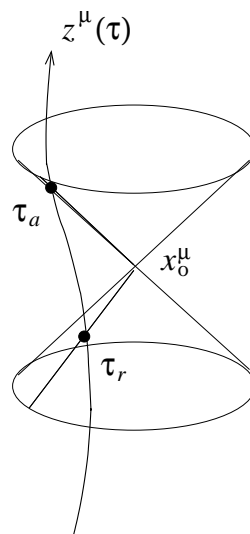


Licence de Physique et un peu après...



Relativité

dynamique, électromagnétisme, gravitation,
cosmologie, mécanique quantique

Bertrand Berche
Laboratoire de Physique des Matériaux,
Université Henri Poincaré, Nancy 1

Relativité

Sommaire

Chapitre 1 : Le principe de relativité	1
1. Perspective historique	1
1.1. Du principe d'inertie au principe de relativité	1
1.2. Dynamique et gravitation newtoniennes	2
1.2.1. <i>Les lois du mouvement</i>	2
1.2.2. <i>Gravitation</i>	3
1.2.3. <i>Invariance galiléenne et principe d'inertie</i>	4
2. Structure métrique de l'espace-temps newtonien	5
2.1. L'espace physique	5
2.2. Le temps	5
2.3. Espace-temps classique	5
2.4. Référentiels inertiels et gravitation	6
2.4.1. <i>Gravitation et forces d'inertie</i>	7
2.4.2. <i>Caractère local d'un référentiel inertiel en présence d'un champ de gravitation</i>	8
3. Invariance de la vitesse de la lumière dans les référentiels inertiels	8
3.1. Expérience de Michelson et Morley	8
3.2. Effet Doppler classique	11
3.3. Statut des équations de Maxwell, programme de la théorie de la relativité	12
3.4. Conséquences immédiates de l'invariance de c	13
4. Le principe de relativité d'Einstein	15
4.1. Les postulats de la relativité restreinte	15
4.2. Intervalle quadridimensionnel	16
4.2.1. <i>La notion d'intervalle</i>	16
4.2.2. <i>Invariance de l'intervalle</i>	16

4.3. Structure causale de l'espace-temps de Minkowski	17
Chapitre 2 : Transformation de Lorentz	21
1. Transformation de Lorentz des coordonnées d'espace-temps	21
1.1. Etablissement de la transformation de Lorentz	21
1.2. Transformation de Lorentz pour une vitesse de direction quelconque . .	23
1.3. Loi de composition des vitesses	23
1.4. Rapidité et rotation	25
2. Quelques conséquences de la transformation de Lorentz	27
2.1. Intervalle invariant	27
2.2. Contraction des longueurs	27
2.3. Dilatation du temps	27
3. Vérifications expérimentales	28
3.1. L'expérience du Mont Washington	28
3.2. L'effet Doppler relativiste	29
4. Calcul k	30
4.1. Principe de calcul k	30
4.2. Effet Doppler	31
4.3. Paradoxe des jumeaux	32
5. Généralisation à des transformations non linéaires	34
Chapitre 3 : Formalisme tensoriel	35
1. Quadri-vecteurs et tenseurs	35
1.1. Quadri-vecteurs contravariants et covariants	35
1.2. Le tenseur métrique	40
1.3. Tenseurs	41
2. Exemples	46
2.1. Temps propre	46
2.2. Quadri-vitesse	46
2.3. Quadri-accelération	47
2.4. Quadri-courant	48
3. Intérêt du formalisme tensoriel en relativité	50
Chapitre 4 : Dynamique relativiste	53
1. Quantité de mouvement et énergie	53
1.1. Préambule	53
1.2. Quadri-vecteur énergie-impulsion	54
1.3. Digression : masse au repos, masse en mouvement	56
1.4. Particules de masse nulle	57
2. Cinématique des réactions entre particules	58

2.1. Terminologie	58
2.2. Le problème de la conservation de la masse	58
2.3. Notion de masse d'un système composé	59
2.4. Désintégrations de particules	60
2.5. Collisions élastiques	61
2.6. Collisions inélastiques et seuil de réaction	63
3. Forces conservatives et équations du mouvement	66
3.1. Notion de force	66
3.1.1. <i>Expression de la force et de la relation fondamentale de la dynamique</i>	66
3.1.2. <i>Quadri-force</i>	67
3.2. Mouvement d'une charge dans un champ magnétique uniforme	67
3.3. Problème de Kepler	69
3.3.1. <i>Les états liés du potentiel newtonien</i>	69
3.3.2. <i>Ecart au comportement newtonien, déplacement du périhélie des planètes</i>	70
3.3.3. <i>Les observations expérimentales dans le cas de Mercure</i>	71
3.3.4. <i>La rosette de Sommerfeld et l'avance du périhélie de Mercure en dynamique relativiste</i>	77
4. Retour aux transformations non linéaires	79

