R}$ , dans lequel elle y apparaît avec une longueur  $\ell < \ell'$ . Sa longueur apparente dans  $\mathcal{R}$  semble tendre vers 0 lorsque  $v \rightarrow c$ .
- L'évènement  $\mathbf{o}_2$  situé dans le futur de  $\mathbf{o}$  peut être atteint suivant les deux histoires différentes suivantes<sup>10</sup> :

<sup>10</sup> Il n'y en a pas que 2, il y en a une infinité...

- en restant immobile dans  $\mathcal{R}$ , on attend simplement qu’il s’écoule 3 unités de temps sans bouger;
- On se déplace à la vitesse  $v$  selon  $(O, \mathbf{e}_1)$ , puis (évènement  $\mathbf{o}_1$ ) après un peu plus d’une unité de temps dans  $\mathcal{R}'$  dans lequel on est immobile, on fait instantanément demi-tour pour se déplacer à la vitesse  $-v$  selon  $(O, \mathbf{e}_1)$  pour rejoindre (c’est l’évènement  $\mathbf{o}_2$ ) l’individu resté immobile sur l’origine qui a simplement vieilli.

En examinant bien les durées écoulées dans chacun des deux référentiels, on aura pas de mal à se convaincre qu’elles sont différentes et qu’il s’est écoulé moins de temps dans  $\mathcal{R}'$  que dans  $\mathcal{R}$ .

Tout étant devenu clair, nous pouvons refermer ici cette parenthèse géométrique pour revenir à des calculs, que l’on peut aussi faire géométriquement (voir [11]).

### 4.3.3 Les 4–vecteurs de la physique

#### Définition

Nous n’avons considéré jusqu’à présent que le 4–vecteur position  $\mathbf{x}$  que l’on peut présenter comme le 4–vecteur fondamental de  $\mathbb{M}^4$ . La généralisation est cependant immédiate : on appelle 4–vecteur physique (ou quadrivecteur), tout ensemble ordonné de 4 quantités physiques, qui se transforme *comme* les composantes de la 4–position lors d’un changement de référentiel.

Puisqu’il y a au moins une contrainte à respecter, force est de constater que toute quantité physique réunissant 4 nombres n’est pas forcément un quadrivecteur.

Nous avons vu que les composantes contravariantes de la 4–position  $x^\mu$  se décomposent en une composante temporelle  $x^0 = ct$  et une composante spatiale rassemblée en un 3–vecteur  $\vec{r}$  précisant la position de l’évènement considéré dans l’espace rapporté à une base. En abusant du langage on pourra écrire  $x^\mu = (ct, \vec{r})^\top$ , avec un tel choix la composante covariante de  $\mathbf{x}$  s’écrit alors  $x_\mu = (ct, -\vec{r})$ .

#### Temps propre

Considérons deux jumeaux :

- l’un est fixe dans le référentiel galiléen  $\mathcal{R} = (O, \mathbf{e}_{\mu=0, \dots, 3})$  et repéré le vecteur  $\mathbf{x} = ct\mathbf{e}_0 + x\mathbf{e}_1$ ;

- l'autre est attaché à l'origine du référentiel  $\mathcal{R}' = (O', \mathbf{e}_{\mu=0,\dots,3})$  galiléen en translation rectiligne uniforme à la vitesse  $\mathbf{u} = v\mathbf{e}_1$ . On choisit  $t = 0$  de telle manière qu'à cet instant  $O = O'$ .

Le jumeau mobile attaché à  $\mathcal{R}'$  est repéré dans ce référentiel par le vecteur  $\mathbf{x}' = ct'\mathbf{e}_0$  car nous l'avons placé en l'origine pour laquelle  $x' = 0$ . La transformation de Lorentz adaptée s'écrit  $\mathbf{x}' = \mathbf{L}(\mathbf{x})$  et se traduit donc par le système d'équations

$$\begin{bmatrix} ct' \\ 0 \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{bmatrix} \begin{bmatrix} ct \\ x \end{bmatrix} \quad (4.63)$$

On en tire trivialement  $x = \beta ct$  puis  $t' = \gamma t(1 - \beta^2) = t/\gamma$ . Le temps  $t'$  est le temps propre du voyageur, on le note généralement  $\tau$ , il subit donc une contraction du facteur  $\gamma > 1$ . C'est la fameuse contraction du temps subie par tous les voyageurs relativistes<sup>11</sup>. Les intervalles de temps propres sont des invariants de Lorentz. Ce n'est pas difficile à vérifier !

Démonstration : Considérons un intervalle infinitésimal de temps propre  $d\tau = dt/\gamma$ , il vient directement

$$d\tau^2 = \frac{dt^2}{\gamma^2} = dt^2 \left(1 - \frac{v^2}{c^2}\right) = \frac{1}{c^2} \left[ c^2 dt^2 - dt^2 \left(\frac{d\mathbf{x}}{dt}\right)^2 \right] = \frac{d\mathbf{x} \cdot d\mathbf{x}}{c^2} \quad (4.64)$$

soit, pour des évènements de genre temps<sup>12</sup>

$$d\tau = \frac{\sqrt{d\mathbf{x} \cdot d\mathbf{x}}}{c}$$

l'élément d'intervalle d'espace-temps  $ds = \sqrt{d\mathbf{x} \cdot d\mathbf{x}}$  et la vitesse de la lumière étant des invariants de Lorentz, il en va de même pour l'intervalle de temps propre. On dit que  $d\tau$ , ou simplement  $\tau$  d'ailleurs, est un 4-scalaire.  $\square$

<sup>11</sup> Un calcul du même tonneau montre qu'une tige de longueur  $\ell'$  se déplaçant avec  $\mathcal{R}'$  a pour longueur  $\ell = \ell'/\gamma$  dans  $\mathcal{R}$ .

En effet, pour mesurer la tige dans  $\mathcal{R}$ , il suffit d'étaler une règle dans  $\mathcal{R}$  puis faire une photo de la tige au moment où elle passe devant la règle. Les extrémités  $x_1$  et  $x_2$  de la tige sont donc mesurées au même instant  $t$  : celui de la photo. La formule de transformation des longueurs donne  $x'_{1,2} = \gamma(x_{1,2} - \beta ct)$ , la longueur propre de la tige est par définition  $\ell' = x'_2 - x'_1$ , le calcul montre alors que  $\ell' = \gamma(x_2 - x_1)$ .

C'est la fameuse contraction relativiste des longueurs !

<sup>12</sup> Pour les vecteurs de genre espace il suffit de mettre un  $-$  sous la racine.

## La 4–vitesse et la 4–accélération

En mécanique classique, le vecteur vitesse d'un point est simplement la dérivée de son vecteur position par rapport au temps. La vitesse mesure donc la variation instantanée de position pendant une variation infinitésimale de temps. Cette séparation de l'espace et du temps ne convient plus du tout à nos nouvelles exigences relativistes. Pour construire une vitesse qui soit un 4–vecteur de  $\mathbb{M}^4$ , la quadrivitesse ou 4–vitesse, on peut mesurer la variation d'intervalle de longueur d'espace-temps au cours du temps propre soit

$$\mathbf{u} = \frac{d\mathbf{x}}{d\tau} \quad (4.65)$$

Puisque  $d\tau$  est un 4–scalaire  $\mathbf{u}$  est automatiquement un 4–vecteur.

Nous avons vu que  $dx^\mu = (cdt, dx, dy, dz)^\top$  et que  $d\tau = dt/\gamma$ , la composante contravariante de la quadrivitesse sera donc

$$u^\mu = \left( \gamma c, \gamma \frac{dx}{dt}, \gamma \frac{dy}{dt}, \gamma \frac{dz}{dt} \right)^\top = (\gamma c, \gamma \vec{v})^\top \quad (4.66)$$

Une remarque s'impose alors : la 4–vitesse est un vecteur de genre temps de longueur constante, en effet

$$\mathbf{u} \cdot \mathbf{u} = u^\mu u_\mu = \gamma^2 (c^2 - v^2) = \gamma^2 c^2 (1 - \beta^2) = c^2 \quad (4.67)$$

En multipliant la 4–vitesse par la masse, on définit la 4–quantité de mouvement  $\mathbf{p} = m\mathbf{u}$ , dont la composante contravariante est donc

$$p^\mu = \left( \gamma mc, \gamma m \frac{dx}{dt}, \gamma m \frac{dy}{dt}, \gamma m \frac{dz}{dt} \right)^\top. \quad (4.68)$$

La quadrivitesse est une fonction du temps propre, sa dérivée par rapport à cette variable, dont la composante contravariante s'écrit

$$\Gamma^\mu = \frac{du^\mu}{d\tau} = \gamma \frac{du^\mu}{dt} \quad (4.69)$$

est appelée 4–accélération.

Une petite propriété, la quadrivitesse et la quadri-accélération sont orthogonales :  $\mathbf{u} \cdot \mathbf{\Gamma} = 0$ . Au moins deux démonstrations sont possibles :

1. Un calcul explicite montre la nullité du produit scalaire, mais ce n'est pas drôle !

2. On sait que  $u^\mu u_\mu = c^2$ , ainsi

$$\frac{dc^2}{d\tau} = 0 = \frac{d}{d\tau} (u^\mu u_\mu) = \Gamma^\mu u_\mu + u^\mu \Gamma_\mu \quad (4.70)$$

mais

$$u^\mu \Gamma_\mu = \eta^{\beta\mu} u_\beta \eta_{\mu\alpha} \Gamma^\alpha = u_\beta \Gamma^\alpha \delta^\beta_\alpha = u_\alpha \Gamma^\alpha \quad (4.71)$$

d'où finalement

$$\frac{d}{d\tau} (u^\mu u_\mu) = 2\Gamma^\mu u_\mu = 0 \quad \square \quad (4.72)$$

On peut même aller encore plus vite sans composante...

### Le 4–vecteur densité de courant

La densité de courant relativiste est un 4–vecteur construit à partir de la densité spatiale de charges  $\rho(t, \vec{r})$ . Les charges en question peuvent être de toutes nature !

La charge étant généralement une quantité conservée, on peut écrire une équation de continuité ou conservation

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \vec{\nabla} \cdot (\rho\vec{v}) = 0 \quad (4.73)$$

une telle relation s'écrit donc en coordonnées cartésiennes (en introduisant un facteur  $c$  dans le terme de dérivée temporelle)

$$\frac{\partial}{\partial(ct)} (c\rho) + \frac{\partial}{\partial x} (\rho v_x) + \frac{\partial}{\partial y} (\rho v_y) + \frac{\partial}{\partial z} (\rho v_z) = 0 \quad (4.74)$$

en introduisant de façon formelle l'opérateur

$$\partial_\mu := \frac{\partial}{\partial x^\mu} \quad (4.75)$$

l'équation de continuité (4.74) se réécrit simplement

$$\partial_\mu \left( \rho \frac{dx^\mu}{dt} \right) = 0 \quad (4.76)$$

Nous sommes donc tout naturellement conduits à introduire la quantité

$$J^\mu := \rho \frac{dx^\mu}{dt} \quad (4.77)$$

qui s'interprète comme la composante contravariante de la densité de courant relativiste. Pour que les choses naturelles soient tout à fait correctes, il faut tout de même montrer que cette définition est cohérente et que  $\mathbf{J}$  est bien un 4-vecteur.

Pour cela partons d'un constat : la charge  $Q$  contenue dans un élément de volume spatial  $dV$  est la même dans tous les référentiels, avec les notations habituelles, nous aurons donc

$$Q = \rho dV = \rho' dV' \quad (4.78)$$

Le nombre  $Q$  est donc ce qu'il convient d'appeler dans notre contexte un 4-scalaire (quantité invariante sous un changement de référentiel). Par définition,  $d\mathbf{x}$  est un 4-vecteur, de composante contravariante  $dx^\mu$ . Le produit d'un 4-vecteur par un 4-scalaire, étant selon les préceptes de l'algèbre linéaire un 4-vecteur, la quantité  $\rho dV dx^\mu$  est donc la composante contravariante d'un 4-vecteur, ainsi il en va de même pour

$$\rho dV dt \frac{dx^\mu}{dt} \quad (4.79)$$

nous y sommes presque ...!

L'élément de volume de l'espacetemps  $d\Omega = c dt dV$  est un 4-scalaire, tout simplement parce que la transformation de Lorentz préserve l'élément de longueur de  $\mathbb{M}^4$ . Ainsi, la quantité

$$\frac{1}{d\Omega} \rho dV dt \frac{dx^\mu}{dt} = \frac{1}{c} J^\mu \quad (4.80)$$

est bien la composante contravariante d'un quadrivecteur, ce qui prouve bien que la nature fait bien les choses.

Notons pour en finir avec ce 4-vecteur que sa composante contravariante s'écrit comme chacun avait pu le remarquer au vu de la relation (4.77)

$$J^\mu = \left( \rho c, \vec{j} \right)^\top \quad (4.81)$$

où  $\vec{j}$  est le vecteur courant de la mécanique classique. En utilisant l'opérateur (4.75) de dérivation dans  $\mathbb{M}^4$  sur lequel nous reviendrons plus tard, l'équation de continuité (4.73) s'écrit donc tout simplement

$$\partial_\mu J^\mu = 0 \quad (4.82)$$

C'est plus simple !

## Le 4–vecteur d’onde

Considérons à présent une onde monochromatique de pulsation  $\omega$  et de vecteur d’onde  $\vec{k}$ . Les champs et potentiels associés à cette onde sont tous proportionnels à  $e^{i\varphi}$  où

$$\varphi = \omega t - \vec{k} \cdot \vec{r} \quad (4.83)$$

désigne la phase de l’onde.

Montrons que cette dernière ne dépend pas du référentiel choisi, et constitue donc un 4–scalaire.

Supposons pour cela que l’onde électromagnétique considérée se propage selon la direction du vecteur unitaire  $\vec{u}_k = \vec{k} / \|\vec{k}\|$ . Le vecteur d’onde s’écrit donc

$$\vec{k} = \frac{\omega}{c} \vec{u}_k = \frac{2\pi}{\lambda} \vec{u}_k = 2\pi \frac{\nu}{c} \vec{u}_k \quad (4.84)$$

où  $\lambda$  et  $\nu$  sont respectivement la longueur d’onde et la fréquence de l’onde considérée. Appelons  $t = 0$  l’instant où une crête référence passe en l’origine de la droite de vecteur directeur  $\vec{u}_k$ . Pour atteindre le point désigné par le vecteur  $\vec{r}$  sur la droite, cette crête mettra un temps  $t_1 = \vec{u}_k \cdot \vec{r} / c$ . Le temps

$$T = t - t_1 = t - \frac{\vec{u}_k \cdot \vec{r}}{c} \quad (4.85)$$

correspond à celui écoulé depuis que la crête est passée en  $\vec{r}$ . La quantité  $\nu T$  représente alors le nombre de crêtes qui sont passées en  $\vec{r}$  depuis le passage de notre crête référence (elle-même passée en  $t_1$ ), ce nombre est évidemment le même quel que soit le référentiel dans lequel on se place !

Or

$$\nu T = \nu \left( t - \frac{\vec{u} \cdot \vec{r}}{c} \right) = \frac{1}{2\pi} \left( \omega t - \frac{\omega}{c} \vec{u} \cdot \vec{r} \right) \quad (4.86)$$

soit

$$\nu T = \frac{1}{2\pi} \varphi \quad (4.87)$$

ce qui prouve bien que  $\varphi$  est un 4–scalaire.  $\square$

Revenons à notre propos, en *posant*

$$k^\mu = \left( \frac{\omega}{c}, \vec{k} \right)^\top \quad \text{soit} \quad k_\mu = \left( \frac{\omega}{c}, -\vec{k} \right) \quad (4.88)$$

la relation (4.83) peut s’écrire

$$\varphi = k_\mu x^\mu = k^\mu x_\mu = \mathbf{k} \cdot \mathbf{x} \quad (4.89)$$

La dernière égalité sous-entend que  $\mathbf{k} \in \mathbb{M}^4$  : c'est bien le cas car son produit scalaire dans cet espace de Minkowski avec  $\mathbf{x} \in \mathbb{M}^4$  donne bien un 4-scalaire, la phase ! Ce 4-vecteur est le 4-vecteur d'onde.

Quelques remarques d'ordre plus ou moins physique :

- Pour une onde plane dans le vide, tout le monde sait que  $\omega^2 - k^2 c^2 = 0$ , dans notre formalisme cette relation s'écrit

$$k_\mu k^\mu = 0 \quad (4.90)$$

qui met en évidence le fait que le 4-vecteur d'onde est de genre lumière, c'est le cas de tous les vecteurs associés à des informations se propageant à la vitesse de la lumière.

- Nous avons, par un moyen détourné, montré que le 4-vecteur d'onde est un quadrivecteur. Lors d'un changement de référentiel galiléen  $\mathcal{R} \rightarrow \mathcal{R}'$  nous pouvons donc écrire  $\mathbf{k}' = \mathbf{L}(\mathbf{k})$ , soit pour des composantes contravariantes

$$k'^\mu = L^\mu_\nu k^\nu \quad (4.91)$$

en explicitant cette transformation (toujours spéciale selon  $Ox$  pour simplifier les notations), si  $v$  est la vitesse de  $O'$  dans  $\mathcal{R}$  nous aurons

$$\begin{cases} \omega' = \gamma (\omega - vk_x) \\ k'_x = \gamma (k_x - \omega v/c^2) \\ k'_y = k_y \\ k'_z = k_z \end{cases} \quad (4.92)$$

La première de ces équations est la forme relativiste de l'effet Doppler, si  $\gamma \approx 1$  i.e.  $v \ll c$ , on retrouve le résultat connu de tous  $\omega' = \omega - vk$ , mais uniquement vrai en physique classique !

La deuxième relation est quant à elle complètement nouvelle et indique un changement relativiste de la direction de propagation de l'onde. Cet effet qui devra être pris en compte dans la navigation spatiale à grande vitesse est connu sous le nom d'aberration<sup>13</sup> relativiste de la lumière.

<sup>13</sup> L'aberration des étoiles a été découverte en 1728 par James Bradley. Il a remarqué que, pour observer des étoiles fixes, l'angle auquel on doit pointer son télescope varie avec la saison. James Bradley comprit que cette aberration était due au fait que la Terre tourne autour du Soleil. Le changement de la position apparente des étoiles dépend du rapport entre la vitesse de la lumière et la vitesse de la Terre autour du Soleil, ce qui permit à James Bradley de calculer la vitesse de la lumière (qui était connue depuis les travaux de Römer en 1675).

### 4.3.4 Les Tenseurs

La notion de produit tensoriel d'espace vectoriels est fondamentale en mathématique. Elle est très couramment utilisée dans de nombreux domaines de la physique. Elle est omniprésente en relativité !

Il est donc temps de la définir précisément et de clarifier les notations que nous avons déjà utilisées depuis le début de cet ouvrage avec des lettres grasses  $\mathbf{x}$  ou  $\mathbf{L}$ , des indices en «haut» ou en «bas», différents produits scalaires, etc.

Faisons un peu le point !

**Définition 4.1.** Soit  $E$  un  $\mathbb{K}$ -espace vectoriel de dimension finie et  $E^*$  l'ensemble des applications linéaires continues de  $E \rightarrow \mathbb{K}$ , i.e son dual. On appelle produit tensoriel de  $E^*$  par lui-même, l'ensemble noté

$$\mathcal{T}_2 = E^* \otimes E^*$$

de toutes les applications linéaires continues de  $E \times E \rightarrow \mathbb{K}$ .

Un élément de  $\mathcal{T}_2$  est appelé tenseur d'ordre 2. Un tenseur d'ordre 2 est donc tout simplement une application linéaire qui transforme deux vecteurs de  $E$  en un scalaire de  $\mathbb{K}$ . Ce n'est pas si compliqué que ça !

#### Tenseurs d'ordre 1

En étendant la définition 4.1 vers le bas, chaque vecteur  $\mathbf{x}$  de  $E$  peut être considéré comme un tenseur d'ordre 1. En effet, considérons

$$\forall \mathbf{x} \in E, \quad \exists ! \varphi_{\mathbf{x}}, \forall \mathbf{y} \in E : \begin{cases} E \rightarrow \mathbb{K} \\ \mathbf{y} \mapsto \varphi_{\mathbf{x}}(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y} \end{cases}$$

L'application  $\varphi_{\mathbf{x}} \in E^*$  est par définition un tenseur d'ordre 1. Ainsi, doter un espace vectoriel du produit scalaire  $\cdot$ , permet d'associer de façon bijective un vecteur à un tenseur d'ordre 1.

Pour représenter  $\forall \mathbf{x} \in E$ , il suffit de dotter  $E$  d'une base, ici  $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ , qui permet d'écrire de façon univoque  $\mathbf{x} = x^1 \mathbf{e}_1 + \dots + x^n \mathbf{e}_n$ . Le  $n$ -uplet  $(x^1, \dots, x^n)^\top$  souvent présenté en colonne<sup>14</sup>, correspond à ce que nous avons appelé la *composante contravariante* du vecteur  $\mathbf{x}$ , par abus de langage c'est donc aussi celle du tenseur  $\mathbf{x}$ .

<sup>14</sup> Cet aspect des choses est indiqué par le symbole de transposition  $\top$ .

Attendu que l'on a mis en correspondance bijective  $\mathbf{x}$  et  $\varphi_{\mathbf{x}}$  via le produit scalaire, on peut également représenter  $\mathbf{x}$  par les composantes de  $\varphi_{\mathbf{x}}$  dans la base de  $B^*$  de  $E^*$  soit plus explicitement  $\varphi_{\mathbf{x}} = x_1\varphi_{e_1} + \cdots + x_n\varphi_{e_n}$ . Le  $n$ -uplet  $(x_1, \dots, x_n)$  présenté en ligne correspond à ce que nous avons appelé la *composante covariante* du vecteur  $\mathbf{x}$ , c'est donc aussi celle du tenseur  $\mathbf{x}$ .

En résumé le vecteur  $\mathbf{x}$  et l'application linéaire  $\varphi_{\mathbf{x}}$ , que l'on appelle parfois une forme linéaire ou un covecteur, sont deux tenseurs d'ordre 1.

## Tenseurs d'ordre 2, et plus

Un endomorphisme linéaire de  $E$ , c'est-à-dire pour les non-poètes une application linéaire de  $E \rightarrow E$ , peut être associé à un tenseur d'ordre 2. C'est la raison pour laquelle de nombreux ouvrages parlent de tenseurs lorsqu'ils présentent simplement les matrices représentatives de ces endomorphismes dans certaines bases...

Utilisons cette application linéaire très simple pour comprendre ce qui se passe. Faisons agir un endomorphisme  $\mathbf{F}$  de  $E$  sur un quelconque de ses éléments  $\mathbf{x}$ , nous récolterons un vecteur  $\mathbf{y} = \mathbf{F}(\mathbf{x}) \in E$ . Nous aurons alors beau jeu de faire intervenir la forme  $\varphi_{\mathbf{y}}$  qui, agissant sur un vecteur quelconque  $\mathbf{z} \in E$ , permettra d'obtenir le scalaire  $f = \varphi_{\mathbf{y}}(\mathbf{z}) \in \mathbb{K}$ .

L'endomorphisme  $\mathbf{F}$  alimenté par un vecteur  $\mathbf{x}$  et un covecteur  $\varphi_{\mathbf{y}}$  permet d'obtenir un scalaire  $f$ . On peut donc associer un tenseur d'ordre 2 à  $\mathbf{F}$ . Ce que nous venons de dire n'est qu'un cas particulier car comme tous les tenseurs d'ordre 2,  $\mathbf{F}$  possède de multiples natures décrites par ses multiples composantes. Au lieu d'en parler écrivons-les

$$f = F_i^j x^i y_j = F^i_j x_i y^j = F_{ij} x^i y^j = F^{ij} x_i y_j$$

On pourrait donner des noms différents à tous les tenseurs différents qui se cachent derrière toutes ces composantes. Ce n'est pas la peine lorsque l'on décrit l'action d'un tenseur en donnant ses composantes. On dira par exemple que  $F_{ij}$  est la composante complètement covariante du tenseur  $\mathbf{F}$ . Sous cette forme, il transforme deux vecteurs en un scalaire : il s'agit d'un élément de l'ensemble des applications linéaires de  $E \times E \rightarrow \mathbb{K}$ , i.e.  $\mathcal{T}_2$  comme on peut le vérifier dans la définition 4.1.

En dimension finie, on sait que  $E = (E^*)^*$ . Ainsi, en appliquant la définition 4.1 avec  $E$  ou  $E^*$ , on comprend que  $F^{ij}$  est la composante complètement contravariante d'un tenseur qui vit dans  $\mathcal{T}^2 = E \otimes E$ . Les composantes dites mixtes  $F^i_j$  et  $F_i^j$  sont celles de tenseurs qui vivent respectivement dans  $\mathcal{T}^1_1$  et  $\mathcal{T}_1^1$ .

On comprend pourquoi il est absolument primordial de bien séparer les indices et de ne jamais les superposer : on écrira jamais une composante sous la forme  $F_j^i$  car on ne sait plus à qui on fait allusion !

La seule exception que nous nous permettrons est le tenseur de Kronecker dont toutes les formes de composantes ont les mêmes valeur :

$$1 \leq i, j \leq n \quad \delta_i^j = \delta_j^i = \delta_{ij} = \delta^{ij} = \delta_j^i = \delta_i^j = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

En conclusion, s'il est donc possible de représenter un endomorphisme par la composante mixte d'un tenseur d'ordre 2, réciproquement, comme nous l'avions vu pour les vecteurs, la notion de tenseur d'ordre 2 est bien plus générale que celle d'endomorphisme.

Chacun aura enfin compris que l'on peut sans problème généraliser la notion de produit tensoriel d'un nombre  $p > 2$  d'espaces vectoriels : on remplace  $E$  par  $E \times E$ ,  $E \times E \times E$ , etc.

### Remarque fondamentale

Nous venons de voir que la notion de tenseur est fondamentalement liée à la notion de produit scalaire. On peut ainsi définir les deux classes de tenseurs que nous connaissons pour le moment :

- Mécanique classique

Les vecteurs sont des éléments de  $E = \mathbb{R}^n$ , le produit scalaire rapporté à la base canonique  $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  est dit euclidien

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \quad \mathbf{x} \cdot \mathbf{y} = \delta_{ij} x^i y^j$$

Les tenseurs éléments de  $\mathbb{R}^n$  ou de produits tensoriels de  $\mathbb{R}^n$  muni de ce produit scalaire sont appelés tenseurs euclidiens.

- Relativité restreinte

Nous avons vu que pour répondre aux contraintes expérimentales imposées par les expériences de Michelson et Morley, la nature a choisi localement<sup>15</sup> d'utiliser  $\mathbb{M}^4$  pour représenter les 4-vecteurs de la physique. Le produit scalaire est celui de Minkowski :

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{M}^4 \quad \mathbf{x} \cdot \mathbf{y} = \eta_{\mu\nu} x^\mu y^\nu$$

où  $\eta_{\mu\nu}$  a été défini en (4.28) pour les besoins de la cause relativiste, et par application de notre convention des indices grecs qui courent de 0 à 3. Les tenseurs associés sont appelés tenseurs minkowskiens.

<sup>15</sup> Vous verrez pourquoi plus tard quand nous ferons de la relativité générale...

Tout le monde comprend maintenant que les tenseurs  $\delta$  et  $\eta$ , d'ordre 2, qui définissent respectivement les produits scalaires dans les espaces euclidiens et minkowskien. Ils permettent de calculer tout aussi respectivement des intervalles d'espace et d'espace-temps. On les appelle des tenseurs métriques.

Le cas particulier des espaces de Minkowski de dimension 4 est spécifique à la relativité restreinte. La relativité générale sera celui moins particuliers des espaces de Riemann de dimension 4, mais ceci est un autre histoire que nous étudierons plus tard.

## Retour sur les notations et vocabulaire

Comme notre fidèle lecteur l'aura remarqué, nous sommes déjà familiers de nombreux tenseurs minkowskiens :

- d'ordre 0 comme la phase d'une onde monochromatique, le temps propre  $\tau$ , l'intervalle d'espace-temps  $\ell^2 = c^2 t^2 - \vec{r} \cdot \vec{r}$  et tous les autres 4-scalaires que nous avons déjà rencontrés;
- d'ordre 1 comme tous les 4-vecteurs  $\mathbf{x}$ ,  $\mathbf{u}$ ,  $\mathbf{\Gamma}$ ,  $\mathbf{J}$  et  $\mathbf{k}$ ;
- d'ordre 2 comme le tenseur de changement de base  $\mathbf{L}$ , la métrique euclidienne  $\delta$  i.e. le tenseur de Kronecker d'ordre 2, et celle de Minkowski  $\eta$  que l'on utilise aussi en relativité restreinte pour faire monter et/ou descendre les indices des composantes de tenseurs.

Il en existe bien d'autre mais la règle que nous nous sommes proposés est de les construire au fur et à mesure de nos besoins. On comprend maintenant pourquoi nous avons choisi de représenter les vecteurs comme les matrices avec cette notation grasse italique. Nous pourrions regretter d'avoir conservé les flèches pour les vecteurs de  $\mathbb{R}^3$  car ils sont aussi associés à des tenseurs d'ordre 1, mais si on note tout de la même manière, on ne s'y retrouve plus... De plus, en mécanique classique, nous avons besoin de différentier les coordonnées  $\mathbf{q}$ , vitesses  $\dot{\mathbf{q}}$  et impulsion généralisées  $\mathbf{p}$  vivant dans  $\mathbb{R}^s$  pour des système possédant  $s$  degrés de liberté, des vecteurs position  $\vec{r}$  et vitesse  $\vec{v}$  habituels de la mécanique qui vivent habituellement dans  $\mathbb{R}^3$ . De même nous avons depuis le début utilisé une notation fine pour les scalaires de toute sortes. C'est un choix éditorial, il est maintenant plus clair.

Pour être le plus précis possible dans notre notation, on explicitera en général une composante particulière d'un tenseur (comme ce fut pratiquement toujours le cas jusqu'à présent), par exemple au lieu de parler de la composante 2 fois covariante et 1 fois contravariante du tenseur  $\mathbf{T}$ , on écrira plus explicitement mais aussi plus précisément soit  $T_{\mu}^{\nu}{}_{\rho}$ , soit  $T_{\mu\rho}{}^{\nu}$  ou soit

encore  $T^\nu_{\mu\rho}$  ces différentes écritures décrivant a priori des objets différents, souvenons-nous par exemple que  $L^\mu_\nu$  et  $L_\mu^\nu$  sont inverses l'un de l'autre...

Un tenseur de rang 2 est dit symétrique si et seulement si

$$T^{\mu\nu} = T^{\nu\mu} \quad (4.93)$$

on en déduit alors immédiatement que

$$T_{\mu\nu} = T_{\nu\mu} \quad \text{et} \quad T_\mu^\nu = T^\nu_\mu := T_\mu^\nu \quad (4.94)$$

dans cette dernière relation nous avons introduit une nouvelle notation : deux indices «superposés» mentionnent une symétrie sur ces derniers. Notons enfin que la symétrie n'implique en aucun cas que  $T_\mu^\nu$  soit égal à  $T_\nu^\mu$  et comme indiqué un peu plus haut nous éviterons ce genre de notation ambiguë.

Pour les tenseurs d'ordre supérieur, on parlera de symétrie en spécifiant les couples d'indices impliqués.

Un tenseur de rang 2 est dit antisymétrique si  $T^{\mu\nu} = -T^{\nu\mu}$ , ses termes diagonaux  $T^{\mu\mu}$  sont alors évidemment nuls.

On appelle trace  $T$  d'un tenseur  $\mathbf{T}$  de rang 2, le 4-scalaire  $T = \text{tr}(\mathbf{T}) = T_\mu^\mu = T^\mu_\mu$ . De façon plus générale, on appelle contraction d'un tenseur une quantité telle que  $T_\mu^\nu$  où deux indices identiques apparaissent en haut et en bas dans une même composante de tenseur, selon nos principes, il y a alors sommation sur cet indice (ici  $\nu$ ). Ainsi un tenseur d'ordre  $n$  contracté sur un indice devient un tenseur d'ordre  $n - 2$ . Pour contracter un tenseur on utilise le tenseur métrique, comme sur les exemples ci-dessous

$$T = T_\mu^\mu = T_\mu^\nu \eta_{\mu\nu} = T_{\mu\nu} \eta^{\mu\nu} \quad \text{ou} \quad X^\alpha = T_{\mu\nu}^\alpha \eta^{\mu\nu}$$

Un tenseur bien pratique pour écrire des formules compactes et le tenseur de Levi-Civita : il s'agit d'un tenseur de rang 4 noté  $\varepsilon$  complètement antisymétrique. Parmi les 256 valeurs de ses composantes les seules non nulles sont :

- celles issues d'une permutation paire de 0123 qui valent 1 (il y en a 12)
- celles issues d'une permutation impaire de 0123 qui valent  $-1$  (il y en a 12 aussi !)

On peut alors montrer facilement que

$$\varepsilon^{\mu\nu\rho\sigma} = -\varepsilon_{\mu\nu\rho\sigma} \quad \text{et} \quad \varepsilon^{\mu\nu\rho\sigma} \varepsilon_{\mu\nu\rho\sigma} = -24$$

Essayez pour voir !

## Changement de base

Un dernier point important avant de revenir à la relativité.

Lorsque l'on décide de changer de passer de la base  $B = \{e_1, \dots, e_n\}$  à la base  $B' = \{e'_1, \dots, e'_n\}$ , il est nécessaire d'indiquer la bijection  $\mathbf{P}$  qui permet de le faire. Cette bijection décrit la façon d'obtenir chaque vecteur de  $B'$  en fonction de ceux de  $B$ . On peut donner la représentation de  $\mathbf{P}$  dans la base  $B$  sous la forme d'une matrice de passage bien connue des étudiants de premier cycle. D'une manière plus générale et efficace, on peut associer  $\mathbf{P}$  à un tenseur d'ordre 2. Les composantes de ce tenseur permettront alors d'écrire les composantes d'un vecteur ou d'un covecteur dans la nouvelle base en fonction de celles dans l'ancienne :  $\mathbf{x}' = \mathbf{P}(\mathbf{x})$ . En terme de composantes nous avons

$$x'^i = P^i_j x^j = P^{ij} x_j \quad \text{ou} \quad x'_i = P_i^j x_j = P_{ij} x^j \quad (4.95)$$

On montre assez facilement qu'un objet  $\mathbf{x}$ , possédant  $n$  composantes, est un tenseur d'ordre 1 de  $\mathbb{R}^n$  si et seulement si lors d'un changement de base quelconque décrit par  $\mathbf{P}$ , les composantes de  $\mathbf{x}$  se transforment de la manière décrite par les relations (4.95).

Cette dernière remarque est donc une façon de caractériser un tenseur d'ordre 1. Elle se généralise à un ordre quelconque. L'action d'un tenseur  $\mathbf{T}$  d'ordre  $k$  est complètement définie par la donnée de son action sur les  $k$  vecteurs de  $E$  qui lui procurent son ordre : par cette phrase on entend qu'un tenseur d'ordre  $k$ , i.e. agissant sur  $k$  vecteurs de  $E$  ou  $E^*$ , est complètement décrit si l'on connaît les  $kn$  réels qui définissent ses composantes dans la base  $B$  choisie pour  $E$ . Chacun de ces  $k$  vecteurs de  $E$  ou  $E^*$  verra ses composantes affectées par un changement de base...

Pour écrire les composantes  $\mathbf{T}'$  du tenseur  $\mathbf{T}$  dans la nouvelle base il faudra ainsi faire intervenir autant de fois le tenseur de passage  $\mathbf{P}$  que l'ordre de  $\mathbf{T}$  le nécessite. Sur des exemples de divers ordres on aura donc

$$\begin{aligned} T'_{ij} &= P_i^k P_j^\ell T_{k\ell} \quad \text{ou} \quad T'^i_j = P_i^k P^j_\ell T_k^\ell \\ T'_{abc} &= P_a^i P_b^j P_c^k T_{ijk} \quad \text{etc.} \end{aligned}$$

Un fois dans la nouvelle base on peut faire monter ou descendre les indices en utilisant toujours le même tenseur métrique puisque ses composantes sont les mêmes dans toutes les bases...

Il faut par contre utiliser la métrique associée à l'espace que nous considérons et définir le tenseur de passage. En relativité restreinte tout le monde aura compris qu'il s'agit du tenseur de Lorentz  $\mathbf{L}$  et de la métrique de Minkowski  $\boldsymbol{\eta}$ .

### 4.3.5 Calcul vectoriel différentiel en relativité restreinte

Il est bien connu des amateurs d'«électromagnétisme classique», si tant est qu'il soit possible d'accoler ces deux mots, que l'utilisation astucieuse du calcul vectoriel différentiel et notamment du 3–vecteur

$$\vec{\nabla} = \frac{\partial}{\partial \vec{r}} = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)^\top$$

peut contribuer à rendre agréables certaines manipulations. Nous allons voir plus loin qu'en fait le véritable opérateur différentiel de l'électromagnétisme est la généralisation quadridimensionnelle de  $\vec{\nabla}$ .

#### Le vecteur 4–nabla

A partir du quadrivecteur position nous avons pu définir l'opération de dérivée spatio-temporelle, lors de la nouvelle écriture (4.75) de l'équation de conservation nous avons posé

$$\partial_\mu := \frac{\partial}{\partial x^\mu} \quad (4.96)$$

il pourrait sembler qu'il ne s'agisse là que d'une simple notation, en fait nous allons montrer que tout comme  $\vec{\nabla}$  permet de fabriquer des 3–vecteurs ou des 3–scalaires,  $\partial_\mu$  peut être considéré comme la composante covariante d'un 4–vecteur appelé 4–nabla qui permet d'en faire autant !

Soit  $f(\mathbf{x})$  un champ scalaire de  $\mathbb{M}^4$ , c'est-à-dire une fonction de  $\mathbb{M}^4 \rightarrow \mathbb{R}$ . Sa différentielle à l'ordre 1, notée traditionnellement  $df$ , est simplement le terme d'ordre 1 du développement limité de  $f(\mathbf{x} + d\mathbf{x}) - f(\mathbf{x})$ , on a donc avec nos chères notations

$$df = \frac{\partial f}{\partial x^\mu} dx^\mu = \partial_\mu f dx^\mu = d\mathbf{x} \cdot \boldsymbol{\partial}(f) \quad (4.97)$$

La quantité  $df$  étant un 4–scalaire et  $d\mathbf{x}$  étant un 4–vecteur, la quantité  $\boldsymbol{\partial}(f)$ , ou tout simplement  $\boldsymbol{\partial}f$ , est donc promue au rang de 4–vecteur. On l'appelle le 4–gradient de  $f$ .

En utilisant la définition d'un tenseur par transformation de composante on peut vérifier cette propriété : Un champ scalaire est par définition invariant par changement de base, nous aurons donc  $f = f(\mathbf{x}) = f(\mathbf{x}') = f'$  et par ailleurs

$$\frac{\partial f'}{\partial x'} = \frac{\partial f}{\partial x'} = \frac{\partial \mathbf{x}}{\partial x'} \frac{\partial f}{\partial \mathbf{x}} \quad (4.98)$$

En relativité restreinte, l'expression de  $\mathbf{x}'$  en fonction de  $\mathbf{x}$  est donnée par la transformation de Lorentz. C'est une application linéaire et avec les notations que nous avons développées on a

$$\frac{\partial \mathbf{x}}{\partial \mathbf{x}'} = \mathbf{L}(\mathbf{x}) \quad \text{et par ailleurs} \quad \boldsymbol{\partial} = \frac{\partial}{\partial \mathbf{x}} \quad (4.99)$$

En récrivant l'équation (4.98) en composantes et en éliminant le facteur  $f = f'$  de ces produits il vient

$$\partial'_\mu = L_\mu{}^\nu \partial_\nu \quad (4.100)$$

qui montre encore une fois que  $\boldsymbol{\partial}$  peut bien être manipulé comme un vecteur. Ses composantes sont subtiles. Si  $\partial_\mu$  dérive par rapport aux composantes contravariantes. En faisant agir le tenseur métrique, ce n'est qu'un jeu d'enfant de faire apparaître  $\partial^\mu = \eta^{\mu\nu} \partial_\nu$  l'opérateur de dérivation par rapport aux coordonnées covariantes

$$\partial^\mu = \frac{\partial}{\partial x_\mu} \quad (4.101)$$

On remarque dans ces écritures qu'un indice en bas au dénominateur devient un indice en haut, et réciproquement. On dirait de la musique<sup>16</sup>...

### Analyse vectorielle

Pour un champ scalaire  $f(x^\nu)$ , nous avons vu précédemment que l'on pouvait introduire un 4-gradient, dont la composante covariante s'écrit

$$\partial_\mu f = \left( \frac{1}{c} \frac{\partial f}{\partial t}, +\vec{\nabla} f \right) \quad (4.102)$$

et la composante contravariante

$$\partial^\mu f = \left( \frac{1}{c} \frac{\partial f}{\partial t}, -\vec{\nabla} f \right)^\top \quad (4.103)$$

On notera que les signes + et - sont échangés par rapport à ceux définis pour la 4-position (voir 4.3.3), encore une subtilité due au produit scalaire de Minkowski.

<sup>16</sup> On peut y voir aussi une prémonition biblique relativiste en citant l'évangile selon Matthieu : 23-12 " Qui s'élèvera sera abaissé, qui s'abaissera sera élevé."

Pour un champ de 4-vecteurs  $\mathbf{A}(\mathbf{x})$  dont on peut écrire la composante contravariante sous la forme  $A^\mu = (A^0, \vec{A})^\top$ , nous pouvons définir une 4-divergence : il s'agit du 4-scalaire

$$\boldsymbol{\partial} \cdot \mathbf{A} = \partial_\mu A^\mu = \partial^\mu A_\mu = \frac{1}{c} \frac{\partial A^0}{\partial t} + \vec{\nabla} \cdot \vec{A} \quad (4.104)$$

qui permettra donc d'écrire succinctement toutes les équations de conservation. Nous pouvons aussi définir un 4-rotationnel  $T$  du vecteur  $A$ , qui sera maintenant un tenseur complètement antisymétrique de rang 2, dont la composante contravariante sera

$$T^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \quad (4.105)$$

et la covariante

$$T_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (4.106)$$

attention à l'ordre d'intervention.

On peut enfin remarquer que l'opérateur d'Alembertien de nos chères ondes, prend ici un statut de «laplacien» de l'espacetemps en effet

$$\square = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta = \partial_\mu \partial^\mu . \quad (4.107)$$

L'utilisation combinée de tous ces opérateurs conduit à ce que l'on appelle l'analyse vectorielle et s'avère, comme chacun le sait, très utile dès qu'une onde ou un lagrangien surgit...

Dans ce contexte indiciel, une question importante se pose alors : les expressions  $\partial_\mu A^\nu$  ou  $\partial^\mu A_\nu$  correspondent-elles à des composantes tensorielles? Pour le savoir, vérifions si elles se transforment correctement. On a  $\partial'_\mu = L_\mu^\alpha \partial_\alpha$  et  $A'^\nu = L^\nu_\beta A^\beta$  ainsi

$$\begin{aligned} \partial'_\mu A'^\nu &= L_\mu^\alpha \partial_\alpha (L^\nu_\beta A^\beta) \\ &= A^\beta L_\mu^\alpha \partial_\alpha (L^\nu_\beta) + L_\mu^\alpha L^\nu_\beta \partial_\alpha A^\beta \end{aligned} \quad (4.108)$$

La transformation de Lorentz est linéaire : les composantes  $L^\nu_\beta$  sont des constantes qui ne dépendent pas de  $\mathbf{x}$ , ses dérivées par rapport à cette variable sont donc nulles, ainsi

$$\partial'_\mu A'^\nu = L_\mu^\alpha L^\nu_\beta \partial_\alpha A^\beta \quad (4.109)$$

et l'on assiste à une transformation correcte, qui confère donc le statut de tenseur de rang 2 à l'objet dont la composante mixte est  $\partial_\alpha A^\beta$ .

On pourrait chercher une écriture en gras italique du tenseur dont la composante  $\partial_\alpha A^\beta$  en fonction notamment de  $\boldsymbol{\partial}$  et  $\mathbf{A}$ . Mais cela se compliquerait très vite... On préfère utiliser une notation à virgule très pratique.

## Dériver en ponctuant ses phrases...

Afin de faire profiter notre valeureux lecteur de la plénitude de la notation indicielle, nous devons lui confier une convention très souvent utilisée : puisque  $\partial_\alpha A^\beta$  est un tenseur, nous ne conserverons que l'indice de dérivation et sa place en indiquant qu'il s'agit d'une dérivée en le précédant d'une virgule. Nous placerons cet indice de dérivation «au bout» de la série d'indices. Ainsi nous noterons<sup>17</sup>

$$\begin{aligned} \partial_\alpha A^\beta &:= A^\beta_{,\alpha} & \text{et} & & \partial^\alpha A^\beta &:= A^{\beta,\alpha} \\ \partial_\alpha A_\beta &:= A_{\beta,\alpha} & \text{et} & & \partial^\alpha A_\beta &:= A_\beta{}^{,\alpha} \end{aligned}$$

pour les tenseurs de rang plus élevé et les dérivations d'ordre supérieur, on pourra donc parler par exemple du tenseur de rang 4

$$T_{\alpha\ \mu,\nu}{}^\beta := \frac{\partial T_{\alpha\ \mu}{}^\beta}{\partial x^\nu} \quad \text{ou bien} \quad T^{\alpha\beta,\nu}{}_{,\mu} := \frac{\partial^2 T^{\alpha\beta}}{\partial x_\nu \partial x^\mu}$$

La notation devient très subtile mais d'une redoutable efficacité. Par exemple la composante covariante du 4-gradient du 4-scalaire  $\varphi(\mathbf{x})$  s'écrira  $\varphi_{,\mu}$  qui du coup devient un vecteur et n'est pas écrit en gras italique, il s'agit d'une exception... La 4-divergence de ce vecteur s'écrira  $\varphi_{,\mu}{}^{,\mu}$  on notera ici aussi que la superposition des indices de dérivation ne pose pas de problème grâce au théorème de Schwartz.

## Remarque de nature conclusive

Le fait que  $\partial_\alpha A^\beta$  soit un tenseur n'est pas une simple anecdote !

En effet, ces quantités qui sont des dérivées de composantes de vecteurs interviennent constamment en physique; sans le statut de tenseur les équations associées seraient dépendantes du choix du référentiel ! Dans un état d'esprit relativiste une telle particularité serait intolérable.

Comme nous le verrons plus loin, en relativité générale lorsque nous écrirons des changements de référentiels quelconques (et donc non linéaires), il faudra généraliser la notion de dérivée pour sauvegarder son statut tensoriel, et garder «covariantes» les équations de la physique : nous introduirons alors une dérivée covariante, mais ceci est toujours une autre histoire...

<sup>17</sup> Attention à l'ordre des indices !



## Principe de moindre action et relativité

Nous proposons dans un premier temps de construire une action et de la faire varier de façon intuitive. Nous ferons à la fin de ce chapitre un récapitulatif afin d'aller un peu plus vite en profitant de l'expérience acquise.

### 5.1 Construction de l'action

#### 5.1.1 Particule libre

En mécanique classique, l'action était définie par une relation de la forme

$$S = \int \mathcal{L} dt . \quad (5.1)$$

Il est facile de vérifier que ni le lagrangien  $\mathcal{L}$  ni  $dt$  ne sont des 4-scalaires, que seraient donc des équations du mouvement obtenues à partir d'une variation de leur produit ? Tout simplement des équations dépendant du référentiel considéré, comme le sont la plupart des équations classiques. Pour obtenir des équations relativistes et donc covariantes, il faut considérer un lagrangien qui soit un 4-scalaire, la plus simple de toutes les possibilités est une constante multipliée par l'élément d'intervalle d'espace-temps, soit

$$S_l = \int \alpha ds . \quad (5.2)$$

Nous savons par ailleurs que

$$ds^2 = d\mathbf{x} \cdot d\mathbf{x} = c^2 dt^2 - d\mathbf{r}^2 = c^2 dt^2 \left( \sqrt{1 - \frac{v^2}{c^2}} \right)^2 \quad (5.3)$$

ainsi l'action la plus simple pour une particule libre s'écrit

$$S_l = \int \alpha c dt \sqrt{1 - \frac{v^2}{c^2}} \quad (5.4)$$

Dans la limite des vitesses faibles devant  $c$ , si tout se passe bien, nous devons retrouver l'action de la particule libre classique (à une jauge  $\phi(t)$  près), i.e.

$$S_{cl} = \int \frac{1}{2} m v^2 dt \quad (5.5)$$

en faisant un développement limité pour  $\beta = v/c$  petit devant 1, on obtient donc à partir de (5.4)

$$S = \int \alpha c dt \sqrt{1 - \frac{v^2}{c^2}} \xrightarrow{\beta \ll 1} \int \alpha c dt \left(1 - \frac{1}{2} \frac{v^2}{c^2}\right) = \phi(t) - \int \frac{1}{2} \alpha \frac{v^2}{c} dt \quad (5.6)$$

l'identification avec l'action classique et la liberté de jauge conduisent donc à

$$\alpha = -mc. \quad (5.7)$$

L'action covariante la plus simple d'une particule libre relativiste sera donc

$$S_l = -mc \int ds \quad (5.8)$$

### 5.1.2 Particule soumise à des forces

Considérons à présent le cas d'une particule soumise à des forces. Que faire sinon l'hypothèse que ces forces sont représentées par un 4-vecteur  $\mathbf{A}(\mathbf{x})$  dont la composante covariante s'écrit  $A_\mu$ . Il faut à présent former une action, celle-ci doit être invariante lors d'une transformation de Lorentz (covariante), encore une fois le plus simple est d'écrire la densité de lagrangien  $\mathcal{L}$  comme un produit scalaire, soit

$$S_i = -\kappa \int \mathbf{A} \cdot d\mathbf{x} = -\kappa \int A_\mu dx^\mu \quad (5.9)$$

la constante  $-\kappa$  décrit l'intensité des forces en présence et leurs unités (le signe  $-$  est conventionnel).

En vertu de la propriété d'additivité, l'action totale d'une particule soumise à des forces sera donc

$$S = -mc \int ds - \kappa \int A_\mu dx^\mu \quad (5.10)$$

qu'il ne reste plus qu'à faire varier pour obtenir l'équation de la dynamique relativiste.

## 5.2 Equation de la dynamique relativiste

### 5.2.1 Action toujours varie, bien fol est qui s'y fie

Le principe de moindre action auquel nous voulons toujours croire nous indique que le mouvement d'une particule relativiste soumise à des forces sera celui produisant le minimum de la fonctionnelle

$$S = -mc \int ds - \kappa \int A_\mu dx^\mu \quad (5.11)$$

Nous restons en composante dans cette phase intuitive. La variable du problème d'optimisation posé est ici  $x^\mu$ , la variation associée sera notée  $\delta x^\mu$ . La variation d'action correspondante est donc

$$\delta S = - \int [mc \delta ds + \kappa \delta A_\mu dx^\mu + \kappa A_\mu \delta dx^\mu] . \quad (5.12)$$

Attendu que

$$ds = \sqrt{dx_\mu dx^\mu} \quad (5.13)$$

la formule de dérivée des applications composées nous donne

$$\delta ds = \frac{1}{2} \frac{\delta(dx_\mu dx^\mu)}{\sqrt{dx_\mu dx^\mu}} = \frac{1}{2} \frac{\delta(dx_\mu dx^\mu)}{ds} = \frac{dx^\mu}{ds} \delta dx_\mu = \frac{dx_\mu}{ds} \delta dx^\mu \quad (5.14)$$

où l'on a utilisé le fait que  $\delta(dx_\mu dx^\mu) = 2dx_\mu \delta dx^\mu = 2dx^\mu \delta dx_\mu$ .

En faisant apparaître la quadrivitesse

$$\frac{dx_\mu}{ds} = \frac{1}{c} \frac{dx_\mu}{d\tau} = \frac{1}{c} u_\mu \quad (5.15)$$

il vient finalement

$$\delta ds = \frac{1}{c} u_\mu \delta dx^\mu \quad (5.16)$$

ainsi

$$\delta S = - \int [m u_\mu \delta dx^\mu + \kappa \delta A_\mu dx^\mu + \kappa A_\mu \delta dx^\mu] \quad (5.17)$$

en permutant les variations  $\delta$  et  $d$ , il reste

$$\delta S = - \int [m u_\mu d\delta x^\mu + \kappa \delta A_\mu dx^\mu + \kappa A_\mu d\delta x^\mu] \quad (5.18)$$

en intégrant par parties le premier et le troisième terme, il vient

$$\delta S = - \int [-m d u_\mu \delta x^\mu + \kappa \delta A_\mu dx^\mu - \kappa d A_\mu \delta x^\mu] \quad (5.19)$$

les termes de bord introduits par cette intégration sont nuls. En effet, comme en mécanique classique, les conditions aux limites sur  $x^\mu$  imposent la connaissance de  $x^\mu(1)$  et  $x^\mu(2)$ , on cherche la trajectoire suivie en imposant les deux bornes, ainsi

$$\delta x^\mu(1) = \delta x^\mu(2) = 0 \quad (5.20)$$

Un jeu d'écriture permet alors

$$\delta A_\mu = \frac{\partial A_\mu}{\partial x^\nu} \delta x^\nu = A_{\mu,\nu} \delta x^\nu \quad \text{et} \quad d A_\mu = \frac{\partial A_\mu}{\partial x^\nu} dx^\nu = A_{\mu,\nu} dx^\nu \quad (5.21)$$

ainsi

$$\delta S = - \int [-m d u_\mu \delta x^\mu + \kappa A_{\mu,\nu} \delta x^\nu dx^\mu - \kappa A_{\mu,\nu} dx^\nu \delta x^\mu] \quad (5.22)$$

il suffit alors de remarquer que par définition de la quadrivitesse

$$dx^\mu = u^\mu d\tau \quad (5.23)$$

pour obtenir

$$\delta S = - \int [-m d u_\mu \delta x^\mu + \kappa A_{\mu,\nu} \delta x^\nu u^\mu d\tau - \kappa A_{\mu,\nu} u^\nu d\tau \delta x^\mu] \quad (5.24)$$

en inversant les indices muets  $\mu$  et  $\nu$  du terme central, la variation s'écrit finalement

$$\delta S = - \int \left[ -m \frac{d u_\mu}{d\tau} + \kappa A_{\nu,\mu} u^\nu - \kappa A_{\mu,\nu} u^\nu \right] d\tau \delta x^\mu \quad (5.25)$$

pour obtenir un minimum de l'action,  $\delta S$  doit être nul quelle que soit la variation  $\delta x^\mu$ . L'équation du mouvement s'écrit donc

$$m \frac{d u_\mu}{d\tau} = \kappa (A_{\nu,\mu} - A_{\mu,\nu}) u^\nu \quad (5.26)$$

l'on voit sourdre ainsi le tenseur<sup>1</sup>  $\mathbf{F}(\mathbf{x})$  dont la composante complètement covariante s'écrit

<sup>1</sup> Le fait que  $F_{\mu\nu}$  soit la composante d'un tenseur est automatiquement dû à sa définition.

$$\begin{aligned}
F_{\mu\nu} &:= A_{\nu,\mu} - A_{\mu,\nu} \\
&= \partial_\mu A_\nu - \partial_\nu A_\mu \\
&= \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}
\end{aligned} \tag{5.27}$$

qui permet d'écrire l'équation du mouvement d'une particule relativiste sous la forme

$$m \frac{d u_\mu}{d\tau} = \kappa F_{\mu\nu} u^\nu . \tag{5.28}$$

En introduisant la composante covariante de la 4-force

$$G_\mu := \kappa F_{\mu\nu} u^\nu \tag{5.29}$$

L'équation (5.28) prend une forme qui n'est pas sans rappeler un principe fondamental de la dynamique

$$m \frac{d u_\mu}{d\tau} = G_\mu \tag{5.30}$$

mais en dimension 4 et manifestement covariante ....

Il ne reste plus qu'à spécifier notre problème, en précisant la nature de la force  $\mathbf{A}(\mathbf{x})$ , pour pouvoir faire de la dynamique en relativité restreinte.

Avant de préciser la force, étudions les propriétés du nouveau venu : le tenseur d'ordre 2  $\mathbf{F}(\mathbf{x})$  appelé tenseur champ, ou tenseur de Faraday.

## 5.2.2 Propriétés du tenseur champ

### Généralités

Par définition, le tenseur champ  $\mathbf{F}(\mathbf{x})$  tel que  $F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}$  est anti-symétrique, en conséquence, ses termes diagonaux sont tous nuls ainsi que sa trace

$$F_{\mu\mu} = 0 \quad \text{et} \quad F_\mu{}^\mu = F^\mu{}_\mu = 0 . \tag{5.31}$$

### Composantes

Afin de décrire de façon moins abstraite ce tenseur, exhibons ses composantes en fonction de celles de  $\mathbf{A}(\mathbf{x})$ , pour cela posons

$$A^\mu = \left( \frac{V(\mathbf{x})}{c}, \vec{A}(\mathbf{x}) \right)^\top \tag{5.32}$$

où, pour quelques instants encore,  $V = V(\mathbf{x})$  est un champ scalaire quelconque et  $\vec{A}(\mathbf{x})$  est un champ de 3-vecteur tout aussi quelconque, le facteur

$c$  ici présent ne nuit en rien à la généralité de notre propos. La composante covariante de  $\mathbf{A}(\mathbf{x})$  s'écrit donc

$$A_\mu = \left( \frac{V(\mathbf{x})}{c}, -\vec{A}(\mathbf{x}) \right) \quad (5.33)$$

Ainsi pour les indices spatiaux  $i, j = 1, 2, 3$  nous avons :

- Partie temporelle

$$F_{0i} = \partial_0 A_i - \partial_i A_0 = -\frac{\partial(\vec{A}_i)}{\partial ct} - \frac{1}{c} \frac{\partial V}{\partial x^i} = -\frac{1}{c} \left[ \frac{\partial \vec{A}}{\partial t} + \vec{\nabla}(V) \right] \cdot \vec{e}_i$$

ainsi en introduisant le 3-vecteur

$$\vec{E} = -\vec{\nabla}V - \frac{\partial \vec{A}}{\partial t} \quad (5.34)$$

nous pouvons simplement écrire

$$F_{0i} = \frac{E_i}{c} \quad (5.35)$$

les lecteurs les plus malins auront tout de suite remarqué que le changement de signe pour les composantes co et contravariantes ne s'appliquent pas (c'est la moindre des choses) aux 3-vecteurs.

- Partie spatiale

$$F_{ij} = \partial_i A_j - \partial_j A_i$$

Il s'agit manifestement, et au signe près, des composantes d'un 3-rotationnel. En effet si  $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$  constitue une base de l'hyperplan spatial de  $\mathbb{M}^4$  un simple calcul montre que pour  $i \neq j \neq k = 1, 2, 3$  modulo 3 on peut écrire

$$F_{ij} = (-1)^{i+j} \left( \vec{\nabla} \wedge \vec{A} \right) \cdot \vec{e}_k \quad (5.36)$$

ainsi en introduisant le 3-vecteur

$$\vec{B} = \vec{\nabla} \wedge \vec{A} \quad (5.37)$$

il vient

$$\begin{aligned} F_{12} &= (-1)^3 \left( \vec{\nabla} \wedge \vec{A} \right) \cdot \vec{e}_3 = -B_3 \\ F_{13} &= (-1)^4 \left( \vec{\nabla} \wedge \vec{A} \right) \cdot \vec{e}_2 = B_2 \\ F_{23} &= (-1)^5 \left( \vec{\nabla} \wedge \vec{A} \right) \cdot \vec{e}_1 = -B_1 \end{aligned} \quad (5.38)$$

en récapitulant tous les résultats obtenus et leur antisymétrie, nous pouvons écrire la composante complètement covariante du tenseur champ sous la forme matricielle suivante ( $\mu$  est ici un indice de ligne et  $\nu$  un indice de colonne)

$$F_{\mu\nu} = \begin{bmatrix} 0 & E_1/c & E_2/c & E_3/c \\ -E_1/c & 0 & -B_3 & B_2 \\ -E_2/c & B_3 & 0 & -B_1 \\ -E_3/c & -B_2 & B_1 & 0 \end{bmatrix} \quad (5.39)$$

### 5.2.3 Premier groupe d'équations de Maxwell

Comme nous l'avons judicieusement remarqué plus haut,  $F_{\mu\nu}$  est un 4-rotationnel, par conséquent si tout se passe comme en dimension 3, sa 4-divergence s'annule. Pour ceux qui ne seraient pas encore convaincus des vertus multidimensionnelles de l'analyse tensorielle, vérifions-le ensemble.

La 4-divergence de  $F_{\mu\nu}$  est un tenseur de rang 3 qui s'écrit tout naturellement

$$\begin{aligned} \partial_\rho F_{\mu\nu} + \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} &= \partial_\rho \partial_\mu A_\nu - \partial_\rho \partial_\nu A_\mu \\ &+ \partial_\mu \partial_\nu A_\rho - \partial_\mu \partial_\rho A_\nu \\ &+ \partial_\nu \partial_\rho A_\mu - \partial_\nu \partial_\mu A_\rho \\ &= 0 \end{aligned} \quad (5.40)$$

par simple application du théorème de Schwarz !

Notons avant d'aller plus loin que la relation (5.40) symbolise tout de même la bagatelle de  $4^3 = 64$  équations ! Beaucoup d'entre elles sont cependant triviales :

- Si les trois indices sont égaux :  $\rho = \mu = \nu$ , l'équation (5.40) se réduit à l'égalité fondamentale  $0 = 0$ , car les termes diagonaux de  $F_{\mu\nu}$  sont tous nuls comme son antisymétrie nous l'avait fait remarquer.
- Si deux des trois indices sont égaux :  $\mu = \nu \neq \rho$ , l'équation (5.40) s'écrit

$$\partial_\rho F_{\mu\mu} + \partial_\mu (F_{\mu\rho} + F_{\rho\mu}) = 0 \quad (5.41)$$

qui est encore une équation triviale compte-tenu de l'antisymétrie de  $F_{\mu\nu}$ .

- Il est donc nécessaire d'avoir  $\mu \neq \nu \neq \rho$  pour espérer obtenir une équation non triviale.

L'équation (5.40) est obtenue pour le choix ordonné  $[\rho, \mu, \nu]$ , pour le choix  $[\rho, \nu, \mu]$  nous obtenons

$$\partial_\rho F_{\nu\mu} + \partial_\nu F_{\mu\rho} + \partial_\mu F_{\rho\nu} = 0 \Leftrightarrow -(\partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} + \partial_\mu F_{\nu\rho}) = 0 \quad (5.42)$$

qui est donc identique à (5.40). Quelques dernières manipulations montrent aisément que tous les choix ordonnés d'indices distincts conduisent toujours à cette même équation.

Il n'existe donc que 4 versions distinctes :

1.  $[\rho, \mu, \nu] = [1, 2, 3]$  qui s'écrit

$$\begin{aligned} & \partial_1 F_{23} + \partial_2 F_{31} + \partial_3 F_{12} = 0 \\ \Leftrightarrow & \partial_1 (-B_1) + \partial_2 (-B_2) + \partial_3 (-B_3) = 0 \\ \Leftrightarrow & -\vec{\nabla} \cdot \vec{B} = 0 \end{aligned} \quad (5.43)$$

2.  $[\rho, \mu, \nu] = [0, 1, 2]$  qui s'écrit donc

$$\begin{aligned} & \partial_0 F_{12} + \partial_1 F_{20} + \partial_2 F_{01} = 0 \\ \Leftrightarrow & \frac{\partial(-B_3)}{\partial(ct)} + \frac{\partial(-E_2/c)}{\partial x^1} + \frac{\partial(E_1/c)}{\partial x^2} = 0 \\ \Leftrightarrow & \left( \vec{\nabla} \wedge \vec{E} \right) \cdot \vec{e}_3 = -\frac{\partial \vec{B}}{\partial t} \cdot \vec{e}_3 \end{aligned} \quad (5.44)$$

et ainsi de suite ...

3.  $[\rho, \mu, \nu] = [0, 1, 3]$  donne

$$\left( \vec{\nabla} \wedge \vec{E} \right) \cdot \vec{e}_2 = -\frac{\partial \vec{B}}{\partial t} \cdot \vec{e}_2 \quad (5.45)$$

4.  $[\rho, \mu, \nu] = [0, 2, 3]$  donne

$$\left( \vec{\nabla} \wedge \vec{E} \right) \cdot \vec{e}_1 = -\frac{\partial \vec{B}}{\partial t} \cdot \vec{e}_1 \quad (5.46)$$

Le fait que  $F_{\mu\nu}$  soit un 4-rotationnel s'exprime donc de façon tensorielle sous la forme

$$\partial_\rho F_{\nu\mu} + \partial_\nu F_{\mu\rho} + \partial_\mu F_{\rho\nu} = 0 \quad (5.47)$$

ou bien de façon équivalente mais plus longue en utilisant des 3-vecteurs

$$\vec{\nabla} \wedge \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad \text{et} \quad \vec{\nabla} \cdot \vec{B} = 0 \quad (5.48)$$

Dans le contexte des forces électromagnétiques ces équations sont bien sûr 2 des 4 équations de Maxwell qui apparaissent ici comme jaillissant d'un principe de moindre action relativiste.

### 5.2.4 Force de Lorentz

Revenons à notre équation de la dynamique relativiste

$$m \frac{d u_\mu}{d\tau} = \kappa F_{\mu\nu} u^\nu \quad (5.49)$$

la 4-vitesse s'écrit simplement

$$u^\mu = \frac{dx^\mu}{d\tau} = \gamma \frac{dx^\mu}{dt} = (\gamma c, \gamma \vec{v})^\top \quad \text{soit} \quad u_\mu = (\gamma c, -\gamma \vec{v}) \quad (5.50)$$

La composante temporelle de l'équation de la dynamique s'écrit donc

$$m \frac{d(\gamma c)}{d\tau} = \kappa F_{0\nu} u^\nu \quad (5.51)$$

comme  $F_{00}$  est nul, il ne reste plus que

$$m \frac{d(\gamma c)}{d\tau} = \kappa F_{0i} u^i = \kappa \gamma \frac{\vec{E} \cdot \vec{v}}{c} \quad (5.52)$$

en introduisant l'énergie relativiste  $\mathcal{E} = m\gamma c^2$ , il vient

$$\frac{d\mathcal{E}}{d\tau} = \kappa \gamma \vec{E} \cdot \vec{v}$$

on note donc que la variation d'énergie est purement d'origine électrique. Les composantes spatiales livrent quant à elles :

1. Pour  $\mu = 1$

$$m \frac{d(\gamma v_1)}{d\tau} = \kappa F_{1\nu} u^\nu \quad (5.53)$$

c'est-à-dire

$$m \frac{d(\gamma v_1)}{dt} = \kappa (-\gamma E_1 + (-B_3) \gamma v_2 + B_2 \gamma v_3) \quad (5.54)$$

$$\Leftrightarrow m \frac{d(\gamma \vec{v})}{dt} \cdot \vec{e}_1 = -\kappa \gamma (\vec{E} + \vec{v} \wedge \vec{B}) \cdot \vec{e}_1 \quad (5.55)$$

2. pour  $\mu = 2, 3$  on trouve immédiatement

$$m \frac{d(\gamma \vec{v})}{dt} \cdot \vec{e}_\mu = -\kappa \gamma (\vec{E} + \vec{v} \wedge \vec{B}) \cdot \vec{e}_\mu \quad (5.56)$$

Ainsi, en prenant  $\kappa = q$  le module de la charge électrique de l'électron, ces trois équations s'écrivent

$$m \frac{d(\gamma \vec{v})}{dt} = -q\gamma \left( \vec{E} + \vec{v} \wedge \vec{B} \right) . \quad (5.57)$$

Elles expriment la force subie par un électron dans un champ électromagnétique : la force de Lorentz. Comme le champ électromagnétique, cette force est une manifestation purement relativiste ! Bien entendu, lorsque  $\gamma \rightarrow 1$ , on retrouve la formule empirique de la force de Lorentz présentée dans les études classiques.

### 5.3 Sources du champ : 2<sup>eme</sup> groupe d'équations de Maxwell

Nous avons souligné que l'interaction entre *une* particule et un champ était décrite par l'action

$$S_i = -q \int \mathbf{A} \cdot d\mathbf{x} = -q \int A_\mu dx^\mu \quad (5.58)$$

Si nous considérons à présent une *assemblée* de charges en mouvement, il est bien connu (loi d'induction de Faraday) que ces dernières créent un champ supplémentaire de même nature que celui qui est à l'origine de leur mouvement. Il est donc nécessaire, pour une description complète, dans un premier temps de généraliser l'action d'interaction à une assemblée de charge et non plus une charge unique, puis dans un deuxième temps de la compléter avec un terme décrivant l'auto-interaction du champ.

#### 5.3.1 Action d'interaction entre un champ et une assemblée de charges en mouvement

La généralisation à un nombre fini  $k = 1, \dots, N$  de charges est immédiate. Le principe de superposition du lagrangien auquel nous croyons, nous permet d'écrire que

$$S_i = - \sum_{k=1}^N q_k \int A_\mu dx^\mu \quad (5.59)$$

En repérant ces charges dans l'espace à l'instant  $t$  par un 3-vecteur  $\vec{r}_k$ , on peut introduire leur densité volumique  $\rho(\vec{r}, t)$  grâce à la relation

$$q_k = \int \rho(\vec{r}', t) \delta(\vec{r}_k - \vec{r}') d\vec{r}' \quad (5.60)$$

où  $d\vec{r}'$  représente un élément de volume de l'hyperplan spatial de  $\mathbb{M}^4$ . Ainsi, on pourra écrire (au moins au sens des distributions)

$$S_i = - \int dV \rho A_\mu dx^\mu \quad (5.61)$$

Il n'est alors qu'astucieux de faire apparaître la densité de courant  $J^\mu = \rho dx^\mu/dt$  (cf. équation (4.77)), pour obtenir

$$S_i = - \int A_\mu J^\mu dV dt = - \int \mathbf{A} \cdot \mathbf{J} dV dt \quad (5.62)$$

Sous cette forme, l'action d'interaction prend en compte l'assemblée de charges et son mouvement via la densité de courant. L'élément d'intégration est à présent  $dV dt$ , il s'agit là (au facteur  $c$  près) de l'élément de volume de  $\mathbb{M}^4$ , on dit alors que le terme  $\mathbf{A} \cdot \mathbf{J}$  est une densité de lagrangien : c'est une histoire d'unités de mesure.

### 5.3.2 Action d'auto-interaction du champ

Imaginons une partie de l'espacetemps dans laquelle il n'y aurait pas de particules, i.e.  $\mathbf{J}(\mathbf{x}) \equiv \mathbf{0}$  car  $\rho(\mathbf{x}) \equiv 0$ , mais seulement un champ : on parle de champ libre, tout comme une particule est dite libre quand elle n'est soumise à aucune force et donc à aucun champ.

Dans l'état actuel de notre avancement, une façon simple de construire l'action d'un champ libre est de la bâtir à partir du tenseur champ  $\mathbf{F}(\mathbf{x})$  qui ne dépend pas de  $\mathbf{J}(\mathbf{x})$ .

Cette action doit être invariante de Lorentz (on dit covariante), c'est pourquoi le plus simple est de choisir un 4-scalaire<sup>2</sup>. Il n'existe que 2 familles de 4-scalaires distincts non triviaux formés uniquement à base de  $F^{\mu\nu}$  (c'est un excellent exercice de calcul manipulations d'indices que de le démontrer)

$$F^{\mu\nu} F_{\mu\nu} \quad \text{et} \quad \varepsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta} \quad (5.63)$$

C'est un autre exercice simple de démontrer que le scalaire  $\varepsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta}$  se met sous la forme d'une 4-divergence : il entre donc dans la liberté de

<sup>2</sup> Si nous parvenons à nos fins avec un 4-scalaire, nous aurons donné du sens à la simplicité, sinon nous essaierons quelque chose de plus compliqué ...

jauge de l'action. Ainsi, il ne reste plus que  $F^{\mu\nu}F_{\mu\nu}$  pour bâtir une action raisonnable. Nous essayerons donc

$$S_c = k \int F^{\mu\nu} F_{\mu\nu} dV dt \quad (5.64)$$

où la constante  $k$  (comme précédemment  $\alpha$  et  $\kappa$ ) reste à fixer. L'action totale de l'électromagnétisme s'écrit donc

$$S = \underbrace{-mc \int ds}_{\text{particule libre}} - \underbrace{\int A_\mu J^\mu dV dt}_{\text{interaction champ-particule}} + k \underbrace{\int F^{\mu\nu} F_{\mu\nu} dV dt}_{\text{champ libre}} \quad (5.65)$$

qu'il ne reste plus qu'à faire varier ...

Pour trouver les équations de la dynamique de l'évènement  $\mathbf{x} \in \mathbb{M}^4$ , nous avons fixé les conditions aux limites de son mouvement, i.e.  $\mathbf{x}(1)$  et  $\mathbf{x}(2)$  et cherché le minimum de l'action pour des variations  $\mathbf{x} \rightarrow \mathbf{x} + \delta\mathbf{x}$ .

Un autre type de variation est possible pour cette action : le tenseur champ  $\mathbf{F}$  ne dépend de  $\mathbf{x}$  qu'à travers le 4-potential  $\mathbf{A}$ . Pour un courant de particules fixé, on peut donc envisager calculer la variation de l'action  $\delta S$  correspondant à une variation  $\mathbf{A} \rightarrow \mathbf{A} + \delta\mathbf{A}$  en imposant bien sur les conditions aux limites fixées  $\mathbf{A}(1)$  et  $\mathbf{A}(2)$ .

Ce calcul est simple à débiter, toujours en composante pour le moment, il donne

$$\delta S = 0 - \int dV dt [J^\mu \delta A_\mu - k \delta (F^{\mu\nu} F_{\mu\nu})] = 0 \quad (5.66)$$

quelques manipulations montrent que

$$\delta (F^{\mu\nu} F_{\mu\nu}) = 2F^{\mu\nu} \delta F_{\mu\nu} \quad (5.67)$$

de plus, toujours en supposant que les variations  $\partial_\mu$  et  $\delta$  commutent,

$$\delta F_{\mu\nu} = \partial_\mu (\delta A_\nu) - \partial_\nu (\delta A_\mu) \quad (5.68)$$

ainsi

$$\begin{aligned} \delta (F^{\mu\nu} F_{\mu\nu}) &= 2F^{\mu\nu} \partial_\mu (\delta A_\nu) - 2F^{\mu\nu} \partial_\nu (\delta A_\mu) \\ &= 2F^{\nu\mu} \partial_\nu (\delta A_\mu) - 2F^{\mu\nu} \partial_\nu (\delta A_\mu) \\ &= -4F^{\mu\nu} \partial_\nu (\delta A_\mu) \end{aligned} \quad (5.69)$$

le principe de moindre action s'écrit donc

$$\delta S = 0 - \int dV dt [J^\mu \delta A_\mu + 4k F^{\mu\nu} \partial_\nu (\delta A_\mu)] = 0 \quad (5.70)$$

la désormais traditionnelle intégration par parties<sup>3</sup> du second terme de l'intégrand donne

$$\delta S = 0 - \int dV dt [J^\mu \delta A_\mu - 4k \delta A_\mu \partial_\nu F^{\mu\nu}] = 0 \quad (5.71)$$

le terme de bord est toujours nul eu égard au fait que le champ est fixé sur les extrémités du domaine de variation :  $\delta A_\mu(1) = \delta A_\mu(2) = 0$ . Finalement, il reste

$$\delta S = 0 \Leftrightarrow \int dV dt [J^\mu - 4k \partial_\nu F^{\mu\nu}] \delta A_\mu = 0 \quad (5.72)$$

La variation d'action se devant d'être nulle pour toutes les variations  $\delta A_\mu$ , le champ correspondant à un extremum de l'action vérifie donc l'équation

$$J^\mu = 4k \partial_\nu F^{\mu\nu} \quad (5.73)$$

Deux choses restent à faire : écrire en terme des observables  $\vec{E}$  et  $\vec{B}$  les équations correspondantes et identifier la constante  $k$ . Pour ce faire nous devons expliciter la composante complètement contravariante du tenseur champ. Lors de l'écriture explicite de l'équation de la dynamique relativiste nous nous étions servis de sa composante complètement covariante, pour passer de l'une à l'autre il suffit de faire agir le tenseur métrique :  $F^{\mu\nu} = \eta^{\nu\alpha} \eta^{\mu\beta} F_{\alpha\beta}$ , ce calcul donne immédiatement (cf. équations (4.33) et (5.39))

$$F^{\mu\nu} = \begin{bmatrix} 0 & -E_1/c & -E_2/c & -E_3/c \\ E_1/c & 0 & -B_3 & B_2 \\ E_2/c & B_3 & 0 & -B_1 \\ E_3/c & -B_2 & B_1 & 0 \end{bmatrix} \quad (5.74)$$

Lors de cet abaissement d'indices  $\vec{E} \rightarrow -\vec{E}$  et  $\vec{B} \rightarrow \vec{B}$ . La composante contravariante du 4-courant est comme nous l'avons longuement évoqué  $J^\mu = (\rho c, \vec{j})^\top$ , où  $\vec{j}$  est le simple 3-courant. Ainsi la composante temporelle ( $\mu = 0$ ) de l'équation (5.73) s'écrit

$$\begin{aligned} J^0 = 4k \partial_\nu F^{0\nu} &\Leftrightarrow \rho c = 4k \left( \frac{\partial 0}{\partial ct} + \frac{\partial (-E_1/c)}{\partial x^1} + \frac{\partial (-E_2/c)}{\partial x^2} + \frac{\partial (-E_3/c)}{\partial x^3} \right) \\ &\Leftrightarrow \vec{\nabla} \cdot \vec{E} = -\frac{\rho c^2}{4k} \end{aligned} \quad (5.75)$$

<sup>3</sup> qui en dimension plus grande que l'unité s'appelle plutôt formule de Green ...

Pour une assemblée de charges statiques de densité spatiale  $\rho(\vec{r})$ , on a  $\vec{E} = -\vec{\nabla}(V)$  avec

$$V = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' = \frac{1}{4\pi\epsilon_0} \rho(\vec{r}) * \frac{1}{|\vec{r}|} \quad (5.76)$$

ainsi la divergence de  $\vec{E}$  est instantanément reliée au laplacien de  $V$  (c'est l'équation de Poisson), un simple calcul de distributions montre en effet que

$$\vec{\nabla} \cdot \vec{E} = \vec{\nabla} \cdot [-\vec{\nabla}(V)] \quad (5.77)$$

$$= -\Delta \left( \frac{1}{4\pi\epsilon_0} \rho(\vec{r}) * \frac{1}{|\vec{r}|} \right) = \frac{\rho(\vec{r})}{\epsilon_0} * \Delta \left( \frac{-1}{4\pi|\vec{r}|} \right) \quad (5.78)$$

$$= \frac{\rho(\vec{r})}{\epsilon_0} * \delta(\vec{r}) \quad (5.79)$$

la distribution de Dirac étant l'élément neutre de l'algèbre de convolution, il ne reste plus que

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \quad (5.80)$$

Notre formulation variationnelle de l'électromagnétisme prend donc tout son sens en fixant

$$-\frac{1}{4k} = \mu_0 \quad (5.81)$$

où l'on a introduit la nouvelle constante  $\mu_0$  qui est donc telle que

$$\mu_0 \epsilon_0 c^2 = 1 \quad (5.82)$$

Mais revenons à notre propos et tirons le parti de la composante spatiale de notre nouvelle équation (5.73).

Pour  $\mu = 1$ , il vient

$$\begin{aligned} J^1 = 4k \partial_\nu F^{1\nu} &\Leftrightarrow \vec{j} \cdot \vec{e}_1 = -\frac{1}{\mu_0} \left( \frac{\partial(E_1/c)}{\partial ct} + \frac{\partial(0)}{\partial x^1} + \frac{\partial(-B_3)}{\partial x^2} + \frac{\partial(B_2)}{\partial x^3} \right) \\ &\Leftrightarrow \left( \vec{\nabla} \wedge \vec{B} \right) \cdot \vec{e}_1 = \left( \mu_0 \vec{j} + \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t} \right) \cdot \vec{e}_1 \end{aligned} \quad (5.83)$$

pour  $\mu = 2$  et  $3$ , on obtient les deux autres composantes de la fameuse équation

$$\vec{\nabla} \wedge \vec{B} = \mu_0 \vec{j} + \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t} \quad (5.84)$$

Dès lors que nous savons à présent comment obtenir les équations de Maxwell via une généralisation 4-dimensionnelle du principe de moindre action ad-joint d'un principe de covariance, nous ne pouvons que nous incliner devant le génie de ce monsieur qui les a déduites de considérations physiques.

Le travail dans les espaces de Minkowski a été fait par ce dernier juste avant sa mort en 1909, à 44 ans, d'une crise d'appendicite...

### 5.3.3 Invariance de jauge

Il est remarquable de noter que la définition du tenseur champ  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$  laisse la place à une transformation non triviale sur  $A_\mu$  qui ne modifie pas  $F_{\mu\nu}$ . En effet, en posant

$$A_\mu^* = A_\mu + \partial_\mu \psi \quad (5.85)$$

où  $\psi$  est une fonction quelconque de  $x^\mu$ , on a

$$F_{\mu\nu}^* = \partial_\mu A_\nu - \partial_\nu A_\mu + \partial_\mu \partial_\nu \psi - \partial_\nu \partial_\mu \psi = F_{\mu\nu} \quad (5.86)$$

par application du théorème de Schwarz. Cette liberté est bien connue en électromagnétisme. Reprenons en effet les notations vectorielles, si, comme nous l'avions posé (cf. equation (5.32))  $A_\mu = (V/c, -\vec{A})$  alors

$$A_\mu^* = \left( \frac{V}{c} + \frac{1}{c} \frac{\partial \psi}{\partial t}, -\vec{A} + \vec{\nabla} \psi \right) =: \left( \frac{V^*}{c}, -\vec{A}^* \right). \quad (5.87)$$

Les 3-vecteurs champs se déduisent comme nous l'avons vu de ces deux potentiels (voir (5.34) et (5.37))

$$\vec{E}^* = -\vec{\nabla} V^* - \frac{\partial \vec{A}^*}{\partial t} = -\vec{\nabla} V - \frac{\partial \vec{A}}{\partial t} - \vec{\nabla} \frac{\partial \psi}{\partial t} + \frac{\partial \vec{\nabla} \psi}{\partial t} = \vec{E} \quad (5.88)$$

et

$$\vec{B}^* = \vec{\nabla} \wedge \vec{A}^* = \vec{\nabla} \wedge \vec{A} - \vec{\nabla} \wedge \vec{\nabla} \psi = \vec{B} \quad (5.89)$$

Les champs  $\vec{E}$  et  $\vec{B}$  ne sont pas non plus affectés par la transformation (5.85) qu'Hermann Weyl nomma transformation de jauge. Dans le formalisme lagrangien de la relativité, cette transformation apporte encore un peu d'eau à un moulin qui en avait déjà reçu beaucoup en mécanique classique.

<sup>4</sup> Pour utiliser une expression consacrée qui signifie invariant sous la transformation (5.85)

En effet, l'action de la particule libre qui ne contient pas de terme  $A_\mu$  est donc «invariante de jauge»<sup>4</sup>. Il en va de même pour le terme d'auto-interaction du champ dont l'action ne fait intervenir que  $F_{\mu\nu}$  qui est lui-même invariant de jauge (cf (5.86)). L'action  $S_i$  décrivant l'interaction entre les particules et le champ se transforme quant à elle en

$$S_i^* = - \left[ \int A_\mu J^\mu dV dt + \int \partial_\mu \psi J^\mu dV dt \right] = S_i - \int \partial_\mu \psi J^\mu dV dt \quad (5.90)$$

une intégration du terme supplémentaire fournit

$$S_i^* = S_i + \int \psi \partial_\mu J^\mu dV dt \quad (5.91)$$

le terme de bord est nul car il n'y a pas de charges à l'infini. En nous souvenant de l'équation de conservation de la charge écrite sous forme quadri-vectorielle  $\partial_\mu J^\mu = 0$ , on trouve bien que  $S_i^* = S_i$ . On retrouve ici une nouvelle facette insoupçonnée du théorème de Noether : l'invariance de jauge des équations de l'électromagnétisme ( $S_i^* = S_i$ ) est équivalente à la conservation de la charge ( $\partial_\mu J^\mu = 0$ ).

## 5.4 Les équations de Lagrange en relativité restreinte

Tout est maintenant plus ou moins construit. Il est temps de revenir sur ce que nous avons fait d'une manière plus formelle et plus directe afin que chacun puisse comprendre que la physique est belle !

### 5.4.1 Equation de la dynamique d'une particule relativiste

Considérons une particule repérée par sa position  $\mathbf{x} \in \mathbb{M}^4$  qui évolue dans l'espace-temps selon un certain paramétrage  $\lambda$  qui peut-être par exemple son temps propre, mais pas que... Sa ligne d'univers  $\mathcal{L}$  est l'ensemble des positions successives de  $\mathbf{x}$  au gré de l'évolution sur toutes les valeurs possibles de  $\lambda$ . En choisissant par exemple de représenter  $\mathbf{x}$  par sa composante contravariante nous aurons

$$\mathcal{L} : x^\mu = x^\mu(\lambda) \text{ avec } \lambda \in \mathbb{R} \text{ et } \mu \in (0, 1, 2, 3)$$

En généralisant la notation  $\cdot$  de la mécanique classique nous pouvons écrire

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}}{d\lambda} \text{ et donc par exemple } \dot{x}^\mu = \frac{dx^\mu}{d\lambda}$$

Le lagrangien  $\mathcal{L}$  de cette particule est un champ scalaire des variables indépendantes  $\mathbf{x}(\lambda)$  et  $\dot{\mathbf{x}}(\lambda)$ . Si  $\mathcal{L}$  dépend en plus de  $\lambda$ , on pourra sans peine généraliser le cas traité en mécanique classique (voir 2.102), en introduisant un lagrangien étendu.

L'action de cette particule est alors donnée par l'intégrale

$$S = \int_{\lambda_1}^{\lambda_2} \mathcal{L}[\mathbf{x}(\lambda), \dot{\mathbf{x}}(\lambda)] d\lambda \quad (5.92)$$

qui sera minimale le long de  $\mathcal{L}$ .

Le lagrangien  $\mathcal{L}[\mathbf{x}(\lambda), \dot{\mathbf{x}}(\lambda)]$  étant un scalaire pour que ces équations soient les mêmes dans tous les référentiels, c'est-à-dire indépendante du paramétrage  $\lambda$ . Cette dernière contrainte impose que  $\mathcal{L}$  soit une fonction homogène de degré 1 pour sa variable  $\dot{\mathbf{x}}$ . Ce qui signifie que  $\forall k \in \mathbb{R}$ ,  $\mathcal{L}(\mathbf{x}, k\dot{\mathbf{x}}) = k\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$ , et ce qui implique par le théorème d'Euler sur les fonctions homogènes que

$$\dot{\mathbf{x}} \cdot \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}} = \mathcal{L} \quad (5.93)$$

La condition d'extremum pour l'action (5.92) conduit aux équations de Lagrange qui s'écrivent en relativité restreinte

$$\frac{d}{d\lambda} \left( \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}} \right) - \frac{\partial \mathcal{L}}{\partial \mathbf{x}} = 0$$

La ligne d'univers d'une particule est donc définie par 4 équations de Lagrange en relativité au lieu de 3 en mécanique classique, on pouvait s'y attendre ! Cependant, la contrainte (5.93) est indépendante des équations de Lagrange. Ce sont donc bien 3 équations qui décrivent le mouvement : *la relativité n'ajoute pas de degré de liberté !*

Pour obtenir les équations du mouvement, il suffit de faire la variation  $\mathbf{x}(\lambda) \rightarrow \mathbf{x}(\lambda) + \delta \mathbf{x}(\lambda)$  telle que  $\delta \mathbf{x}(\lambda_1) = \delta \mathbf{x}(\lambda_2) = \mathbf{0}$  et d'écrire que la variation correspondante  $\delta S = 0$ . Cette condition assure l'extremum, pour le minimum il faut encore montrer que la variation de  $S$  à l'ordre 2 est positive...

Par exemple, si cette particule de masse  $m$  évolue dans un potentiel décrit par le 4-vecteur  $\mathbf{P} = \mathbf{P}(\mathbf{x})$  et qu'elle y est sensible par l'intermédiaire de son scalaire de couplage  $\chi$  alors son mouvement dans ce potentiel est décrit par l'action

$$S = \int_{\lambda_1}^{\lambda_2} \mathcal{L} d\lambda \text{ avec } \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = -mc\sqrt{\dot{\mathbf{x}} \cdot \dot{\mathbf{x}}} - \chi \mathbf{P}(\mathbf{x}) \cdot \dot{\mathbf{x}} \quad (5.94)$$

La recherche d'un extremum de cette action ( $\delta S = 0$ ) conduit à l'équation du mouvement, que l'on pourrait même appeler de façon nostalgique le principe fondamental de la dynamique relativiste

$$m\Gamma^\mu = \chi W_{\mu\nu} u^\nu \quad \text{avec } W_{\mu\nu} = P_{\nu,\mu} - P_{\mu,\nu}$$

Cette équation qui relie la 4–accélération  $\Gamma$  à la 4–vitesse  $\mathbf{u}$  via le potentiel  $\mathbf{P}$  est une généralisation relativiste de la fameuse loi de Newton. Nous l'avons trouvée dans le cadre d'un calcul laborieux, voir (5.28), elle vient ici en quelques lignes dans un calcul de dérivée sans composantes que nous laissons au lecteur. Le résultat final est délicat à exprimer sans composantes car  $P_{\nu,\mu}$  et son compère ne s'expriment pas simplement en terme des 4–vecteurs  $\mathbf{P}$  et  $\partial$ .

Il existe bien d'autres lagrangien pour des particules relativistes, à vous d'en créer ! On peut également écrire des équations pour des champs tensoriels...

## 5.4.2 Les équations du champ

### Quelques généralités

Les équations de Lagrange que nous venons de reformuler utilisent des variables vectorielles que sont les positions et les vitesses généralisées  $\mathbf{x}$  et  $\dot{\mathbf{x}}$ , qui sont les variables indépendantes du lagrangien.

Nous savons bien par l'expérience que certains mouvements sont à l'origine de certains champs : en fait, en théorie des champs ces derniers sont indissociablement rattachés au mouvement à travers un principe variationnel... La physique est bien faite !

Pour obtenir de telles équations il faut considérer non plus des variations de la position  $\mathbf{x}$  dans un champ modélisé par un tenseur d'ordre  $n$  que l'on appellera  $\mathbf{C}$ , mais le contraire.

Ce champ est induit par un courant de particules modélisé par un 4–courant  $\mathbf{J}(\mathbf{x})$ . La composante contravariante de ce 4–courant est comme nous l'avons vu  $J^\mu = [\rho(\mathbf{x})c, \rho(\mathbf{x})\vec{v}]^\top$  où  $\rho(\mathbf{x})$  est le champ scalaire décrivant la densité volumique des particules à chaque instant et  $\vec{v}$  la 3–vitesse de ces particules.

Le lagrangien du champ  $\mathbf{C}$  est alors une fonction de  $\mathbf{C}$  lui-même est de ses variations que l'on note  $\partial_\mu \mathbf{C}$  ainsi on notera  $\mathcal{L} = \mathcal{L}(\mathbf{C}, \partial_\mu \mathbf{C})$ . La notation  $\partial_\mu \mathbf{C}$  est symbolique car elle mélange un tenseur et une composante, elle est cependant efficace pour écrire les équations du champ. On peut être

plus explicite et faire apparaître les composantes de  $\mathbf{C}$  et  $\partial_\mu \mathbf{C}$  qui sont des tenseurs d'ordre différent...

L'action se calcule par une intégration de ce lagrangien sur toute la partie  $\vartheta$  de l'espacetemps contenant le courant  $\mathbf{J}(\mathbf{x})$ , pour que tout cela ne diverge pas on considère physiquement que le flux de  $\mathbf{J}$  à travers la surface  $\Sigma$  de  $\vartheta$  est nul. On a donc

$$S = \int_{\vartheta} \mathcal{L}(\mathbf{C}, \partial_\mu \mathbf{C})$$

On appelle parfois ce genre de lagrangien une densité de lagrangien car ses variables sont des champs, mais restons calme !

Les équations de Lagrange sont celles qui correspondent à un extremum de cette action, on montre assez facilement qu'elles s'écrivent

$$\frac{\partial \mathcal{L}}{\partial \mathbf{C}} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \mathbf{C})} \right) = 0$$

On appelle ces équations les équations du champ : les degrés de liberté  $x^\mu$  de la particule sont remplacés par les composantes du champ  $\mathbf{C}$  et le paramètre  $\lambda$  par les coordonnées de  $\mathbf{x}$  dans l'espacetemps restreint à  $\vartheta$ .

### Quelques exemples

Le cas le plus simple est celui d'un champ scalaire :  $\mathbf{C} = \phi(\mathbf{x}) \in \mathbb{R}$  et d'une théorie de type Klein-Gordon dont le lagrangien est de la forme

$$\mathcal{L}(\phi, \partial_\mu \phi) = -\frac{1}{2} \left[ \boldsymbol{\partial}(\phi) \cdot \boldsymbol{\partial}(\phi) + \frac{\phi^2}{\ell^2} \right]$$

où  $\ell$  est une longueur. En théorie quantique des champs  $\ell = \frac{\hbar}{mc}$  si  $m$  est la masse de la particule considérée. L'équation du champ de ce lagrangien scalaire est l'équation de Klein-Gordon

$$\square \phi = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \Delta \phi = \frac{1}{\ell^2} \phi$$

Un cas un peu moins simple est celui dans lequel le champ  $\mathbf{C}$  est un 4-vecteur et son lagrangien, dit de Proca, donné par

$$\mathcal{L}(\mathbf{C}, \partial_\mu \mathbf{C}) = \left( \mathbf{J} - \frac{\xi}{2} \mathbf{C} \right) \cdot \mathbf{C} - \frac{\Lambda}{4} \eta^{\mu\rho} \eta^{\nu\sigma} (\partial_\mu C_\nu - \partial_\nu C_\mu) (\partial_\rho C_\sigma - \partial_\sigma C_\rho)$$

Les paramètres réels  $\xi$  et  $\Lambda$  sont fixés par les caractéristiques physiques du champ considéré. Les équations de ce champ généralisent les équations de Maxwell, nous laisserons les lecteurs courageux les écrire. Ce n'est pas si compliqué...

## 5.5 Le théorème de Noether en théorie des champs

Pour terminer ce chapitre, examinons ce que devient le théorème de Noether en relativité des champs, car en relativité des particules rien n'est changé...

### 5.5.1 Le théorème pour un champ scalaire

On note  $\mathbf{x}$  un évènement de l'espacetemps. On considère une théorie dont la densité de lagrangien  $\mathcal{L}$  dépend uniquement d'un champ scalaire  $\phi(\mathbf{x})$  et de ses dérivées  $\partial_\mu \phi$ , soit  $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi)$ . On suppose que ce lagrangien est invariant sous l'action d'un groupe de transformation  $\mathbb{G}$ . C'est-à-dire qu'il existe une fonction  $g_\sigma$  telle que

$$\mathbb{G} = \{g_\sigma, \sigma \in \mathbb{R}\} \begin{cases} g_0 = \text{Id} \\ \forall (\sigma_1, \sigma_2) \in \mathbb{R}^2, & g_{\sigma_1} \circ g_{\sigma_2} = g_{\sigma_1 + \sigma_2} \\ \forall \sigma \in \mathbb{R}, & \mathcal{L}(g_\sigma(\phi), \partial_\mu [g_\sigma(\phi)]) = \mathcal{L}(\phi, \partial_\mu \phi) \end{cases} \quad (5.95)$$

Il n'y a pas de problème de notation car  $g$  est une fonction et pas un tenseur, cette fonction dépend du paramètre réel  $\sigma$  qui n'est donc pas un indice.

On définit le *courant de Noether* associée au groupe de symétrie  $\mathbb{G}$  par le 4-vecteur dont la composante contravariante est donnée par la relation

$$j^\mu = \left. \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \frac{dg_\sigma(\phi)}{d\sigma} \right|_{\sigma=0}$$

On montre assez facilement que la symétrie de  $\mathcal{L}$  sous l'action de  $\mathbb{G}$  implique la conservation du courant de Noether le long de la ligne d'Univers de  $\mathbf{x}$ . Ce résultat constitue un cas particulier du théorème de Noether pour le cas des champs scalaires. On en acceptera sans peine une généralisation à un champ tensoriel d'ordre quelconque.

### 5.5.2 Un exemple pour champ de 4-vecteurs

En théorie quantique des champs, le vide peut être assimilé à une région de l'espacetemps  $\mathcal{V}$  dans laquelle règne un champ électromagnétique décrit

par le 4-potential  $\mathbf{A}(\mathbf{x})$  défini par la relation (5.32), mais dans laquelle le 4-courant est identiquement nul :  $\mathbf{J}(\mathbf{x}) \equiv 0$ . Le lagrangien associé au vide s'écrit

$$\mathcal{L}(A_\mu, \partial_\nu A_\mu) = -\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} \quad \text{avec } F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

On considère le groupe de symétrie engendré par la fonction

$$g_\sigma : \begin{array}{l} \mathbb{M}_{4,\mathbb{R}} \rightarrow \mathbb{M}_{4,\mathbb{R}} \\ \mathbf{y} \mapsto \mathbf{y} + \sigma \boldsymbol{\partial} \psi \end{array}$$

où  $\psi$  est un champ scalaire appelé jauge électromagnétique, ses seules propriétés sont d'être dérivable et de s'annuler sur le bord de  $\vartheta$ . On parle de symétrie de jauge...

En généralisant la notion de courant de Noether on définit maintenant le courant de jauge électromagnétique par la relation

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \mathbf{A})} \frac{dg_\sigma(\mathbf{A})}{d\sigma} \Big|_{\sigma=0}$$

Le théorème de Noether assure ici que dans toutes les jauges électromagnétiques

$$\int_{\vartheta} j^\mu = 0$$

Voilà qui n'a rien de surprenant !



**Théorie du champ de gravitation : Relativité  
générale**



## Principe d'équivalence

---

Il n'est pas anodin de remarquer que les champs de gravitation jouissent d'une propriété extraordinaire : tous les corps s'y déplacent de la même manière indépendamment de leur masse. Par exemple dans un champ de gravitation uniforme  $\vec{g}$  (celui régnant dans tous les laboratoires de physique normaux), un corps ponctuel de masse  $m$  constante, voit son accélération  $\vec{a}$  vérifier l'équation

$$m\vec{a} = m\vec{g} \quad (6.1)$$

Très tôt dans sa scolarité, on n'hésite pas à simplifier dans cette équation ce facteur  $m$  qui ne semble donc pas intervenir dans le mouvement<sup>1</sup>. Une remarque doit cependant être faite : dans le terme de gauche de l'équation (6.1), le coefficient  $m$  est associé à une quantité inertielle : plus  $m$  est grand, plus il est difficile de mettre en mouvement le corps associé. On parle donc de masse inertielle,  $m_i$ . Dans le terme de droite par contre, le coefficient  $m$  décrit un couplage avec l'interaction gravitationnelle, on pourrait parler de charge gravitationnelle, on dit plutôt masse grave,  $m_g$ . En simplifiant l'équation (6.1), on fait donc l'hypothèse que  $m_i = m_g$ .

L'égalité de ces deux termes pourtant de nature différente est déjà identifiée par Galilée, qui au moins par la pensée, laisse choir des corps du haut de la tour de Pise pour s'en convaincre. Plus tard, en 1889, le Hongrois Eötvös utilisant un pendule de torsion montra que  $m_i/m_g = 1$  au milliardième près ! En 1964, Dicke poussa la précision jusqu'à  $10^{-11}$  et l'on envisage même de satelliser un gyroscope pour prouver l'équivalence entre  $m_i$  et  $m_g$  jusqu'à  $10^{-15}$  !

---

<sup>1</sup> Si le champ n'est plus uniforme le vecteur constant  $\vec{g}$  se met à dépendre de la position, mais la force subie par la masse  $m$  reste toujours donnée par la même relation.

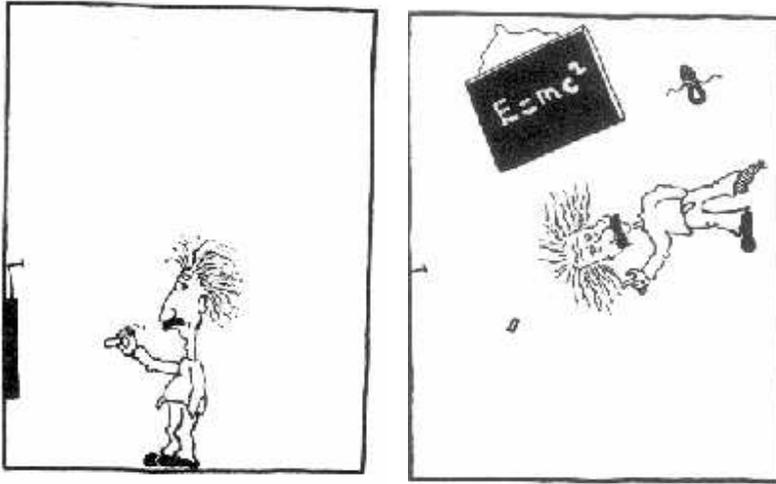


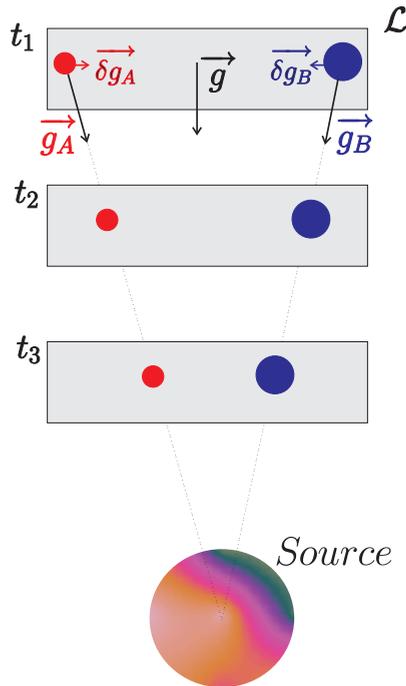
Fig. 6.1. Albert Einstein dans un ascenseur

Il pourrait paraître exagéré de chercher une telle précision, nous allons voir qu'en fait c'est de l'équivalence de ces deux termes que naît la nouvelle théorie de la relativité générale.

Si  $m_i = m_g$ , la relation (6.1) nous indique qu'il y a équivalence entre un champ de gravitation homogène et une accélération absolue : c'est la fameuse expérience de l'ascenseur d'Einstein. Dans une cabine d'ascenseur au repos (voir figure 6.1 de gauche), un observateur est capable de mettre en évidence la présence d'un champ de gravitation dirigé vers le bas (il suffit de laisser tomber un objet ...). Dans la même cabine en chute libre (voir figure 6.1 de droite), le même observateur n'est plus en mesure de détecter la présence du champ qui est pourtant toujours bien là ! Pour résumer, si  $m_i = m_g$ , tout se passe comme si un champ de gravitation uniforme était équivalent à une accélération absolue. Il n'est cependant pas très difficile de comprendre que cette expérience<sup>2</sup> n'a de valeur que locale. Comme nous l'avons dit plus haut, si le champ de gravitation est constant, ce n'est que parce que l'on considère un ascenseur (ou plus généralement un laboratoire) petit devant la source du champ de gravitation (en l'occurrence la Terre). En observant à divers instants ( $t_1 < t_2 < t_3$ ) la chute libre d'un ascenseur de taille comparable à celle de la Terre dans le champ de gravitation de celle-ci (voir figure 6.2), il n'est pas difficile de mettre en évidence la présence du champ de gravitation. Celle-ci se manifeste en effet, sous la forme d'une "force" différentielle (cer-

<sup>2</sup> qu'il ne faut surtout pas essayer de reproduire en grandeur nature.

tains disent “de marée”) qui tend à faire se rapprocher les objets en chute libre dans la cabine. Cette force résulte directement de la non-uniformité du champ de gravitation, toutefois les propriétés de continuité de celui-ci nous permettent toujours d'écrire : “*Un champ de gravitation peut être localement compensé par une accélération absolue*” ou bien “*D'un point de vue local, il n'y a pas de différence entre un champ de gravitation et une accélération absolue*”. Il s'agit là de ce que l'on appelle le principe d'équivalence faible.



**Fig. 6.2.** Référentiel inertiel local

Un référentiel galiléen est un référentiel dans lequel un système abandonné au repos et ne subissant pas de forces extérieures, reste dans cet état et suit un mouvement rectiligne et uniforme. Un référentiel en chute libre dans un champ de gravitation strictement uniforme est donc un référentiel galiléen. Par essence, un champ de gravitation ne peut être partout constant : il faudrait qu'il soit nul partout ou bien associé à une masse infinie. De ce constat nous sommes donc forcés de conclure qu'il n'existe pas un référentiel galiléen (on dit aussi inertiel) global couvrant tout l'univers ! La relativité restreinte nous a appris que la propriété principale de l'espace-temps associé à un référentiel inertiel global est sa métrique de Minkowski - caractéristique

d'une variété plate. Si l'on veut prendre en compte le champ de gravitation dans une théorie plus générale, il semble donc qu'il faille abandonner la métrique de Minkowski pour une métrique plus générale, et donc notre variété plate pour une variété plus générale mais localement minkowskienne !

Le principe de relativité restreinte imposait une même forme pour toutes les équations écrites dans des référentiels *inertiels*. Le nouveau principe qui va nous guider pour écrire les équations les plus générales sera donc tout simplement : “*Les équations de la physique s'écrivent de la même façon dans tous les référentiels*”.

## Application du principe d'équivalence

### 7.1 Mouvement géodésique

Considérons une particule libre de toute force autre que la gravitation (seule force que l'on ne peut écranter). En vertu du principe d'équivalence, il existe un référentiel inertiel local associé à des coordonnées  $\xi^\alpha$  dans lequel les équations du mouvement sont celles d'une ligne droite dans l'espace-temps, c'est-à-dire

$$\frac{d^2 \xi^\alpha}{d\tau^2} = 0 \quad (7.1)$$

où  $d\tau$  est le temps propre relié à l'élément d'intervalle spatio-temporel et à la métrique par l'équation

$$ds^2 = c^2 d\tau^2 = \eta_{\alpha\beta} d\xi^\alpha d\xi^\beta.$$

La métrique qui intervient ici est celle de  $\mathbb{M}_{4,\mathbb{R}}$  car les coordonnées sont inertielles.

Considérons à présent un autre système de coordonnées  $x^\mu$ , quelconque celui-ci. Les coordonnées inertielles sont fonction de ce nouveau système  $\xi^\alpha = \xi^\alpha(x^\mu)$ , ainsi l'équation du mouvement s'écrit

$$\begin{aligned} 0 &= \frac{d}{d\tau} \left( \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{dx^\mu}{d\tau} \right) \\ &= \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{d^2 x^\mu}{d\tau^2} + \frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \end{aligned} \quad (7.2)$$

en multipliant cette dernière relation par  $\frac{\partial x^\lambda}{\partial \xi^\alpha}$  et en se souvenant de l'identité

$$\frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial x^\lambda}{\partial \xi^\alpha} = \delta_\mu^\lambda \quad (7.3)$$

nous avons donc

$$0 = \frac{d^2 x^\lambda}{d\tau^2} + \frac{\partial x^\lambda}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \quad (7.4)$$

la quantité

$$\Gamma_{\mu\nu}^\lambda := \frac{\partial x^\lambda}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\nu} \quad (7.5)$$

est bien sûr indépendante de  $\alpha$ ; on l'appelle symbole de connection affine. Grâce à elle, l'équation de la chute libre dans un référentiel quelconque s'écrit donc

$$\frac{d^2 x^\lambda}{d\tau^2} + \Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0 \quad (7.6)$$

cette équation est appelée équation des géodésiques. Comme le montre la relation de départ (7.1), la solution de cette équation est la courbe  $x^\nu(\tau)$  d'accélération nulle écrite dans un système de coordonnées quelconque.

Si nous ne considérons que des référentiels en translation uniforme les uns par rapport aux autres (et donc échangeables par une transformation de Lorentz), alors les fonctions  $\xi^\alpha(x^\mu)$  sont linéaires, leurs dérivées secondes s'annulent ainsi que les symboles de connection affine. Une interprétation du principe d'équivalence permet donc de comprendre que l'apparition de symboles de connection affine est une conséquence directe de la non-uniformité du champ de gravitation, en poussant l'analogie on peut même identifier ces deux notions<sup>1</sup>.

## 7.2 Connexion affine et tenseur métrique

Dans le système de coordonnées inertielles, nous avons  $ds^2 = \eta_{\alpha\beta} d\xi^\alpha d\xi^\beta$  dans un système de coordonnées quelconque nous aurons

$$\begin{aligned} ds^2 &= \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\nu} dx^\mu dx^\nu \\ &:= g_{\mu\nu} dx^\mu dx^\nu \end{aligned} \quad (7.7)$$

où l'on a fait apparaître le tenseur métrique

<sup>1</sup> L'implication constitue le principe d'équivalence faible et l'identification constitue le principe d'équivalence fort.

$$g_{\mu\nu} := \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\nu} \quad (7.8)$$

Attendu que  $\eta_{\alpha\beta}$  est constant, en dérivant cette équation par rapport à  $x^\lambda$  il vient

$$\partial_\lambda g_{\mu\nu} = \eta_{\alpha\beta} \left( \frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\lambda} \frac{\partial \xi^\beta}{\partial x^\nu} + \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial^2 \xi^\beta}{\partial x^\nu \partial x^\lambda} \right) \quad (7.9)$$

en se rappelant la définition (7.5) du symbole de connexion affine, on peut constater que

$$\partial_\lambda g_{\mu\nu} = \eta_{\alpha\beta} \left( \Gamma_{\lambda\mu}^\rho \frac{\partial \xi^\alpha}{\partial x^\rho} \frac{\partial \xi^\beta}{\partial x^\nu} + \Gamma_{\lambda\nu}^\rho \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\rho} \right) \quad (7.10)$$

qui s'écrit encore en utilisant la définition (7.8) du tenseur métrique

$$\partial_\lambda g_{\mu\nu} = \Gamma_{\lambda\mu}^\rho g_{\rho\nu} + \Gamma_{\lambda\nu}^\rho g_{\rho\mu} \quad (7.11)$$

en se livrant à un petit jeu de permutation d'indices on obtient directement

$$\begin{aligned} \partial_\mu g_{\lambda\nu} &= \Gamma_{\mu\lambda}^\rho g_{\rho\nu} + \Gamma_{\mu\nu}^\rho g_{\rho\lambda} \\ \text{et} \\ \partial_\nu g_{\mu\lambda} &= \Gamma_{\nu\mu}^\rho g_{\rho\lambda} + \Gamma_{\nu\lambda}^\rho g_{\rho\mu} \end{aligned} \quad (7.12)$$

le symbole de connexion affine a ses deux indices du bas qui commutent (la définition (7.5) est explicite à ce sujet), ainsi

$$\begin{aligned} \partial_\lambda g_{\mu\nu} + \partial_\mu g_{\lambda\nu} - \partial_\nu g_{\lambda\mu} &= \Gamma_{\lambda\mu}^\rho g_{\rho\nu} + \Gamma_{\lambda\nu}^\rho g_{\rho\mu} + \Gamma_{\mu\lambda}^\rho g_{\rho\nu} + \Gamma_{\mu\nu}^\rho g_{\rho\lambda} \\ &\quad - \Gamma_{\nu\lambda}^\rho g_{\rho\mu} - \Gamma_{\nu\mu}^\rho g_{\rho\lambda} \\ &= 2\Gamma_{\lambda\mu}^\rho g_{\rho\nu} \end{aligned} \quad (7.13)$$

multipliant cette expression par la métrique inverse il vient

$$\begin{aligned} (\partial_\lambda g_{\mu\nu} + \partial_\mu g_{\lambda\nu} - \partial_\nu g_{\lambda\mu}) g^{\nu\sigma} &= 2\Gamma_{\lambda\mu}^\rho g_{\rho\nu} g^{\nu\sigma} \\ &= 2\Gamma_{\lambda\mu}^\rho \delta_\rho^\sigma \end{aligned} \quad (7.14)$$

nous avons donc explicité le symbole de connexion affine en fonction de la métrique et de ses dérivées premières, en notation à virgule, on a

$$\Gamma_{\lambda\mu}^\sigma = \frac{1}{2} (g_{\mu\nu,\lambda} + g_{\lambda\nu,\mu} - g_{\lambda\mu,\nu}) g^{\nu\sigma} \quad (7.15)$$

Pour faire le lien entre ce petit calcul et la physique, rappelons-nous que le principe d'équivalence (dans sa version forte) ne stipulait pas autre chose que l'équivalence entre le fait que les  $\Gamma_{\lambda\mu}^\sigma$  soient non nuls et la non-uniformité intrinsèque d'un champ de gravitation. En poussant l'équivalence encore plus loin, nous pouvons au vu de l'équation (7.15) identifier la non-uniformité du champ gravitationnel avec le fait que le tenseur métrique n'est pas constant.

### 7.3 La connexion affine est-elle un tenseur ?

Nous avons relié la connexion affine aux dérivées premières du tenseur métrique, et constaté qu'elle possédait une symétrie sur ses indices du bas, mais est-ce un tenseur ?

Pour le savoir, il faut lui faire passer le test du changement de repère. Soit donc  $\xi^\alpha(x^\mu)$  le référentiel inertiel local dans lequel par définition

$$\Gamma_{\mu\nu}^\lambda = \frac{\partial x^\lambda}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\nu} \quad (7.16)$$

passant de  $x^\mu$  à  $x'^\mu$  nous avons

$$\begin{aligned} \Gamma'_{\mu\nu}{}^\lambda &= \frac{\partial x'^\lambda}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x'^\mu \partial x'^\nu} \\ &= \frac{\partial x'^\lambda}{\partial x^\rho} \frac{\partial x^\rho}{\partial \xi^\alpha} \frac{\partial}{\partial x'^\mu} \left( \frac{\partial \xi^\alpha}{\partial x^\sigma} \frac{\partial x^\sigma}{\partial x'^\nu} \right) \\ &= \frac{\partial x'^\lambda}{\partial x^\rho} \frac{\partial x^\rho}{\partial \xi^\alpha} \left( \frac{\partial^2 \xi^\alpha}{\partial x^\sigma \partial x^\tau} \frac{\partial x^\tau}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} + \frac{\partial \xi^\alpha}{\partial x^\sigma} \frac{\partial^2 x^\sigma}{\partial x'^\nu \partial x'^\mu} \right) \end{aligned} \quad (7.17)$$

en faisant apparaître la connexion dans le second membre, il vient

$$\begin{aligned} \Gamma'_{\mu\nu}{}^\lambda &= \Gamma_{\sigma\tau}^\rho \frac{\partial x^\tau}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} \frac{\partial x'^\lambda}{\partial x^\rho} + \frac{\partial x'^\lambda}{\partial x^\rho} \frac{\partial x^\rho}{\partial \xi^\alpha} \frac{\partial \xi^\alpha}{\partial x^\sigma} \frac{\partial^2 x^\sigma}{\partial x'^\nu \partial x'^\mu} \\ &= \Gamma_{\sigma\tau}^\rho \frac{\partial x^\tau}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} \frac{\partial x'^\lambda}{\partial x^\rho} + \frac{\partial x'^\lambda}{\partial x^\rho} \delta_\sigma^\rho \frac{\partial^2 x^\sigma}{\partial x'^\nu \partial x'^\mu} \\ &= \Gamma_{\sigma\tau}^\rho \frac{\partial x^\tau}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} \frac{\partial x'^\lambda}{\partial x^\rho} + \frac{\partial x'^\lambda}{\partial x^\sigma} \frac{\partial^2 x^\sigma}{\partial x'^\nu \partial x'^\mu} \end{aligned} \quad (7.18)$$

les tenseurs (car ils en sont trivialement) de la forme  $\frac{\partial x^\tau}{\partial x'^\mu}$  ou bien  $\frac{\partial x'^\lambda}{\partial x^\rho}$  sont des généralisations non linéaires de nos chères transformations de Lorentz, nous leur consacrons donc une notation :

$$\mathcal{A}^\alpha{}_\beta := \frac{\partial x'^\alpha}{\partial x^\beta} \quad \text{et} \quad \mathcal{A}_\alpha{}^\beta := \frac{\partial x^\beta}{\partial x'^\alpha} \quad (7.19)$$

Nul n'est besoin de rappeler à notre cher lecteur devenu familier de ce genre de notations, que l'ordre des  $\alpha, \beta$  est fondamental, ni la petite règle que

nous avons évoquée pour les tenseurs de Lorentz qui permet justement de ne pas se tromper dans l'ordre des indices : sur ce genre de tenseur la sommation a toujours lieu avec l'indice le plus à droite. Pour se rappeler l'ordre d'intervention du prime, c'est-à-dire savoir s'il faut primer le  $x$  du haut ou celui du bas, il faut se rappeler qu'en mathématique, physique où ailleurs seul le premier est primé ... Ainsi c'est la variable associée au premier indice apparaissant dans  $\mathcal{A}$  qui reçoit la prime.

Dans ce jargon, la dernière des relations (7.18) s'écrit simplement

$$\Gamma'^{\lambda}_{\mu\nu} = \mathcal{A}_{\mu}{}^{\tau} \mathcal{A}_{\nu}{}^{\sigma} \mathcal{A}^{\lambda}{}_{\rho} \Gamma^{\rho}_{\sigma\tau} + \mathcal{A}^{\lambda}{}_{\sigma} \frac{\partial^2 x^{\sigma}}{\partial x'^{\mu} \partial x'^{\nu}} \quad (7.20)$$

le premier terme de cette relation est donc manifestement covariant (il correspond à une transformation normale de tenseur une fois contravariant et deux fois covariant), la présence du second terme, que nous nous efforçons de qualifier d'inhomogène, ne permet donc pas au symbole de connexion affine d'accéder au statut tant recherché de tenseur.

Ceci n'est pas une anecdote et constitue donc une bonne mise en garde envers celui qui croirait que tout est pour le mieux dans le meilleur des mondes.

## 7.4 Dérivée covariante

### 7.4.1 Nécessité d'une nouvelle dérivée

La connexion affine n'est donc pas un tenseur. Rappelons-nous que nous avons déjà «failli» connaître pareille désillusion : en relativité restreinte, nous avons remarqué (cf. remarque fondamentale de la page 131) que la dérivée d'un vecteur  $\partial_{\mu} A^{\nu}$  n'était un tenseur que grâce au fait que la transformation de Lorentz était linéaire.

Qu'en est-il en relativité générale où toutes les transformations sont permises?

Pour la composante contravariante d'un vecteur nous avons la transformation

$$V'^{\mu} = \mathcal{A}^{\mu}{}_{\nu} V^{\nu} \quad \text{avec rappelons-le} \quad \mathcal{A}^{\mu}{}_{\nu} = \frac{\partial x'^{\mu}}{\partial x^{\nu}} \quad (7.21)$$

en dérivant cette expression par rapport à  $x'^{\lambda}$  il vient

$$\begin{aligned}
 \frac{\partial V'^{\mu}}{\partial x'^{\lambda}} &= \frac{\partial V^{\nu}}{\partial x'^{\lambda}} \frac{\partial x'^{\mu}}{\partial x^{\nu}} + V^{\nu} \frac{\partial^2 x'^{\mu}}{\partial x^{\nu} \partial x'^{\lambda}} \\
 &= \frac{\partial x'^{\mu}}{\partial x^{\nu}} \frac{\partial x^{\rho}}{\partial x'^{\lambda}} \frac{\partial V^{\nu}}{\partial x^{\rho}} + V^{\nu} \frac{\partial^2 x'^{\mu}}{\partial x^{\nu} \partial x'^{\lambda}} \\
 &= \mathcal{A}^{\mu}_{\nu} \mathcal{A}_{\lambda}^{\rho} \frac{\partial V^{\nu}}{\partial x^{\rho}} + V^{\nu} \frac{\partial^2 x'^{\mu}}{\partial x^{\nu} \partial x'^{\lambda}}
 \end{aligned} \tag{7.22}$$

Tout comme pour la connexion affine, si le premier terme du membre de droite de cette dernière relation est manifestement covariant, il persiste un terme qui fait perdre, comme attendu par l'analyse faite en relativité restreinte, le précieux statut de tenseur à la dérivée d'un vecteur.

Les équations de la physique, grandes consommatrices de ce genre d'objet, ne seraient donc pas covariantes !

Il n'en est bien sûr pas question et deux possibilités s'offrent à nous :

- Premièrement nous pourrions ne chercher à écrire que des équations covariantes mais faisant intervenir des objets non covariants. Au prix d'habiles écritures, ceci doit être possible dans certains cas particuliers, mais paraît très difficile à envisager dans un contexte général;
- Une deuxième possibilité consiste à modifier la notion de dérivée pour l'adapter au caractère général covariant de la relativité générale : c'est cette deuxième voie que nous suivrons.

### 7.4.2 Construction

Avec les deux échecs de la dérivée et de la connexion, formons une victoire !

En effet, multipliant terme à terme (7.20) par (7.21), il vient

$$\Gamma'^{\lambda}_{\mu\nu} V'^{\nu} = \left( \mathcal{A}_{\mu}^{\tau} \mathcal{A}_{\nu}^{\sigma} \mathcal{A}^{\lambda}_{\rho} \Gamma^{\rho}_{\sigma\tau} + \mathcal{A}^{\lambda}_{\rho} \frac{\partial^2 x^{\rho}}{\partial x'^{\nu} \partial x'^{\mu}} \right) \mathcal{A}^{\nu}_{\eta} V^{\eta} \tag{7.23}$$

Un petit calcul intermédiaire permet de montrer que<sup>2</sup>

<sup>2</sup> ce n'est qu'une sorte d'intégration par parties :

$$\begin{aligned}
 \mathcal{A}^{\lambda}_{\rho} \frac{\partial^2 x^{\rho}}{\partial x'^{\nu} \partial x'^{\mu}} &= \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial^2 x^{\rho}}{\partial x'^{\nu} \partial x'^{\mu}} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial}{\partial x'^{\mu}} \left( \frac{\partial x^{\rho}}{\partial x'^{\nu}} \right) \\
 &= \frac{\partial}{\partial x'^{\mu}} \left( \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\rho}}{\partial x'^{\nu}} \right) - \frac{\partial x^{\rho}}{\partial x'^{\nu}} \frac{\partial}{\partial x'^{\mu}} \left( \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \right) = \frac{\partial \delta^{\lambda}_{\nu}}{\partial x'^{\mu}} - \frac{\partial x^{\rho}}{\partial x'^{\nu}} \frac{\partial^2 x'^{\lambda}}{\partial x^{\rho} \partial x'^{\mu}} \\
 &= 0 - \frac{\partial x^{\rho}}{\partial x'^{\nu}} \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \frac{\partial^2 x'^{\lambda}}{\partial x^{\sigma} \partial x^{\rho}} = -A_{\nu}^{\rho} A_{\mu}^{\sigma} \frac{\partial^2 x'^{\lambda}}{\partial x^{\sigma} \partial x^{\rho}} \quad \square
 \end{aligned} \tag{7.24}$$

$$\mathcal{A}^\lambda{}_\rho \frac{\partial^2 x^\rho}{\partial x'^\nu \partial x'^\mu} = -\mathcal{A}_\nu{}^\rho \mathcal{A}_\mu{}^\sigma \frac{\partial^2 x'^\lambda}{\partial x^\rho \partial x^\sigma} \quad (7.25)$$

de sorte que

$$\Gamma'^{\lambda}{}_{\mu\nu} V'^{\nu} = \mathcal{A}_\mu{}^\tau \mathcal{A}_\nu{}^\sigma \mathcal{A}^\lambda{}_\rho \mathcal{A}^\nu{}_\eta V^\eta \Gamma_{\sigma\tau}^\rho - \mathcal{A}_\nu{}^\rho \mathcal{A}_\mu{}^\sigma \mathcal{A}^\nu{}_\eta V^\eta \frac{\partial^2 x'^\lambda}{\partial x^\rho \partial x^\sigma} \quad (7.26)$$

notant que  $\mathcal{A}_\nu{}^\sigma \mathcal{A}^\nu{}_\eta = \delta_\eta^\sigma$  d'une part et que  $\mathcal{A}_\nu{}^\rho \mathcal{A}^\nu{}_\eta = \delta_\eta^\rho$  d'autre part, il reste

$$\Gamma'^{\lambda}{}_{\mu\nu} V'^{\nu} = \mathcal{A}_\mu{}^\tau \mathcal{A}^\lambda{}_\rho V^\sigma \Gamma_{\sigma\tau}^\rho - \mathcal{A}_\mu{}^\sigma V^\rho \frac{\partial^2 x'^\lambda}{\partial x^\rho \partial x^\sigma} \quad (7.27)$$

que nous conservons bien au chaud !

Pour le 4–gradient de la composante contravariante du vecteur  $V$  nous avons (cf. (7.22) en inversant les indices  $\mu$  et  $\lambda$  pour des raisons futures)

$$\frac{\partial V'^{\lambda}}{\partial x'^{\mu}} = \mathcal{A}^\lambda{}_\nu \mathcal{A}_\mu{}^\rho \frac{\partial V^\nu}{\partial x^\rho} + \frac{\partial^2 x'^\lambda}{\partial x^\nu \partial x'^\mu} V^\nu \quad (7.28)$$

au prix de deux petites réorganisations :

1. Premier terme : renomination d'indices muets

$$\mathcal{A}^\lambda{}_\nu \mathcal{A}_\mu{}^\rho \frac{\partial V^\nu}{\partial x^\rho} = \mathcal{A}^\lambda{}_\rho \mathcal{A}_\mu{}^\sigma \frac{\partial V^\rho}{\partial x^\sigma} \quad (7.29)$$

2. Second terme : cette dérivée seconde est mixte ( $\partial x^\nu \partial x'^\mu$ ) il faut faire disparaître la dérivée par rapport à  $x'^\mu$  pour la rendre compatible avec la dérivée seconde intervenant dans (7.27). Ce ne sont que de sordides manipulations :

$$\frac{\partial^2 x'^\lambda}{\partial x^\nu \partial x'^\mu} = \frac{\partial}{\partial x'^\mu} \left( \frac{\partial x'^\lambda}{\partial x^\nu} \right) = \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial}{\partial x^\sigma} \left( \frac{\partial x'^\lambda}{\partial x^\nu} \right) = \mathcal{A}_\mu{}^\sigma \frac{\partial^2 x'^\lambda}{\partial x^\nu \partial x^\sigma} \quad (7.30)$$

on a donc, en renommant l'indice muet  $\nu$  en  $\rho$

$$\frac{\partial^2 x'^\lambda}{\partial x^\nu \partial x'^\mu} V^\nu = \frac{\partial^2 x'^\lambda}{\partial x^\rho \partial x^\sigma} \mathcal{A}_\mu{}^\sigma V^\rho \quad (7.31)$$

On obtient donc finalement

$$\frac{\partial V'^{\lambda}}{\partial x'^{\mu}} = \mathcal{A}^\lambda{}_\rho \mathcal{A}_\mu{}^\sigma \frac{\partial V^\rho}{\partial x^\sigma} + \frac{\partial^2 x'^\lambda}{\partial x^\rho \partial x^\sigma} \mathcal{A}_\mu{}^\sigma V^\rho \quad (7.32)$$

et en mélangeant (7.27) et (7.32), les termes inhomogènes disparaissent dans l'harmonie la plus totale en donnant

$$\begin{aligned}
 \frac{\partial V'^{\lambda}}{\partial x'^{\mu}} + \Gamma'^{\lambda}_{\mu\nu} V'^{\nu} &= \mathcal{A}^{\lambda}_{\rho} \mathcal{A}_{\mu}^{\sigma} \frac{\partial V^{\rho}}{\partial x^{\sigma}} + \mathcal{A}_{\mu}^{\tau} \mathcal{A}^{\lambda}_{\rho} V^{\sigma} \Gamma^{\rho}_{\sigma\tau} \\
 &= \mathcal{A}^{\lambda}_{\rho} \mathcal{A}_{\mu}^{\sigma} \left( \frac{\partial V^{\rho}}{\partial x^{\sigma}} + \Gamma^{\rho}_{\alpha\sigma} V^{\alpha} \right)
 \end{aligned} \tag{7.33}$$

la quantité

$$\frac{\partial V^{\rho}}{\partial x^{\sigma}} + \Gamma^{\rho}_{\alpha\sigma} V^{\alpha} \tag{7.34}$$

est donc la composante mixte d'un tenseur d'ordre 2 qui généralise la notion de dérivée. On l'appelle dérivée covariante, en généralisant les notations  $(\partial)$  et  $(,)$  on la note

$$\frac{\partial V^{\rho}}{\partial x^{\sigma}} + \Gamma^{\rho}_{\alpha\sigma} V^{\alpha} = \begin{cases} \partial_{\sigma} V^{\rho} + \Gamma^{\rho}_{\alpha\sigma} V^{\alpha} := D_{\sigma} V^{\rho} \\ V^{\rho}_{, \sigma} + \Gamma^{\rho}_{\alpha\sigma} V^{\alpha} := V^{\rho}_{; \sigma} \end{cases} \tag{7.35}$$

on remarque immédiatement que dans le cas de changement de coordonnées linéaires (relativité restreinte), les termes de connexion s'annulent et  $D_{\sigma} \equiv \partial_{\sigma}$ .

Un calcul en tout point identique permet d'obtenir la dérivée covariante de la composante covariante d'un vecteur, on trouve

$$D_{\sigma} V_{\rho} := V_{\rho; \sigma} = \frac{\partial V_{\rho}}{\partial x^{\sigma}} - \Gamma^{\alpha}_{\rho\sigma} V_{\alpha} \tag{7.36}$$

On peut ainsi généraliser à la dérivée covariante d'un tenseur  $n$  fois covariant et  $p$  fois contravariant : on rajoute au terme de «dérivée normale» des «termes connectés» affublés d'un signe  $-$  pour chaque indice covariant et d'un signe  $+$  pour chaque indice contravariant, par exemple

$$D_{\rho} T^{\mu\nu}_{\sigma} := T^{\mu\nu}_{\sigma; \rho} = T^{\mu\nu}_{\sigma, \rho} + \Gamma^{\mu}_{\alpha\rho} T^{\alpha\nu}_{\sigma} + \Gamma^{\nu}_{\alpha\rho} T^{\mu\alpha}_{\sigma} - \Gamma^{\alpha}_{\sigma\rho} T^{\mu\nu}_{\alpha} \tag{7.37}$$

### 7.4.3 Propriétés

Cette nouvelle dérivée adaptée à la relativité générale possède les propriétés essentielles de tout opérateur de dérivation :

- Linéarité : pour tout couple de réels  $(a, b)$ , et pour toute somme compatible de composantes de tenseurs, on a (ici pour des tenseurs de rang 2)

$$(a A^{\mu}_{\nu} + b B^{\mu}_{\nu})_{; \sigma} = a A^{\mu}_{\nu; \sigma} + b B^{\mu}_{\nu; \sigma} \tag{7.38}$$

- Identité de Leibniz du produit : par exemple

$$(A^\mu{}_\nu B^\rho)_{;\sigma} = A^\mu{}_{\nu;\sigma} B^\rho + A^\mu{}_\nu B^\rho{}_{;\sigma} \quad (7.39)$$

La dérivée covariante du tenseur métrique est nulle. Cette propriété qui s'écrit prosaïquement

$$g_{\mu\nu} \in \ker(D_\lambda) \quad (7.40)$$

se démontre de multiples façons dont nous expliciterons deux versions instructives :

1. Le principe d'équivalence nous indique qu'il existe toujours un système de coordonnées dans lequel l'espace est localement plat et donc dans lequel la métrique se réduit à  $\eta_{\mu\nu}$  dont la dérivée et les connexions induites sont nulles. Or, un tenseur - en l'occurrence  $g_{\mu\nu;\lambda}$  - dont toutes les composantes sont nulles dans un système de coordonnées est nul dans tous les autres.  $\square$
2. Un simple calcul est convaincant mais peu instructif :  $g_{\mu\nu}$  est la composante deux fois covariante d'un tenseur de rang deux, ainsi

$$g_{\mu\nu;\lambda} = g_{\mu\nu,\lambda} - \Gamma_{\lambda\mu}^\alpha g_{\alpha\nu} - \Gamma_{\lambda\nu}^\alpha g_{\mu\alpha} \quad (7.41)$$

mais nous avons justement vu un peu plus haut (voir (7.11)) que

$$g_{\mu\nu,\lambda} := \partial_\lambda g_{\mu\nu} = \Gamma_{\lambda\mu}^\rho g_{\rho\nu} + \Gamma_{\lambda\nu}^\rho g_{\rho\mu} \quad \dots \quad \square \quad (7.42)$$

#### 7.4.4 Dérivée covariante le long d'une courbe

Nous avons défini la dérivée covariante pour des tenseurs eux-mêmes définis sur tout l'espace-temps. Intéressons-nous à des tenseurs définis uniquement sur une courbe, par exemple le vecteur position  $Z^\mu := x^\mu(\tau)$ .

Pour un tel tenseur, lors d'un changement de repère on a

$$Z'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} Z^\nu = \mathcal{A}^\mu{}_\nu Z^\nu \quad (7.43)$$

dérivant cette expression par rapport au temps propre il vient

$$\frac{dZ'^\mu}{d\tau} = \mathcal{A}^\mu{}_\nu \frac{dZ^\nu}{d\tau} + \frac{\partial^2 x'^\mu}{\partial x^\nu \partial x^\lambda} \frac{dx^\lambda}{d\tau} Z^\nu \quad (7.44)$$

cette expression est donc manifestement non covariante. Grâce à notre expérience fraîchement acquise, nous appellerons naturellement "dérivée covariante le long d'une courbe paramétrée par  $\tau$ ", l'opération

$$\frac{DZ^\mu}{D\tau} := \frac{dZ^\mu}{d\tau} + \Gamma_{\lambda\nu}^\mu \frac{dx^\lambda}{d\tau} Z^\nu \quad (7.45)$$

Cette intuition naturelle est confirmée par un calcul qui montre en effet que cette quantité est bien la composante contravariante d'un vecteur qui vérifie donc

$$\frac{DZ^\mu}{D\tau} = \mathcal{A}^\mu{}_\nu \frac{DZ^\nu}{D\tau} \quad (7.46)$$

Cette nouvelle opération est fondamentale, elle permet de comprendre la véritable nature physique de la gravitation (dont nous commençons à nous douter ...).

En mécanique classique, une particule libre se propage en ligne droite et à vitesse constante, l'équation de son mouvement est

$$\frac{d\vec{v}}{dt} = 0 \quad (7.47)$$

Nous avons compris maintenant que la non-uniformité du champ de gravitation conduisait à une équation de *chute* libre, qui devait remplacer (7.47), et qui était de la forme (cf. (7.1))

$$\frac{d^2 x^\lambda}{d\tau^2} + \Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0 . \quad (7.48)$$

Dans un contexte relativiste, la vitesse  $\vec{v}$  doit être remplacée par la quadri-vitesse dont la composante contravariante s'écrit

$$u^\mu = \frac{dx^\mu}{d\tau} \quad (7.49)$$

Il est alors facile de constater que l'équation des géodésiques (7.48) s'écrit très simplement en utilisant la nouvelle dérivée covariante, en effet

$$\frac{d^2 x^\lambda}{d\tau^2} + \Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0 \quad \Leftrightarrow \quad \frac{Du^\lambda}{D\tau} = 0 \quad (7.50)$$

et tout s'éclaire !

Tout comme Galilée s'exclamait «le mouvement est comme rien» en constatant la covariance classique entre deux référentiels en translation uniforme l'un par rapport à l'autre, en relativité générale le mouvement est aussi comme rien, en ce sens qu'une particule libre suit un mouvement géodésique. Mais dans l'espacetemps relativiste, les géodésiques sont courbées, le principe d'équivalence nous indique que cette courbure est inhérente au fait qu'un champ de gravitation est intrinsèquement non uniforme.

D'un point de vue beaucoup plus formel, cette petite aventure permet de construire un principe de correspondance entre la relativité restreinte et la relativité générale :

On vérifie en effet que la transformation

$$\begin{aligned} \eta_{\mu\nu} &\rightarrow g_{\mu\nu} \\ \partial_\mu &\rightarrow D_\mu \quad \text{ou bien} \quad , \rightarrow ; \end{aligned}$$

permet d'obtenir la version générale d'une équation de relativité restreinte.

### 7.5 Déviation géodésique : Courbure

Considérons dans l'espace-temps, deux géodésiques voisines (cf. figure 7.1) .

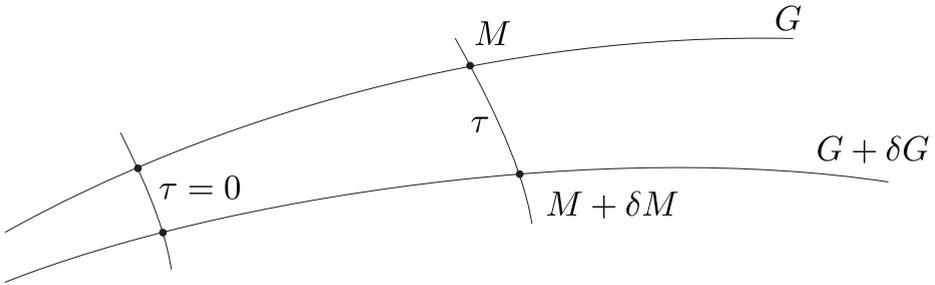


Fig. 7.1. Deux géodésiques voisines

Soit  $G$  l'ensemble des points  $M$  de coordonnées  $x^\mu(\tau)$  qui vérifient l'équation

$$\frac{d^2 x^\lambda}{d\tau^2} + \Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0$$

et  $G + \delta G$  l'ensemble des points  $M + \delta M$  ayant la même abscisse curviligne  $\tau$  et donc de coordonnées  $y^\mu(\tau) = x^\mu(\tau) + \delta x^\mu(\tau)$  qui vérifient l'équation

$$\frac{d^2 (y^\lambda)}{d\tau^2} + \Gamma_{\mu\nu}^\lambda (y^\rho) \frac{d(y^\mu)}{d\tau} \frac{d(y^\nu)}{d\tau} = 0 \tag{7.51}$$

la différence au premier ordre en  $\delta x^\mu$  entre ces deux équations décrivant des courbes voisines, fournit au prix d'un petit calcul

$$0 = \frac{d^2 \delta x^\mu}{d\tau^2} + (\partial_\rho \Gamma_{\lambda\nu}^\mu) \delta x^\rho \frac{dx^\lambda}{d\tau} \frac{dx^\nu}{d\tau} + 2\Gamma_{\lambda\nu}^\mu \frac{dx^\nu}{d\tau} \frac{d(\delta x^\lambda)}{d\tau} \tag{7.52}$$

ou bien en faisant intervenir des dérivées covariantes le long de la trajectoire

$$\frac{D^2(\delta x^\lambda)}{D\tau^2} = (\partial_\kappa \Gamma_{\mu\nu}^\lambda - \partial_\mu \Gamma_{\nu\kappa}^\lambda + \Gamma_{\eta\kappa}^\lambda \Gamma_{\mu\nu}^\eta - \Gamma_{\nu\eta}^\lambda \Gamma_{\mu\kappa}^\eta) \delta x^\mu \frac{dx^\kappa}{d\tau} \frac{dx^\nu}{d\tau} \quad (7.53)$$

dans cette relation, et sortant d'on ne sait où, apparaît une quantité a priori fort peu sympathique

$$-R^\lambda{}_{\nu\mu\kappa} := \partial_\kappa \Gamma_{\mu\nu}^\lambda - \partial_\mu \Gamma_{\nu\kappa}^\lambda + \Gamma_{\eta\kappa}^\lambda \Gamma_{\mu\nu}^\eta - \Gamma_{\nu\eta}^\lambda \Gamma_{\mu\kappa}^\eta . \quad (7.54)$$

Les plus courageux vérifieront que cet énergomène est la composante d'un tenseur de rang 4, et qu'il possède la propriété extraordinaire suivante

$$\begin{aligned} [D_\alpha, D_\beta] A_\mu &:= D_\alpha D_\beta A_\mu - D_\beta D_\alpha A_\mu := A_{\mu;\alpha;\beta} - A_{\mu;\beta;\alpha} \\ &= R^\lambda{}_{\mu\alpha\beta} A_\lambda \end{aligned} \quad (7.55)$$

cette propriété géométrique rend compte du fait que le transport parallèle le long d'une courbe fermée ne laisse pas invariant un vecteur si la courbe n'est pas plane ! On peut s'en convaincre en effectuant le transport parallèle d'un vecteur de Marseille à Pékin en passant par le pôle nord et en revenant à Marseille, comme sur la figure .

Cette caractéristique est bien connue des géomètres qui connaissent déjà l'énergomène en question : la courbure de Riemann-Christoffel.

L'équation de déviation géodésique s'écrit donc

$$\frac{D^2(\delta x^\lambda)}{D\tau^2} = -R^\lambda{}_{\nu\mu\kappa} \delta x^\mu \frac{dx^\kappa}{d\tau} \frac{dx^\nu}{d\tau} \quad (7.56)$$

le tenseur de courbure y apparaît comme un terme semblable à une force qui viendrait rappeler la particule sur la géodésique  $G$  lorsqu'elle s'écarte de la trajectoire naturelle de chute libre.

L'équivalence entre courbure de l'espacetemps et champ de gravitation est à présent complète !

## 7.6 Propriétés de la courbure de Riemann-Christoffel

Les propriétés du tenseur de courbure sont des éléments techniques et pratiques essentiels de la relativité générale.



Fig. 7.2. Transport parallèle d'un vecteur sur une surface courbe

### 7.6.1 Symétrie

La propriété de symétrie du tenseur métrique implique la symétrie des indices covariants de la connexion affine. Une utilisation astucieuse du théorème de Schwarz montre alors que le tenseur de Riemann-Christoffel présente l'antisymétrie suivante :

$$R^\lambda_{\nu\mu\kappa} = -R^\lambda_{\nu\kappa\mu} \quad (7.57)$$

cette propriété entraîne la relation de cyclicité suivante

$$R^\lambda_{\nu\mu\kappa} + R^\lambda_{\kappa\nu\mu} + R^\lambda_{\mu\kappa\nu} = 0 \quad (7.58)$$

### 7.6.2 Courbure complètement covariante

Il est bien souvent très utile de travailler avec la composante complètement covariante du tenseur de courbure. Elle s'obtient en faisant agir le tenseur métrique

$$R_{\mu\nu\alpha\beta} = g_{\sigma\mu} R^{\sigma}{}_{\nu\alpha\beta} \quad (7.59)$$

c'est-à-dire en remontant à la définition

$$\begin{aligned} -R_{\lambda\mu\nu\kappa} = & \frac{1}{2} [g_{\nu\lambda,\mu,\kappa} - g_{\mu\nu,\lambda,\kappa} - g_{\lambda\kappa,\nu,\mu} + g_{\mu\kappa,\nu,\lambda}] \\ & + g_{\eta\sigma} [\Gamma_{\lambda\nu}^{\eta} \Gamma_{\mu\kappa}^{\sigma} - \Gamma_{\lambda\kappa}^{\eta} \Gamma_{\mu\nu}^{\sigma}] \end{aligned} \quad (7.60)$$

Cette dernière relation implique la symétrie par paire d'indices

$$R_{\lambda\mu\nu\kappa} = R_{\nu\kappa\lambda\mu} \quad (7.61)$$

l'antisymétrie

$$R_{\lambda\mu\nu\kappa} = -R_{\mu\lambda\nu\kappa} = -R_{\lambda\mu\kappa\nu} = R_{\mu\lambda\kappa\nu} \quad (7.62)$$

la cyclicité

$$\begin{aligned} R_{\lambda\mu\nu\kappa} + R_{\lambda\kappa\mu\nu} + R_{\lambda\nu\kappa\mu} &= 0 \\ \iff \varepsilon^{\lambda\mu\nu\kappa} R_{\lambda\mu\nu\kappa} &= 0 \end{aligned} \quad (7.63)$$

Deux indices égaux annulent la composante du tenseur

$$R_{\lambda\lambda\nu\kappa} = R_{\nu\kappa\mu\mu} = 0 \quad (7.64)$$

Notons enfin la propriété qui se révélera fondamentale : les dérivées covariantes du Riemann-Christoffel permettent de former l'identité de Bianchi

$$R_{\mu\nu\alpha\beta;\lambda} + R_{\mu\nu\lambda\alpha;\beta} + R_{\mu\nu\beta\lambda;\alpha} = 0 \quad (7.65)$$

### 7.6.3 Contractions : Tenseur de Ricci et courbure scalaire

En contractant sur le premier et le troisième indice<sup>3</sup> de la composante complètement covariante du Riemann-Christoffel, on obtient un tenseur de rang 2

$$R_{\nu\beta} = g^{\mu\alpha} R_{\mu\nu\alpha\beta} \quad (7.66)$$

appelé tenseur de Ricci, que l'on peut aussi écrire directement à partir des connexions

$$-R_{\nu\beta} = \Gamma_{\nu\lambda,\beta}^{\lambda} - \Gamma_{\nu\beta,\lambda}^{\lambda} + \Gamma_{\nu\lambda}^{\sigma} \Gamma_{\beta\sigma}^{\lambda} - \Gamma_{\nu\beta}^{\lambda} \Gamma_{\lambda\sigma}^{\sigma} \quad (7.67)$$

<sup>3</sup> La symétrie du Riemann-Christoffel nous indique que l'on obtient le même résultat en contractant sur le deuxième et le quatrième indice.

Finalement en contractant tous les indices du tenseur de courbure de Riemann-Christoffel on obtient un 4-scalaire définissant la courbure scalaire au point considéré

$$\begin{aligned}
 R &= g^{\mu\nu} R_{\mu\nu} \\
 &= g^{\mu\nu} g^{\alpha\beta} R_{\mu\nu\alpha\beta} \\
 &= g^{\mu\nu} R^{\beta}_{\mu\beta\nu}
 \end{aligned} \tag{7.68}$$

A titre d'exercice de manipulation, on pourra montrer que la courbure scalaire d'une sphère de rayon  $r$  est une constante  $R = 2/r^2$ .

On rappelle que l'élément de longueur sur la sphère est donné par

$$ds^2 = r^2 (d\theta^2 + \sin^2 \theta d\varphi^2)$$

on a donc les composantes covariantes de la métrique, desquelles nous pouvons déduire les connexions, puis le Ricci, puis la courbure, essayez pour voir...

#### 7.6.4 Une remarque remarquable

L'identité de Bianchi (7.65) stipule que

$$R_{\lambda\mu\nu;\eta} + R_{\lambda\mu\eta;\nu} + R_{\lambda\mu\kappa;\nu} = 0 \tag{7.69}$$

Attendu que

$$g_{\mu\nu} \in \ker(D_\lambda) \tag{7.70}$$

la dérivée covariante du tenseur métrique est nulle, en contractant chacun des termes de l'identité de Bianchi par  $g^{\lambda\nu}$  il vient

- $g^{\lambda\nu} R_{\lambda\mu\nu;\eta} = (g^{\lambda\nu} R_{\lambda\mu\nu})_{;\eta} = R_{\mu\kappa;\eta}$
- $g^{\lambda\nu} R_{\lambda\mu\eta;\nu} = (g^{\lambda\nu} R_{\lambda\mu\eta})_{;\nu} = (-g^{\lambda\nu} R_{\lambda\mu\nu\eta})_{;\kappa} = -R_{\mu\eta;\kappa}$
- $g^{\lambda\nu} R_{\lambda\mu\kappa;\nu} = R^{\nu}_{\mu\kappa\eta;\nu}$

la relation (7.69) devient donc

$$R_{\mu\kappa;\eta} - R_{\mu\eta;\kappa} + R^{\nu}_{\mu\kappa\eta;\nu} = 0 \tag{7.71}$$

contractons à présent par  $g^{\mu\kappa}$  il vient

- $g^{\mu\kappa} R_{\mu\kappa;\eta} = (g^{\mu\kappa} R_{\mu\kappa})_{;\eta} = R_{;\eta}$
- $g^{\mu\kappa} R_{\mu\eta;\kappa} = (g^{\mu\kappa} R_{\mu\eta})_{;\kappa} = R^{\kappa}_{\eta;\kappa} (= R^{\nu}_{\eta;\nu})$

- $g^{\mu\kappa} R^\nu_{\mu\kappa\eta;\nu} = (g^{\mu\kappa} R^\nu_{\mu\kappa\eta})_{;\nu} = (-g^{\mu\kappa} R^\nu_{\mu\eta\kappa})_{;\nu} = -R^\nu_{\eta;\nu}$

la relation (7.71) devient donc

$$R_{;\eta} - 2R^\kappa_{\eta;\kappa} = 0 \tag{7.72}$$

il suffit à présent de réorganiser la dérivée covariante du scalaire de courbure et de remonter tous les indices, i.e.

$$\begin{aligned} R_{;\eta} - 2R^\kappa_{\eta;\kappa} = 0 &\Leftrightarrow \delta_\eta^\kappa R_{;\kappa} - 2R^\kappa_{\eta;\kappa} = 0 \\ \Leftrightarrow (R^\kappa_\eta - \frac{1}{2}\delta_\eta^\kappa R)_{;\kappa} = 0 &\Leftrightarrow g^{\eta\nu} (R^\kappa_\eta - \frac{1}{2}\delta_\eta^\kappa R)_{;\kappa} = 0 \\ \Leftrightarrow [g^{\eta\nu} (R^\kappa_\eta - \frac{1}{2}\delta_\eta^\kappa R)]_{;\kappa} = 0 &\Leftrightarrow [R^{\kappa\nu} - \frac{1}{2}g^{\kappa\nu} R]_{;\kappa} = 0 \end{aligned} \tag{7.73}$$

en introduisant le tenseur de rang 2 symétrique

$$G^{\mu\nu} := R^{\mu\nu} - \frac{1}{2}g^{\mu\nu} R \tag{7.74}$$

appelé tenseur d'Einstein, notre double contraction de l'identité de Bianchi montre donc que

$$G^{\mu\nu} \in \ker(D_\mu) \tag{7.75}$$

la dérivée covariante du tenseur d'Einstein est nulle.

E. Cartan montrera dans les années 1930 qu'outre le tenseur métrique, le tenseur d'Einstein est le seul tenseur d'ordre 2 symétrique construit uniquement à partir du tenseur métrique et de ses dérivées premières et secondes dont la divergence covariante soit nulle !

Le noyau de la divergence covariante était donc cerné !

Pour notre part, nous sommes enfin prêts pour dériver efficacement les équations générales du champ de gravitation : les équations d'Einstein.

## Equations d'Einstein

---

### 8.1 Remarques préliminaires

Nous avons vu comment les équations de l'électromagnétisme pouvaient être issues d'un principe de moindre action. Cette action faisait intervenir une densité de lagrangien  $\mathcal{L}$  scalaire (pour être naturellement covariant) et contenant en elle les sources du mouvement (i.e. les interactions) ou des champs en présence :  $\mathcal{L}_e = \mathcal{L}_l + \mathcal{L}_i + \mathcal{L}_c$ .

La nouveauté en relativité générale réside dans le fait que le champ de gravitation est identifié – via le principe d'équivalence – à la courbure de l'espacetemps.

Toute présence de matière, plus généralement d'énergie ou d'impulsion, est inévitablement associée en relativité restreinte à un équivalent massique. En plus de la densité de lagrangien  $\mathcal{L}_m$  associée à cette énergie impulsion, il semble donc nécessaire en relativité générale d'introduire une densité de lagrangien  $\mathcal{L}_g$  décrivant la courbure induite (ou pas) par la présence de la matière.

L'action totale s'écrit donc

$$S = \int (\mathcal{L}_m + \mathcal{L}_g) d\mathbf{x} = S_m + S_g \quad (8.1)$$

L'action  $S_m$  décrit la nature des forces et de la matière en présence, l'action  $S_g$  décrit quant à elle l'état de courbure de l'espacetemps correspondant.

Comme nous l'avons dit à de multiples reprises le terme sous l'intégrale "doit" être un scalaire : il s'agit en effet de la plus simple des façons pour assurer la covariance générale des équations<sup>1</sup> que nous déduirons de la variation de l'action sous un changement de coordonnées.

<sup>1</sup> On entend ici par covariance générale, l'invariance de la forme des équations dans des change-

Lors d'une transformation quelconque (et plus seulement une transformation linéaire et unitaire comme l'était la transformée de Lorentz de la relativité restreinte), l'élément de volume  $d^4x$  n'est plus covariant. En relativité générale, cet élément d'intégration doit prendre en compte la variation de volume spatio-temporel, il doit donc faire intervenir un déterminant !

En effet, lors d'une transformation  $nvec{x} \rightarrow nvec{x}'$  la métrique devient

$$g'_{\mu\nu} = \mathcal{A}_\mu{}^\rho \mathcal{A}_\nu{}^\sigma g_{\rho\sigma} = \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma} \quad (8.2)$$

en prenant le déterminant de chacun des deux membres de cette équation, il vient

$$g' = \left| \frac{\partial \mathbf{x}}{\partial \mathbf{x}'} \right|^2 g \quad (8.3)$$

où nous avons posé,  $g = \det g_{\mu\nu}$ ,  $g' = \det g'_{\mu\nu}$  et  $|\partial \mathbf{x} / \partial \mathbf{x}'|$  le jacobien de la transformation.

Muni de ces notations, il est facile de voir que  $d\mathbf{x}$  n'est pas covariant. En effet par définition

$$d\mathbf{x} = \left| \frac{\partial \mathbf{x}}{\partial \mathbf{x}'} \right| d\mathbf{x}' \quad (8.4)$$

la relation (8.3) peut donc toujours s'écrire

$$\left| \frac{\partial \mathbf{x}}{\partial \mathbf{x}'} \right| = \sqrt{\frac{|g'|}{|g|}} = \sqrt{\frac{-g'}{-g}} \quad (8.5)$$

le déterminant de la métrique est bien négatif : dans le référentiel inertiel local il est en effet le produit de trois plus et d'un moins !

Il est alors facile de remarquer que

$$\sqrt{-g} d\mathbf{x} = \sqrt{-g} \left| \frac{\partial \mathbf{x}}{\partial \mathbf{x}'} \right| d\mathbf{x}' = \sqrt{-g} \sqrt{\frac{-g'}{-g}} d\mathbf{x}' = \sqrt{-g'} d\mathbf{x}' \quad (8.6)$$

est un élément de volume covariant. C'est celui que nous choisirons pour nos intégrations spatio-temporelles en relativité générale.

## 8.2 L'action de courbure et sa variation

Eu égard aux quelques remarques préliminaires du paragraphe précédent, l'action de courbure peut s'écrire

$$S_g = -\frac{1}{2\chi} \int \mathcal{L}_g \sqrt{-g} d\mathbf{x} \quad (8.7)$$

la constante  $\chi$  nous permettra de fixer les unités de nos nouvelles équations<sup>2</sup>, la densité  $\mathcal{L}_g$  doit être un 4-scalaire décrivant l'état de courbure de l'espace-temps. L'intuition<sup>3</sup> immédiate est alors d'identifier le lagrangien de courbure et le scalaire de la même chose. Essayons donc

$$\mathcal{L}_g = R = g^{\mu\nu} R_{\mu\nu} = g^{\mu\nu} g^{\alpha\beta} R_{\alpha\mu\beta\nu} = g^{\mu\nu} R^\beta_{\mu\beta\nu} \quad (8.8)$$

ainsi

$$S_g = -\frac{1}{2\chi} \int g^{\mu\nu} R_{\mu\nu} \sqrt{-g} d\mathbf{x} \quad (8.9)$$

qu'il ne reste plus qu'à faire varier. La variable de notre problème est la métrique (le champ de gravitation), la variation s'écrit donc

$$\delta S_g = -\frac{1}{2\chi} \int \left[ R_{\mu\nu} \sqrt{-g} + g^{\mu\nu} R_{\mu\nu} \frac{\delta(\sqrt{-g})}{\delta g^{\mu\nu}} + g^{\mu\nu} \sqrt{-g} \frac{\delta R_{\mu\nu}}{\delta g^{\mu\nu}} \right] \delta g^{\mu\nu} d\mathbf{x} \quad (8.10)$$

Explicitons chacun des deux derniers termes :

1. Et soudain surgit, la différentielle d'un déterminant ...

La variation  $\delta(\sqrt{-g})$  est la variation d'une application composée, on a donc

$$\delta(\sqrt{-g}) = \delta\left([-\det g_{\mu\nu}]^{\frac{1}{2}}\right) = \frac{1}{2} \frac{\delta[-\det g_{\mu\nu}]}{\sqrt{-g}} \quad (8.11)$$

on se souvient de la différentielle du déterminant<sup>4</sup>

ments de repères quelconques.

<sup>2</sup> le 2 et le signe  $-$  sont purement conventionnels et permettent des simplifications futures.

<sup>3</sup> Cette intuition prend le statut de seule issue possible depuis que Cartan et Weyl (voir l'appendice II de [24]) ont montré que le seul 4-scalaire que l'on peut former avec la métrique et ses dérivées premières et secondes (i.e. les connexions) est précisément le scalaire de courbure.

<sup>4</sup> L'application déterminant de l'ensemble des matrices carrées d'ordre  $n$  dans  $\mathbb{C}$ , est évidemment continue et dérivable car elle est polynomiale. Sa différentielle en  $A$  au point  $H$  est par définition

$$D(\det(A))(H) = \lim_{t \rightarrow 0} \frac{\det(A + tH) - \det(A)}{t} \quad (8.12)$$

Calculons dans un premier temps, la différentielle du déterminant en l'identité  $I$ .

On sait que

$$\det(H - tI) = (-1)^n (t^n - \text{tr}(H)t^{n-1} + O(t^{n-2})) \quad (8.13)$$

ainsi

$$\det(I + tH) = 1 + \text{tr}(H)t + O(t^2) \quad (8.14)$$

on en déduit que

$$\delta[-\det A] = -D(\det A)(\delta A) = \text{tr}(-\tilde{A}^\top \delta A) \quad (8.18)$$

si  $A$  est inversible, la transposée de la comatrice  $\tilde{A}^\top$  se simplifie en  $\det A A^{-1}$ , dans ce cas nous avons donc

$$\delta[-\det A] = \det A \text{tr}(-A^{-1} \delta A).$$

Il ne reste plus qu'à choisir  $A = g_{\mu\nu}$ , soit  $A^{-1} = g^{\mu\nu}$  et l'on obtient

$$\begin{aligned} \delta(\sqrt{-g}) &= \frac{1}{2\sqrt{-g}} \text{tr} [(-g) g^{\mu\nu} \delta g_{\nu\rho}] \\ &= \frac{\sqrt{-g}}{2} g^{\mu\nu} \delta g_{\nu\mu} \end{aligned} \quad (8.19)$$

un dernier petit tour de passe-passe indiciel pour constater que

$$\begin{aligned} g^{\mu\nu} g_{\nu\mu} = \text{tr}(\delta_\nu^\mu) = 4 &\Rightarrow \delta(g^{\mu\nu} g_{\nu\mu}) = g_{\nu\mu} \delta g^{\mu\nu} + g^{\mu\nu} \delta g_{\nu\mu} = 0 \\ &\Rightarrow g^{\mu\nu} \delta g_{\nu\mu} = -g_{\mu\nu} \delta g^{\mu\nu} \end{aligned} \quad (8.20)$$

et finalement

$$\delta(\sqrt{-g}) = -\frac{\sqrt{-g}}{2} g_{\mu\nu} \delta g^{\mu\nu} \quad (8.21)$$

2.  $g^{\mu\nu} \delta R_{\mu\nu}$  nécessite plus d'attention.

Dans le but de vérifier la perspicacité de notre lecteur (s'il en reste !), nous lui laissons démontrer les deux «petits» résultats intermédiaires suivants :

- Une mise en allure :

$$\delta R_{\mu\nu} = (\delta \Gamma_{\mu\nu}^\lambda)_{;\lambda} - (\delta \Gamma_{\mu\lambda}^\lambda)_{;\nu} \quad (8.22)$$

---


$$D(\det(I))(H) = \lim_{t \rightarrow 0} \frac{\det(I + tH) - \det(I)}{t} = \text{tr}(H) \quad (8.15)$$

Pour toute matrice inversible  $B$ , il suffit ensuite d'écrire  $\det(B + tH) = \det(B) \det(I + tB^{-1}H)$ , pour conclure que

$$D(\det(B))(H) = \det(B) \text{tr}(B^{-1}H) \quad (8.16)$$

en écrivant que  $\det(B) B^{-1} = \tilde{B}^\top$  transposée de la comatrice de  $B$ , on peut conclure «par densité» : l'application  $\det$  est continue, l'ensemble des matrices dont le déterminant est non nul est dense, ainsi pour toute matrice  $A$

$$D(\det(A))(H) = \text{tr}(\tilde{A}^\top H) \quad (8.17)$$

où l'on a lâchement introduit le tenseur<sup>5</sup> de rang 3

$$\delta I_{\mu\nu}^{\lambda} := \frac{1}{2} g^{\lambda\rho} \left[ (\delta g_{\rho\mu})_{;\nu} + (\delta g_{\rho\nu})_{;\mu} - (\delta g_{\mu\nu})_{;\rho} \right] \quad (8.23)$$

- Un résultat général : pour tout vecteur la divergence covariante est donnée par

$$V^{\mu}_{;\mu} = \frac{1}{\sqrt{-g}} (\sqrt{-g} V^{\mu})_{;\mu} := \frac{1}{\sqrt{-g}} \frac{\partial (\sqrt{-g} V^{\mu})}{\partial x^{\mu}} \quad (8.24)$$

Ces deux résultats s'associent fort bien : utilisant (8.22) et encore une fois le fait que la dérivée covariante du tenseur métrique est nulle, il vient

$$g^{\mu\nu} \delta R_{\mu\nu} = (g^{\mu\nu} \delta I_{\mu\nu}^{\lambda})_{;\lambda} - (g^{\mu\nu} \delta I_{\mu\lambda}^{\lambda})_{;\nu} \quad (8.25)$$

chacun de ces deux termes est une divergence covariante. En effet, compte tenu des diverses positions indicielles on peut toujours poser

$$a^{\lambda} := g^{\mu\nu} \delta I_{\mu\nu}^{\lambda} \quad \text{et} \quad b^{\nu} := g^{\mu\nu} \delta I_{\mu\lambda}^{\lambda}$$

ainsi

$$g^{\mu\nu} \delta R_{\mu\nu} = (a^{\lambda})_{;\lambda} - (b^{\nu})_{;\nu}$$

et en utilisant (8.24) il vient

$$\begin{aligned} \int g^{\mu\nu} \sqrt{-g} \delta R_{\mu\nu} d\mathbf{x} &= \int \frac{\partial (\sqrt{-g} a^{\nu})}{\partial x^{\nu}} d\mathbf{x} - \int \frac{\partial (\sqrt{-g} b^{\lambda})}{\partial x^{\lambda}} d\mathbf{x} \\ &= [\sqrt{-g} a^{\nu}]_{Bord} - [\sqrt{-g} b^{\lambda}]_{Bord} \\ &= [\sqrt{-g} g^{\mu\nu} \delta I_{\mu\nu}^{\lambda}]_{Bord} - [\sqrt{-g} g^{\mu\nu} \delta I_{\mu\lambda}^{\lambda}]_{Bord} \end{aligned} \quad (8.26)$$

Sur le bord du domaine d'intégration (c'est-à-dire loin de la source de gravitation), l'espace est plat. Ainsi  $[g^{\mu\nu}]_{Bord} = [\eta^{\mu\nu}]$  et par conséquent  $[\delta g^{\mu\nu}]_{Bord} = [\delta \eta^{\mu\nu}] = 0$ . En observant (cf. (8.23)) dans le blanc des yeux, on constate alors avec bonheur que

$$[\sqrt{-g} g^{\mu\nu} \delta I_{\mu\nu}^{\lambda}]_{Bord} = 0 \quad \text{et} \quad [\sqrt{-g} g^{\mu\nu} \delta I_{\mu\lambda}^{\lambda}]_{Bord} = 0$$

Exit le terme en variation du Ricci !

<sup>5</sup> Mais oui c'en est un ! Il est facile de le voir sur la définition ! On constate donc fort à propos

Nous pouvons à présent revenir à la variation de notre action de courbure. Nous l'avons laissée dans l'état (8.10), en incorporant les deux résultats précédents il vient

$$\begin{aligned}\delta S_g &= -\frac{1}{2\chi} \int \left[ R_{\mu\nu} \sqrt{-g} - \frac{1}{2} g^{\mu\nu} R_{\mu\nu} g_{\mu\nu} \sqrt{-g} + 0 \right] \delta g^{\mu\nu} d\mathbf{x} \\ &= -\frac{1}{2\chi} \int \left[ R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right] \sqrt{-g} \delta g^{\mu\nu} d\mathbf{x}\end{aligned}\tag{8.27}$$

C'est la jubilation : qui ne reconnaît pas dans cette relation le fameux tenseur d'Einstein

$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}\tag{8.28}$$

dont nous avons signalé (cf. 7.6.4) l'appartenance au noyau de la dérivation covariante.

## 8.3 L'action de matière et sa variation

### 8.3.1 Le tenseur énergie-impulsion

Conformément aux remarques préliminaires, l'action décrivant la matière et ses interactions est de la forme

$$S_m = \int \mathcal{L}_m \sqrt{-g} d\mathbf{x}\tag{8.29}$$

Nous avons traité en détail le cas du champ électromagnétique dont l'action s'écrit

$$S_e = -mc \int ds - \int A_\mu J^\mu dV dt - \frac{1}{4\mu_0} \int F^{\mu\nu} F_{\mu\nu} dV dt\tag{8.30}$$

Dans le cas général  $\mathcal{L}_m$  décrit formellement le contenu matériel de la région dont la nature géométrique (courbure) est décrite par la densité de lagrangien  $\mathcal{L}_g$ . La variation correspondante s'écrit tout aussi formellement

$$\delta S_m = \int \left[ \sqrt{-g} \delta(\mathcal{L}_m) + \mathcal{L}_m \delta(\sqrt{-g}) \right] d\mathbf{x}\tag{8.31}$$

---

que  $\delta\Gamma_{\mu\nu}^\lambda$  est un tenseur alors que  $\Gamma_{\mu\nu}^\lambda$  n'en est pas un, tout ceci est bien inquiétant !

notre petite escapade chez la différentielle du déterminant et l'explicitation de la variation du lagrangien donnent alors

$$\begin{aligned}\delta S_m &= \int \left[ \sqrt{-g} \frac{\partial \mathcal{L}_m}{\partial g^{\mu\nu}} \delta g^{\mu\nu} - \mathcal{L}_m \frac{\sqrt{-g}}{2} g_{\mu\nu} \delta g^{\mu\nu} \right] d\mathbf{x} \\ &= \int \left[ \frac{\partial \mathcal{L}_m}{\partial g^{\mu\nu}} - \frac{1}{2} g_{\mu\nu} \mathcal{L}_m \right] \sqrt{-g} \delta g^{\mu\nu} d\mathbf{x}\end{aligned}\quad (8.32)$$

En introduisant le tenseur

$$T_{\mu\nu} = 2 \left[ \frac{\partial \mathcal{L}_m}{\partial g^{\mu\nu}} - \frac{1}{2} g_{\mu\nu} \mathcal{L}_m \right] \quad (8.33)$$

appelé tenseur énergie-impulsion, la variation de l'action décrivant la matière s'écrit donc

$$\delta S_m = \int \frac{1}{2} T_{\mu\nu} \sqrt{-g} \delta g^{\mu\nu} d\mathbf{x} \quad (8.34)$$

Par définition nous avons donc :

**Définition 8.1.** *Le tenseur énergie-impulsion  $T^{\mu\nu}$  (resp.  $T_{\mu\nu}$ ) est la dérivée fonctionnelle (à un facteur 2 près) de l'action de matière  $S_m$  par rapport à  $g_{\mu\nu}$  (resp.  $g^{\mu\nu}$ )*

$$\text{Si } g_{\mu\nu} \longrightarrow g_{\mu\nu} + \delta g_{\mu\nu} \quad \text{alors } S \longrightarrow S + \delta S \quad (8.35)$$

La fonctionnelle  $\delta S$  est linéaire en  $\delta g_{\mu\nu}$  et l'on a

$$\delta S = \delta \int \mathcal{L}_m \sqrt{-g} d\mathbf{x} = \int \frac{\delta S_m}{\delta g_{\mu\nu}} \delta g_{\mu\nu} \sqrt{-g} d\mathbf{x} \quad (8.36)$$

ainsi

$$\frac{1}{2} T^{\mu\nu} := \frac{\delta S_m}{\delta g^{\mu\nu}} = \frac{\partial \mathcal{L}_m}{\partial g^{\mu\nu}} - \frac{1}{2} g_{\mu\nu} \mathcal{L}_m \quad (8.37)$$

ou bien

$$\frac{1}{2} T^{\mu\nu} := \frac{\delta S_m}{\delta g_{\mu\nu}} = \frac{\partial \mathcal{L}_m}{\partial g_{\mu\nu}} - \frac{1}{2} g^{\mu\nu} \mathcal{L}_m \quad (8.38)$$

Le facteur 1/2 est conventionnel.

Nous nous devons à cet instant d'informer notre lecteur d'un flou artistique régnant dans la littérature : il est clair que la définition du tenseur énergie-impulsion le considère comme une dérivée fonctionnelle de l'action

par rapport à la métrique. Il en résulte une certaine liberté de jauge : si l'on rajoute une divergence covariante<sup>6</sup> à l'action, celle-ci n'est pas changée, mais, attendu que la dérivée fonctionnelle de ce terme supplémentaire n'est pas forcément nulle, le tenseur énergie-impulsion s'en voit modifié d'autant (voir exercice).

C'est ainsi que chez certains auteurs, l'on écrit l'action sous la forme (8.29) habituelle, on dit alors que

$$\delta S_m = \delta \int \mathcal{L}_m \sqrt{-g} d\mathbf{x} = \int \delta (\sqrt{-g} \mathcal{L}_m) d\mathbf{x} \quad (8.39)$$

on invoque alors le fait que le terme  $f = \sqrt{-g} \mathcal{L}_m$  est une fonction de la métrique  $g_{\mu\nu}$  et de sa dérivée (standard)  $g_{\mu\nu,\rho}$  pour écrire la variation sous la forme «traditionnelle» ou canonique

$$\delta S_m = \int \delta [f(g_{\mu\nu}, g_{\mu\nu,\rho})] d\mathbf{x} = \int \left[ \frac{\partial f}{\partial g_{\mu\nu}} \delta g_{\mu\nu} + \frac{\partial f}{\partial g_{\mu\nu,\rho}} \delta g_{\mu\nu,\rho} \right] d\mathbf{x} \quad (8.40)$$

qu'une désormais célèbre intégration par parties transforme en

$$\delta S_m = \int \left[ \frac{\partial f}{\partial g_{\mu\nu}} - \partial_\rho \left( \frac{\partial f}{\partial g_{\mu\nu,\rho}} \right) \right] \delta g_{\mu\nu} d\mathbf{x} \quad (8.41)$$

La très authentique définition du tenseur énergie-impulsion

$$\delta \int \mathcal{L}_m \sqrt{-g} d\mathbf{x} := \frac{1}{2} \int T^{\mu\nu} \delta g_{\mu\nu} \sqrt{-g} d\mathbf{x} \quad (8.42)$$

s'empresse alors de donner

$$T^{\mu\nu} = \frac{2}{\sqrt{-g}} \left\{ \frac{\partial (\sqrt{-g} \mathcal{L}_m)}{\partial g_{\mu\nu}} - \partial_\rho \left[ \frac{\partial (\sqrt{-g} \mathcal{L}_m)}{\partial g_{\mu\nu,\rho}} \right] \right\} \quad (8.43)$$

cette définition est tout aussi correcte que celle que nous avons adoptée (voir définition 8.1, eq. (8.37) ou (8.38)). Il est toutefois de notre devoir de signaler que dans certains cas des termes de divergence issus de la liberté de jauge peuvent apparaître selon que l'on utilise l'une ou l'autre définition.

<sup>6</sup> d'un tenseur qui s'annule sur le bord du système bien entendu !

### 8.3.2 Un petit exemple...

Considérons à titre d'exemple le cas d'une région de l'espace contenant un champ électromagnétique (pas de particules, pas d'interactions).

Les résultats de nos investigations du chapitre (5.64), nous indiquent que le lagrangien s'écrit

$$\mathcal{L}_e = -\frac{1}{4\mu_0} F_{\alpha\beta} F^{\alpha\beta} \quad (8.44)$$

en utilisant la définition 8.1, le tenseur énergie-impulsion correspondant au champ électromagnétique s'écrit donc

$$T_{\mu\nu} = 2 \left[ \frac{\partial \mathcal{L}_e}{\partial g^{\mu\nu}} - \frac{1}{2} g_{\mu\nu} \mathcal{L}_e \right] \quad (8.45)$$

soit

$$T_{\mu\nu} = \left[ -\frac{1}{2\mu_0} \frac{\partial (F_{\alpha\beta} F^{\alpha\beta})}{\partial g^{\mu\nu}} + \frac{1}{4\mu_0} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \right] \quad (8.46)$$

Attendu que  $g^{\mu\nu}$  est ici bien caché dans  $F_{\alpha\beta} F^{\alpha\beta}$ , il faut faire preuve d'espièglerie :

$$\begin{aligned} \frac{\partial (F_{\alpha\beta} F^{\alpha\beta})}{\partial g^{\mu\nu}} &= \frac{\partial (F_{\alpha\beta} F_{\rho\sigma} g^{\rho\alpha} g^{\sigma\beta})}{\partial g^{\mu\nu}} \\ &= F_{\alpha\beta} F_{\rho\sigma} \left[ g^{\sigma\beta} \frac{\partial (g^{\rho\alpha})}{\partial g^{\mu\nu}} + g^{\rho\alpha} \frac{\partial (g^{\sigma\beta})}{\partial g^{\mu\nu}} \right] \end{aligned} \quad (8.47)$$

$$= F_{\alpha\beta} F_{\rho\sigma} \left[ g^{\sigma\beta} \frac{\partial (\delta_\mu^\rho \delta_\nu^\alpha g^{\mu\nu})}{\partial g^{\mu\nu}} + g^{\rho\alpha} \frac{\partial (\delta_\mu^\sigma \delta_\nu^\beta g^{\mu\nu})}{\partial g^{\mu\nu}} \right]$$

soit

$$\frac{\partial (F_{\alpha\beta} F^{\alpha\beta})}{\partial g^{\mu\nu}} = F_{\alpha\beta} F_{\rho\sigma} [\delta_\mu^\rho \delta_\nu^\alpha g^{\sigma\beta} + \delta_\mu^\sigma \delta_\nu^\beta g^{\rho\alpha}] \quad (8.48)$$

et finalement

$$\frac{\partial (F_{\alpha\beta} F^{\alpha\beta})}{\partial g^{\mu\nu}} = 2F_{\mu\sigma} F_\nu^\sigma \quad (8.49)$$

le tenseur énergie-impulsion du champ électromagnétique s'écrit donc

$$T_{\mu\nu} = -\frac{1}{\mu_0} \left[ F_{\mu\alpha} F_\nu^\alpha - \frac{1}{4} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \right] \quad (8.50)$$

sa symétrie provient de sa définition, mais on peut toujours la vérifier sur la formulation (8.50), c'est bien plus drôle !

## 8.4 Equations du champ gravitationnel

### 8.4.1 Courbure = matière

Nous avons écrit l'action totale d'un système sous la forme

$$S = \int (\mathcal{L}_m + \mathcal{L}_g) d\mathbf{x} = S_m + S_g \quad (8.51)$$

où le terme

$$S_m = \int \mathcal{L}_m \sqrt{-g} d\mathbf{x} \quad (8.52)$$

décrit l'état de la matière en présence via sa densité de lagrangien  $\mathcal{L}_m$  (ou son tenseur énergie-impulsion), et le terme

$$S_g = -\frac{1}{2\chi} \int g^{\mu\nu} R_{\mu\nu} \sqrt{-g} d\mathbf{x} \quad (8.53)$$

décrit l'état de courbure de l'espacetemps associé à la distribution de matière.

Les équations du champ de gravitation s'obtiennent simplement par variation de cette action. Les calculs explicitement effectués dans les deux sections précédentes permettent donc d'obtenir (cf. (8.27) et (8.34))

$$\begin{aligned} \delta S &= \delta S_g + \delta S_m \\ &= -\frac{1}{2\chi} \int \left[ R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right] \sqrt{-g} \delta g^{\mu\nu} d\mathbf{x} + \int \frac{1}{2} T_{\mu\nu} \sqrt{-g} \delta g^{\mu\nu} d\mathbf{x} \end{aligned} \quad (8.54)$$

l'extremum de l'action pour des variations arbitraires de la métrique est donc obtenu lorsque

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \chi T_{\mu\nu} \quad (8.55)$$

ces équations sont appelées équations d'Einstein du champ de gravitation. Plusieurs remarques s'imposent !

- La constante  $\chi$  n'est pas fixée par notre théorie, pour l'obtenir il faut encore travailler et imposer que lorsque le champ de gravitation et les vitesses deviennent relativement faibles, les équations d'Einstein doivent redonner la mécanique classique. Ceci fait l'objet du calcul explicite de la section suivante.

- Les équations d'Einstein (8.55) établissent la proportionnalité entre le tenseur énergie-impulsion  $T$  et le tenseur d'Einstein  $G$ . Nous avons établi (cf. (7.75)) le fait que la dérivée covariante de ce dernier était nulle. Les équations d'Einstein stipulent donc aussi que

$$D_\mu T^{\mu\nu} = 0 \text{ ou plus directement } T^{\mu\nu} \in \ker(D_\mu) \quad (8.56)$$

Cette dernière relation est une simple équation de conservation de l'énergie-impulsion.

- La remarque précédente en appelle une autre : lors de nos investigations covariantes, nous avons aussi remarqué que le tenseur métrique était un autre valeureux membre du noyau de la dérivation covariante. Il est donc possible de rajouter aux équations d'Einstein traditionnelles un terme proportionnel à la métrique. On obtient ainsi

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \chi T_{\mu\nu}$$

où  $\Lambda$  est appelée constante cosmologique. Nous aurions tout aussi bien pu présenter ce terme comme entrant dans la liberté de jauge de l'action. Rien ne l'impose a priori, toutefois de récentes observations semblent lui faire une place importante.

- La symétrie et la divergence covariante nulle laissent indépendantes seulement 10 des 16 composantes des tenseurs en jeu dans les équations d'Einstein. Ces dernières sont donc un système de 10 équations différentielles non linéaires couplées. Il est généralement impossible de les résoudre directement. Pour résoudre des problèmes de relativité générale, on se place souvent dans un contexte symétrique qui permet d'écrire la métrique comme une fonction de quelques variables seulement (une ou deux en général), la distribution de matière est très simple (le plus souvent un gaz parfait barotropique dont l'équation d'état est un simple lien de proportionnalité entre densité et pression). Malgré toutes ces hypothèses simplificatrices, les résultats obtenus sont souvent très précis et permettent de rendre compte de phénomènes très fins comme l'avance du périhélie de planètes internes (comme Mercure), ou bien la variation de période du pulsar binaire...
- La déduction des équations (8.55) présentée ici est due à David Hilbert aux environs de 1915, une déduction originale et indépendante plus physique est due à Albert Einstein à la même époque. Cette déduction provient d'une analyse détaillée du passage de l'électrostatique à l'électrodynamique. On pourra consulter les divers ouvrages de la bibliographie ou directement les articles du maître pour partager ses intuitions.

### 8.4.2 Choix de la constante

Pour trouver la constante  $\chi$ , il convient que la théorie du champ de gravitation dont nous venons de construire les équations, redonne dans la limite des champs faibles les équations correspondantes de la mécanique classique. Ce type de limite est clairement obtenu en faisant trois hypothèses :

- Hypothèse 1 : les vitesses sont faibles devant  $c$ .  
Ainsi pour  $k = 1, 2, 3$  nous devons avoir

$$\left| \frac{dx^k}{dt} \right| \ll c \quad (8.57)$$

et comme  $x^0 = ct$ , on a donc

$$\left| \frac{dx^k}{ds} \right| \ll \left| \frac{dx^0}{ds} \right| \quad (8.58)$$

- Hypothèse 2 : la métrique est quasi-minkowskienne.  
Le champ de gravitation étant faible, la métrique de l'espace-temps peut donc s'écrire comme une métrique de Minkowski perturbée

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \quad (8.59)$$

avec

$$|h_{\alpha\beta}| \ll 1 \quad \text{et} \quad |h_{\alpha\beta}| \ll |\eta_{\alpha\beta}| \quad (8.60)$$

- Hypothèse 3 : le champ est stationnaire.  
Le champ gravitationnel étant à l'origine des variations temporelles des coefficients de la métrique (en relativité restreinte ces derniers sont constants), un champ faible aura une influence faible sur la métrique, on imposera notamment que les variations temporelles de la métrique perturbative sont nulles, i.e.

$$\partial_0 h_{\alpha\beta} = 0 \quad (8.61)$$

Sous l'hypothèse 1 et à l'ordre le plus bas, l'équation des géodésiques

$$\frac{d^2 x^\mu}{ds^2} + \Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} = 0 \quad (8.62)$$

devient

$$\frac{d^2 x^\mu}{ds^2} + \Gamma_{00}^\mu \left( \frac{dx^0}{ds} \right)^2 \simeq 0 \quad (8.63)$$

mais nous savons (cf. (7.15)) que

$$\Gamma_{\alpha\beta}^{\mu} = \frac{1}{2}g^{\mu\rho} (\partial_{\alpha}g_{\beta\rho} + \partial_{\beta}g_{\alpha\rho} - \partial_{\rho}g_{\alpha\beta}) \quad (8.64)$$

ainsi sous les hypothèses 2 et 3, il vient

$$\begin{aligned} \Gamma_{00}^{\mu} &= -\frac{1}{2}g^{\mu\rho}\partial_{\rho}g_{00} \\ &= -\frac{1}{2}(\eta^{\mu\rho} + h^{\mu\rho})\partial_{\rho}(\eta_{00} + h_{00}) \\ &\simeq -\frac{1}{2}\eta^{\mu\rho}\partial_{\rho}h_{00} \end{aligned} \quad (8.65)$$

Nos hypothèses simplifient donc l'équation des géodésiques en

$$\frac{d^2x^{\mu}}{ds^2} = \frac{1}{2}\eta^{\mu\rho}\partial_{\rho}h_{00} \left(\frac{dx^0}{ds}\right)^2 \quad (8.66)$$

La stationnarité de la perturbation de métrique annule la composante temporelle de cette dernière équation

$$\frac{d^2x^0}{ds^2} = \frac{1}{2}\eta^{0\rho}\partial_{\rho}h_{00} \left(\frac{dx^0}{ds}\right)^2 = 0 \quad (8.67)$$

ainsi  $dx^0/ds = cste$ . Nous aurions pu nous en douter car  $dx^0 = cdt$  et  $ds = cdt(1 - v^2/c^2)^{1/2}$ , et donc sous nos hypothèses

$$\frac{dx^0}{ds} \simeq 1. \quad (8.68)$$

Les composantes spatiales de (8.66) sont plus intéressantes : notre convention veut que  $\eta^{\alpha\beta} = (+, - - -)$ ; en incorporant le résultat (8.68), nous avons pour  $k = 1, 2, 3$

$$\frac{d^2x^k}{ds^2} = -\frac{1}{2}\vec{\nabla}h_{00}\cdot\vec{e}_k \quad (8.69)$$

comme ici  $ds^2 \approx c^2dt^2$ , la partie spatiale de l'équation des géodésiques s'écrit donc finalement

$$\frac{d^2x^k}{dt^2} = -\frac{c^2}{2}\vec{\nabla}h_{00}\cdot\vec{e}_k \quad (8.70)$$

Si seules des forces gravitationnelles dérivant d'un potentiel  $\psi$  sont présentes, le principe fondamental de la dynamique nous dit qu'en mécanique classique

$$k = 1, 2, 3 \quad \frac{d^2 x^k}{dt^2} = -\vec{\nabla}\psi \cdot \vec{e}_k \quad (8.71)$$

par identification nous avons donc

$$h_{00} = cste + \frac{2\psi}{c^2} \quad (8.72)$$

d'où

$$g_{00} = \eta_{00} + \frac{2\psi}{c^2} + cste . \quad (8.73)$$

Loin de la faible source de gravitation, les effets du champ sur la métrique s'estompent complètement. A l'infini, dans notre convention, nous devons avoir  $g_{00} = \eta_{00} = 1$ . Dans ces régions éloignées, le potentiel de gravitation doit s'annuler et nous obtenons ainsi le lien entre la métrique et le potentiel de gravitation dans la limite minkowskienne :

$$g_{00} = 1 + \frac{2\psi}{c^2}$$

Sous nos hypothèses de stationnarité, il est facile de voir que les autres composantes de la métrique sont égales à celles de la relativité restreinte, i.e.  $g_{ik} = \eta_{ik}$ .

Connaissant les composantes de la métrique, nous pouvons calculer les connexions affines, puis les composantes du tenseur de Riemann puis celles du Ricci et en déduire le scalaire de courbure...

Un calcul un peu fastidieux donne

$$R^{00} - \frac{1}{2}g^{00}R = +\frac{2}{c^2}\Delta\psi$$

par ailleurs lorsque  $\psi \ll c^2$  (limite newtonienne) il est facile de voir que la composante purement temporelle du tenseur énergie-impulsion se réduit à

$$T^{00} \simeq \rho c^2$$

$\rho$  désignant la densité de masse. La composante purement temporelle des équations d'Einstein nous dit alors

$$R^{00} - \frac{1}{2}Rg^{00} = \chi T^{00}$$

soit

$$\frac{2}{c^2}\Delta\psi = \chi\rho c^2$$

l'équation classique correspondante est bien évidemment l'équation de Poisson,  $\Delta\psi = 4\pi G\rho$ , il convient donc de choisir

$$\chi = \frac{8\pi G}{c^4}$$

pour retrouver la mécanique classique à partir de la relativité générale.

Les équations d'Einstein s'écrivent donc finalement

$$R^{\mu\nu} - \frac{1}{2}Rg^{\mu\nu} = \frac{8\pi G}{c^4}T^{\mu\nu}$$

en prenant la trace de cette dernière formulation, on en trouve une équivalente

$$R - \frac{1}{2}\text{tr}(\check{g})R = \frac{8\pi G}{c^4}T$$

soit

$$-R = \frac{8\pi G}{c^4}T$$

que l'on réintroduit dans l'équation de départ pour obtenir

$$R^{\mu\nu} = \frac{8\pi G}{c^4} \left( T^{\mu\nu} - \frac{1}{2}g^{\mu\nu}T \right)$$

qui ne nécessite plus la contraction du Ricci.



Mécanique quantique et autres crochets



## Fondements de la mécanique quantique.

### 9.1 Postulats de la mécanique quantique

Le moins que l'on puisse dire c'est qu'il existe plusieurs manières de présenter les postulats de la mécanique quantique.

Celle que nous avons choisi de présenter ici est basée, nous semble-t-il, sur un minimum d'hypothèses et tente de proposer un « raisonnement » *avant* d'énoncer le postulat correspondant. On peut aussi faire l'inverse et demander au lecteur d'accepter un énoncé *ex nihilo* pour le justifier par la suite. C'est sûrement une question de goût... Pour des raisons de concision intellectuelle dans un exposé nécessairement introductif à ce niveau, certains postulats présentés indépendamment dans certains ouvrages avancés ont été ici regroupés. Que les puristes nous excusent...

La seule chose dont on peut être sûr est que ces postulats ne sont souvent et *a priori* pas intuitifs. Ils ont été le fruit d'une longue réflexion d'une trentaine d'année au début du vingtième siècle et ont été motivés par des *résultats expérimentaux* pour la plupart inattendus. Ils nécessitent un profond changement de point de vue sur la réalité physique à l'échelle microscopique.

#### 9.1.1 Espace des états de la mécanique quantique

L'une des propriétés fondamentale d'un système quantique est le fait que son état dépend intimement du processus de mesure qui permet de le définir. À l'échelle atomique il semble clair en effet que le processus expérimental qui conduit à l'affirmation « *Telle particule est à tel endroit à tel instant* » ou bien « *À tel instant, telle particule possède telle vitesse* », interagit fortement avec la dite particule. Cette interaction nécessite quelques éclaircissements, mais il semble raisonnable de penser que les moyens expérimentaux que l'on doit utiliser pour pouvoir être aussi affirmatif, mettent en jeu d'autres

particules et sont donc d'un certain point de vue du même ordre de grandeur que l'objet à mesurer. En poussant plus loin le raisonnement ne peut-on pas penser que *c'est le processus de mesure qui conduit au résultat observé...* Si l'on suit cette idée, qu'en est-il alors de la réalité ou de l'existence d'un objet quantique (quanton) indépendamment de sa mesure ou bien en dehors de celle-ci !

Nous pourrions discuter longtemps de ce problème d'ontologie quantique, il est préférable dans un premier temps d'y couper court en faisant un certain nombre de constats qui deviendront vite des postulats.

L'état d'un système physique décrit tous les aspects de ce système, dans le but de prévoir les résultats des expériences que l'on peut réaliser. L'état d'un système classique constitué de  $N$  particules est parfaitement déterminé par la donnée de  $N$  vecteurs positions et  $N$  vecteurs vitesses. En mécanique quantique le statut de la mesure et de son interaction avec le quanton fait que la seule chose que l'on soit en mesure d'affirmer est que l'on peut être sûr de trouver un système dans l'état où l'a mis sa mesure au moment où on le mesure... Cette affirmation est en fait tautologique !

Pour rendre les choses plus concrètes, supposons que l'on puisse décrire l'état d'un système par une fonction  $\Psi$  des variables canoniques de position ou d'impulsion. Pour simplifier, nous parlerons de «l'état  $\Psi$ ». Si le résultat d'une mesure donne l'état  $\Psi$ , on peut affirmer sans sourciller que la probabilité de trouver l'état  $\Psi$  sachant que l'on est dans l'état  $\Psi$  est égale à 1...

Plus généralement, on peut imaginer que l'on puisse calculer la probabilité de transition de l'état  $\Phi$  vers l'état  $\Psi$ . Le premier axiome de la mécanique quantique consiste en l'affirmation que c'est en fait la seule chose que l'on puisse connaître a priori.

Pour la suite, le plus simple est de supposer que l'ensemble des états d'un système quantique forme un espace vectoriel :

- on est capable d'additionner deux états et la somme de deux états est un nouvel état ;
- on sait donner un sens au fait de multiplier un état par quelque chose qui ressemble à un nombre.

En physique cette hypothèse est appelée principe de superposition. On peut donner un sens plus ou moins expérimental à cette structure d'espace vectoriel. L'état  $\Theta = \alpha\Phi + \beta\Psi$  où  $\alpha$  et  $\beta$  sont deux constantes sera tel que la mesure de  $\Theta$  donnera soit  $\Phi$  soit  $\Psi$  avec des probabilités dépendant des poids  $\alpha$  et  $\beta$ . Une particularité remarquable est le fait que si l'on définit

l'état  $\Theta$  comme la superposition de  $\Phi$  sur lui-même, soit  $\Theta = \alpha\Phi + \beta\Phi$ ; la mesure de  $\Theta$  donnera toujours  $\Phi$  indépendamment des valeurs de  $\alpha$  et  $\beta$ <sup>1</sup>. En termes vectoriel, seule la direction du vecteur associé à l'état quantique est donc importante. On lit parfois *rayon quantique* en lieu et place d'état quantique. En conséquence, on choisira de représenter un état quantique par un vecteur unitaire, ce choix est conventionnel.

L'idée géniale des concepteurs de la mécanique quantique fût de comprendre que l'on pouvait munir cet espace vectoriel des états quantiques d'une norme construite sur la probabilité de transition d'un état vers un autre...

Compte-tenu de la remarque initiale, si l'on sait que l'on se trouve dans l'état  $\Psi$  et si on ne le perturbe pas, on est sûr d'y rester :

$$P(\Psi \rightarrow \Psi) = 1$$

on pourrait donc associer la probabilité  $P(\Psi \rightarrow \Psi)$  à un certain produit scalaire noté  $(\Psi, \Psi)$  tel que  $P(\Psi \rightarrow \Psi) = (\Psi, \Psi)$ . Un produit scalaire n'étant pas forcément positif une généralisation à la probabilité de transition  $P(\Psi \rightarrow \Phi)$  de l'état  $\Psi$  vers l'état  $\Phi$  serait alors problématique. Il apparaît donc naturel de définir cette probabilité comme le carré du module de ce produit scalaire, ainsi

$$P(\Psi \rightarrow \Psi) = |(\Psi, \Psi)|^2 = 1 \quad \text{et} \quad P(\Psi \rightarrow \Phi) = |(\Psi, \Phi)|^2 \leq 1$$

L'introduction de ce carré permet de voir que le corps des scalaires permettant la construction de l'espace vectoriel des états d'un système quantique peut être celui des nombres complexes. Il ne faut pas se priver de cette opportunité<sup>2</sup>.

À chaque instant la fonction  $\Psi$  associe donc, par exemple, les coordonnées canoniques  $\mathbf{q}$  d'un système physique à un nombre complexe. Ce nombre complexe  $\Psi(\mathbf{q}, t)$  définit l'état quantique du système considéré. Un état quantique est donc un vecteur que l'on peut représenter par une fonction à valeurs complexes, ce constat permet de préciser le produit scalaire que l'on peut utiliser pour rendre tout cela opérationnel. Pour tout état  $\Psi$  écrira que

$$(\Psi, \Psi) = \int \overline{\Psi(\mathbf{q}, t)} \Psi(\mathbf{q}, t) d\mathbf{q} = \int |\Psi(\mathbf{q}, t)|^2 d\mathbf{q}$$

Si l'on a pris soin de choisir le représentant unitaire de l'état  $\Psi$  on a bien  $P(\Psi \rightarrow \Psi) = |(\Psi, \Psi)|^2 = 1$ . La quantité réelle  $|\Psi(\mathbf{q}, t)|^2 d\mathbf{q}$  représente donc

<sup>2</sup> En fait, on peut montrer que le corps des réels serait insuffisant pour la construction de la

la probabilité pour qu'au temps  $t$  le système soit trouvé (lors d'une mesure) dans un élément de volume  $d\mathbf{q}$  situé autour du point  $\mathbf{q}$ . En outre, la probabilité de transition de l'état  $\Psi$  vers l'état  $\Phi$  sera donnée par

$$P(\Psi \rightarrow \Phi) = |(\Psi, \Phi)|^2 = \left| \int \overline{\Psi(\mathbf{q}, t)} \Phi(\mathbf{q}, t) d\mathbf{q} \right|^2$$

Dans certains cas il se peut que la fonction  $\Psi(\mathbf{q}, t)$  représentant l'état du système ne soit pas de carré sommable<sup>3</sup>. La notion d'état quantique associée à de tels systèmes est alors plus délicate à interpréter. En y réfléchissant bien, de tels systèmes apparaissent bien souvent trop idéalisés<sup>4</sup>. Il n'en demeure pas moins que si pour de tels systèmes la quantité  $|\Psi(\mathbf{q}, t)|^2$  ne peut plus représenter une probabilité, le rapport de deux valeurs de cette fonction en différents points de l'espace des configurations du système permet de d'obtenir une probabilité relative entre ces deux « états ».

Notons finalement que l'ensemble  $\mathbb{L}^2(\mathbb{R})$  des fonctions de  $\mathbb{R} \rightarrow \mathbb{C}$  de carré sommable muni de ce produit scalaire forme un espace de Hilbert. En termes clairs cela signifie que cet espace est complet et que toutes les suites de Cauchy sont convergentes. En termes moins clairs cela signifie aussi que l'on peut associer de façon bijective chaque vecteur de cet espace à un vecteur de son espace dual. Ce n'est pas si évident que cela en dimension infinie, et d'une importance colossale dans la manipulation des états quantiques.

**POSTULAT 1 : L'état d'un système quantique est représenté par un vecteur, normé à l'unité d'un espace de Hilbert. La probabilité de transition entre deux états est donnée par le carré du module du produit scalaire entre ces deux états. Le produit scalaire qui permet de définir la norme sur cet espace est**

$$(\Psi, \Phi) = \int \overline{\Psi(\mathbf{q}, t)} \Phi(\mathbf{q}, t) d\mathbf{q} \quad (9.1)$$

### 9.1.2 Grandeurs physiques

#### Valeurs propres, états propres

Considérons une certaine grandeur physique  $A$  pour laquelle il est possible de mesurer une valeur par l'expérience. Cette valeur sera donc en nombre réel.

théorie quantique.

<sup>3</sup> On le verra par exemple pour une particule libre non confinée.

<sup>4</sup> Par exemple, est-ce qu'expérimentalement une particule peut se trouver libre de toute force

On appellera une telle grandeur une observable. On peut toujours considérer l'ensemble  $\mathbb{S}$  des résultats possibles de la mesure de  $A$ . En mécanique quantique  $\mathbb{S}$  est appelé spectre de  $A$ , et les éléments de ce spectre, qui correspondent donc à l'un des résultats potentiels de la mesure de  $A$ , sont appelés valeurs *propres*<sup>5</sup> de  $A$ . Notons d'emblée que  $\mathbb{S}$  peut être discret si ses éléments sont dénombrables, c'est-à-dire en bijection avec  $\mathbb{N}$ , ou continu si ses éléments ne sont pas dénombrables, c'est-à-dire en bijection avec  $\mathbb{R}$ . Considérons pour simplifier le cas où  $\mathbb{S}$  est discret, en repoussant à plus tard le cas continu ! Nous avons donc  $\mathbb{S} = \{a_1, a_2, \dots, a_n, \dots\}$  tout en étant discret, cet ensemble ne contient pas forcément un nombre fini d'éléments. Une autre remarque qui peut paraître triviale mais qui aura son importance, il semble évident que chacun des résultats d'une mesure est un nombre réel :  $a_{n=1, \dots} \in \mathbb{R}$ .

Désignons à présent par  $\Psi_n$  l'état correspondant à celui dans lequel la mesure de  $A$  donne  $a_n$ . On parle de l'état propre associé à la valeur propre  $a_n$ . De manière générale, nous pouvons affirmer que si un système quantique se trouve dans un état arbitraire  $\Psi$  alors la mesure de sa grandeur  $A$  ne peut donner que l'un des éléments de son spectre. Le principe de superposition nous indique alors que l'état  $\Psi$  s'écrit nécessairement comme une combinaison linéaire des états  $\Psi_n$ . Ainsi

$$\exists \alpha_{1, \dots} \in \mathbb{C} \text{ tel que } \Psi = \sum_n \alpha_n \Psi_n \quad (9.2)$$

En termes probabilistes, nous avons

$$\begin{aligned} P(A = a_n) &= 1 \text{ si } \Psi = \Psi_n \text{ c'est-à-dire si } \alpha_n = 1 \text{ et } \alpha_{m \neq n} = 0 \\ P(A = a_n) &= 0 \text{ si } \alpha_n = 0 \end{aligned}$$

Les nombres complexes  $\alpha_n$  représentent les poids respectifs des différents états conduisant aux différentes valeurs de la grandeur  $A$ . Il paraît clair que le nombre réel positif  $|\alpha_n|^2$  est inférieur ou égal à 1, il s'interprète comme la probabilité d'obtenir le résultat  $a_n$  lors de la mesure de  $A$  dans l'état  $\Psi$ . Cette interprétation n'est possible que si, comme nous l'avons précisé au paragraphe précédent, nous prenons soin de choisir chacun des états de telle manière qu'il soit unitaire; ainsi pour chaque entier  $n$

$$(\Psi, \Psi) = \int |\Psi(\mathbf{q}, t)|^2 d\mathbf{q} = (\Psi_n, \Psi_n) = \int |\Psi_n(\mathbf{q}, t)|^2 d\mathbf{q} = 1$$

---

sur une droite infinie ...

<sup>5</sup> Oublions un instant ce que nous savons par ailleurs sur la notion de valeur propre en algèbre

Dans ce cas nous aurons

$$\sum_n |\alpha_n|^2 = (\Psi, \Psi) = 1$$

Si nous ne prenons pas cette précaution de normalisation, nous aurons simplement

$$\sum_n |\alpha_n|^2 = \sum_n \alpha_n \bar{\alpha}_n = \int \overline{\Psi(\mathbf{q}, t)} \Psi(\mathbf{q}, t) d\mathbf{q} \quad (9.3)$$

et l'interprétation probabiliste des  $\alpha_n$  n'est alors plus aussi directe; la relation ci-dessus est cependant plus générale.

Prenons la relation (9.2) et conjuguons la, il vient

$$\bar{\Psi} = \overline{\sum_n \alpha_n \Psi_n} = \sum_n \bar{\alpha}_n \bar{\Psi}_n$$

en multipliant cette relation par  $\Psi$  et en intégrant il vient

$$\int \overline{\Psi(\mathbf{q}, t)} \Psi(\mathbf{q}, t) d\mathbf{q} = \sum_n \bar{\alpha}_n \int \overline{\Psi_n(\mathbf{q}, t)} \Psi(\mathbf{q}, t) d\mathbf{q}$$

en utilisant la relation (9.3) on trouve donc

$$\sum_n \alpha_n \bar{\alpha}_n = \sum_n \bar{\alpha}_n \int \overline{\Psi_n(\mathbf{q}, t)} \Psi(\mathbf{q}, t) d\mathbf{q}$$

soit

$$\alpha_n = \int \overline{\Psi_n(\mathbf{q}, t)} \Psi(\mathbf{q}, t) d\mathbf{q} = (\Psi_n, \Psi) \quad (9.4)$$

il ne reste plus qu'à remplacer dans cette relation  $\Psi$  par son expression donnée par la relation (9.2) et l'on obtient

$$\begin{aligned} \alpha_n &= \int \overline{\Psi_n(\mathbf{q}, t)} \left( \sum_m \alpha_m \Psi_m(\mathbf{q}, t) \right) d\mathbf{q} \\ &= \sum_m \alpha_m \int \overline{\Psi_n(\mathbf{q}, t)} \Psi_m(\mathbf{q}, t) d\mathbf{q} \\ &= \sum_m \alpha_m (\Psi_n, \Psi_m) \end{aligned}$$

On déduit sans peine de cette relation que

$$(\Psi_n, \Psi_m) = \delta_{mn} = \begin{cases} 0 & \text{si } m \neq n \\ 1 & \text{si } m = n \end{cases}$$

les états propres sont donc orthonormés.

## Opérateur associé à une grandeur physique

L'interprétation probabiliste des valeurs propres présentée lors de la section précédente permet de définir la valeur moyenne  $\langle A \rangle$  de la grandeur  $A$ . De manière évidente, une valeur moyenne est la somme des valeurs possibles pondérée par la probabilité correspondante. Ainsi nous prendrons

$$\langle A \rangle = \sum_n a_n |\alpha_n|^2 = \sum_n a_n \alpha_n \overline{\alpha_n} \quad (9.5)$$

Admettons que l'on puisse associer un opérateur linéaire  $\hat{A}$  à la grandeur physique  $A$ .

Admettons aussi, au moins dans un premier temps, que l'image par cet opérateur d'un état quantique soit encore un état quantique. Une définition simple de cet opérateur pourrait être que le produit scalaire  $(\Psi, \hat{A}\Psi)$  soit égal à la valeur moyenne  $\langle A \rangle$  de la grandeur physique  $A$ . Sous cette hypothèse nous aurons

$$\langle A \rangle = \sum_n a_n \alpha_n \overline{\alpha_n} = \int \overline{\Psi(\mathbf{q}, t)} \hat{A}\Psi(\mathbf{q}, t) d\mathbf{q}$$

en utilisant alors l'expression (9.4) pour déterminer  $\overline{\alpha_n}$  il vient

$$\int \overline{\Psi(\mathbf{q}, t)} \sum_n a_n \alpha_n \Psi_n(\mathbf{q}, t) d\mathbf{q} = \int \overline{\Psi(\mathbf{q}, t)} \hat{A}\Psi(\mathbf{q}, t) d\mathbf{q}$$

si tout se passe bien, nous pouvons identifier

$$\hat{A}\Psi(\mathbf{q}, t) = \sum_n \alpha_n a_n \Psi_n(\mathbf{q}, t)$$

Cette relation est vraie pour un état  $\Psi$  quelconque, dans le cas particulier où  $\Psi = \Psi_n$  qui correspond comme nous l'avons déjà indiqué à  $\alpha_n = 1$  et  $\alpha_{m \neq n} = 0$ , elle devient

$$\hat{A}\Psi_n = a_n \Psi_n$$

et tout s'éclaire: l'état  $\Psi_n$  est représenté par le vecteur propre de l'opérateur  $\hat{A}$  associé à la valeur propre  $a_n$ . Dans cet état le résultat de la mesure de  $A$  donne la valeur  $a_n$ .

Le produit scalaire permet de définir l'adjoint  $\hat{T}^\dagger$  d'un opérateur  $\hat{T}$ . De façon générale la définition est la suivante

$$(f, \hat{T}^\dagger g) = (\hat{T} f, g)$$

dans le cas de nos états quantiques représentés par des fonction complexes et avec le produit scalaire (9.1) nous aurons

$$(\Phi, \hat{T}^\dagger \Psi) = (\hat{T} \Phi, \Psi) = \overline{(\Psi, \hat{T} \Phi)}$$

Il est clair que la valeur moyenne  $\langle A \rangle$  définie par la relation (9.5) est un nombre réel et vérifie donc  $\langle A \rangle = \overline{\langle A \rangle}$ . La conséquence pour l'opérateur  $\hat{A}$  associé à la grandeur physique  $A$  est immédiate. Pour tout état quantique  $\Psi$  la relation (9.6) donne successivement

$$\langle A \rangle = (\Psi, \hat{A} \Psi) = \overline{\langle A \rangle} = \overline{(\Psi, \hat{A} \Psi)} = (\hat{A} \Psi, \Psi) = (\Psi, \hat{A}^\dagger \Psi)$$

Nous arrivons donc à la conclusion que l'opérateur  $\hat{A}$  associé à la grandeur physique  $A$  est nécessairement auto-adjoint (on lit aussi hermitien)

$$\hat{A} = \hat{A}^\dagger$$

On progresse ...

**POSTULAT 2 : Chaque observable  $A$  est associée à un opérateur linéaire  $\hat{A}$  auto-adjoint. Le résultat de la mesure de  $A$  donne une valeur propre de  $\hat{A}$ , c'est un nombre réel. On peut décomposer chaque état normalisé du système sur un ensemble complet de vecteurs propres orthonormés de  $\hat{A}$ . La valeur moyenne  $\langle A \rangle$  de la grandeur physique  $A$  est donnée par**

$$\langle A \rangle = (\Psi, \hat{A} \Psi) \tag{9.6}$$

La théorie des opérateur linéaires nous apprend que deux opérateurs qui commutent peuvent être diagonalisés dans la même base<sup>6</sup>. Avec des mots cela signifie que la base propre  $\{\Psi_{n=1, \dots}\}$  qui diagonalise  $\hat{A}$ , c'est-à-dire celle dans laquelle pour chaque valeur de  $n$ , il existe une valeur (propre)  $a_n$  telle que  $\hat{A} \Psi_n = a_n \Psi_n$ , diagonalise aussi l'opérateur  $\hat{B}$ , c'est-à-dire que toujours pour chaque valeur de  $n$ , il existe une valeur  $b_n$  telle que  $\hat{B} \Psi_n = b_n \Psi_n$ . Généralement les valeurs propres  $a_n$  et  $b_n$  sont différentes sans quoi les opérateurs seraient proportionnels...

---

linéaire...

<sup>6</sup> On trouvera la démonstration de ce théorème un peu plus loin, dans le cadre du formalisme

Revenons dans notre état quantique. Pour que le résultat d'une mesure puisse nous apprendre quelque chose sur une grandeur physique  $A$ , nous avons vu qu'il fallait que l'on puisse écrire l'état  $\Psi$  du système comme une superposition des états propres de l'opérateur associé à cette grandeur. Nous avons vu aussi que ces états propres peuvent être représentés par les vecteurs propres de l'opérateur  $\hat{A}$  associé à  $A$ . Cet ensemble forme une base de l'espace des états. S'il existe un autre opérateur  $\hat{B}$  tel que  $[\hat{A}, \hat{B}] = 0$ , les vecteurs propres de  $\hat{A}$  seront aussi ceux de  $\hat{B}$ . En supposant qu'elle existe, la grandeur  $B$  associée à l'opérateur  $\hat{B}$ , pourra donc être mesurée à partir du même état  $\Psi$ . Les grandeurs  $A$  et  $B$  seront dites compatibles, c'est-à-dire mesurables simultanément. Il semble que dans ce cas la mesure de l'une n'a pas d'influence sur la mesure de l'autre.

### 9.1.3 Évolution temporelle

Nous avons vu que la description de l'état d'un système quantique à l'instant  $t$  pouvait être envisagée par la donnée d'une fonction  $\Psi(\mathbf{q}, t)$ . La détermination de cette fonction repose sur un certain nombre d'hypothèses et son interprétation est possible à l'aune du principe de superposition. Si l'état  $\Psi(\mathbf{q}, t)$  se décompose linéairement à l'instant  $t$  sur une certaine famille d'autres états formant une base de l'espace vectoriel de tous les états, les coefficients de cette décomposition permettent de déterminer la probabilité des différents résultats d'une mesure que l'on effectuerait à cet instant.

Comment envisager l'évolution temporelle d'un tel système ?

Si l'évolution temporelle de  $\Psi(\mathbf{q}, t)$ , et donc sa dérivée par rapport à  $t$ , devait s'envisager à travers l'action d'un opérateur non linéaire sur  $\Psi(\mathbf{q}, t)$ , la théorie des équations différentielles nous indique alors que la structure d'espace vectoriel de l'ensemble des états volerait en éclats dans cette évolution. Dans ces conditions, nous devrions alors revoir à chaque instant le lien étroit que nous avons commencé à tisser entre le principe de superposition et le processus de mesure!

Le cadre de la théorie semble donc pour le moment intrinsèquement linéaire, peut-être l'expérience viendra-t-elle briser ce cadre, ou pas ! Dans un premier temps, et suivant le principe du rasoir d'Okam «pluralitas non est ponenda sine necessitate », restons linéaires...

## Quel est l'adjoint de $\widehat{H}$ ?

De façon moins linguistique, l'hypothèse de linéarité s'écrit

$$i\hbar \frac{\partial \Psi}{\partial t} = \widehat{H}\Psi \quad (9.7)$$

où  $\widehat{H}$  est un opérateur linéaire et le facteur  $i\hbar$  conventionnel,  $\hbar$  est un paramètre réel dont la valeur n'est pas fixée pour le moment, nous verrons plus tard que ce paramètre doit être dimensionné. L'appellation  $\widehat{H}$  est un hommage à Lagrange...

Nous savons que  $\Psi$  est représenté par une fonction de carré sommable, ainsi pour un état normé

$$\int \overline{\Psi}\Psi d\mathbf{q} = 1 \quad \Rightarrow \quad \frac{\partial}{\partial t} \left( \int \overline{\Psi}\Psi d\mathbf{q} \right) = \int \left( \Psi \frac{\partial \overline{\Psi}}{\partial t} + \overline{\Psi} \frac{\partial \Psi}{\partial t} \right) d\mathbf{q} = 0$$

En prenant le conjugué de la relation (9.7), on obtient

$$-i\hbar \frac{\partial \overline{\Psi}}{\partial t} = \overline{\widehat{H}\Psi}$$

La dérivée temporelle de la norme de  $\Psi$  s'écrit donc

$$\int \frac{i}{\hbar} \left[ \Psi \overline{\widehat{H}\Psi} - \overline{\Psi} \widehat{H}\Psi \right] d\mathbf{q} = 0$$

en passant à l'adjoint sur le premier terme on trouve

$$\int \Psi \overline{\widehat{H}\Psi} d\mathbf{q} = \left( \overline{\Psi}, \overline{\widehat{H}\Psi} \right) = \left( \widehat{H}\Psi, \Psi \right) = \left( \Psi, \widehat{H}^\dagger \Psi \right) = \int \overline{\Psi} \widehat{H}^\dagger \Psi d\mathbf{q}$$

et donc

$$\int \frac{i}{\hbar} \left[ \overline{\Psi} \widehat{H}^\dagger \Psi - \overline{\Psi} \widehat{H}\Psi \right] d\mathbf{q} = 0$$

soit

$$\frac{i}{\hbar} \int \overline{\Psi} \left[ \widehat{H}^\dagger - \widehat{H} \right] \Psi d\mathbf{q} = 0$$

ce résultat est vrai pour tout état normalisable, il implique donc que l'opérateur  $\widehat{H}$  d'évolution temporelle est auto-adjoint

$$\widehat{H}^\dagger = \widehat{H}$$

Tiens, comme c'est bizarre, cet opérateur serait-il associé à une grandeur physique ?

## Une nouvelle équation ?

Maintenant que nous savons que  $\hat{H}$  est auto-adjoint, étudions la dérivée temporelle de la valeur moyenne d'une grandeur physique. Nous avons vu que l'estimation la plus raisonnable de valeur moyenne  $\langle A \rangle$  des mesures d'une grandeur physique  $A$  s'écrit

$$\langle A \rangle = \int \bar{\Psi} \hat{A} \Psi d\mathbf{q}$$

Cette intégration possède une conséquence claire : la valeur moyenne  $\langle A \rangle$  ne dépend plus de  $\mathbf{q}$  sa seule variable est le temps  $t$ , ainsi

$$\frac{d\langle A \rangle}{dt} = \frac{\partial \langle A \rangle}{\partial t}$$

que nous pouvons aisement calculer

$$\frac{d\langle A \rangle}{dt} = \frac{\partial \langle A \rangle}{\partial t} = \int \left[ \frac{\partial \bar{\Psi}}{\partial t} \hat{A} \Psi + \bar{\Psi} \frac{\partial \hat{A}}{\partial t} \Psi + \bar{\Psi} \hat{A} \frac{\partial \Psi}{\partial t} \right] d\mathbf{q}$$

puis un peu comme tout à l'heure

$$\frac{d\langle A \rangle}{dt} = \frac{\partial \langle A \rangle}{\partial t} = \int \left[ \frac{i}{\hbar} \overline{\hat{H} \Psi} \hat{A} \Psi + \bar{\Psi} \frac{\partial \hat{A}}{\partial t} \Psi - \frac{i}{\hbar} \bar{\Psi} \hat{A} \hat{H} \Psi \right] d\mathbf{q}$$

observons de plus près le premier terme de l'intégrale,  $\hat{H}$  étant auto-adjoint, quelques manipulations donnent

$$\int \overline{\hat{H} \Psi} \hat{A} \Psi d\mathbf{q} = (\hat{H} \Psi, \hat{A} \Psi) = (\Psi, \hat{H}^\dagger \hat{A} \Psi) = (\Psi, \hat{H} \hat{A} \Psi) = \int \bar{\Psi} \hat{H} \hat{A} \Psi d\mathbf{q}$$

ainsi

$$\begin{aligned} \frac{d\langle A \rangle}{dt} &= \int \left[ \frac{i}{\hbar} \bar{\Psi} \hat{H} \hat{A} \Psi + \bar{\Psi} \frac{\partial \hat{A}}{\partial t} \Psi - \frac{i}{\hbar} \bar{\Psi} \hat{A} \hat{H} \Psi \right] d\mathbf{q} \\ &= \int \bar{\Psi} \left[ \frac{i}{\hbar} (\hat{H} \hat{A} - \hat{A} \hat{H}) + \frac{\partial \hat{A}}{\partial t} \right] \Psi d\mathbf{q} \\ &= \int \bar{\Psi} \left[ \frac{i}{\hbar} [\hat{H}, \hat{A}] + \frac{\partial \hat{A}}{\partial t} \right] \Psi d\mathbf{q} \end{aligned}$$

Que l'on peut donc écrire

$$\frac{d\langle A \rangle}{dt} = \left\langle \frac{\partial \hat{A}}{\partial t} - \frac{i}{\hbar} [\hat{A}, \hat{H}] \right\rangle \quad (9.8)$$

Si l'on convient que

$$\frac{d\langle A \rangle}{dt} = \frac{d\langle \hat{A} \rangle}{dt} = \left\langle \frac{d\hat{A}}{dt} \right\rangle$$

l'affaire est dans le sac et l'on retrouve la mécanique classique. En effet, la relation (9.8) devient

$$\left\langle \frac{d\hat{A}}{dt} \right\rangle = \left\langle \frac{\partial \hat{A}}{\partial t} - \frac{i}{\hbar} [\hat{A}, \hat{H}] \right\rangle$$

Ainsi, en se souvenant de notre bonne vieille mécanique classique, dont l'équation fondamentale s'écrivait

$$\frac{d\varphi}{dt} = \frac{\partial \varphi}{\partial t} + \{\varphi, \mathcal{H}\}$$

où  $\varphi$  était une observable classique, fonction du temps ainsi que des coordonnées et impulsions généralisées. La mécanique classique pourrait s'identifier à une vision moyenne des résultats de mesure quantiques en remplaçant le crochet de Poisson  $\{, \}$  par  $-\frac{i}{\hbar} [, ]$  et les observables classiques par des opérateurs auto-adjoints. L'opérateur  $\hat{H}$  d'évolution temporelle correspondrait alors au hamiltonien classique et ses valeurs propres seraient donc l'énergie du système.

**POSTULAT 3 : L'évolution dans le temps de l'état quantique  $\psi$  est régie par l'équation**

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}\Psi$$

où  $\hat{H}$  est l'opérateur associé à l'énergie totale du système. Les opérateurs associés aux grandeurs physiques s'obtiennent à partir de leurs expressions classiques en remplaçant le crochet de poisson  $\{, \}$  par  $-\frac{i}{\hbar} [, ]$ .

Cette équation est appelée équation de , elle a été postulée par son auteur en 1925. Sa correspondance avec l'équation fondamentale de la mécanique classique est l'œuvre de Dirac quelques années plus tard.

### 9.1.4 Représentation spatiale

Appliquons sans tarder ce postulat dans le cas simple des observables position et impulsion.

Nous avons vu en mécanique classique que certains systèmes<sup>7</sup> possédant  $\ell$  degrés de liberté, admettent d'être décrits par des coordonnées canoniques  $\mathbf{q} := \{q_1, \dots, q_\ell\}$  et  $\mathbf{p} := \{p_1, \dots, p_\ell\}$  vérifiant

$$\forall \alpha, \beta = 1, \dots, \ell \quad \begin{cases} \{q_\alpha, p_\beta\} = \delta_{\alpha\beta} \\ \{q_\alpha, q_\beta\} = 0 \\ \{p_\alpha, p_\beta\} = 0 \end{cases}$$

en mécanique quantique, pour chaque valeur de  $\alpha = 1, \dots, \ell$  les observables  $q_\alpha$  et  $p_\alpha$  de ce système seront donc associées à des opérateurs  $\widehat{q}_\alpha$  et  $\widehat{p}_\alpha$  tels que

$$\forall \alpha, \beta = 1, \dots, \ell \quad \begin{cases} [\widehat{q}_\alpha, \widehat{p}_\beta] = i\hbar\delta_{\alpha\beta} \mathbb{I} \\ [\widehat{q}_\alpha, \widehat{q}_\beta] = 0 \\ [\widehat{p}_\alpha, \widehat{p}_\beta] = 0 \end{cases} \quad (9.9)$$

où  $\mathbb{I}$  et  $0$  représentent respectivement l'identité et l'opérateur nul sur l'espace des états. Ces relations sont dénommées relations de commutation canoniques.

Pourrait-on rendre ceci plus concret ? La réponse est positive mais il faut pour cela choisir une représentation du système par cette fameuse fonction  $\Psi$ . Supposons que  $\Psi = \Psi(\mathbf{q}, t)$ , la première relation de commutation canonique s'écrit

$$\widehat{q}_\alpha (\widehat{p}_\beta \Psi(\mathbf{q}, t)) - \widehat{p}_\beta (\widehat{q}_\alpha \Psi(\mathbf{q}, t)) = i\hbar\delta_{\alpha\beta} \Psi(\mathbf{q}, t)$$

Plusieurs possibilités sont envisageable pour résoudre cette équation en  $\widehat{q}_\alpha$  et  $\widehat{p}_\beta$ . Certaines sont très formelles et très jolies ([17] ch. 4), on peut aussi procéder de façon intuitive. On remarque que l'action de ces deux opérateurs est complémentaire, après quelques essais on arrive assez vite à se persuader qu'une solution de la forme

$$\widehat{q}_\alpha = a q_\alpha \mathbb{I} \quad \text{et} \quad \widehat{p}_\beta = b \frac{d}{dq_\beta} \quad \text{avec} \quad (a, b) \in \mathbb{C}^2$$

est envisageable. Il ne reste plus qu'à essayer. En reportant on trouve

$$\widehat{q}_\alpha (\widehat{p}_\beta \Psi(\mathbf{q}, t)) = a q_\alpha \times \left( b \frac{d\Psi(\mathbf{q}, t)}{dq_\beta} \right)$$

<sup>7</sup> Ceux qui acceptent de laisser décrire leur dynamique par un formalisme Hamiltonien. A minima, les systèmes conservatifs ont ce bon goût...

puis

$$\begin{aligned}\widehat{p}_\beta (\widehat{q}_\alpha \Psi (\mathbf{q}, t)) &= b \frac{d}{dq_\beta} (a q_\alpha \times \Psi (\mathbf{q}, t)) \\ &= ab \left( \delta_{\alpha\beta} \Psi (\mathbf{q}, t) + q_\alpha \frac{d\Psi (\mathbf{q}, t)}{dq_\beta} \right)\end{aligned}$$

la relation de commutation s'écrit donc

$$abq_\alpha \frac{d\Psi (\mathbf{q}, t)}{dq_\beta} - ab \delta_{\alpha\beta} \Psi (\mathbf{q}, t) - abq_\alpha \frac{d\Psi (\mathbf{q}, t)}{dq_\beta} = i\hbar \delta_{\alpha\beta} \Psi (\mathbf{q}, t)$$

Ainsi en imposant  $ab = -i\hbar$  le tour est joué... La solution n'est donc pas unique, on prendra

$$\text{Si } \Psi = \Psi (\mathbf{q}, t) \begin{cases} \widehat{q}_\alpha = q_\alpha \mathbb{I} \\ \widehat{p}_\beta = -i\hbar \frac{d}{dq_\beta} \end{cases} \quad (9.10)$$

Dans ce qui précède nous avons représenté l'état quantique  $\Psi$  par une fonction des positions  $\mathbf{q}$  et du temps  $t$ . On parle de la représentation  $q$  de l'état  $\Psi$ . Rien ne nous oblige à un tel choix, l'état d'un système peut aussi concerner les impulsions de celui-ci. Cet état  $\Phi$  sera alors représenté par une fonction complexe des variables  $\mathbf{p}$  et  $t$ , notée<sup>8</sup>  $\Phi (\mathbf{p}, t)$ . Avec les définitions (9.10), on a toujours

$$\begin{aligned}[\widehat{q}_\alpha, \widehat{p}_\beta] \Phi (\mathbf{p}, t) &= q_\alpha \left( -i\hbar \frac{d\Phi (\mathbf{p}, t)}{dq_\beta} \right) - \left[ -i\hbar \frac{d}{dq_\beta} (q_\alpha \Phi (\mathbf{p}, t)) \right] \\ &= 0 + i\hbar \delta_{\alpha\beta} \Phi (\mathbf{p}, t) + i\hbar q_\alpha \frac{d\Phi (\mathbf{p}, t)}{dq_\beta} \\ &= i\hbar \delta_{\alpha\beta} \mathbb{I} [\Phi (\mathbf{p}, t)]\end{aligned}$$

mais l'on peut entrevoir une autre solution. En effet, en posant

$$\text{Si } \Phi = \Phi (\mathbf{p}, t) \begin{cases} \widehat{q}_\alpha = i\hbar \frac{d}{dp_\alpha} \\ \widehat{p}_\beta = p_\beta \mathbb{I} \end{cases} \quad (9.11)$$

<sup>8</sup> Les notations ne sont pas encore très claires au stade où nous en sommes. Le prochain chapitre

on a aussi

$$\begin{aligned}
 [\widehat{q}_\alpha, \widehat{p}_\beta] \Phi(\mathbf{p}, t) &= i\hbar \frac{d}{dp_\alpha} [p_\beta \Phi(\mathbf{p}, t)] - i\hbar p_\beta \frac{d\Phi(\mathbf{p}, t)}{dp_\alpha} \\
 &= i\hbar \delta_{\alpha\beta} \Phi(\mathbf{p}, t) + i\hbar p_\beta \frac{d\Phi(\mathbf{p}, t)}{dp_\alpha} - i\hbar p_\beta \frac{d\Phi(\mathbf{p}, t)}{dp_\alpha} \\
 &= i\hbar \delta_{\alpha\beta} \mathbb{I}[\Phi(\mathbf{p}, t)]
 \end{aligned}$$

On parle alors de la représentation  $p$  de l'état  $\Phi$ . Notons que l'on pourrait définir des représentations mixtes ... Toutes les représentations d'un état quantiques sont équivalentes, seule la base de l'espace des états utilisée varie. Par exemple, la valeur moyenne d'une observable doit être indépendante de la représentation utilisée, ainsi pour la position

$$\langle q_\alpha \rangle = \int \overline{\Psi(\mathbf{q}, t)} \times \widehat{q}_\alpha(\Psi(\mathbf{q}, t)) d\mathbf{q} = \int \overline{\Phi(\mathbf{p}, t)} \times \widehat{q}_\alpha(\Phi(\mathbf{p}, t)) d\mathbf{p}$$

en utilisant les deux expressions de l'opérateur  $\widehat{q}_\alpha$  dans chacune des représentation  $q$  et  $p$  on a donc

$$\int q_\alpha \overline{\Psi(\mathbf{q}, t)} \Psi(\mathbf{q}, t) d\mathbf{q} = i\hbar \int \overline{\Phi(\mathbf{p}, t)} \frac{d\Phi(\mathbf{p}, t)}{dp_\alpha} d\mathbf{p}$$

Le produit en variable  $\mathbf{q}$  devient la dérivée en variable  $\mathbf{p}$  : il s'agit de la propriété fondamentale de la transformée de Fourier. Vérifions qu'il s'agit bien d'elle, pour cela posons

$$\begin{aligned}
 \Phi(\mathbf{p}, t) &= \mathcal{F}(\Psi(\mathbf{q}, t))(\mathbf{p}, t) \\
 &= \frac{1}{(2\pi\hbar)^{\ell/2}} \int d\mathbf{q} \Psi(\mathbf{q}, t) e^{-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{q}}
 \end{aligned}$$

C'est ici que l'on tire tout l'avantage d'avoir introduit une constante  $\hbar$  dont la dimension égale à celle du produit scalaire  $\mathbf{p} \cdot \mathbf{q}$ , soit le joule.seconde, permet d'avoir un nombre dans l'exponentielle...

Dans ces conditions nous avons

$$\begin{aligned}
 \langle q_\alpha \rangle &= i\hbar \int \overline{\Phi(\mathbf{p}, t)} \frac{d\Phi(\mathbf{p}, t)}{dp_\alpha} d\mathbf{p} \\
 &= i\hbar \int d\mathbf{p} \left( \frac{\int d\mathbf{q} \overline{\Psi(\mathbf{q}, t)} e^{\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{q}}}{(2\pi\hbar)^{\ell/2}} \right) \frac{d}{dp_\alpha} \left( \frac{\int d\mathbf{q}' \Psi(\mathbf{q}', t) e^{-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{q}'}}{(2\pi\hbar)^{\ell/2}} \right) \\
 &= \frac{1}{(2\pi\hbar)^\ell} \int d\mathbf{p} \int d\mathbf{q} \overline{\Psi(\mathbf{q}, t)} e^{\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{q}} \int d\mathbf{q}' q'_\alpha \Psi(\mathbf{q}', t) e^{-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{q}'}
 \end{aligned}$$

On fait apparaître une distribution de Dirac dans la dernière intégrale en écrivant que

$$q'_\alpha \Psi(\mathbf{q}', t) e^{-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{q}'} = \int d\mathbf{q}'' q''_\alpha \Psi(\mathbf{q}'', t) e^{-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{q}''} \delta(\mathbf{q}' - \mathbf{q}'')$$

on obtient donc

$$\langle q_\alpha \rangle = \frac{1}{(2\pi\hbar)^\ell} \int d\mathbf{p} \int d\mathbf{q} \overline{\Psi(\mathbf{q}, t)} e^{\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{q}} \int d\mathbf{q}' \int d\mathbf{q}'' q''_\alpha \Psi(\mathbf{q}'', t) e^{-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{q}''} \delta(\mathbf{q}' - \mathbf{q}'')$$

si le théorème de Fubini s'applique, on peut regrouper les exponentielles pour avoir

$$\langle q_\alpha \rangle = \frac{1}{(2\pi\hbar)^\ell} \int d\mathbf{p} \int d\mathbf{q} \int d\mathbf{q}' \int d\mathbf{q}'' \overline{\Psi(\mathbf{q}, t)} q''_\alpha \Psi(\mathbf{q}'', t) e^{\frac{i}{\hbar} \mathbf{p} \cdot (\mathbf{q} - \mathbf{q}'')} \delta(\mathbf{q}' - \mathbf{q}'')$$

on peut alors intégrer sur  $\mathbf{p}$ , on sait bien qu'au sens des distributions

$$\int d\mathbf{p} e^{\frac{i}{\hbar} \mathbf{p} \cdot (\mathbf{q} - \mathbf{q}'')} = \hbar^\ell \int d\mathbf{x} e^{i\mathbf{x} \cdot (\mathbf{q} - \mathbf{q}'')} = (2\pi\hbar)^\ell \delta(\mathbf{q} - \mathbf{q}'')$$

ainsi

$$\langle q_\alpha \rangle = \int d\mathbf{q} \int d\mathbf{q}' \int d\mathbf{q}'' \overline{\Psi(\mathbf{q}, t)} q''_\alpha \Psi(\mathbf{q}'', t) \delta(\mathbf{q} - \mathbf{q}'') \delta(\mathbf{q}' - \mathbf{q}'')$$

l'intégration sur  $\mathbf{q}'$  est immédiate

$$\int d\mathbf{q}' \delta(\mathbf{q}' - \mathbf{q}'') = 1(\mathbf{q}'')$$

Il s'agit de la fonction de  $\mathbf{q}''$  qui vaut toujours 1, on a donc

$$\langle q_\alpha \rangle = \int d\mathbf{q} \int d\mathbf{q}'' \overline{\Psi(\mathbf{q}, t)} q''_\alpha \Psi(\mathbf{q}'', t) \delta(\mathbf{q} - \mathbf{q}'')$$

l'intégration sur  $\mathbf{q}''$  devient alors très simple

$$\int d\mathbf{q}'' q''_\alpha \Psi(\mathbf{q}'', t) \delta(\mathbf{q} - \mathbf{q}'') = q_\alpha \Psi(\mathbf{q}, t)$$

et finalement

$$\langle q_\alpha \rangle = \int d\mathbf{q} \overline{\Psi(\mathbf{q}, t)} q_\alpha \Psi(\mathbf{q}, t)$$

ce qui fallait démontrer...

Nous retiendrons que la fonction de  $\mathbf{q}$  qui représente un état quantique est la transformée de Fourier de la fonction de  $\mathbf{p}$  qui représente le même état.

Il est temps que nous nous dotions d'un formalisme adapté à toutes ces nouvelles choses.

## 9.2 Formulation de Dirac de la mécanique quantique

Cette section reprend les résultats de la précédente afin de présenter l'un des formalismes incontournables de la physique moderne. Nous ne pouvons pas, dans cet ouvrage dont l'un des objectifs est d'apprendre à écrire la physique, passer sous silence cette leçon. Les lecteurs pressés ou connaisseurs peuvent toutefois la sauter en première lecture et nous retrouver directement à la page 229 pour la formulation lagrangienne de la mécanique quantique.

### 9.2.1 Définitions, propriétés

Nous avons vu que l'élément central de la théorie quantique était rassemblé dans la notion d'état quantique  $\Psi$ . Nous avons vu aussi que l'ensemble des états quantiques potentiellement accessibles à un système ne pouvait être qu'un espace vectoriel. Pour signaler la nature vectorielle d'un objet on utilise souvent une notation spécifique qui est souvent une flèche ou l'utilisation d'un caractère gras. Dans le contexte de la mécanique quantique, P. A. M. Dirac utilisa une autre notation et une autre dénomination...

Nous avons supposé que l'état quantique d'un système pouvait être représenté par une fonction du temps et des variables canoniques  $\mathbf{q}, \mathbf{p}$  de ce système. Pour différencier d'une part l'objet mathématique vectoriel «état  $\Psi$ » et d'autre part ses composantes qui constitue l'ensemble des nombres complexes  $\Psi(\mathbf{q}, \mathbf{p}, t)$  pour  $t$  fixé et  $(\mathbf{q}, \mathbf{p})$  variables, Dirac note

$$|\Psi_t\rangle : \text{état d'un système à l'instant } t$$

Le principe de superposition et son interprétation probabiliste nous ont permis de voir que l'ensemble des états d'un système quantique est un espace vectoriel  $\mathbb{H}$  et que le corps associé à la construction de  $\mathbb{H}$  est celui des nombres complexes. Dans le cadre de cette interprétation la probabilité de transition entre deux états  $|\Psi_t\rangle$  et  $|\Phi_{t'}\rangle$  s'écrit comme du carré du module du produit scalaire entre les deux vecteurs caractérisant ces états

$$P(|\Psi_t\rangle \rightarrow |\Phi_{t'}\rangle) = |(|\Psi_t\rangle, |\Phi_{t'}\rangle)|^2$$

les parenthèses et la virgule signifient habituellement le produit scalaire entre les deux vecteurs que l'on note désormais  $|\Psi_t\rangle$  et  $|\Phi_{t'}\rangle$ . Il est bien connu que l'on peut construire une forme linéaire  $F_{|\Psi_t\rangle}$  à partir d'un vecteur  $|\Psi_t\rangle$ , il suffit en effet de poser

---

a pour objectif de clarifier la situation. Patience...

$$F_{|\Psi_t\rangle} (|\Phi_{t'}\rangle) = (|\Psi_t\rangle, |\Phi_{t'}\rangle)$$

de manière formelle on peut écrire  $F_{|\Psi_t\rangle} = (|\Psi_t\rangle, \cdot)$ . Cette forme linéaire est un élément de l'ensemble  $\mathbb{H}^*$  des formes linéaires agissant sur  $\mathbb{H}$ . Cette ensemble est appelé dual de  $\mathbb{H}$ . Il est bien connu que  $\mathbb{H}^*$  est lui aussi un espace vectoriel, la notation introduite par Dirac pour représenter l'élément de  $\mathbb{H}^*$  associé, grâce à  $F_{|\Psi_t\rangle}$ , au vecteur  $|\Psi_t\rangle$  de  $\mathbb{H}$  est la suivante

$$\text{Si } |\Psi_t\rangle \in \mathbb{H}, \text{ alors } F_{|\Psi_t\rangle} = (|\Psi_t\rangle, \cdot) := \langle \Psi_t | \in \mathbb{H}^*$$

Le fait que  $\mathbb{H}$ , muni de la norme associée à cette probabilité de transition, soit un espace de Hilbert assure l'existence d'une relation bijective entre  $|\Psi_t\rangle$  et  $\langle \Psi_t |$  pour tout vecteur  $|\Psi_t\rangle$  de  $\mathbb{H}$ . Avec cette notation le produit scalaire entre deux états devrait être noté  $\langle \Psi_t | \Phi_{t'} \rangle$ , Dirac simplifia l'affaire en supprimant l'une des deux barres : le crochet (*bracket* en anglais) entre  $|\Psi_t\rangle$  et  $|\Phi_{t'}\rangle$  perd une barre, Dirac décide donc de séparer le *bracket* en un *bra* qui vit dans  $\mathbb{H}^*$ , un *ket* qui vit dans  $\mathbb{H}$  et donc de supprimer le  $c$  !

Qui a dit que les anglais n'avaient pas d'humour !

Pour résumer, on peut dire que le ket  $|\Phi_{t'}\rangle$  est l'élément de  $\mathbb{H}$  qui est associé à l'état  $\Phi$  à l'instant  $t'$ , le bra  $\langle \Psi_t |$  est l'élément de  $\mathbb{H}^*$  associé à l'état  $\Psi$  à l'instant  $t$  et la probabilité de transition entre les ces deux états est notée

$$P (|\Psi_t\rangle \rightarrow |\Phi_{t'}\rangle) = |( |\Psi_t\rangle, |\Phi_{t'}\rangle )|^2 = | \langle \Psi_t | \Phi_{t'} \rangle |^2$$

Le fait que le corps des scalaires soit celui des nombres complexes rend le produit scalaire non commutatif, on a dans ce cas

$$( |\Psi_t\rangle, |\Phi_{t'}\rangle ) = \overline{( |\Phi_{t'}\rangle, |\Psi_t\rangle )}$$

soit en notation de Dirac

$$\langle \Psi_t | \Phi_{t'} \rangle = \overline{\langle \Phi_{t'} | \Psi_t \rangle}$$

La raison d'une telle antisymétrie provient du fait que l'on souhaite construire une norme à partir du produit scalaire. Oublions le temps quelques instants : si, par définition, nous souhaitons que  $\langle \Psi | \Psi \rangle$  représente le carré de la norme du vecteur  $|\Psi\rangle$ , il est nécessaire que pour tout complexe  $\lambda$ , la quantité positive  $|\lambda|^2 \langle \Psi | \Psi \rangle$  représente le carré de la norme du vecteur  $\lambda |\Psi\rangle$ . C'est pourquoi, on aura de façon générale

$$\forall (\lambda, \mu) \in \mathbb{C}^2, \quad \langle \lambda \Psi | \mu \Phi \rangle = \bar{\lambda} \mu \langle \Psi | \Phi \rangle$$

et, avec des notations évidentes

$$\begin{aligned} ( |\Psi\rangle, \lambda_1 |\Phi_1\rangle + \lambda_2 |\Phi_2\rangle ) &= \lambda_1 \langle \Psi | \Phi_1 \rangle + \lambda_2 \langle \Psi | \Phi_2 \rangle \\ ( \lambda_1 |\Psi_1\rangle + \lambda_2 |\Psi_2\rangle, |\Phi\rangle ) &= \bar{\lambda}_1 \langle \Psi_1 | \Phi \rangle + \bar{\lambda}_2 \langle \Psi_2 | \Phi \rangle \end{aligned}$$

### 9.2.2 Relation de fermeture et applications

Une base de l'espace  $\mathbb{H}$  des états est un ensemble de vecteurs de  $\mathbb{H}$ ,  $\mathcal{B} = \{|\Psi_n\rangle \mid n \in I\}$  où  $I$  est un ensemble *dénombrable*, tel que

$$\forall |\Psi\rangle \in \mathbb{H}, \text{ la décomposition } |\Psi\rangle = \sum_{n \in I} c_n |\Psi_n\rangle \text{ avec } c_n \in \mathbb{C} \text{ soit unique.}$$

Il est clair que le coefficient complexe  $c_n$  correspond à la projection de  $|\Psi\rangle$  sur la droite engendrée par  $|\Psi_n\rangle$ , dans le formalisme de Dirac nous avons donc

$$c_n = F_{|\Psi_n\rangle}(|\Psi\rangle) = \langle \Psi_n | \Psi \rangle$$

la décomposition de  $|\Psi\rangle$  s'écrira donc

$$|\Psi\rangle = \sum_{n \in I} \langle \Psi_n | \Psi \rangle |\Psi_n\rangle = \sum_{n \in I} |\Psi_n\rangle \langle \Psi_n | \Psi \rangle$$

Cette relation permet de voir que

$$\sum_{n \in I} |\Psi_n\rangle \langle \Psi_n| = \text{Id}$$

où  $\text{Id}$  désigne l'opérateur identité sur  $\mathbb{H}$ . Cette relation est appelée condition de complétude par les mathématiciens, les physiciens parlent plutôt de *relation de fermeture*. Cette relation est un membre éminent de l'arsenal des techniques disponibles pour manipuler les objets quantiques. Considérons à titre d'exemple le fameux théorème de Pythagore, on écrit tour à tour

$$\begin{aligned} \langle \Psi | \Psi \rangle &= \langle \Psi | \text{Id} | \Psi \rangle = \langle \Psi | \left( \sum_{n \in I} |\Psi_n\rangle \langle \Psi_n| \right) | \Psi \rangle = \sum_{n \in I} \langle \Psi | \Psi_n \rangle \langle \Psi_n | \Psi \rangle \\ &= \sum_{n \in I} \overline{\langle \Psi_n | \Psi \rangle} \langle \Psi_n | \Psi \rangle \\ &= \sum_{n \in I} |\langle \Psi_n | \Psi \rangle|^2 = \sum_{n \in I} |c_n|^2 \end{aligned}$$

si l'on considère un état normalisé<sup>9</sup>, on retrouve bien le fait que le carré de chaque  $c_n$  représente la probabilité de transition de l'état  $|\Psi\rangle$  vers l'état  $|\Psi_n\rangle$  comme nous avons essayé de l'expliquer dans les postulats.

<sup>9</sup> C'est-à-dire de norme unité...

### 9.2.3 Opérateurs

La relation de fermeture permet également de manipuler automatiquement les opérateurs associés aux observables quantiques. Soit un opérateur linéaire  $\widehat{A}$  agissant sur  $\mathbb{H}^{10}$ , comment noter le produit scalaire  $(|\Psi\rangle, \widehat{A}|\Phi\rangle)$ ? Une réponse rapide consiste à poser

$$(|\Psi\rangle, \widehat{A}|\Phi\rangle) = \langle\Psi|\widehat{A}|\Phi\rangle \quad (9.12)$$

Cette notation, qui est bel est bien utilisée dans les cours de physique, est cependant ambiguë car elle laisse planer un doute si  $\widehat{A}$  n'est pas auto-adjoint. En effet, si l'on interprète la relation  $\langle\Psi|A|\Phi\rangle$  comme l'action de  $A$  sur  $|\Phi\rangle$  projetée dans la direction indiquée par le vecteur  $|\Psi\rangle$ , on trouve

$$\langle\Psi|\widehat{A}|\Phi\rangle = (|\Psi\rangle, \widehat{A}|\Phi\rangle) = (\widehat{A}^\dagger|\Psi\rangle, |\Phi\rangle)$$

Mais on peut aussi interpréter  $\langle\Psi|\widehat{A}|\Phi\rangle$  comme la projection de  $|\Phi\rangle$  dans la direction indiquée par le vecteur  $\widehat{A}|\Psi\rangle$  soit

$$\langle\Psi|\widehat{A}|\Phi\rangle = (\widehat{A}|\Psi\rangle, |\Phi\rangle) = (|\Psi\rangle, \widehat{A}^\dagger|\Phi\rangle)$$

On remarque donc que si l'opérateur  $\widehat{A}$  est auto-adjoint, i.e.  $\widehat{A} = \widehat{A}^\dagger$ , la notation ne pose pas de problème, dans le cas contraire il faut rajouter une convention et poser

$$\langle\Psi|\widehat{A} := \widehat{A}^\dagger|\Psi\rangle$$

ainsi dans tous les cas

$$\langle\Psi|\widehat{A}|\Phi\rangle = \begin{cases} (|\Psi\rangle, \widehat{A}|\Phi\rangle) \\ (\langle\Psi|\widehat{A}, |\Phi\rangle) = (\widehat{A}^\dagger|\Psi\rangle, |\Phi\rangle) = (|\Psi\rangle, \widehat{A}|\Phi\rangle) \end{cases}$$

suivant que l'on fasse agir  $\widehat{A}$  à droite ou à gauche. Avec cette définition et pour un opérateur quelconque  $\widehat{M}$  nous aurons

$$\langle\Psi|\widehat{M}|\Phi\rangle = (|\Psi\rangle, \widehat{M}|\Phi\rangle) = (\widehat{M}^\dagger|\Psi\rangle, |\Phi\rangle) = \overline{(|\Phi\rangle, \widehat{M}^\dagger|\Psi\rangle)} = \overline{\langle\Phi|\widehat{M}^\dagger|\Psi\rangle}$$

C'est le moment de faire un petit rappel de chose bien connues mais toujours utiles

$$(\widehat{M}^\dagger)^\dagger = \widehat{M}$$

$$\begin{aligned}\forall \lambda \in \mathbb{C}, \quad (\lambda \widehat{M})^\dagger &= \bar{\lambda} \widehat{M}^\dagger \\ (\widehat{M} + \widehat{N})^\dagger &= \widehat{M}^\dagger + \widehat{N}^\dagger \\ (\widehat{M} \widehat{N})^\dagger &= \widehat{N}^\dagger \widehat{M}^\dagger\end{aligned}$$

et de conseiller à ceux qui ne sauraient pas prouver ces affirmations de se réveiller...

### 9.2.4 Observables

Il est temps de reprendre les éléments du postulat 2 dans la notation de Dirac.

Toute observable  $A$  est associée à un opérateur auto-adjoint  $\widehat{A}$ . Les solutions du problème

$$\widehat{A} |\Psi_n\rangle = a_n |\Psi_n\rangle$$

sont les valeurs propres  $a_n$  de  $\widehat{A}$  et les kets propres associés  $|\Psi_n\rangle$ . Nous avons vu qu'en mécanique quantique les valeurs propres de  $\widehat{A}$  correspondait aux différents résultats possibles de la mesure de l'observable  $A$ . Nous avons montré par ailleurs que le fait que ces résultats de mesures soient toujours des nombres réels, avait pour conséquence que  $\widehat{A}$  soit auto-adjoint :  $\widehat{A} = \widehat{A}^\dagger$ , l'ensemble de ces nombres réels, appelé spectre de  $A$ , peut être de deux natures différentes : discret ou continu.

### Spectre discret

Si l'ensemble  $\mathbb{S} = \{a_1, a_2, \dots, a_n, \dots\}$  peut être mis en relation bijective avec  $\mathbb{N}$  ou une partie de  $\mathbb{N}$ , il est alors dit dénombrable, et le spectre de  $\widehat{A}$  qualifié de discret. Les kets propres associés à ces vecteurs propres forment un famille dénombrable de vecteurs indépendants qui engendre l'espace des états, il est clair que cette famille est libre<sup>11</sup> et qu'elle forme donc une base de cet espace. Comme nous l'avons déjà vu cette base est orthonormée si l'on prend soin de normer les différents kets  $|\Psi_n\rangle$ . Avec le formalisme de Dirac cette démonstration devient ludique : en effet, considérons deux kets propres  $|\Psi_n\rangle$  et  $|\Psi_m\rangle$  associés à deux valeurs propres distinctes  $a_n \neq a_m$ , on a respectivement

$$\widehat{A} |\Psi_n\rangle = a_n |\Psi_n\rangle \quad \text{et} \quad \widehat{A} |\Psi_m\rangle = a_m |\Psi_m\rangle$$

<sup>11</sup> C'est une propriété bien connue des familles de vecteurs propres ...

la quantité  $\langle \Psi_m | \widehat{A} | \Psi_n \rangle$  peut être calculée en faisant agir l'opérateur à droite ou à gauche. En n'oubliant pas que  $\widehat{A} = \widehat{A}^\dagger$  et que les valeurs propres sont réelles, on a dans chacun des cas

$$\langle \Psi_m | \widehat{A} | \Psi_n \rangle = \begin{cases} \left( |\Psi_m\rangle, \widehat{A} |\Psi_n\rangle \right) = a_n \left( |\Psi_m\rangle, |\Psi_n\rangle \right) = a_n \langle \Psi_m | \Psi_n \rangle \\ \left( \widehat{A}^\dagger |\Psi_m\rangle, |\Psi_n\rangle \right) = \left( \widehat{A} |\Psi_m\rangle, |\Psi_n\rangle \right) = a_m \langle \Psi_m | \Psi_n \rangle \end{cases}$$

Ces deux résultats doivent être égaux, ainsi

$$a_n \langle \Psi_m | \Psi_n \rangle = a_m \langle \Psi_m | \Psi_n \rangle \Rightarrow (a_n - a_m) \langle \Psi_m | \Psi_n \rangle = 0$$

étant donné que  $a_n \neq a_m$  on a donc  $\langle \Psi_m | \Psi_n \rangle = 0$  : deux kets propres associés à des valeurs propres distinctes sont orthogonaux.

Lorsqu'il existe plusieurs kets propres  $\{ |\Psi_{n1}\rangle, |\Psi_{n2}\rangle, \dots, |\Psi_{n\ell}\rangle, \dots \}$  associés à une même valeur propre  $a_n$ , on dit que cette valeur propre est dégénérée. L'interprétation physique de cette situation est claire : il existe plusieurs états quantiques distincts dans lesquels le résultat d'une mesure de  $A$  donnera la valeur  $a_n$ .

On peut décomposer un état quelconque  $|\Psi\rangle$  sur la base formée par les vecteurs propres de  $\widehat{A}$ . Notons,  $\alpha_{n \in I}$  les coefficients de cette décomposition

$$|\Psi\rangle = \sum_{n \in I} \alpha_n |\Psi_n\rangle$$

La probabilité de transition entre l'état  $|\Psi\rangle$  vers l'état  $|\Psi_m\rangle$  se calcule, comme nous l'avons vu plus haut, par la relation

$$P(|\Psi_t\rangle \rightarrow |\Psi_m\rangle) = |\langle \Psi | \Psi_m \rangle|^2$$

la décomposition du ket  $|\Psi\rangle$  correspond à celle du bra  $\langle \Psi | = \sum_{n \in I} \overline{\alpha_n} \langle \Psi_n |$ , les vecteurs propres étant orthogonaux, le calcul de la probabilité est immédiat, il donne

$$\begin{aligned} P(|\Psi\rangle \rightarrow |\Psi_m\rangle) &= |\langle \Psi | \Psi_m \rangle|^2 \\ &= \left| \sum_{n \in I} \overline{\alpha_n} \langle \Psi_n | \Psi_m \rangle \right|^2 = \left| \sum_{n \in I} \overline{\alpha_n} \delta_{nm} \right|^2 = |\overline{\alpha_m}|^2 = |\alpha_m|^2 \end{aligned}$$

L'état  $|\Psi_m\rangle$  correspond à celui dans lequel le résultat d'une mesure de  $A$  donne  $a_m$ , nous pouvons donc en conclure que

- si la valeur propre  $a_n$  n'est pas dégénérée, il existe un seul vecteur propre  $|\Psi_n\rangle$  correspondant, et

$$P(a_n) = |\langle \Psi | \Psi_n \rangle|^2 = |\alpha_n|^2$$

- si la valeur propre est dégénérée, il existe une collection d'états quantiques  $\{\dots, |\Psi_{n\ell}\rangle, \dots\}$  correspondant à cette valeur propre, chacun de ces états indépendants réalise une mesure de  $A$  ayant  $a_n$  pour résultat. La formule de la probabilité des évènements indépendants assure que

$$P(a_n) = \sum_{\ell} |\langle \Psi | \Psi_{n\ell} \rangle|^2 = \sum_{\ell} |\alpha_{n\ell}|^2$$

N'oublions pas que c'est le fait de mesurer qui va mettre le système dans l'état correspondant au résultat de la mesure. Avant celle-ci, la seule chose à laquelle nous avons accès est la probabilité de transition entre les divers états. Tout ceci commence à prendre forme...

## Spectre continu

Le cas discret de la section précédente est le plus simple, il n'est cependant pas le plus général. Envisageons, par exemple le cas de l'opérateur impulsion  $\hat{P}$  en représentation  $q$ , et plaçons nous pour simplifier en dimension 1 : le système possède donc une coordonnée généralisée  $q$  conjuguée à une impulsion  $p$ . L'opérateur  $\hat{P}$  est selon les postulats associé à la grandeur physique  $p$ . Nous avons vu en 9.1.4 que cet opérateur agissait en dérivant sur un état  $|\Psi\rangle$  en le multipliant par  $-i\hbar$  et en le dérivant par rapport à la variable  $q$ . On montre facilement que cet opérateur est auto-adjoint. Un vecteur propre  $|\pi_a\rangle$  associé à la valeur propre  $a$  de cet opérateur devra vérifier

$$\hat{P}|\pi_a\rangle = -i\hbar \frac{d|\pi_a\rangle}{dq} = a|\pi_a\rangle \quad \forall q \in \mathbb{R} \quad (9.13)$$

la solution de cette équation est immédiate

$$\forall q \in \mathbb{R} \quad |\pi_a\rangle = e^{iaq/\hbar} |\pi_0\rangle$$

ou le ket  $|\pi_0\rangle$  est indépendant de  $q$ . Si  $a \in \mathcal{I}$  où  $\mathcal{I}$  est un intervalle ouvert de  $\mathbb{R}$  l'ensemble de ces vecteurs propres potentiels est donc  $\mathcal{E} = \{|\pi_a\rangle = e^{iaq/\hbar} |\pi_0\rangle, a \in \mathcal{I}\}$ . Chaque valeur de  $a$  définit un élément de  $\mathcal{E}$  qui est lui-même une fonction de  $q$  définie sur  $\mathbb{R}$ , il est donc clair que  $\mathcal{E}$  n'est pas dénombrable et ne peut donc en aucun cas constituer une base de quoi que

ce soit. Il s'agit là d'une différence fondamentale avec le cas discret. Pour marquer cette différence on dit que  $|\pi_a\rangle$  est un pseudo-vecteur propre ou un vecteur propre généralisé de l'opérateur  $\hat{P}$ .

Il n'est pas question de restreindre le domaine d'application de la mécanique quantique au cas des observables associées à des opérateurs ayant un spectre discret ! Il faut simplement faire attention à ce que l'on est capable d'écrire dans le cas d'un spectre continu...

En considérant deux valeurs propres différentes  $a \neq a'$ , et en généralisant les notations précédentes nous aurons toujours

$$\langle \pi_{a'} | \hat{P} | \pi_a \rangle = \begin{cases} a \langle \pi_{a'} | \pi_a \rangle \\ a' \langle \pi_{a'} | \pi_a \rangle \end{cases}$$

suivant que  $\hat{P}$  agisse sur le ket  $|\pi_a\rangle$  ou sur le bra  $\langle \pi_{a'}|$ , il est alors facile d'en conclure que

$$(a - a') \langle \pi_{a'} | \pi_a \rangle = 0$$

et donc

$$\langle \pi_{a'} | \pi_a \rangle = 0 \text{ si } a \neq a'$$

Ce résultat nous permet de tenter de généraliser la relation de fermeture discrète à sa consœur continue

$$\text{Id} = \int_{a \in \mathcal{I}} |\pi_a\rangle \langle \pi_a| da \quad (9.14)$$

Un état quelconque sera ainsi associé au ket

$$|\Psi\rangle = \int_{a \in \mathcal{I}} |\pi_a\rangle \langle \pi_a | \Psi \rangle da \quad (9.15)$$

en projetant sur  $|\pi_{a'}\rangle$ , il vient

$$\langle \pi_{a'} | \Psi \rangle = \int_{a \in \mathcal{I}} \langle \pi_{a'} | \pi_a \rangle \langle \pi_a | \Psi \rangle da$$

C'est à ce moment précis qu'apparaît l'affaire ...

Nous avons vu que la quantité  $\langle \pi_{a'} | \pi_a \rangle$  est nulle si  $a \neq a'$  mais, lorsque l'on calcule l'intégrale sur  $a$  de son produit avec une quantité qui dépend de  $a$ , on trouve la valeur de cette quantité en  $a = a'$ . C'est ainsi que P.A.M. Dirac inventa dans les années 30, la distribution de Dirac !

$$\delta(a - a') = \langle \pi_{a'} | \pi_a \rangle$$

qui est donc telle que pour toute fonction  $\varphi(a)$  définie sur  $\mathcal{I}$  on a

$$\varphi(a) = \int_{a' \in \mathcal{I}} \delta(a - a') \varphi(a') da'$$

en prenant pour  $\varphi$  la fonction qui vaut 1 sur  $\mathcal{I}$  et 0 sur  $\mathbb{R} \setminus \mathcal{I}$ , une intégration par partie montre même que

$$1 = \int_{-\infty}^{+\infty} \delta(a) da$$

On remarque donc que si les vecteurs propres généralisés  $|\pi_a\rangle$  sont bien orthogonaux entre eux, ils ne sont pas normés, ni même normalisables. La seule «valeur» que l'on pourrait attribuer à la quantité  $\langle \pi_a | \pi_a \rangle$  est  $\delta(0)$  qui semble infinie.

Ainsi  $|\pi_a\rangle$  n'est pas un vecteur de l'espace de Hilbert des états quantiques du système considéré, mais la relation de fermeture (9.14) et sa conséquence (9.15) nous montrent comment décomposer tout de même un état quantique sur les éléments de  $\mathcal{E}$ .

### 9.2.5 Fonction d'onde

L'excursion hasardeuse dans le domaine du spectre continu des opérateurs linéaires va nous permettre de donner naissance à une formulation efficace de la mécanique quantique à travers sa représentation sous forme de fonctions. Dans le chapitre introductif sur les postulats de la mécanique quantique, nous avons mentionné le fait qu'il serait sympathique de pouvoir relier la notion abstraite d'état quantique à la notion tout aussi abstraite, mais plus familière, de fonction de carré sommable. Tentons d'éclaircir ce mystère...

Afin de simplifier les notations, considérons un système physique possédant un seul degré de liberté, pour lequel on peut donc définir une coordonnée généralisée  $q$ . Comme nous l'avons vu lors des postulats, cette grandeur physique<sup>12</sup> est associée à un opérateur auto-adjoint  $\hat{Q}$ . Supposons que lors d'une mesure nous trouvions  $q = a$ , selon les principes quantiques nous en déduisons qu'immédiatement après cette mesure, le système se trouve dans l'état  $|\chi_a\rangle$  tel que

$$\hat{Q}|\chi_a\rangle = a|\chi_a\rangle$$

Généralement,  $a$  peut varier continuellement dans un intervalle  $\mathcal{I} \in \mathbb{R}$ . Il est donc clair que  $\hat{Q}$  possède généralement un spectre continu.

<sup>12</sup> Les experts de la mécanique analytique que nous sommes savent maintenant qu'il peut s'agir

## Représentation $q$

Plaçons-nous en représentation  $q$ . Nous avons vu en 9.1.4 que l'action de  $\widehat{Q}$  sur un état revenait à multiplier cet état par la grandeur physique  $q$ , l'équation propre précédente s'écrit donc dans cette représentation

$$\widehat{Q}|\chi_a\rangle = q|\chi_a\rangle = a|\chi_a\rangle \quad \forall q \in \mathbb{R}$$

Le vecteur propre  $|\chi_a\rangle$  qui intervient dans cette équation propre fait donc à nouveau intervenir la distribution de Dirac : il sélectionne la valeur  $q = a$  parmi toutes les valeurs possibles de  $q$ . En utilisant les résultats du paragraphe précédent on écrira donc

$$\forall q \in \mathbb{R} \quad |\chi_a\rangle = \delta(q - a) |\chi_0\rangle$$

Une fois de plus cet état n'est pas dans l'espace des états, mais comme nous l'avons fait plus haut on peut s'en servir pour décomposer les états accessibles au système en écrivant

$$|\Psi\rangle = \int |\chi_a\rangle \langle \chi_a | \Psi \rangle da$$

Pour chaque valeur de  $a$  le vecteur propre généralisé correspondant est une fonction de  $q$ , il en est donc de même pour la quantité  $\langle \chi_a | \Psi \rangle$ . Pour affirmer ce caractère fonctionnel posons

$$\langle \chi_a | \Psi \rangle = \psi_a(q)$$

Le carré de la norme de  $|\Psi\rangle$  s'écrit avec ces notations

$$\begin{aligned} (|\Psi\rangle, |\Psi\rangle) &= \left( |\Psi\rangle, \int |\chi_a\rangle \langle \chi_a | \Psi \rangle da \right) \\ &= \int \langle \Psi | \chi_a \rangle \langle \chi_a | \Psi \rangle da \\ &= \int \overline{\langle \chi_a | \Psi \rangle} \langle \chi_a | \Psi \rangle da \\ &= \int \overline{\psi_a(q)} \psi_a(q) da \\ &= \int |\psi_a(q)|^2 da \end{aligned}$$

et nous donnons donc un sens à la représentation de l'état quantique  $|\Psi\rangle$  par une fonction de la coordonnée généralisée  $q$  : la fonction d'onde  $\psi_a(q)$  est la projection sur un vecteur propre généralisé de l'opérateur position du vecteur de l'espace de Hilbert décrivant le système à l'instant  $t$ . En toute rigueur cette fonction dépend également du temps, nous traiterons ce problème très bientôt.

Cette fonction s'interprète comme nous l'avons vu comme la densité de probabilité associée à la mesure de la grandeur physique  $q$ . Si tous les résultats possibles de la mesure de  $q$  sont envisagés par les variations du paramètre  $a$ , et si l'on a pris soin de normaliser l'état  $|\Psi\rangle$  on retrouve bien

$$1 = \langle\Psi|\Psi\rangle = \int |\psi_a(q)|^2 da = 1$$

De cette manière, en représentation  $q$ , on peut donc «représenter» à chaque instant l'espace des états  $\mathbb{H}$  d'un système quantique par une collection de fonctions d'ondes de carré sommable de la variable  $q$ .

## Représentation $p$

Les mêmes considérations peuvent s'appliquer à ce même opérateur  $\widehat{Q}$ , mais cette fois-ci en représentation  $p$ . Nous avons vu qu'à ce moment là, son action était similaire à celle de  $\widehat{P}$  en représentation  $q$ . En notant toujours  $|\chi_a\rangle$  le vecteur propre généralisé associé à la valeur propre  $a$  de l'opérateur  $\widehat{Q}$ , nous avons maintenant

$$\widehat{Q}|\chi_a\rangle = i\hbar \frac{d|\chi_a\rangle}{dp} = a|\chi_a\rangle \quad \forall p \in \mathbb{R}$$

et nous revoici revenu au paragraphe précédent (équation (9.13)) en changeant un signe et la variable... Les mêmes calculs fournissent alors les mêmes résultats ! On trouve a présent

$$\forall p \in \mathbb{R} \quad |\chi_a\rangle = e^{-iap/\hbar} |\chi_0\rangle$$

ou le ket  $|\chi_0\rangle$  est indépendant de  $p$ . En reprenant les considérations effectuées en représentation  $q$  et pour un état quelconque  $|\Phi\rangle$  on peut définir la fonction d'onde en impulsion

$$\langle\chi_a|\Phi\rangle = \phi_a(p)$$

---

de la position sur un axe, d'un angle, etc...

des calculs déjà faits ou que l'on pourrait refaire dans le nouveau formalisme montrent assez vite que  $\phi_a(p)$  est la transformée de Fourier de  $\psi_a(p)$ .

$$\phi_a(p) = \frac{1}{(2\pi\hbar)^{1/2}} \int dq \psi_a(q) \exp\left(-\frac{i}{\hbar}pq\right)$$

Tout ceci donne un sens aux représentations  $q$  et  $p$  introduites «à la main» à la section 9.1.4.

La généralisation à un système possédant  $\ell$  degrés de liberté est immédiate, la fonction d'onde en position dépend de la variable  $\mathbf{q} = [q_{i=1,\dots,\ell}]^T$  la fonction d'onde en impulsion dépend de la variable  $\mathbf{p} = [p_{i=1,\dots,\ell}]^T$  et la transformée de Fourier se passe dans  $\mathbb{R}^\ell$ ...

### 9.2.6 Commutation, compatibilité et indétermination.

Revenons sur terre et considérons des opérateurs auto-adjoints  $\hat{A}$  et  $\hat{B}$  dont le spectre est discret et non dégénéré : chaque valeur propre est associée à un seul vecteur propre. La famille formée par ces vecteurs propres est une base orthonormée de l'espace des états.

Supposons que les kets propres  $|u_n \in I \subset \mathbb{N}\rangle$  de  $\hat{A}$  soient les mêmes que ceux de  $\hat{B}$ , sans pour autant qu'ils aient les mêmes valeurs propres. Nous avons donc

$$\forall n \in I \subset \mathbb{N}, \begin{cases} \hat{A}|u_n\rangle = a_n|u_n\rangle \text{ avec } a_n \in \mathbb{R} \\ \hat{B}|u_n\rangle = b_n|u_n\rangle \text{ avec } b_n \in \mathbb{R} \end{cases}$$

En utilisant l'orthonormalité des vecteurs propres, on écrit

$$\begin{cases} \hat{A}|u_n\rangle = \sum_m a_m |u_m\rangle \delta_{mn} = \sum_m a_m |u_m\rangle \langle u_m|u_n\rangle \Rightarrow \hat{A} = \sum_m a_m |u_m\rangle \langle u_m| \\ \hat{B}|u_n\rangle = \sum_\ell b_\ell |u_\ell\rangle \delta_{\ell n} = \sum_\ell b_\ell |u_\ell\rangle \langle u_\ell|u_n\rangle \Rightarrow \hat{B} = \sum_\ell b_\ell |u_\ell\rangle \langle u_\ell| \end{cases}$$

On peut à présent évaluer le produit  $\hat{A}\hat{B}$ , il vient

$$\hat{A}\hat{B} = \left( \sum_m a_m |u_m\rangle \langle u_m| \right) \left( \sum_\ell b_\ell |u_\ell\rangle \langle u_\ell| \right) = \sum_{m,\ell} a_m b_\ell |u_m\rangle \langle u_m|u_\ell\rangle \langle u_\ell|$$

les vecteurs propres étant orthonormés, on a

$$\hat{A}\hat{B} = \sum_{m,\ell} a_m b_\ell |u_m\rangle \delta_{m\ell} \langle u_\ell| = \sum_m a_m b_m |u_m\rangle \langle u_m|$$

de la même façon on calcule le produit  $\widehat{B}\widehat{A}$ , et l'on obtient

$$\widehat{B}\widehat{A} = \sum_{\ell} a_{\ell} b_{\ell} |u_{\ell}\rangle \langle u_{\ell}|$$

Soit finalement,

$$\widehat{A}\widehat{B} = \widehat{B}\widehat{A} \text{ soit } [\widehat{A}, \widehat{B}] = 0$$

Réciproquement, supposons que les opérateurs  $\widehat{A}$  et  $\widehat{B}$  commutent. Notons  $|u_{n \in I \subset \mathbb{N}}\rangle$  les kets propres de  $\widehat{A}$ , on a donc

$$\forall n \in I \subset \mathbb{N}, \widehat{A}|u_n\rangle = a_n |u_n\rangle$$

Par hypothèse  $\widehat{A}\widehat{B} = \widehat{B}\widehat{A}$ , pour chaque valeur de  $n \in I$  on peut donc écrire

$$\begin{aligned} \widehat{A}(\widehat{B}|u_n\rangle) &= \widehat{B}(\widehat{A}|u_n\rangle) \\ &= \widehat{B}(a_n |u_n\rangle) \\ &= a_n \widehat{B}|u_n\rangle \end{aligned}$$

cette relation énonce que le ket  $\widehat{B}|u_n\rangle$  est valeur propre de l'opérateur  $\widehat{A}$  avec la valeur propre  $a_n$ . Le ket  $\widehat{B}|u_n\rangle$  appartient donc à l'espace vectoriel engendré par le ket  $|u_n\rangle$ , il existe donc une valeur  $b_n \in \mathbb{C}$  telle que

$$\widehat{B}|u_n\rangle = b_n |u_n\rangle$$

En projetant cette expression sur  $|u_n\rangle$  on obtient

$$\langle u_n | \widehat{B} | u_n \rangle = b_n \langle u_n | u_n \rangle = b_n$$

car  $|u_n\rangle$  est de norme unité, ainsi en conjuguant

$$\overline{b_n} = \overline{\langle u_n | \widehat{B} | u_n \rangle} = \langle u_n | \widehat{B}^{\dagger} | u_n \rangle = \langle u_n | \widehat{B} | u_n \rangle = b_n$$

car  $\widehat{B}$  est auto-adjoint, ainsi  $b_n \in \mathbb{R}$  et nous pouvons conclure que le ket  $|u_n\rangle$  est un ket propre de l'opérateur  $\widehat{B}$ . On note que généralement  $b_n \neq a_n$ .

On peut généraliser, sans trop de peine, ces démonstrations au cas où<sup>1</sup> le spectre est toujours discret mais dégénéré, c'est-à-dire que le sous-espace propre associé à chaque valeur propre n'est pas nécessairement de dimension 1 : pour chaque entier  $n$ , il existe  $k$  kets propres indépendants et orthonormés

$|u_{nk}\rangle$  associés à la valeur propre  $a_n$  de  $\hat{A}$  ou  $b_n$  de  $\hat{B}$ . Si le spectre est continu ou contient une partie continue, on reproduit également les mêmes calculs en considérant la relation de fermeture continue et en utilisant la distribution de Dirac. On peut donc énoncer le résultat fondamental : Deux opérateurs auto-adjoints possèdent la même famille de vecteurs propres si et seulement s'ils commutent.

On dit souvent que l'on peut les diagonaliser dans la même base. Notons finalement, qu'en observant la démonstration on peut même lever la contrainte des opérateurs auto-adjoints...

Ce théorème bien connu des mathématiciens prend dans le contexte de la mécanique quantique une envergure insoupçonnée : dans cette théorie le spectre d'un opérateur auto-adjoint est l'ensemble des valeurs que peut prendre l'observable correspondante dans un processus de mesure. Si l'on place un système dans l'état  $|u_n\rangle$ , le résultat de la mesure respective des grandeurs physiques  $A$  ou  $B$  sera respectivement  $a_n$  ou  $b_n$ . Un seul état permet la mesure de deux grandeurs physiques distinctes. On dit que ces deux grandeurs sont *compatibles*.

Avec des mots, on a envie de dire que la mesure de l'une n'influence pas le résultat de l'autre. C'est en effet bien ce qui se passe ! Considérons un système dans un état quelconque  $|\Psi\rangle$ , admettons que la mesure d'une observable  $A$  donne le résultat  $a_n$ . Dans le cadre des postulats de la mécanique quantique cela signifie que le processus de mesure a placé le système dans l'état  $|u_n\rangle$ . A priori, la seule chose que nous pouvions déterminer était la probabilité de transition de l'état  $|\Psi\rangle$  vers l'état  $|u_n\rangle$ , donnée par la relation

$$P(|\Psi\rangle \rightarrow |u_n\rangle) = |\langle u_n|\Psi\rangle|^2$$

A posteriori, c'est-à-dire après la mesure de  $a_n$  pour  $A$ , nous sommes sûr que le système se trouve dans l'état  $|u_n\rangle$ . On peut donc affirmer avec certitude que l'on va trouver le résultat  $b_n$  lors d'une mesure de l'observable  $B$ . Il n'en serait évidemment pas de même si les opérateurs  $A$  et  $B$  ne commutaient pas... et serait donc incompatibles. On peut ainsi «préparer » la mesure d'une grandeur.

En prenant le contre-pied de ce résultat, on peut tout aussi fièrement annoncer que si deux observables sont incompatibles la mesure de l'une va rendre la mesure successive de l'autre indéterminée. Il s'agit là d'une propriété fondamentale de la mécanique quantique que l'on peut appeler principe d'indétermination de Heisenberg.

D'une manière générale, ce résultat permet d'entrevoir une méthode d'investigation des problèmes quantiques. Celle-ci consistera à chercher des

ensembles d'opérateurs qui commutent deux à deux. On comprend assez bien que certains cas de dégénérescence peuvent induire le fait que le nombre de vecteurs communs ne soit pas suffisant pour constituer une base de l'espace des états. C'est pourquoi, dans la pratique, on se limite à la recherche d'ensemble *complets* d'opérateurs qui commutent deux à deux, c'est-à-dire un ensemble d'opérateurs qui partagent la même famille de vecteurs propres, cette famille formant une base de l'espace des états et sur laquelle il va donc être possible de décomposer un état quelconque du système. Par abus de langage, on parlera aussi d'ensemble complet d'observables qui commutent, un ECOC.

### 9.2.7 Évolution temporelle des observables

Notre troisième postulat de la mécanique quantique nous indique l'équation d'évolution temporelle d'un état  $|\Psi_t\rangle$ .

$$i\hbar \frac{d|\Psi_t\rangle}{dt} = \widehat{H} |\Psi_t\rangle \quad (9.16)$$

où  $\hbar$  est une constante que nous n'avons toujours pas fixée mais qui possède la propriété de rendre le produit  $\mathbf{p}\cdot\mathbf{q}/\hbar$  sans dimension, et  $\widehat{H}$  l'opérateur associé à l'observable énergie. Comme nous l'avons vu, cette équation n'est pas, comme cela est souvent présenté, le résultat d'une approche divinatoire ou heuristique de la physique. Elle est la conséquence de deux partis pris acceptables.

1. si l'équation n'était pas linéaire, nous perdriions la structure d'espace vectoriel nécessaire à la construction de la théorie et à sa justification en terme d'un principe de superposition;
2. l'hypothèse de linéarité de l'opérateur  $\widehat{H}$  permet (cf. 9.1.3) en utilisant la notion de moyenne d'une grandeur physique de retrouver la mécanique classique en tant qu'une certaine moyenne de la mécanique quantique.

L'équation étant posée, «justifiée» et cohérente avec la mécanique classique étudions son contenu physique. Nous envisagerons les deux cas possibles et n'en traiterons qu'un...

#### L'opérateur $\widehat{H}$ dépend du temps

La base de l'espace des états sur lequel on décompose les diverses quantités qui intéressent le mécanicien quantique change au fil du temps. La situation est complexe mais pas désespérée!

## L'opérateur $\widehat{H}$ est indépendant du temps

Il existe une base indépendante du temps sur laquelle on peut effectuer les décompositions des vecteurs associés aux différents états quantiques. Lors de l'évolution temporelle, seules les composantes des vecteurs sur cette base, stationnaire et fondamentale, sont des fonctions du temps. La situation est plus simple mais ce n'est pas si facile que cela...

### *État quelconque*

L'interprétation de la mécanique classique en tant que moyenne de la mécanique quantique nous avait conduit à interpréter l'opérateur auto-adjoint d'évolution temporelle  $\widehat{H}$  comme celui associé à l'énergie contenue dans le système.

Les valeurs propres de  $\widehat{H}$  sont donc les différentes valeurs expérimentales mesurées pour l'énergie du système. Supposons dans un premier temps que ce spectre soit discret :  $\{E_{n \in I \subset \mathbb{N}}\}$  chacune de ces valeurs propres peut être dégénérée et donc associée à une ensemble  $\{|\varphi_{n,k}\rangle\}$  de vecteurs propres. De façon générale, on peut écrire

$$\widehat{H} |\varphi_{n,k}\rangle = E_n |\varphi_{n,k}\rangle \quad (9.17)$$

L'ensemble  $\mathcal{B}_E = \{n \in I \subset \mathbb{N}, k \in I' \subset I, |\varphi_{n,k}\rangle\}$  de tous les vecteurs propres de  $\widehat{H}$  forme une base orthonormée de l'espace des états.

Le fait que  $\widehat{H}$  soit indépendant du temps implique que ses valeurs propres  $E_n$  et tous les vecteurs propres associés  $|\varphi_{n,k}\rangle$  sont indépendants du temps. Considérons un état quelconque du système décrit à l'instant  $t$  par le ket  $|\Psi_t\rangle$ . On peut toujours décomposer cet état sur la base  $\mathcal{B}_E$ . Attendu que  $|\varphi_{n,k}\rangle$  est nécessairement indépendant du temps, il vient

$$|\Psi_t\rangle = \sum_{n,k} c_{n,k}(t) |\varphi_{n,k}\rangle$$

On retrouve ici l'affirmation d'il y a quelques lignes : dans  $\mathcal{B}_E$  seules les composantes de  $|\Psi_t\rangle$  évoluent au cours du temps mais pas  $\mathcal{B}_E$ . Nos investigations précédentes nous permettent maintenant d'affirmer que les fonctions  $c_{n,k}(t)$  sont les projections de  $|\Psi_t\rangle$  sur les vecteurs de la base  $\mathcal{B}_E$  soit

$$c_{n,k}(t) = \langle \varphi_{n,k} | \Psi_t \rangle$$

en dérivant cette expression par rapport au temps il vient

$$\begin{aligned} \frac{dc_{n,k}(t)}{dt} &= \frac{d}{dt} \langle \varphi_{n,k} | \Psi_t \rangle \\ &= \frac{d}{dt} ( | \varphi_{n,k} \rangle , | \Psi_t \rangle ) \\ &= \left( \frac{d | \varphi_{n,k} \rangle}{dt} , | \Psi_t \rangle \right) + \left( | \varphi_{n,k} \rangle , \frac{d | \Psi_t \rangle}{dt} \right) \end{aligned}$$

comme  $|\varphi_{n,k}\rangle$  est indépendant du temps, en utilisant la relation (9.16) et en revenant à la notation de Dirac, il vient

$$\frac{dc_{n,k}(t)}{dt} = \frac{1}{i\hbar} \langle \varphi_{n,k} | \widehat{H} | \Psi_t \rangle$$

en faisant agir l'opérateur auto-adjoint  $\widehat{H}$  sur le bra  $\langle \varphi_{n,k} |$  et en utilisant la relation (9.17), on trouve finalement

$$\begin{aligned} \frac{dc_{n,k}(t)}{dt} &= \frac{E_n}{i\hbar} \langle \varphi_{n,k} | \Psi_t \rangle \\ &= \frac{E_n}{i\hbar} c_{n,k}(t) = -i \frac{E_n}{\hbar} c_{n,k}(t) \end{aligned}$$

cette équation différentielle est à la portée de n'importe quel élève de terminale  $S$ , on trouve

$$c_{n,k}(t) = c_{n,k}(t_0) e^{-i \frac{E_n}{\hbar} (t-t_0)}$$

Nous en déduisons l'expression de l'état  $|\Psi_t\rangle$  sous la forme

$$|\Psi_t\rangle = \sum_{n,k} c_{n,k}(t_0) e^{-i \frac{E_n}{\hbar} (t-t_0)} |\varphi_{n,k}\rangle$$

avec bien évidemment

$$c_{n,k}(t_0) = \langle \varphi_{n,k} | \Psi_{t_0} \rangle$$

Tout ceci peut paraître bien compliqué mais quelques exemples bien choisis pourront illustrer une situation en fait assez sympathique : la connaissance des valeurs propres et des vecteurs propres de l'opérateur  $\widehat{H}$ , malheureusement appelé hamiltonien, suffit à la connaissance de l'état quantique du système à chaque instant.

Dans le cas continu la situation est souvent bien plus sulfureuse mais, en faisant attention, on peut obtenir avec des notations évidentes

$$|\Psi_t\rangle = \sum_k \int dE c_n(E, t_0) e^{-i \frac{E_n}{\hbar} (t-t_0)} |\varphi_{E,k}\rangle$$

*État stationnaire*

Un cas particulièrement intéressant est celui dans lequel l'état quantique considéré est lui-même un vecteur propre du hamiltonien du système. Cet état  $|\Psi_t\rangle$  est donc tel qu'il existe une valeur  $E_m$  dans le spectre de  $\hat{H}$  telle que

$$\hat{H} |\Psi_t\rangle = E_m |\Psi_t\rangle$$

cela signifie que  $|\Psi_t\rangle$  est un élément de l'espace vectoriel engendré par les vecteurs propres<sup>13</sup> de  $\hat{H}$  associés à  $E_m$  et que nous avons notés

$$\{|\varphi_{m,1}\rangle, \dots, |\varphi_{m,k}\rangle, \dots\}$$

Nous pouvons donc écrire à chaque instant

$$|\Psi_t\rangle = \sum_k c_{m,k}(t) |\varphi_{m,k}\rangle$$

la même analyse que celle menée dans le cas général redonne

$$c_{m,k}(t) = c_{m,k}(t_0) e^{-i\frac{E_m}{\hbar}(t-t_0)} \quad \text{avec } c_{m,k}(t_0) = \langle \varphi_{m,k} | \Psi_{t_0} \rangle$$

et donc

$$\begin{aligned} |\Psi_t\rangle &= \sum_k c_{m,k}(t_0) e^{-i\frac{E_m}{\hbar}(t-t_0)} |\varphi_{m,k}\rangle \\ &= e^{-i\frac{E_m}{\hbar}(t-t_0)} \sum_k c_{m,k}(t_0) |\varphi_{m,k}\rangle \end{aligned}$$

on en déduit immédiatement que

$$|\Psi_{t_0}\rangle = \sum_k c_{m,k}(t_0) |\varphi_{m,k}\rangle$$

ainsi

$$|\Psi_t\rangle = |\Psi_{t_0}\rangle e^{-i\frac{E_m}{\hbar}(t-t_0)}$$

les kets  $|\Psi_t\rangle$  et  $|\Psi_{t_0}\rangle$  ne diffèrent donc que d'un facteur de phase de module unité. La probabilité de transition d'un état quelconque  $|\Phi\rangle$  vers l'un ou l'autre des états  $|\Psi_t\rangle$  ou  $|\Psi_{t_0}\rangle$  est donc la même. On dit que les états  $|\Psi_t\rangle$  et  $|\Psi_{t_0}\rangle$  sont indiscernables d'un point de vue quantique.

<sup>13</sup> Si cette valeur propre n'est pas dégénérée, cet espace est une droite et  $|\Psi_t\rangle$  est simplement

L'état du système est indépendant du temps, il est qualifié de *stationnaire*.

Un dernier mot sur  $\hbar$  : nous avons déjà identifié que cette constante réelle devait être dimensionnée en Joule×seconde pour adimensionner l'argument de l'exponentielle dans la transformée de Fourier de la fonction d'onde en variable  $\mathbf{q}$ . On retrouve à présent ce résultat de manière encore plus explicite dans l'argument de  $\exp[-iE_m(t-t_0)/\hbar]$ . La valeur de cette constante n'est toujours pas fixée par la théorie elle le sera expérimentalement.

## 9.3 La mécanique quantique par le théorème de Noether

### 9.3.1 Le lagrangien quantique

La théorie du champ de Schrödinger a été mise sur pied par Richard Feynman au milieu du XX<sup>e</sup> siècle. Elle illustre la puissance de la théorie des champs tout en lui conférant un aspect magique. Cette aspect mystérieux réside dans l'écriture du lagrangien qui ne répond qu'à peu de règles outre celle de fournir le bon résultat. Les symétries que doit posséder ce lagrangien sont alors un guide essentiel dans sa construction.

La mécanique quantique non relativiste n'est quant à elle pas aussi magique que certains auteurs veulent bien le laisser croire. Comme nous l'avons repris dans les deux sections précédentes, dans sa quantification canonique, Dirac en 1927 explique comment on peut construire l'équation de Schrödinger à partir de la mécanique classique de façon complètement logique en changeant simplement d'algèbre. Dans cette quantification canonique, les crochets de Poisson classiques deviennent les commutateurs quantiques. Les grandes nouveautés quantiques sont en fait les axiomes d'état et de mesure, l'évolution de ces états n'a quant à elle rien de nouveau, puisqu'elle se trouve dans le droit fil des équations de la physique. Il n'est donc pas étonnant que l'on puisse étendre naturellement le formalisme de moindre action classique à la mécanique quantique.

En suivant Feynman<sup>14</sup>, et sans autre forme de procès, le lagrangien quantique s'écrit

$$L(\psi, \dot{\psi}) = \int d\mathbf{q} \left[ \frac{i\hbar}{2} (\psi^* \dot{\psi} - \dot{\psi}^* \psi) - \frac{\hbar^2}{2m} \nabla\psi \cdot \nabla\psi^* - V\psi\psi^* \right] \quad (9.18)$$

proportionnel à  $|\varphi_m\rangle$  qui est le seul vecteur propre.

<sup>14</sup> Pour trouver les éléments de la construction de ce lagrangien nous préférons renvoyer le lecteur

Les variables canoniques indépendantes sont la fonction  $\psi(\mathbf{q})$  et sa dérivée temporelle  $\dot{\psi}(\mathbf{q}) = \frac{\partial \psi}{\partial t}(\mathbf{q})$ . Ces fonctions sont à valeurs complexes,  $\mathbf{q}$  est le vecteur coordonnées généralisée qui inclut le temps. La fonction  $V(\mathbf{q})$  décrit le potentiel présent dans la région considérée.

La forme (9.18) est complètement symétrique, on peut en signaler une forme qui l'est moins en écrivant

$$\begin{aligned} L(\psi, \dot{\psi}) &= \int d\mathbf{q} \left[ \frac{i\hbar}{2} \left( 2\psi^* \dot{\psi} - \frac{d}{dt}(\psi^* \psi) \right) - \frac{\hbar^2}{2m} \nabla \psi \cdot \nabla \psi^* - V \psi \psi^* \right] \\ &= \int d\mathbf{q} \left[ i\hbar \psi^* \dot{\psi} - \frac{\hbar^2}{2m} \nabla \psi \cdot \nabla \psi^* - V \psi \psi^* \right] + \frac{d\phi}{dt} \end{aligned}$$

où nous avons posé  $\phi(\psi) = \frac{i\hbar}{2} \int dx \psi^* \psi$ . Sans préjuger de sa valeur, il est clair que  $\phi(\psi)$  une jauge au sens défini depuis le début de cet ouvrage (voir section 2.2.2 page 29). On peut donc tout aussi bien considérer le lagrangien

$$L'(\psi, \dot{\psi}) = \int d\mathbf{q} \left[ i\hbar \psi^* \dot{\psi} - \frac{\hbar^2}{2m} \nabla \psi \cdot \nabla \psi^* - V \psi \psi^* \right] \quad (9.19)$$

pour décrire l'évolution d'un système quantique, c'est ce que nous ferons dans la seconde partie de cette section. Dans la première partie nous utiliserons par contre l'expression (9.18) complètement symétrique.

L'action de Schrödinger est définie naturellement comme

$$S(t_1, t_2) = \int dt L(\psi, \dot{\psi}) \quad (9.20)$$

$$= \int d\mathbf{q} \left[ \frac{i\hbar}{2} (\psi^* \dot{\psi} - \dot{\psi}^* \psi) - \frac{\hbar^2}{2m} \nabla \psi \cdot \nabla \psi^* - V \psi \psi^* \right] \quad (9.21)$$

Comme nous l'avons déjà précisé, tout le génie du physicien consiste à trouver le bon lagrangien, même s'il ne dispose pour cela que de peu de règles plus ou moins formelles (principe de covariance, de simplicité, symétries, etc.). En cas de succès, comme le faisait remarquer Feynman lui-même, il est souvent récompensé par le prix Nobel!

Une fois le lagrangien écrit la procédure est toujours la même : variation de l'action avec contraintes aux bords, intégration par parties avec annulation des termes tout intégrés grâce aux contraintes, factorisation de l'élément variationnel et identification de l'équation d'évolution. C'est parti !

---

vers la source originale [6] ou le livre de synthèse [3]. Ce type de construction relève en fait du réel travail du physicien...

On cherche les fonctions  $\psi(\mathbf{q})$  qui minimisent l'action (9.21) et qui sont complètement déterminées en  $\mathbf{q}_1$  et  $\mathbf{q}_2$ , ce que l'on peut écrire  $\delta\psi|_{\mathbf{q}_1} = \delta\psi|_{\mathbf{q}_2} = 0$ . Il vient

$$\delta S = \int d\mathbf{q} \left[ \frac{i\hbar}{2} \left( \delta\psi^* \dot{\psi} + \psi^* \delta\dot{\psi} - \delta\dot{\psi}^* \psi - \dot{\psi}^* \delta\psi \right) - \frac{\hbar^2}{2m} (\nabla\delta\psi \cdot \nabla\psi^* + \nabla\psi \cdot \nabla\delta\psi^*) - V (\delta\psi\psi^* + \psi\delta\psi^*) \right]$$

Une intégration par parties temporelle sur les deux termes centraux de la première partie de cette variation permet d'obtenir

$$\delta S = \int d\mathbf{q} \left[ i\hbar \left( \dot{\psi} \delta\psi^* - \dot{\psi}^* \delta\psi \right) - \frac{\hbar^2}{2m} (\nabla\delta\psi \cdot \nabla\psi^* + \nabla\psi \cdot \nabla\delta\psi^*) - V (\delta\psi\psi^* + \psi\delta\psi^*) \right]$$

Le terme de bord est nul grâce aux conditions imposées à la variation.

Une intégration par parties spatiale (formule de Green) sur les deux produits scalaires de gradient donne alors

$$\delta S = \int d\mathbf{q} \left[ i\hbar \left( \dot{\psi} \delta\psi^* - \dot{\psi}^* \delta\psi \right) + \frac{\hbar^2}{2m} (\delta\psi \Delta\psi^* + \delta\psi^* \Delta\psi) - V (\delta\psi\psi^* + \psi\delta\psi^*) \right]$$

Le terme de bord est nul car on fait l'hypothèse que  $\psi$ ,  $\psi^*$  et leurs gradients sont nuls sur le bord du domaine spatial<sup>15</sup>. On peut à présent factoriser  $\delta\psi$  d'une part et  $\delta\psi^*$  d'autre part pour obtenir

$$\delta S = \int d\mathbf{q} \left[ \delta\psi^* \left( i\hbar\dot{\psi} + \frac{\hbar^2}{2m}\Delta\psi - V\psi \right) + \left( -i\hbar\dot{\psi}^* + \frac{\hbar^2}{2m}\Delta\psi^* - V\psi^* \right) \delta\psi \right]$$

Si l'on veut avoir  $\delta S = 0$  pour des variations  $\delta\psi$  arbitraires il convient donc d'écrire

$$i\hbar\dot{\psi} = -\frac{\hbar^2}{2m}\Delta\psi + V\psi$$

Ce qui constitue l'équation de Schrödinger, sa conjuguée permettant d'annuler le second morceau.

<sup>15</sup> A ce stade c'est bien une hypothèse car nous n'avons pas précisé plus avant les propriétés de la fonction  $\psi(x, t)$  dans cette version lagrangienne de la mécanique quantique.

### 9.3.2 La symétrie du lagrangien quantique

Il est clair que la transformation

$$\psi \rightarrow g_\sigma(\psi) = e^{i\sigma}\psi \quad (9.22)$$

laisse invariant les lagrangiens quantiques (9.18) et (9.19). Cette transformation est continue et l'ensemble  $\mathbb{G} = \{g_\sigma, \sigma \in \mathbb{R}\}$  forme évidemment un groupe de symétrie du système défini par ces lagrangiens. Elle est appelée symétrie de jauge (voir exercice). Nous pouvons donc utiliser le théorème de Noether afin de déterminer la grandeur conservée associée à cette symétrie. Utilisons pour simplifier les calculs la version non symétrique (9.19) sans pour autant continuer à indiquer le 'superfétatoire. Ce n'est pas bien compliqué; il suffit simplement de bien traiter la forme intégrale du lagrangien quantique. Celle-ci fait en effet apparaître une densité de lagrangien  $\mathcal{L}(\psi, \dot{\psi})$  telle que

$$L(\psi, \dot{\psi}) = \int \mathcal{L}(\psi, \dot{\psi}) d\mathbf{q} \text{ avec } \mathcal{L}(\psi, \dot{\psi}) = i\hbar\psi^*\dot{\psi} - \frac{\hbar^2}{2m}\nabla\psi \cdot \nabla\psi^* - V\psi\psi^*$$

La symétrie  $g_\sigma$  porte en fait sur la densité de lagrangien, on a donc

$$\int \mathcal{L}[\psi, \dot{\psi}] d\mathbf{q} = \int \mathcal{L}\left[g_\sigma(\psi), \frac{dg_\sigma(\psi)}{dt}\right] d\mathbf{q}$$

L'équivalent de la relation (2.90) page 52 s'écrit maintenant

$$0 = \int \left[ \frac{\partial \mathcal{L}}{\partial \psi} \frac{\partial g_\sigma(\psi)}{\partial \sigma} + \frac{\partial \mathcal{L}}{\partial \dot{\psi}} \frac{\partial}{\partial \sigma} \left( \frac{dg_\sigma(\psi)}{dt} \right) \right] d\mathbf{q}$$

Les équations de Lagrange portent également sur la densité de lagrangien et s'écrivent  $\frac{\partial \mathcal{L}}{\partial \psi} = \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\psi}} \right)$ , on a donc

$$0 = \int \left[ \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\psi}} \right) \frac{\partial g_\sigma(\psi)}{\partial \sigma} + \frac{\partial \mathcal{L}}{\partial \dot{\psi}} \frac{\partial}{\partial \sigma} \left( \frac{dg_\sigma(\psi)}{dt} \right) \right] d\mathbf{q} = \frac{d}{dt} \left[ \int \frac{\partial g_\sigma(\psi)}{\partial \sigma} \frac{\partial \mathcal{L}}{\partial \dot{\psi}} d\mathbf{q} \right]$$

La quantité conservée du théorème de Noether dans le contexte d'un lagrangien continu s'écrit donc

$$I = \int \left. \frac{\partial g_\sigma(\psi)}{\partial \sigma} \right|_{\sigma=0} \frac{\partial \mathcal{L}}{\partial \dot{\psi}} d\mathbf{q} = cste$$

Dans le cas particuliers du groupe de symétrie engendré par la transformation (9.22), on a donc

$$I = \int d\mathbf{q} \left. \frac{\partial g_\sigma(\psi)}{\partial \sigma} \right|_{\sigma=0} \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = \int d\mathbf{q} (i\psi) (i\hbar\psi^*)$$

soit

$$\int |\psi(\mathbf{q})|^2 d\mathbf{q} = cste \quad \text{et} \quad \frac{d}{dt} \left( \int |\psi(\mathbf{q})|^2 d\mathbf{q} \right) = 0$$

En choisissant la constante de manière à doter  $\psi(\mathbf{q})$  d'une norme euclidienne unité, on retrouve l'interprétation probabiliste de cette fonction et l'interprétation de Copenhague de la mécanique quantique. Si l'on sait construire le lagrangien (9.18) à partir de considérations générales, le principe de moindre action et le théorème de Noether font une partie du reste...

Le moment est venu pour établir un second bilan intermédiaire de nos investigations. Nous avons vu ici comment l'écriture d'un lagrangien quantique suffisait dans un premier temps pour fournir l'équation de Schrödinger et dans un second permettait grâce au théorème de Noether de comprendre la nature de sa solution  $\psi(\mathbf{q})$  et de la relier en profondeur aux symétries du système. Ce type de méthode a été généralisé depuis plusieurs dizaines d'années comme le principal outil d'investigation en physique théorique.



## La physique non-dissipative en quelques crochets

Outre le plaisir, l'objectif est ici de montrer comment une grande partie des équations de la physique peuvent se mettre sous une forme algébrique unique dite formulation covariante des crochets de Poisson ou tout simplement symplectique. On trouvera des détails et des références dans l'article fondateur de ce genre d'activité [18].

Une condition essentielle pour parvenir à ce tour de force est de considérer des systèmes non dissipatifs. L'équation unique, permettant d'engendrer les équations de la physique s'écrit alors sous la forme

$$(F, S) = 0$$

où  $F$  est une fonction arbitraire des champs intervenant dans le problème et  $S$  est une intégrale d'action, la parenthèse vient ici différencier les contextes physiques étudiés, mais dans tous les cas elle permet de définir une algèbre de Lie (cf. Définition 3.2). Cette opération est construite à partir d'un crochet de Poisson. Deux types de variables interviennent dans cette formulation :

- Les théories de champs purs : les fonctions  $F$  et  $S$  dépendent de champs  $\varphi$  et de leurs variables conjuguées  $\pi$ .
- Les théories de champs moyens : la seule variable est alors une fonction de distribution  $f$  – la densité de probabilité de présence dans l'espace des phases, par exemple. Ce type de variable n'a pas de grandeur conjuguée.

Le plus simple est de voir cette formulation à l'ouvrage en passant en revue les diverses théories concernées.

### 10.1 Mécanique classique

On considère un système mécanique possédant  $s$  degrés de liberté. Ce système est associé à des coordonnées et des impulsions généralisées  $(\mathbf{q}, \mathbf{p})$ ,

chacune de ces grandeurs est une fonction du temps  $t$  qui peut également être une variable canonique en étendant le lagrangien, un autre paramètre d'évolution  $\lambda$  doit alors être utilisé. Le système est de plus décrit par un hamiltonien  $\mathcal{H} = \mathcal{H}(\mathbf{q}, \mathbf{p})$ . On définit habituellement l'action de ce système par l'intégrale

$$S^{cl}(\mathbf{q}, \mathbf{p}) = \int (\mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{H}) dt \tag{10.1}$$

Pour tout couple  $(A, B)$  de fonctions dérivables des variables  $(\mathbf{q}, \mathbf{p})$ , on écrit alors la parenthèse mécanique sous la forme

$$(A, B)^{cl}(\mathbf{q}, \mathbf{p}) := \int \{A, B\} dt = \int \left( \frac{\partial A}{\partial \mathbf{q}} \cdot \frac{\partial B}{\partial \mathbf{p}} - \frac{\partial A}{\partial \mathbf{p}} \cdot \frac{\partial B}{\partial \mathbf{q}} \right) dt$$

dans cette définition  $\{, \}$  est le crochet de Poisson habituel.

Nous pouvons maintenant récrire les équations fondamentales de la dynamique classique d'un système hamiltonien de masses ponctuelles :

**Théorème 10.1.** *Si  $F$  est une fonction quelconque de  $(\mathbf{q}, \mathbf{p})$ , les équations de Hamilton sont une conséquence de l'équation*

$$(F, S^{cl})^{cl} = 0$$

La preuve de cette affirmation est le simple résultat d'une intégration par parties. Utilisons les coordonnées covariantes et contravariantes des positions et des impulsions généralisées ainsi que les règles de sommation des indices répétés haut et bas. Il vient

$$\int \left( \frac{\delta F}{\delta q^i} \frac{\delta S^{cl}}{\delta p_i} - \frac{\delta F}{\delta q^i} \frac{\delta S^{cl}}{\delta p_i} \right) dt = 0$$

cette relation étant vérifiée pour toute fonction  $F(\mathbf{q}, \mathbf{p})$ , elle implique pour chaque  $i = 1, \dots, s$  deux relations indépendantes

$$\frac{\delta S^{cl}}{\delta p_i} = \frac{\delta S^{cl}}{\delta q^i} = 0 \tag{10.2}$$

- La première de ces relations s'écrit directement (cf. 10.1)

$$\forall i = 1, \dots, s \quad 0 = \frac{\delta S^{cl}}{\delta p_i} = \int \left( \dot{q}^i - \frac{\delta \mathcal{H}}{\delta p_i} \right) dt$$

cette dernière relation valant pour chaque instant  $t$  elle implique donc les premières équations de Hamilton

$$\forall i = 1, \dots, s \quad \dot{q}^i = \frac{\delta \mathcal{H}}{\delta p_i}$$

- La deuxième relation dans l'équation (10.2) mérite une intégration par parties (sur le premier terme ...)

$$\begin{aligned} \forall i = 1, \dots, s \quad 0 &= \frac{\delta S}{\delta q^i} = \int \left( p_j \frac{\delta \dot{q}_j}{\delta q^i} - \frac{\delta \mathcal{H}}{\delta q^i} \right) dt \\ &= \int \left( -\dot{p}_j \delta_i^j - \frac{\delta \mathcal{H}}{\delta q^i} \right) dt + [p_j \delta_i^j] \end{aligned}$$

une fois annulé le terme de bord – sans commentaire – il ne reste plus que les deuxièmes équations de Hamilton

$$\forall i = 1, \dots, s \quad \dot{p}_i = -\frac{\delta \mathcal{H}}{\delta q^i}$$

Ceci clôt la preuve du théorème 10.1 et assure la formulation symplectique de la mécanique classique.

## 10.2 Électromagnétisme

En utilisant les notations introduites dans le chapitre sur la relativité restreinte et l'électromagnétisme (cf. Chap. 5), on définit l'action suivante

$$\begin{aligned} S^{el} &= \int \mathcal{L} d\mathbf{x} := \int \left( \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - A_\mu J^\mu \right) d\mathbf{x} \\ &= \int \left( \pi^{\mu\nu} \partial_\nu A_\mu - \mathcal{H}[A, \pi] \right) d\mathbf{x} = S^{el}[A, \pi] \end{aligned} \quad (10.3)$$

avec  $\mathcal{H}[A, \pi] = \frac{1}{4} \pi^{\mu\nu} \pi_{\mu\nu} + A_\mu J^\mu$ . Les variables sont ici le champ  $A_\mu(\mathbf{x})$  et sa quantité conjuguée  $\pi_{\mu\nu}(\mathbf{x})$  définie comme dans le cas classique par (voir 3.2)

$$\pi_{\mu\nu} := \frac{\delta \mathcal{L}}{\delta (\partial_\nu A_\mu)} \quad (10.4)$$

Pour un champ vectoriel  $V^\nu(\mathbf{x})$  quelconque<sup>1</sup>, et pour tout couple  $(F, G)$  des variables  $[A, \pi]$ , on définit le crochet de Lie électromagnétique par

$$(F, G)_V^{el}[A, \pi] := \int \left( \frac{\delta F}{\delta A_\mu} \frac{\delta G}{\delta \pi^{\mu\nu}} - \frac{\delta G}{\delta A_\mu} \frac{\delta F}{\delta \pi^{\mu\nu}} \right) V^\nu d\mathbf{x}$$

On a alors le théorème suivant :

<sup>1</sup> Ce terme est présent pour assurer que l'intégrand soit un scalaire.

**Théorème 10.2.** *Si  $F$  est une fonction quelconque de  $[A, \pi]$ , les équations de Maxwell sont une conséquence de l'équation*

$$(F, S^{el})_V^{el} = 0$$

Quoiqu'un peu plus technique, la démonstration de ce théorème est dans le droit fil de celle effectuée dans le cas classique. La généralité de  $F$  et de  $V$  imposent

$$\frac{\delta S^{el}}{\delta \pi^{\mu\nu}} = \frac{\delta S^{el}}{\delta A_\mu} = 0$$

- La première de ces relations s'écrit

$$\frac{\delta}{\delta \pi^{\mu\nu}} \int \left( \pi^{\mu\nu} \partial_\nu A_\mu - \frac{1}{4} \pi^{\mu\nu} \pi_{\mu\nu} - A_\mu J^\mu \right) d\mathbf{x} = 0$$

les variables  $A$  et  $\pi$  sont indépendantes par hypothèse, un petit calcul que nous détaillons pour les non-initiés donne alors

$$\begin{aligned} 0 &= \int \left( \partial_\nu A_\mu - \frac{1}{4} \pi_{\mu\nu} - \frac{1}{4} \pi_{\rho\sigma} \frac{\delta (\pi_{\mu\nu} \eta^{\mu\rho} \eta^{\nu\sigma})}{\delta \pi^{\mu\nu}} \right) d\mathbf{x} \\ &= \int \left( A_{\mu,\nu} - \frac{1}{2} \pi_{\mu\nu} \right) d\mathbf{x} \end{aligned}$$

comme chaque fois dans ce contexte, c'est donc l'intégrand qui doit être nul, ainsi

$$A_{\mu,\nu} = \frac{1}{2} \pi_{\mu\nu} \quad \text{ou bien} \quad A_{\nu,\mu} = \frac{1}{2} \pi_{\nu\mu} \quad (10.5)$$

Le tenseur  $\pi_{\mu\nu}$  est antisymétrique comme en attestent les définitions (10.3) et (10.4), ainsi la différence des deux relations (10.5) fournit

$$\pi_{\mu\nu} = A_{\mu,\nu} - A_{\nu,\mu} =: F_{\mu\nu}$$

Le tenseur  $\pi_{\mu\nu}$  est donc un 4-rotationnel, dont la nullité de la divergence donne le premier groupe d'équations de Maxwell, dit homogène.

$$\pi_{\mu\nu,\rho} + \pi_{\rho\mu,\nu} + \pi_{\nu\rho,\mu} = 0$$

- La deuxième relation mérite à son tour une intégration par parties : nous avons

$$\frac{\delta}{\delta A_\mu} \int \left( \pi^{\mu\nu} A_{\mu,\nu} - \frac{1}{4} \pi^{\mu\nu} \pi_{\mu\nu} - A_\mu J^\mu \right) d\mathbf{x} = 0$$

c'est sur le premier membre qu'il faut intervenir : en séparant la dérivée de  $A_\mu$  on obtient

$$0 = \int \left( \frac{\delta(-\pi^{\mu\nu},\nu A_\mu)}{\delta A_\mu} - J^\mu \right) d\mathbf{x} + \frac{\delta}{\delta A_\mu} \int (\pi^{\mu\nu} A_\mu),\nu d\mathbf{x}$$

Le deuxième terme est nul pour une quantité de raisons, la relation ci-dessus implique donc

$$\pi^{\mu\nu},\nu = -J^\mu$$

ce qui est bien la seconde équation de Maxwell, dite inhomogène.

Les équations de Maxwell covariantes adoptent donc bien une posture symplectique.

### 10.3 Relativité générale

En utilisant les notations désormais habituelles, introduites quelques pages plus haut, on définit l'action

$$S^{rg} = \int \mathcal{L}_m \sqrt{-g} d\mathbf{x} - (2\chi)^{-1} \int g^{\alpha\beta} R_{\alpha\beta} \sqrt{-g} d\mathbf{x}$$

avec

$$R_{\alpha\beta} := \pi^\lambda_{\alpha\beta,\lambda} - \pi^\lambda_{\alpha\lambda,\beta} - \pi^\lambda_{\nu\beta} \pi^\nu_{\lambda\alpha} + \pi^\lambda_{\nu\lambda} \pi^\nu_{\alpha\beta}$$

Le crochet de Lie de la relativité générale est de la même forme que celui de l'électromagnétisme, pour tout champ vectoriel  $V^\mu$  et pour tout couple  $(F, G)$  fonction de la métrique  $g$  et de sa quantité conjuguée  $\pi$ , on pose

$$(F, G)_V^{rg} [g, \pi] := \int \left( \frac{\delta F}{\delta g^{\alpha\beta}} \frac{\delta G}{\delta \pi^\mu_{\alpha\beta}} - \frac{\delta G}{\delta g^{\alpha\beta}} \frac{\delta F}{\delta \pi^\mu_{\alpha\beta}} \right) V^\mu d\mathbf{x}$$

On a alors le théorème suivant :

**Théorème 10.3.** *Si  $F$  est une fonction quelconque de  $[g, \pi]$ , les équations d'Einstein sont une conséquence de l'équation*

$$(F, S^{rg})_V^{rg} = 0$$

La démonstration de ce théorème repose sur les deux équations qu'il implique, à savoir

$$\frac{\delta S^{rg}}{\delta \pi^\mu_{\alpha\beta}} = \frac{\delta S^{rg}}{\delta g^{\alpha\beta}} = 0 \quad (10.6)$$

- Quelques lignes d’algèbre permettent de se convaincre que la première de ces équations donne directement

$$(g^{\alpha\beta} \sqrt{-g})_{;\mu} + \pi_{\mu\nu}^{\alpha} g^{\nu\beta} \sqrt{-g} + \pi_{\mu\nu}^{\beta} g^{\nu\alpha} \sqrt{-g} - \pi_{\mu\nu}^{\nu} g^{\alpha\beta} \sqrt{-g} = 0 \quad (10.7)$$

comme en atteste la définition de la dérivée covariante, ceci signifie que  $\pi_{\mu\nu}^{\alpha}$  est le symbole de connexion affine,  $\pi_{\mu\nu}^{\alpha} = \Gamma_{\mu\nu}^{\alpha}$ , ceci pris en compte, la relation (10.7) s’écrit donc

$$(g^{\alpha\beta} \sqrt{-g})_{;\mu} = 0$$

qui permet d’obtenir l’identité de Bianchi (7.65), qui se révèle donc être une des deux équations fondamentales de la relativité générale.

- La deuxième équation de la relation (10.6) nécessite un ajustement habituel. Il convient en effet de définir le tenseur énergie-impulsion comme

$$T_{\alpha\beta} := \frac{\delta(\mathcal{L}_{mat} \sqrt{-g})}{\delta g^{\alpha\beta}}$$

pour voir que

$$\frac{\delta S^{rg}}{\delta g^{\alpha\beta}} = 0 \Rightarrow G_{\alpha\beta} := R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R_{\lambda}^{\lambda} = \chi T_{\alpha\beta}$$

fameuse équation d’Einstein et donc deuxième équation fondamentale de la relativité générale.

Et une théorie de plus dans l’escarcelle des crochets !

## 10.4 Physique statistique et mécanique des fluides

Considérons à présent un système physique décrit par sa fonction de distribution dans l’espace des phases  $f(\mathbf{q}, \mathbf{p}, t)$ , c’est-à-dire un système composé d’un grand nombre de particules dont la probabilité d’avoir pour l’une d’elles<sup>2</sup>, à l’instant  $t$ , l’impulsion  $\mathbf{p} \in \Delta_p$  et de se trouver dans le volume  $\Delta_q$  est donnée par l’intégrale

$$P_t(\mathbf{q} \in \Delta_q, \mathbf{p} \in \Delta_p) = \int_{(\mathbf{q}, \mathbf{p}) \in (\Delta_q, \Delta_p)} f(\mathbf{q}, \mathbf{p}, t) d\Gamma$$

<sup>2</sup> représentative de toutes les autres : une particule test en quelque sorte

Où l'on a noté  $d\mathbf{\Gamma}$  l'élément de volume de l'espace des phases. Considérons de plus que le système admet une représentation hamiltonienne, c'est-à-dire qu'il existe une fonction  $\mathcal{H}(\mathbf{q}, \mathbf{p})$  qui permet d'écrire

$$\left\{ \begin{array}{l} \frac{dq_i}{dt} = \frac{\partial \mathcal{H}}{\partial p^i} \\ \frac{dp^i}{dt} = -\frac{\partial \mathcal{H}}{\partial q_i} \end{array} \right. \Rightarrow \forall g(\mathbf{q}, \mathbf{p}, t) \left\{ \begin{array}{l} \frac{dg}{dt} = \frac{\partial g}{\partial t} + \frac{\partial g}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p^i} - \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial g}{\partial p^i} \\ = \frac{\partial g}{\partial t} + \{g, \mathcal{H}\} \end{array} \right. \quad (10.8)$$

On définit alors l'action

$$S^{ms}[f] = \int f(\mathbf{q}, \mathbf{p}, t) \mathcal{H} d\mathbf{\Gamma} dt$$

Le crochet de Lie adapté aux systèmes possédant un continuum de degrés de liberté, introduit par l'école chapeauté par les auteurs de l'article fondateur [18], fait intervenir la notion de dérivée  $A'$  de la fonctionnelle  $A[f]$  :

$$\forall A[f] \text{ dérivable} \quad \left. \frac{dA[f + \lambda \delta f]}{d\lambda} \right|_{\lambda=0} = \int A' \delta f d\mathbf{\Gamma}$$

que l'on notera  $A' = \frac{\delta A}{\delta f}$

Si  $A[f]$  est une fonctionnelle,  $A'$  est une fonction. Nantis de ces précisions nous pouvons écrire le crochet de Lie de la physique statistique. Pour tout couple  $(A, B)$  de fonctionnelles de  $f$  on pose

$$(A, B)^{ms}[f] = \int f \{A', B'\}_4 d\mathbf{\Gamma} dt \quad (10.9)$$

$$=: \int f \left[ \{A', B'\} + \frac{\partial A'}{\partial t} \frac{\partial B'}{\partial \mathcal{H}} - \frac{\partial A'}{\partial \mathcal{H}} \frac{\partial B'}{\partial t} \right] d\mathbf{\Gamma} dt \quad (10.10)$$

avec ces notations on a alors le théorème suivant :

**Théorème 10.4.** *L'équation fondamentale de la physique statistique des systèmes non dissipatifs<sup>3</sup> est une conséquence de la relation*

$$\forall F[f], \quad (F, S^{ms})^{ms}[f] = 0 \quad (10.11)$$

<sup>3</sup> Par système non dissipatif, on sous-entend sournoisement que le hamiltonien du système ne dépend pas explicitement du temps.

Dans de pareilles situations, rien ne vaut un petit calcul pour se convaincre de cette affirmation...

Ce n'est qu'un simple exercice de calcul variationnel de voir que

$$S' = \frac{\delta S^{ms}}{\delta f} = \mathcal{H}$$

Ainsi, l'équation (10.11) donne

$$0 = (F, S^{ms})^{ms} [f] = \int f \left[ \{F', \mathcal{H}\} + \frac{\partial F'}{\partial t} \frac{\partial \mathcal{H}}{\partial \mathcal{H}} - \frac{\partial F'}{\partial \mathcal{H}} \frac{\partial \mathcal{H}}{\partial t} \right] d\mathbf{\Gamma} dt$$

Si le système est non dissipatif  $\mathcal{H}$  ne dépend pas explicitement du temps, il reste donc

$$0 = \int f \left[ \frac{\partial F'}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p^i} - \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial F'}{\partial p^i} + \frac{\partial F'}{\partial t} \right] d\mathbf{\Gamma} dt$$

une intégration par parties sur chaque terme et l'application du théorème de Schwarz sur  $\mathcal{H}$  donnent alors, en notant  $TB$  les termes de bords tout intégrés

$$\begin{aligned} 0 &= TB - \int F' \left[ \frac{\partial f}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p^i} - \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial f}{\partial p^i} + \frac{\partial f}{\partial t} \right] d\mathbf{\Gamma} dt \\ &= TB - \int F' \left[ \{f, \mathcal{H}\} + \frac{\partial f}{\partial t} \right] d\mathbf{\Gamma} dt \end{aligned}$$

Intéressons-nous pour une fois aux termes de bord. En supposant le domaine d'intégration non borné dans l'espace et dans le temps, ils s'écrivent

$$TB = \left[ \int f F' \frac{\partial \mathcal{H}}{\partial p^i} d\mathbf{p} dt \right]_{q_i \rightarrow -\infty}^{q_i \rightarrow +\infty} - \left[ \int f F' \frac{\partial \mathcal{H}}{\partial q_i} d\mathbf{q} dt \right]_{p_i \rightarrow -\infty}^{p_i \rightarrow +\infty} + \left[ \int f F' d\mathbf{q} dp \right]_{t \rightarrow -\infty}^{t \rightarrow +\infty}$$

il n'est alors que temps d'utiliser les équations de Hamilton pour avoir

$$\begin{aligned} TB &= \left[ \int f F' \frac{dq_i}{dt} d\mathbf{p} dt \right]_{q_i \rightarrow -\infty}^{q_i \rightarrow +\infty} + \left[ \int f F' \frac{dp^i}{dt} d\mathbf{q} dt \right]_{p_i \rightarrow -\infty}^{p_i \rightarrow +\infty} + \left[ \int f F' d\mathbf{q} dp \right]_{t \rightarrow -\infty}^{t \rightarrow +\infty} \\ &= \left[ \int f F' q_i d\mathbf{p} \right]_{t, q_i \rightarrow -\infty}^{t, q_i \rightarrow +\infty} + \left[ \int f F' p^i d\mathbf{q} \right]_{t, p_i \rightarrow -\infty}^{t, p_i \rightarrow +\infty} + \left[ \int f F' d\mathbf{q} dp \right]_{t \rightarrow -\infty}^{t \rightarrow +\infty} \\ &\quad - \left[ \int q_i \left( \int \frac{d(f F')}{dt} d\mathbf{p} \right) dt \right]_{q_i \rightarrow -\infty}^{q_i \rightarrow +\infty} - \left[ \int p^i \left( \int \frac{d(f F')}{dt} d\mathbf{q} \right) dt \right]_{p_i \rightarrow -\infty}^{p_i \rightarrow +\infty} \end{aligned}$$

Si le système possède un état d'équilibre unique, ou bien si le volume total de l'espace des phases est conservé :  $\frac{df}{dt} = 0$ , on doit alors pouvoir montrer que  $TB = 0$ , ce qui reste à faire (il s'agit là de l'un des enjeux de la physique statistique !)

Dans ces conditions

$$\begin{aligned} \forall F[f] , \quad (F, S^{ms})^{ms} [f] &= 0 \\ \Rightarrow \{f, \mathcal{H}\} + \frac{\partial f}{\partial t} &= 0 \end{aligned}$$

Cette dernière relation est l'équation de Boltzmann sans collision, elle est à la base de la physique statistique non dissipative.

À titre d'exemple, et sans entrer dans le détail des calculs, on peut déduire les équations de l'hydrodynamique en prenant des moments de l'équation de Boltzmann sans collision pour un ensemble de particules identiques de masse  $m$  (on pourra consulter [20] pour plus de détails)

Une simple intégration donne l'équation dite de continuité

$$\begin{aligned} 0 &= \int \left[ \{f, \mathcal{H}\} + \frac{\partial f}{\partial t} \right] d\mathbf{p} \\ \Rightarrow \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \bar{\mathbf{v}}) &= 0 \end{aligned}$$

où l'on a posé

$$\begin{aligned} \rho &= m \nu \quad \text{avec } \nu = \int f d\mathbf{p} \\ \text{et} \\ \bar{\mathbf{v}} &= \bar{\mathbf{p}}/m \quad \text{avec } \bar{\mathbf{p}} = \frac{1}{\nu} \int \mathbf{p} f d\mathbf{p} \end{aligned}$$

Un calcul du même acabit donne ensuite l'équation d'Euler

$$\begin{aligned} 0 &= \int \mathbf{p} \left[ \{f, \mathcal{H}\} + \frac{\partial f}{\partial t} \right] d\mathbf{p} \\ \Rightarrow \frac{\partial \bar{\mathbf{v}}}{\partial t} + (\bar{\mathbf{v}} \cdot \operatorname{grad}) \bar{\mathbf{v}} &= -\frac{1}{\nu} \operatorname{grad} P - \operatorname{grad} \psi \end{aligned}$$

où l'on a défini le tenseur de pression cinétique

$$P_{ij} := \frac{1}{m^2} \left[ \int p_i p_j f d\mathbf{p} - \int p_i f d\mathbf{p} \int p_j f d\mathbf{p} \right]$$

et  $\psi$  désigne le potentiel duquel dérivent les forces considérées dans le système.

Voilà tout un pan de la physique qui tombe ainsi sous notre coupe !



De la théorie à la pratique



## Exercices

### 11.1 Mécanique classique

#### 11.1.1 Juste pour vérifier ...

Soit  $x$  une variable réelle,  $y(x)$  une fonction  $C^1$  sur  $[a, b] \subset \mathbb{R}$  et  $y'$  sa dérivée. Montrer que la quantité

$$\mathcal{F} = \int_a^b F(x, y, y') dx$$

est extrémale si

$$\frac{d}{dx} \left( \frac{\partial \mathcal{F}}{\partial y'} \right) = \frac{\partial \mathcal{F}}{\partial y}$$

#### 11.1.2 Pour se faire plaisir

Soit un système mécanique bidimensionnel décrit par le schéma de la figure 11.1.

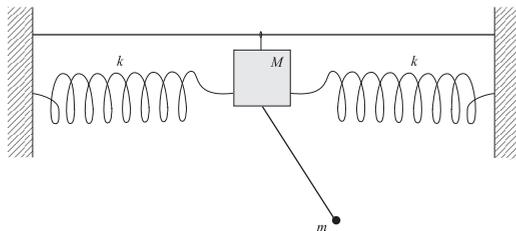
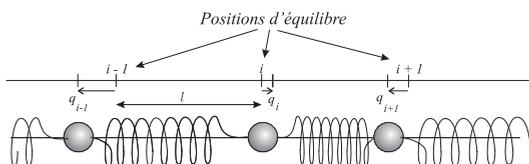


Fig. 11.1. Oscillateur mobile

1. Combien de degrés de liberté possède ce système ?
2. Ecrire son lagrangien
3. Ecrire les équations du mouvement et étudier les petites oscillations.

### 11.1.3 La chaîne d'oscillateurs couplés

On dispose  $N$  masses identiques  $m_{i=1,\dots,N}$  reliées par des ressorts identiques de raideur  $k$  sur un cerceau. Le nombre  $N$  est tel que le rayon du cerceau est très grand devant la taille des ressorts. On fera donc l'hypothèse que ces derniers ne se déforment que dans une seule direction. A l'équilibre toutes les masses sont distantes d'une longueur  $l$  identique. On utilisera les notations de la figure 11.2.



**Fig. 11.2.** Chaîne d'oscillateurs couplés

1. Ecrire le lagrangien du système, on négligera le poids des particules.
2. En déduire les équations du mouvement de chaque masse
3. On cherche des solutions de la forme

$$q_{n=1,\dots,N}(t) = \text{Re} [a_n \exp(i(kl - \omega t))] \quad (k, \omega) \in \mathbb{R}$$

Ecrire la relation de dispersion vérifiée par  $\omega$  et  $k$  et étudier les divers modes de propagation de telles ondes dans ce système.

### 11.1.4 Le brachystochrone

Sous l'action d'un champ de pesanteur uniforme un point matériel (perle) de masse  $m$  glisse sans frottements sur un fil rigide situé dans un plan vertical. A  $t = 0$ , la perle est immobile en  $(x, y) = (0, 0)$ .

1. Quel temps  $T$  met la perle pour atteindre un point quelconque du fil ?
2. Ecrire l'équation différentielle vérifiée par  $y(x)$  pour que  $T$  soit minimum.
3. Résoudre cette équation et montrer que  $T_{\min}$  est atteint pour un fil en forme de branche cycloïdale.

## 4. Application

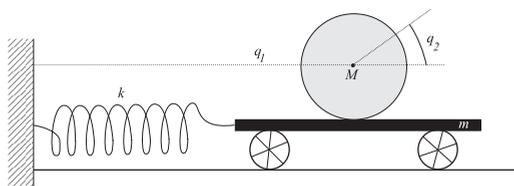
- a) Ecrire l'équation du mouvement d'une perle pesante dans un champ gravitationnel constant le long d'un fil cycloïdal.

$$x = a(\theta - \sin \theta) \quad y = a(1 + \cos \theta)$$

- b) En posant  $u = \cos(\theta/2)$ , résoudre l'équation du mouvement et montrer que la période du pendule cycloïdal est indépendante de son amplitude (pendule libre).

## 11.1.5 Avec contraintes

On considère un disque homogène de masse  $M$  roulant sans glisser sur un chariot mobile de masse  $m$  lui-même arrimé à un mur fixe par un ressort de raideur  $k$  comme représenté sur la figure 11.3.



**Fig. 11.3.** Roue, chariot, ressort ...

1. Ecrire et résoudre les équations du mouvement du disque.
2. Faire la même chose avec un disque inhomogène ...

## 11.1.6 Noether très simple

Dans un espace à 2 dimensions, une particule de masse  $m$  repérée par ses coordonnées cartésiennes  $(x, y)$  est soumise à un potentiel  $V(x, y) = V(x - 2y)$

1. Ecrire le lagrangien du système et montrer qu'il existe un groupe de symétrie qui le laisse invariant
2. Déterminer la grandeur conservée associée à cette symétrie

### 11.1.7 De l'harmonique à Kepler

On considère une particule de masse  $m$  repérée par le vecteur  $\vec{r}(t) \in \mathbb{R}^3$ , évoluant dans le potentiel  $V(\vec{r}) = \frac{1}{2}\omega^2 r^2$  avec  $r = |\vec{r}|$ . Ce potentiel est en fait un potentiel de gravitation de champ moyen créée par une assemblée de particules massive entourant la masse  $m$  mais qu'elle n'influence pas (étoile dans les régions centrales d'une galaxie).

1. Montrer que la particule évolue dans un plan que l'on caractérisera.
2. On considère les coordonnées polaires  $(r, \theta)$  dans ce plan. Ecrire le Lagrangien  $\mathcal{L}(r, \dot{r}, \theta, \dot{\theta})$  puis montrer que le Hamiltonien  $\mathcal{H}(r, \theta, p_r, p_\theta)$  de ce système s'écrit

$$\mathcal{H} = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{1}{2}m\omega^2 r^2 \quad (11.1)$$

Que constatez-vous ?

3. On effectue le changement de variable  $x = \frac{r^2}{\ell}$  où  $\ell$  est une longueur arbitraire dont le rôle est de conférer à  $x$  la dimension d'une longueur tout comme  $r$ . On effectue la transformation  $(r, p_r) \xrightarrow{T} (x, p_x)$ . Déterminer l'expression de l'impulsion  $p_x$  de manière à ce que la transformation  $T$  soit canonique. Ecrire le nouvel hamiltonien.
4. On note  $E^h$  l'énergie totale de la particule. Montrer que

$$\frac{p_x^2}{2m} + \frac{p_\theta^2}{8mx^2} - \frac{\mu}{x} = E^k$$

ou l'on exprimera les constantes  $\mu > 0$  et  $E^k < 0$  en fonction de  $\ell$ ,  $E^h$ ,  $m$  et  $\omega^2$ .

5. Ecrire le hamiltonien  $\mathcal{K}$  d'une particule de masse  $m$  dans un potentiel keplerien  $V(\vec{x}) = -\frac{\mu}{x}$  en utilisant les coordonnées polaires  $(x, \varphi)$ .
6. En déduire la construction d'une orbite d'énergie positive dans un potentiel harmonique en fonction d'une orbite elliptique dans un potentiel képlerien.

### 11.1.8 Lagrange, Hamilton et le champ magnétique

1. Montrer que

$$\mathcal{L} = \mathcal{L}(\vec{r}, \dot{\vec{r}}, t) = -mc^2 \left( 1 - \frac{\dot{\vec{r}} \cdot \dot{\vec{r}}}{c^2} \right) + q_e \left( A(\vec{r}, t) \cdot \dot{\vec{r}} - U(\vec{r}, t) \right)$$

est le lagrangien d'une particule de charge  $q_e$  dans un champ électromagnétique décrit par les potentiels scalaire  $U(\vec{r}, t)$  et vecteur  $A(\vec{r}, t)$ .

2. En déduire le hamiltonien d'une telle particule, on vérifiera qu'il n'est généralement pas conservé.

## 11.2 Relativités

### 11.2.1 Quelques manipulations

1. Montrer que l'équation des géodésiques correspond à un extremum de l'action

$$S = \int ds \quad \text{avec} \quad ds^2 = g_{\mu\nu} dx^\mu dx^\nu$$

2. Montrer que

$$V^\mu{}_{;\mu} = \frac{(\sqrt{-g}V^\mu)_{,\mu}}{\sqrt{-g}}$$

et que pour tout tenseur  $A$  d'ordre 2 antisymétrique

$$A^{\mu\nu}{}_{;\mu} = \frac{(\sqrt{-g}A^{\mu\nu})_{,\mu}}{\sqrt{-g}} \quad \text{et} \quad A_{\mu\nu;\lambda} = A_{\mu\nu,\lambda} + A_{\lambda\mu,\nu} + A_{\nu\lambda,\mu}$$

### 11.2.2 Electromagnétisme tensoriel

1. Montrer que le tenseur énergie-impulsion associé à la densité de lagrangien du champ électromagnétique pur

$$\mathcal{L} = -\frac{1}{4\mu_o} F_{\alpha\beta} F^{\alpha\beta}$$

est donné en relativité restreinte par

$$T^{\alpha\beta} = -\frac{1}{\mu_o} \left( F^\alpha{}_\gamma F^{\beta\gamma} - \frac{1}{4} \eta^{\alpha\beta} F_{\mu\nu} F^{\mu\nu} \right)$$

2. Calcul du tenseur énergie-impulsion

- Montrer que  $F_{\mu\nu} F^{\mu\nu} = 2 \left( \vec{B}^2 - \vec{E}^2/c^2 \right)$  où  $\vec{E}$  et  $\vec{B}$  sont respectivement les champs électrique et magnétique associés à  $F^{\mu\nu}$
- Calculer  $F_{\mu\gamma} F^{\beta\gamma}$
- en déduire que le tenseur énergie-impulsion du champ électromagnétique s'écrit

$$T^{\alpha\beta} = \begin{bmatrix} u & \vec{H}^T/c \\ \vec{H}/c & M \end{bmatrix}$$

où  $u$  est un 4-scalaire,  $\vec{H}$  un tri-vecteur et  $M$  un tenseur de  $\mathbb{R}^3 \otimes \mathbb{R}^{3*}$  dont on précisera pour chacun l'expression des composantes en fonction de  $\vec{E}, \vec{B}, c, \mu_o$  ou  $\varepsilon_o$ .

3. En utilisant les 2 équations de Maxwell montrer que

$$T^{\alpha\beta}_{,\beta} = -F^\alpha_\gamma J^\gamma$$

4. Interpréter les composantes temporelle et spatiale de l'équation ci-dessus.

### 11.2.3 Énergie-impulsion d'un fluide parfait

Un fluide parfait de densité d'énergie  $\varepsilon(t)$  et de pression  $P(t)$  à l'instant  $t$  est décrit par une densité de lagrangien de la forme

$$\mathcal{L} = 3P - \varepsilon$$

1. Montrer que le tenseur énergie-impulsion d'un tel fluide parfait est donné par

$$T^{\mu\nu} = P g^{\mu\nu} + (P + \varepsilon) \frac{u^\mu u^\nu}{c^2}$$

où  $u^\mu$  représente le champ de quadrivitesse dans le fluide. (On fera preuve d'imagination ...)

2. L'équation de conservation de l'énergie du fluide s'écrit en relativité générale  $D_\mu T^{\mu\nu} = 0$ . Montrer que cette équation peut s'écrire sous la forme

$$g^{\mu\nu} \partial_\mu P + \frac{1}{\sqrt{-g}} \partial_\mu [(P + \varepsilon) u^\mu u^\nu \sqrt{-g}]$$

3. On fait l'hypothèse que la métrique est séparée en temps et espace, et donc que  $\sqrt{-g} = V(t) \phi(x_1, x_2, x_3)$  où  $\phi$  est une fonction spatiale et  $V(t)$  représente le volume du fluide à l'instant  $t$ . Montrer qu'en coordonnées comobiles  $u^\mu = (1, 0, 0, 0)^T$ , l'équation de conservation de l'énergie correspond à une évolution adiabatique du fluide

$$TdS = PdV + d(\varepsilon V) = 0$$

### 11.2.4 Un point commun entre l'électromagnétisme et la relativité générale

#### 1. ELECTROMAGNETISME

- a) Montrer qu'une transformation de jauge sur le 4-potentiel

$$A_\mu^* = A_\mu + \partial_\mu \psi \quad (11.2)$$

où  $\psi$  est un 4-scalaire, laisse invariante les équations de Maxwell.

- b) Montrer qu'il existe toujours une jauge  $\psi_L(x^\mu)$  telle que  $A_\mu^*$  vérifie la relation  $\partial^\mu A_\mu^* = 0$ . Cette jauge est appelée jauge de Lorentz.  
 c) Dans la jauge de Lorentz, montrer que l'équation d'évolution pour le 4-potentiel s'écrit

$$\left( \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \right) A^\mu = \mu_o J^\mu \quad (11.3)$$

où  $\Delta$  est l'opérateur laplacien de l'espace habituel,  $\mu_o$  la perméabilité du vide et  $J^\mu$  le 4-courant.

#### 2. RELATIVITE GENERALE

- a) En utilisant les identités de Bianchi, montrer que

$$R^{\alpha\beta\gamma\lambda}_{;\lambda} = R^{\gamma\alpha;\beta} - R^{\gamma\beta;\alpha} \quad (11.4)$$

- b) Montrer que l'on peut écrire les équations d'Einstein sous la forme

$$R_{\mu\nu} = \chi \left( T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right) \quad (11.5)$$

où  $T$  est le scalaire d'énergie-impulsion dont on précisera l'expression.

- c) En utilisant les questions précédentes, montrer que l'on peut écrire les équations d'Einstein sous la forme

$$R^{\alpha\beta\gamma\lambda}_{;\lambda} = \chi Z^{\alpha\beta\gamma} \quad (11.6)$$

où  $Z^{\alpha\beta\gamma}$  est appelé tenseur de courant gravitationnel, on précisera sa valeur en fonction du tenseur métrique (et de ses dérivées covariantes), et du scalaire d'énergie-impulsion.

- d) Calculer  $Z^{\alpha\beta\gamma}_{;\gamma}$ . Qu'en concluez-vous ?  
 e) Quel est l'équivalent électromagnétique de l'équation (11.6) ?

3. L'électromagnétisme est régi par deux équations de Maxwell. Outre l'équation d'Einstein (11.6), quelle équation est l'autre relation régissant la relativité générale ?
4. Conséquences :
- Quelles sont les conséquences physiques contenues dans les équations de Maxwell ?
  - Quelles conséquences possibles envisagez-vous pour l'équation (11.6) ?
5. Application au fluide parfait :
- Un fluide parfait de vitesse de pression  $P$  et de densité d'énergie  $\epsilon$  est décrit par un tenseur énergie-impulsion de la forme

$$T^{\mu\nu} = P g^{\mu\nu} + (\epsilon + P) \frac{u^\mu u^\nu}{c^2}$$

où  $u^\mu$  représente le champ de quadrivitesse du fluide.

- Déterminer le scalaire d'énergie-impulsion d'un fluide parfait.
- En déduire le tenseur de courant gravitationnel associé à un tel fluide (on simplifiera le calcul au maximum de ses possibilités ...)

### 11.2.5 Le champ de Schwarzschild

On désigne par *métrique de Schwarzschild*, la métrique qui règne autour d'une masse  $M$  stationnaire à symétrie sphérique. On admettra que les symétries mentionnées (stationnarité et isotropie) impliquent qu'il existe un système de coordonnées  $x^\mu = (t, r, \theta, \varphi)^T$  tel que l'élément d'intervalle d'espace-temps s'écrive

$$ds^2 = e^{x(r)} dt^2 - e^{y(r)} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2) \quad (11.7)$$

$x(r)$  et  $y(r)$  sont des fonctions que les seules considérations de symétrie ne suffisent pas à préciser.

- Détermination de la métrique
  - Préciser les composantes co- et contravariantes du tenseur métrique.
  - Montrer que les seuls symboles de connexion affine non nuls sont (en notant ' pour la dérivée par rapport à  $r$ )

$$\Gamma_{11}^0 = e^{x-y} x'/2 \quad \Gamma_{11}^1 = y'/2 \quad \Gamma_{22}^1 = -r e^{-y} \quad \Gamma_{33}^1 = -r \sin^2 \theta e^{-y}$$

$$\Gamma_{12}^2 = \Gamma_{13}^3 = 1/r \quad \Gamma_{10}^0 = x'/2 \quad \Gamma_{33}^1 = -\sin \theta \cos \theta \quad \Gamma_{23}^3 = \cot \theta$$

- c) Vérifier que le tenseur de Ricci est diagonal et que ses composantes non nulles sont

$$\left[ \begin{array}{l} R_{00} = \left( \frac{x''}{2} - \frac{x'(y' - x')}{4} - \frac{x'}{r} \right) e^{x-y} \\ R_{11} = -\frac{x''}{2} + \frac{1}{4}x'(y' - x') + \frac{y'}{r} \\ R_{22} = 1 - e^{-y} - \frac{re^{-y}}{2}(x' - y') \\ R_{33} = \sin^2 \theta R_{22} \end{array} \right.$$

- d) On suppose que la masse  $M$  est entourée d'un vide classique (fluide parfait de pression nulle et de densité d'énergie nulle), écrire les équations d'Einstein dans cette région. Conseil : on évitera de calculer le scalaire de courbure ...
- e) Intégrer les équations d'Einstein pour  $x(r)$  et  $y(r)$ .
- f) Montrer que pour un choix approprié de la variable temporelle (une constante d'intégration ...), la métrique (11.7) est minkowskienne pour  $r \rightarrow +\infty$ . Quelle est alors l'expression de la métrique à l'ordre suivant en  $1/r$ ? Comparer avec la métrique de l'approximation newtonienne de la relativité générale

$$ds^2 = dt^2 - dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2)$$

- g) On donnera une interprétation physique à l'autre constante d'intégration.
- h) Qu'observe-t-on lorsque  $r \rightarrow 2GM/c^2$  ?
2. Rougissement gravitationnel : un photon de fréquence  $\nu$  est émis à une distance  $r_1$  du centre de  $M$  dans la direction radiale. Un observateur fixe qui se maintient au point d'émission mesure la fréquence de ce photon. Qu'observe-t-il à partir de l'instant d'émission ?  
Conseil : on exprimera  $\nu$  en fonction du 4-vecteur d'onde  $k^\mu$ , de l'élément de position  $dx^\mu$  et de l'intervalle de temps propre  $d\tau$  ...

### 11.2.6 Une approche lagrangienne du problème des 2 corps : classique et relativiste

Nous considérons dans tout ce problème un espace infini contenant uniquement 2 corps homogènes à symétrie sphérique de masses  $M$  et  $m$  telles que  $M \gg m$ .

## 1. Problème classique

- a) Montrer que le mouvement de  $m$  dans le champ gravitationnel de  $M$  ne possède que 2 degrés de liberté.
- b) Montrer qu'en variables polaires dans le plan orthogonal au moment cinétique de  $m$ , le lagrangien de celui-ci s'écrit (un point représente la dérivée totale par rapport au temps)

$$\mathcal{L} = \mathcal{L}(r; \theta, \dot{r}, \dot{\theta}) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{GmM}{r} \quad (11.8)$$

## c) Equations du mouvement

- i. Montrer que la quantité  $A := r^2 \dot{\theta}$  est une constante du mouvement (constante des aires)
- ii. Montrer que la variable  $u = 1/r$  vérifie l'équation différentielle

$$\frac{d^2u}{d\theta^2} + u = K_o$$

où  $K_o$  est une constante que l'on précisera en fonction de  $G$ ,  $M$  et  $A$ .

- iii. Indiquer la solution de cette équation.

## 2. Problème relativiste

La relativité générale nous indique que  $m$  évolue dans le champ de *Schwarzschild* créé par  $M$  (voir problème précédent), l'élément d'intervalle d'espace-temps s'écrit ici

$$ds^2 = f(r) c^2 dt^2 - \frac{dr^2}{f(r)} - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2)$$

avec

$$f(r) = 1 + \frac{2\psi(r)}{c^2} \quad \text{et} \quad \psi(r) = -\frac{GM}{r}$$

On admettra que le lagrangien classique devient en relativité générale et pour ce problème

$$\mathcal{L} = \mathcal{L}(t, r; \theta, t', r', \theta') = \frac{1}{2} \frac{m}{f(r)} r'^2 + \frac{1}{2} m r^2 \theta'^2 - \frac{1}{2} m c^2 f(r) t'^2$$

où le  $'$  désigne maintenant la dérivée totale par rapport à  $s$ . Pour chaque coordonnée généralisée  $q^\mu = \{t, r; \theta, \varphi\}$ , l'équation de Lagrange correspondante est donc

$$\left( \frac{\partial \mathcal{L}}{\partial q^{\mu'}} \right)' - \frac{\partial \mathcal{L}}{\partial q^\mu} = 0$$

- a) Montrer que  $d\varphi = 0$  et que l'élément de longueur d'espace-temps s'écrit en fait pour ce problème

$$ds^2 = f(r) c^2 dt^2 - \frac{dr^2}{f(r)} - r^2 d\theta^2 \quad (11.9)$$

- b) Montrer que les équations de Lagrange pour les variables  $t$  et  $\theta$  conduisent respectivement à la conservation, le long de la trajectoire paramétrée par  $s$ , de 2 quantités  $k_1$  et  $k_2$  que l'on déterminera. En déduire la généralisation relativiste de la loi des aires.

$$r^2 \frac{d\theta}{dt} = -c^2 f(r) \frac{k_2}{k_1}$$

- c) A partir de la relation (11.9), montrer que l'on peut obtenir

$$\frac{1}{f(r)} \left( \frac{dr}{dt} \right)^2 + r^2 \left( \frac{d\theta}{dt} \right)^2 = f(r) c^2 \left[ 1 - \frac{c^2 f(r)}{k_1^2} \right]$$

- d) Montrer alors que la fonction  $u = 1/r$  vérifie, en variable  $\theta$ , l'équation différentielle

$$\frac{d^2 u}{d\theta^2} + u = K_1 u^2 + K_2 \quad (11.10)$$

où  $K_1 = K_1(G, M, c)$  et  $K_2 = K_2(G, M, c, k_2)$  sont 2 constantes que l'on déterminera.

- e) On cherche une solution de l'équation (11.10) sous la forme

$$u = \frac{GM}{A^2} [1 + e \cos((1 - \gamma)\theta)]$$

où  $e$  et  $\gamma$  sont 2 constantes et  $A$  désigne toujours la constante des aires classique.

- i. Montrer que ceci est possible pour un choix convenable de  $k_2$  et si l'on suppose  $e \ll 1$  ainsi que  $\gamma \ll 1$ .
  - ii. Exprimer dans ces conditions  $\gamma$  en fonction des constantes  $G, M, A$  et  $c$ . Comment s'interprète la quantité  $ck_2$ ?
- f) Application numérique

Les observations montrent que le périhélie de la planète Mercure avance de 574 secondes d'arc par siècle par rapport à la solution classique du problème des 2 corps qui le prévoit fixe. La prise en compte des diverses corrections classiques (aplatissement du soleil, pression de radiation, ...) n'introduit qu'une avance de 531 secondes d'arc par

siècle. Dans le modèle que nous venons d'établir, l'avance du périhélie est donnée par  $\delta\omega = 2\pi\gamma$  par révolution. Calculer la valeur de cette avance en seconde d'arc par siècle

On donne

$$\begin{aligned} G &= 6.72 \times 10^{-11} \text{kg.m}^3.\text{s}^{-1} & T_M &= 87.969 \text{ j} = 7.60 \times 10^6 \text{ s} \\ M_\odot &= 1.989 \times 10^{30} \text{kg} & a_M &= 5.787 \times 10^{10} \text{ m} \\ c &= 2.2279 \times 10^8 \text{m.s}^{-1} & e_M &= 0.2056 \end{aligned}$$

où  $T_M, a_M$  et  $e_M$  sont respectivement la période orbitale, le demi-grand axe et l'excentricité de l'ellipse dans la solution du problème classique des 2 corps sans perturbation appliquées à Mercure. On rappelle qu'un siècle compte 24 années bissextiles ...

## 11.3 Autour du théorème de Noether en théorie des champs

### 11.3.1 Le courant de Noether

On note  $x$  un évènement de l'espace-temps. On considère un système dont la densité de lagrangien  $\mathcal{L}$  dépend uniquement de  $i = 1, \dots, k$  fonctions  $\phi_i(x)$  et de leurs dérivées premières  $\partial_\mu\phi_i$  soit

$$\mathcal{L} = \mathcal{L}(\phi_1, \dots, \phi_k, \partial_\mu\phi_1, \dots, \partial_\mu\phi_k)$$

Dans ce contexte les fonctions  $\phi_i(x)$  sont appelées des champs scalaires. On suppose que ce lagrangien est invariant sous l'action d'un groupe de transformation  $\mathbb{G}$ . C'est à dire qu'il existe une fonction  $g_\sigma$  telle que

$$\begin{cases} g_0 = Id \\ \forall (\sigma_1, \sigma_2) ; g_{\sigma_1} \circ g_{\sigma_2} = g_{\sigma_1 + \sigma_2} \\ \forall \sigma ; \mathcal{L}(\dots, g_\sigma(\phi_i), \dots, \dots, \partial_\mu g_\sigma(\phi_i), \dots) = \mathcal{L}(\dots, \phi_i, \dots, \dots, \partial_\mu\phi_i, \dots) \end{cases}$$

Démontrer que la quantité

$$j^\mu = \sum_i \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_i)} \frac{\partial g_\sigma[\phi_i]}{\partial \sigma} \Big|_{\sigma=0}$$

est conservée, c'est-à-dire que sa divergence est nulle...

Cette quantité est appelée *courant de Noether* associée au groupe de symétrie  $\mathbb{G}$ .

Suggestion : on pourra montrer que  $\partial_\mu j^\mu = \frac{\partial \mathcal{L}}{\partial \sigma}$

### 11.3.2 Symétrie de jauge en électromagnétisme

On note  $x$  un évènement de l'espace-temps. Le champ de Maxwell est un système constitué d'un champ pur en interaction avec lui-même...

Son action est celle du chapitre 5, voir relation (??) page ??, sans particule ni courant de particule ( $J^\mu = 0$ ) soit

$$S = S(A_\mu) = \int \mathcal{L}_c d^4x \quad \text{avec} \quad \mathcal{L}_c = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (11.11)$$

où  $F_{\mu\nu}$  est le champ de Faraday associé au potentiel vecteur  $A_\mu$ .

1. Vérifier que l'action est invariante lors de la transformation

$$A_\mu(x) \longrightarrow g_\sigma[A_\mu(x)] = A_\mu(x) + \sigma \partial_\mu \alpha(x) \quad (11.12)$$

où  $\alpha$  est une fonction arbitraire de  $x$  et  $\sigma$  un paramètre réel de la transformation.

2. Calculer le courant de Noether  $j^\nu$  associé à la transformation (11.12) et en déduire qu'il existe une infinité de charges intégrées sur tout l'espace de la forme  $Q = \int j^0 d^3x$  qui sont conservées.

### 11.3.3 Symétrie infinitésimale

On considère un système dont la densité de lagrangien  $\mathcal{L}$  dépend uniquement de  $i = 1, \dots, k$  fonctions  $\phi_i(x)$  et de leurs dérivées premières  $\partial_\mu \phi_i$ . Dans ce contexte ces fonctions sont appelées des champs. On considère une transformation infinitésimale des champs  $\phi_i$  :

$$\phi_i(x) \longrightarrow \phi_i^*(x) = \phi_i(x) + \varepsilon \delta \phi_i(x) \quad \text{avec} \quad \varepsilon \ll 1 \quad (11.13)$$

On dit que le système présente une symétrie continue globale sous cette transformation si sa densité de lagrangien reste invariante à une dérivée totale près. C'est-à-dire s'il existe un vecteur  $k^\mu$  tel que

$$\mathcal{L}(x) \longrightarrow \mathcal{L}^*(x) = \mathcal{L}(x) + \varepsilon \partial_\mu k^\mu + \mathcal{O}(\varepsilon^2) \quad (11.14)$$

Ce vecteur  $k^\mu$  est le courant associé à cette transformation. Il est fonction des champs  $\phi_i(x)$  et de leurs dérivées.

L'action  $S = \int \mathcal{L} d^4x$  est invariante pour cette transformation si l'on a convenablement choisi les conditions aux limites.

En considérant une variation des champs  $\phi^i(x)$  donnée par (11.14) où le paramètre  $\varepsilon$  possède maintenant une dépendance arbitraire en la position  $x$ , montrer que le courant

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^i)} \delta \phi^i - k^\mu \quad (11.15)$$

est conservée pour toute solution des équations du mouvement, c'est-à-dire que  $\partial_\mu j^\mu = 0$  (on pourra utiliser le principe de moindre action). En déduire que la charge  $\mathcal{Q} = \int d^3x : j_0$  associée à cette symétrie est conservée au cours du temps.

### 11.3.4 Invariance par translation

Les deux exercices précédents ont mis en jeu deux formes de courant différentes. Il s'agit en fait du même objet. Dans les deux cas, la relation donnant le courant de Noether  $j^\mu$  est

$$\partial_\mu j^\mu = \delta \mathcal{L}$$

$\delta \mathcal{L}$  désignant la variation de la densité lagrangienne sous la symétrie considérée. Dans l'exercice précédent, cette variation était elle-même donnée par la dérivée totale  $\varepsilon \partial_\mu k^\mu$  ce qui nous a permis de définir une nouvelle quantité conservée,

$$J^\mu = j^\mu - k^\mu$$

tandis que dans l'exercice traitant du champ de Maxwell, cette variation était nulle et l'équation de conservation s'écrivait

$$\partial_\mu j^\mu = 0$$

Soit à présent un système invariant par translation. Déterminer la variation infinitésimale du lagrangien sous la transformation  $\phi(x) \rightarrow \phi(x^\mu + \varepsilon \xi^\mu)$  et en déduire le courant de Noether conservé  $j_\xi^\mu$  associé à cette translation de vecteur  $\xi^\mu$  grâce aux indications ci-dessus. Calculer le tenseur énergie-impulsion  $T_{\mu\nu}$  tel que  $j_\xi^\mu = \xi_\nu T^{\mu\nu}$ .

### 11.3.5 Tenseur de moment angulaire

On suppose maintenant que le système est également invariant sous les transformations de Lorentz,  $x^\mu \rightarrow x^\mu + \omega^{\mu\nu} x_\nu$  où  $\omega^{\mu\nu} = -\omega^{\nu\mu}$  est un facteur

infinitésimal. Montrer que le courant de Noether associé à cette transformation est donné par  $j_\omega^\mu = \omega_{\nu\rho} M^{\mu\nu\rho}$ , où  $M^{\mu\nu\rho}$  est le tenseur de moment angulaire,

$$M^{\mu\nu\rho} = x^\rho T^{\mu\nu} - x^\nu T^{\mu\rho} \quad (11.16)$$

### 11.3.6 Non unicité du tenseur énergie-impulsion

1. Montrer qu'en général, on peut ajouter à  $T^{\mu\nu}$  un terme  $\partial_\rho U^{\mu\nu\rho}$  où  $U^{\mu\nu\rho} = -U^{\rho\nu\mu}$  est un tenseur arbitraire antisymétrique en ses premiers et derniers indices, sans compromettre l'équation de conservation  $\partial_\mu T^{\mu\nu} = 0$ , ni modifier l'impulsion  $P^\mu = \int d^n x T_{0\mu}$ .
2. Montrer que la conservation du tenseur de moment angulaire (11.16) requiert que le tenseur énergie-impulsion soit symétrique.

### 11.3.7 Tenseur énergie impulsion du champ scalaire

On considère un champ scalaire réel d'action  $\mathcal{S} = \int [\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi)] d^4x$ .

1. Calculer le tenseur énergie-impulsion grâce à la formule de Noether obtenue à l'exercice 11.3.4 en considérant un lagrangien invariant par translation.
2. Montrer que l'on peut ajouter à  $T_{\mu\nu}$  un combinaison linéaire particulière

$$\alpha\partial_\mu\phi\partial_\nu\phi + \beta\eta_{\mu\nu}\phi\partial_\sigma\partial^\sigma\phi + \gamma\phi\partial_\mu\partial_\nu\phi$$

sans modifier l'équation de conservation  $\partial_\mu T^{\mu\nu} = 0$ . Montrer qu'on peut ainsi obtenir un tenseur énergie-impulsion de trace nulle,  $T^{\mu\nu}\eta_{\mu\nu} = 0$ .

### 11.3.8 Tenseur énergie impulsion du champ de Maxwell

1. Calculer le tenseur énergie-impulsion  $T_{\mu\nu}$  du champ de Maxwell grâce à la formule de Noether obtenue à l'exercice 11.3.4 en considérant un lagrangien invariant par translation.
2. Montrer que l'on peut lui ajouter un terme  $\partial_\rho U^{\mu\nu\rho}$  de sorte qu'il devienne symétrique et invariant de jauge.
3. Vérifiez que le tenseur énergie-impulsion du champ de Maxwell en dimension 4 est de trace nulle,  $T^{\mu\nu}\eta_{\mu\nu} = 0$ . Pouvez-vous expliquer cette observation par une symétrie particulière du lagrangien de Maxwell?

### 11.3.9 Généralisation du tenseur énergie impulsion

On considère l'action

$$\mathcal{S} = \int \left[ \frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) \right] \sqrt{-g} d^4x$$

où  $g = \det g_{\mu\nu}$  le tenseur métrique généralisé. On considère également une variation infinitésimale de la métrique au voisinage du tenseur de Minkowski,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$$

1. En utilisant l'expression du tenseur énergie-impulsion  $T_{\mu\nu}$  obtenu par la formule de Noether que nous rappelons ici,  $T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \partial_\nu \phi - \eta_{\mu\nu} \mathcal{L}$ , montrer que cette fluctuation infinitésimale résulte au premier ordre en une variation de l'action de la forme  $\mathcal{S} = \int h_{\mu\nu} T^{\mu\nu} d^4x$ .
2. Plus généralement, le tenseur énergie-impulsion est défini comme la variation de la densité lagrangienne par rapport à la métrique,

$$T^{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta \mathcal{S}}{\delta g^{\mu\nu}}$$

Calculer ce tenseur pour le champ de Maxwell d'action

$$\mathcal{S} = \int [F_{\mu\nu} F_{\rho\sigma} g^{\mu\rho} g^{\nu\sigma}] \sqrt{-g} d^4x$$

et comparer au tenseur obtenu par la méthode Noether.

---

# Bibliographie

1. Vladimir Igorevitch Arnol'd. *Equations différentielles ordinaire*. Moscou : Mir ; Paris : Ellipses, 1988.
2. Vladimir Igorevitch Arnol'd. *Mathematical methods of classical mechanics, 2nd edition*. New York : Springer-Verlags, 1989.
3. Laurie M. Brown. *Feynman's Thesis — A New Approach to Quantum Theory*. World Scientific, 2005.
4. Moshe Carmeli. *Classical fields : general relativity and gauge theory*. New York : John Wiley & Sons, 1982.
5. Paul Adrien Maurice Dirac. *Les principes de la mécanique quantique*. Presses polytechniques et universitaires romandes (réédition de l'ouvrage original de 1927), 2009.
6. Richard P. Feynmann. Space-time approach to non-relativistic quantum mechanics. *Reviews of Modern Physics*, 20:367–387, 1948.
7. Claude Gignoux and Bernard Silvestre-Brac. *Mécanique, de la formulation lagrangienne au chaos hamiltonien*. EDP Sciences, collection Grenoble Sciences, 2002.
8. Claude Gignoux and Bernard Silvestre-Brac. *Problèmes corrigés de mécanique et résumés de cours : de Lagrange à Hamilton*. EDP Sciences, collection Grenoble Sciences, 2004.
9. Domenico Giulini. *Algebraic and geometric structures of special relativity*. Lectures notes in Physics, Vol. 702, Springer, 2006.
10. Herbert Goldstein. *Mécanique classique*. Paris : Presses universitaires de France, 1964.
11. Ericourgoulhon. *Relativité restreinte, des particules à l'astrophysique*. Collection Savoirs Actuels, EDP Sciences (Les Ulis) / CNRS Editions (Paris), 2010.
12. Rémi Hakim. *Gravitation relativiste*. Paris : InterEditions : CNRS éditions, 1994.
13. John David Jackson. *Classical electrodynamics third edition*. New York : John Wiley & Sons, 1999.
14. Lev Davidovitch Landau and Evgueni Mikhailovitch Lifchits. *Physique théorique - Tome 1 - Mécanique 5e édition*. Moscou : Mir ; Paris : Ellipses, 1994.
15. Lev Davidovitch Landau and Evgueni Mikhailovitch Lifchits. *Physique théorique - Tome 2 - Théorie des champs 5e édition*. Moscou : Mir ; Paris : Ellipses, 1994.
16. Jean-Marc Levy-Leblond and Françoise Balibar. *Quantique : Rudiments*. Paris : Dunod, 1984.
17. Jean-Marc Levy-Leblond, Françoise Balibar, Alain Laverne, and Dominique Mouhanna. *Quantique : Éléments*. Non publié, 2016.
18. J.E. Marsden, R. Montgomery, P.J. Morrison, and W.B. Thompson. Covariant poisson brackets for classical field. *Annals of Physics*, 169:29–47, 1986.
19. Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler. *Gravitation*. San Francisco : W.H. Freeman, 1973.

20. Jérôme Perez. *Gravitation Classique, Seconde Edition*. Les presses de l'Ensta, Palaiseau, 2012.
21. Jean-Michel Raimond. *Théorie classique des champs*. Notes de cours du Magistère Inter-Universitaire de Physique, 2002.
22. Yves Talpaert. *Leçons et applications de géométrie différentielle et de mécanique analytique*. Toulouse : Cépaduès, 1993.
23. Steven Weinberg. *Gravitation and cosmology : principles and applications of the general theory of relativity*. New York : John Wiley & Sons, 1972.
24. Hermann Weyl. *Temps, espace, matière : leçons sur la théorie de la relativité générale, 4e éd.* Paris : A. Blanchard, 1979.

---

# Index

- 4–force, 137
- énergie
  - cinétique, 12, 32
  - potentielle, 14
- énergie-impulsion, tenseur, 183
- équation
  - de Liouville, 92
  - de Schrödinger, 206
  - de Vlasov, 93
  - des géodésiques, 162
  - fondamentale de la mécanique classique, 65
- équations
  - d'Einstein, 186
- équivalence, principe d', 159
- éther, 98
  
- accélération
  - 4–accélération, 117
  - définition, 25
- action, 44, 47, 239, 241
  - classique, 27
  - d'une particule libre relativiste, 134
  - de Feynmann, 229
  - de matière, 182
- ailleurs, 110
- algèbre, définition, 66
- angle-action
  - variable, 79
  
- Bernoulli, Daniel, 17
- Bianchi, identité de , 174
- Binet, Jacques-Philippe , 20
- Boltzmann, équation de, 94
- Boltzmann, Ludwig, 92
- Bonaparte, Napoléon, 22
- Borda, Jean-Charles de , 19
  
- bra, 212
- bracket, 212
  
- calorique, 98
- canonique
  - transformation, 72
- champ
  - de vecteurs, 89
  - tenseur, 137
  - tenseur, composante contravariante, 145
  - tenseur, composante covariante, 139
- chronologique, cône, 110
- cinétique
  - énergie, 12
- Clairaut, Alexis, 7
- classique
  - équation fondamentale de la mécanique, 65
- collision, opérateur de, 93
- composante, 107
- connection affine, 162
- constante
  - cosmologique, 187
- contrainte
  - de roulement sans glissement, 44
  - holonome, 41
  - rhéonome, 44
  - scléronome, 44
- contravariante, composante, 106
- coordonnée
  - coordonnée généralisée, 12
- cosmologique, constante, 187
- Coulomb, Charles de, 19
- courant
  - 4–courant, 119
- courbure
  - action de,, 178

- scalaire de , 175
- covariante
  - dérivée, 165
- covariante, composante, 106
- crochet, 237
- crochet de Poisson, introduction historique, 17
- crochets de Poisson
  - théorème des,, 86
- cyclique, coordonnée, 71
  
- d'Alembert, Jean le Rond, 17
- d'Alembert, Jean le rond, 7
- dérivée, covariante, 165
- degré
  - de liberté, définition, 26
- Dicke, Robert, 157
- diophantienne, condition, 85
- Dirac, Paul, 218
- divergence, 181
- Doppler, effet relativiste, 121
- dynamique
  - principe fondamental relativiste, 137
  
- Ecoc, 225
- Einstein
  - équation d', 186
  - convention d', 105
  - tenseur d', 176, 182
- elliptiques
  - éléments, 8
  - éléments osculateurs , 8
- Eotvos, Roland, 157
- espace, 25, 50, 185
  - des configurations, 50
  - des phases, 50
  - homogène, 53
  - isotrope, 54
- espace, genre, 110
- Euler
  - équation de la mécanique des fluides, 93
- Euler, Léonard, 14
  
- Faraday
  - loi d'induction de, 142
  - tenseur de,, 137
  - tenseur, composante contravariante, 145
  - tenseur, composante covariante, 139
- fermeture, relation de, 213
- Feynman, action de, 229
- Feynmann, Richard, 229
- Fick, loi de, 94
  
- flot d'un champ de vecteurs, 89
- force
  - définition, 3
- Fourier, loi de, 94
  
- génératrice, fonction, 72
- géodésiques, équation des, 162
- Galilée
  - principe de,, 31
  - transformation de,, 31
- Galilée, transformation de, 99
- Gauss, Karl Friedrich, 85
- Green, fonction de, 93
  
- Hamilton, 236, 237
  - fonction de, hamiltonien, 59
- hamiltonien, 59
  - électromagnétique classique, 62
- hamiltonien, nomination historique, 16
- Hilbert, David, 187
- Hilbert, espace de , 198
- holonomes, contrainte, 41
- Hooke, Robert, 3
- Huygens, Christiaan, 15
- huygensien, 15
  
- intégrable
  - système intégrable, définition, 77
- intégrables, système, 77
  
- Jacobi
  - identité de Jacobi, 65
- Jacobi, Karl Gustav, 63
- jauge, 31, 134
  - de Lagrange, 7
  - liberté de, classique, 29
  - transformation de,, 49
- jauge, symétrie de, 232
- jauge, transformation de, 147
  
- KAM, théorème, 84
- ket, 212
- Kolmogorov, Andrei, 85
- Kronecker, symbole, 106
- Kronecker, tenseur de, 124
  
- lagrangien, 50
- Lagrange, 29, 44
  - équation de, introduction historique, 14
  - équations planétaires de , 10
  - fonction de, lagrangien, 27
  - matrice de , 10

- parenthèses de, 10  
 Lagrange-Poisson  
   théorème de, 66  
 lagrangien  
   électromagnétique classique, 49  
   étendu, 55  
   d'une particule classique, 27  
   de Feynmann, 229  
   particule libre classique, 31  
 Laplace, Pierre-Simon, 7  
 Laplace, Pierre-Simon, 19  
 Lavoisier, Antoine, 19, 98  
 Legendre  
   transformation de, 59  
 Legendre, transformation de, 15  
 Leibniz, règle de, 65  
 Levi-Civita, tenseur de, 126  
 liberté  
   degrés de, définition, 26  
 Lie  
   algèbre de Lie, 66  
 Liouville  
   équation de, 92  
   théorème de, 89  
 Lorentz, 47, 102  
   force de, 142  
 Lorentz, groupe de, 102  
 Lorentz, transformation de, 99  
 lumineux, cône, 110  
  
 métrique, tenseur, 162  
 métrique, tenseur en relativité restreinte, 105  
 Mac Laurin, Colin, 3  
 matière, action de, 182  
 Maxwell  
   équation de (second groupe), 147  
   équations de (premier groupe), 140  
 Maxwell, James Clerk, 97  
 Mercure, 257  
 Michelson, expérience de, 98  
 Mirabeau, Honoré Gabriel Riqueti comte de,  
   18  
  
 nabla, 4-vecteur, 128  
 Noether  
   couple de, 70  
   théorème classique de, 51  
   théorème de (quantique), 232  
 Noether, Emmy, 51  
  
 observable, 202  
 Ohm, loi d', 94  
  
 Okam, principe d', 203  
 onde  
   4-vecteur d'onde, 120  
  
 parenthèses de Lagrange, 10  
 phlogistique, 98  
 Poincaré, groupe de, 102  
 Poincaré, Henri, 63  
 Poisson, 236  
   équation de, 191  
   équation de, 93, 146  
   crochet de, définition, 64  
 Poisson, Siméon Denis, 20  
 Poisson, Siméon-Denis, 16  
 position  
   4-vecteur, 115  
 potentiel  
   de champ moyen, 93  
 potentielle  
   énergie, 14  
   pression, 147  
   principe, 162  
   d'équivalence, 159  
   de relativité généralisée, 160  
   des travaux virtuels, 37  
   du rasoir d'Okam, 203  
   fondamental de la dynamique relativiste,  
   137  
 propagateur, 68  
 propres  
   valeur, 199  
  
 résolvante, 68  
 rapidité, 101  
 rhéonome, contrainte, 44  
 Ricci, tenseur de, 174  
 Riemann-Christoffel, courbure de, 172  
  
 scalaire  
   4-scalaire, 119  
 Schrödinger, équation de, 206  
 scléronome, contrainte, 44  
 stationnaire, état, 229  
  
 temps, 30  
 temps propre, 116  
 temps, genre, 110  
 tenseur, 124, 187  
   d'énergie-impulsion, 183  
   d'Einstein, 176  
   de Ricci, 174  
 tenseur, définition, 122

tensoriel, produit définition, 122  
théorème  
  de Lagrange-Poisson, 66  
  de Liouville des physiciens, 89  
  de Noether (quantique), 232  
  des crochets de Poisson, 86  
  KAM, 84  
tore de KAM, 85  
trésor, de l'Ensta, 22  
variable

  angle-action, 79  
vecteur  
  champ de , 89  
vitesse  
  4-vitesse, 117  
  définition, 25  
  vitesse généralisée, 12  
Vlasov, équation de, 93  
Weyl, Hermann, 147  
Wronski, Josef Hoëné, 6

# Ouvrages parus aux Presses de l'ENSTA

## Collection «Les cours »

1. Alain Bovis, *Hydrodynamique navale : le sous-marin*. 2<sup>ème</sup> édition, 2016. ISBN 978-2-7225-0953-5.
2. Jean-François Delmas, *Introduction au calcul des probabilités et à la statistique*. 4<sup>ème</sup> édition. 2016. ISBN 978-2-7225-0952-8.
3. Gabriel Chardin, Pascal Debu, Jérôme Pérez, *Éléments de physique statistique*. 2<sup>ème</sup> édition 2014. ISBN 978-2-7225-0948-1
4. Jean Pierre Bessis (sous la direction de), *Philosophie et religions du Livre*. 2013. ISBN 978-2-7225-0930-6.
5. Jean-François Delmas, *Introduction au calcul des probabilités et à la statistique*. 2<sup>ème</sup> édition. 2013. ISBN 978-2-7225-0943-6.
6. Frédéric Jean, *Stabilité et commande des systèmes dynamiques : cours et exercices corrigés*, 2011, ISBN 978-2-7225-0936-8.
7. Jérôme Perez, *Gravitation classique : problème à N corps, de 2 à l'infini*, 2<sup>ème</sup> édition, 2011. ISBN 978-2-7225-0932-0.
8. Eliane Bécache, Patrick Ciarlet, Christophe Hazard, Eric Lunéville, *La méthode des éléments finis. Tome II : Compléments*, 2010. ISBN 978-2-7225-0923-8.
9. Michèle Fieux, *L'océan planétaire*, 2010. ISBN 978-2-7225-0915-3.
10. Alexis Rosenbaum, *Leçons d'introduction à la philosophie des sciences*, 2009. ISBN 978-2-7225-0918-4.
11. Patrick Ciarlet, Eric Lunéville, *La méthode des éléments finis. Tome I : Concepts généraux*, 2009. ISBN 978-2-7225-0917-7.
12. Alain Bovis, *Hydrodynamique navale : théories et modèles*, 2009. ISBN 978-2-7225-0916-0.
13. Jérôme Perez, *Gravitation classique : problème à N corps, de 2 à l'infini*, 2008. ISBN 978-2-7225-0914-6.
14. Jérôme Perez, *Théorie des champs classiques*, 2008. ISBN 978-2-7225-0913-9.

## Collection «Les invitations »

1. Patrick Rimond, *Portraits*, 2009. ISBN 978-2-7225-0920-7.
2. Bruno Hilaire, *Trente-deux*, 2009. ISBN 978-2-7225-0919-1.

## Collection «Les actes »

### • Les colloques

1. Diogo Queiros-Conde, Michel Feidt (Sous la direction de) , *Constructal theory and multi-scale geometries: Theory and applications in energetics, chemical engineering and materials*, 2009. ISBN 978-2-7225-0921-4.

## • Les tables rondes de l'ENSTA

1. Georges Chapouthier, Jean-Gabriel Ganascia, Lionel Naccache, Pascal Picq, *Que reste-t-il du propre de l'Homme* 2012. ISBN 978-2-7225-0938-2.
2. Jonny Wilkinson, Etienne Klein, Jean Iliopoulos, *Rugby quantique*, 2011. ISBN 978-2-7225-0928-3.
3. Jean Claude Ameisen, Etienne Klein et Dominique Leglu, *Ecrire la science*, 2010. ISBN 978-2-7225-0924-5.
4. André comte-Sponville, François Euvé, Guillaume Lecointre, *Dieu et la science*, 2011. ISBN 978-2-7225-0927-6.

## • Les Nouvelles avancées

1. Concours de nouvelles présidé par Jérôme Ferrari, Etienne Klein et al., *Dans la peau d'Archimède, Einstein et les autres*, 2016, ISBN 978-2-7225-0954-2.
2. Concours de nouvelles présidé par Cédric Villani, Fatou Diome et al., *Comme c'est curieux, comme c'est bizarre et quelle coïncidence!*, 2015, ISBN 978-2-7225-0950-4.
3. Concours de nouvelles présidé par Cédric Villani, Fatou Diome et al., *Mélange des genres : panique chez les taxons*, 2014, ISBN 978-2-7225-0945-0.
4. Concours de nouvelles présidé par Cédric Villani et al. *Hors-la-loi : quand la pomme ne tombe plus*, 2013, ISBN 978-2-7225-0941-2.
5. Concours de nouvelles présidé par Alain Thébault et al. *Eaux d'ici, eaux de là : 2012, l'odyssée de l'eau*, 2012, ISBN 978-2-7225-0939-9.
6. Concours de nouvelles présidé par Andrea H. Japp et al., *Un bug, un clic, un imprévu... et tout bascule*, 2011. ISBN 978-2-7225-0929-0.
7. Concours de nouvelles présidé par Pierre Bordage et al., *2084 : le meilleur ou le pire des mondes ?*, 2010. ISBN 978-2-7225-0925-2.

## • Hors collection

1. Jean-Claude Dern, Jean-Marc Quenez, Philip Wilson (sous la direction de) *Compendium of ship hydrodynamics : Practical Tools and Applications*, 2016. ISBN 978-2-7225-0949-8.

## **Les presses de l'ENSTA**

imprimé en France

Dépôt légal :

**Couverture** : Création ENSTA-ParisTech