Chapitre 5 : Formulation covariante de l'électromagnétisme et de l'électrodynamique 81

1. Quadrivecteur potentiel électromagnétique	81
1.1. Equations de propagation des ondes	81
1.2. Jauge de Lorentz et équations de propagation en notation covariante	82
1.3. Transformation des potentiels	83
2. Tenseur champ électromagnétique et équations de Maxwell	83
2.1. Définition du tenseur champ électromagnétique	83
2.2. Changement de référentiel inertiel	85
2.3. Equations de Maxwell covariantes	87
2.4. Force de Lorentz	88
3. Applications	89
3.1. Champ d'une charge ponctuelle en mouvement arbitraire	89
3.1.1. <i>Quadripotentiel A^μ</i>	89
3.1.2. <i>Composantes ordinaires des champs \mathbf{E} et \mathbf{B} de la charge en mouvement arbitraire</i>	92
3.1.3. <i>Cas limites</i>	96
3.1.4. <i>Tenseur champ électromagnétique $F^{\mu\nu}$</i>	96
3.2. Application au cas d'une charge ponctuelle en mouvement uniforme	97
3.3. Calcul approché du champ électrique rayonné	99

Chapitre 6 : Formulation lagrangienne de la relativité restreinte 103

1. Le principe de moindre action pour la particule libre	103
1.1. Le principe de moindre action en physique classique	104

1.2. Le principe de moindre action pour la particule libre en relativité	105
1.2.1. <i>L'action de la particule libre en dynamique relativiste</i>	105
1.2.2. <i>L'action en notation tridimensionnelle</i>	106
1.2.3. <i>Le principe de moindre action dans le formalisme de Minkowski</i>	107
2. Particule chargée sous champ	108
2.1. Action invariante	109
2.2. Calcul variationnel en notation tridimensionnelle	109
2.3. Calcul variationnel dans le formalisme de Minkowski	110
2.4. Equations d'Euler-Lagrange covariantes	111
2.5. Invariance de jauge	113
2.6. Densité lagrangienne	113
2.7. Généralisation	115
Chapitre 7 : Géométrie des surfaces courbes	117
1. Le cinquième postulat d'Euclide	117
2. Courbure des surfaces bidimensionnelles	119
2.1. Géodésiques et vols intercontinentaux	120
2.2. Le théorème de Pythagore et l'équation métrique	120
2.3. Mesure de la courbure d'une surface bidimensionnelle	125
2.4. Courbure de Gauss et courbure moyenne	127
2.5. Courbure et équation métrique : l'excellent théorème de Gauss	131
Chapitre 8 : Le principe d'équivalence	135
1. Les expériences d'Eötvös-Dicke	135
2. Principe d'équivalence faible et principe d'équivalence fort	140
3. Interactions fondamentales et principe d'équivalence	142
4. Les constantes fondamentales	143
5. Le principe de Mach	146
Chapitre 9 : Gravitation relativiste	151
1. Le décalage spectral gravitationnel	151
2. Courbure de l'espace-temps en présence d'un champ de gravitation statique à symétrie sphérique	153
3. La métrique de Schwarzschild	155
3.1. Tenseur métrique pour un champ statique à symétrie sphérique	155
3.2. Equation géodésique dans la métrique de Schwarzschild	156
3.2.1. <i>Retour à l'action classique</i>	157
3.2.2. <i>Le principe de moindre action en relativité générale</i>	158
4. Les trois tests classiques de la théorie de la relativité générale	162
4.1. Vérification du décalage spectral gravitationnel	162
4.2. Avance du périhélie de Mercure	163
4.3. Déflexion des rayons lumineux	166

4.4. Un quatrième test, retard des échos radar	169
5. Les équations d'Einstein	171
5.1. Tenseur de Riemann, tenseur énergie impulsion, tenseur d'Einstein	171
5.2. Limite newtonienne de l'équation d'Einstein	173
5.3. Action d'Einstein - Hilbert	175
5.4. Conclusion	176

Chapitre 10 : Cosmologie relativiste 178

1. Introduction	178
2. Récession des galaxies, hypothèse de Hubble	179
3. Cosmologie newtonienne	181
3.1. Principe cosmologique	181
3.2. Loi de Hubble et ballon de baudruche	181
3.3. Dynamique de l'Univers plat	182
4. Cosmologie relativiste	184
4.1. L'Univers comme un fluide	184
4.2. Le modèle de Robertson et Walker	184
4.2.1. La métrique de Robertson et Walker	184
4.2.2. Référentiel comobile	185
4.2.3. Paramètres mesurables	186
4.3. Dynamique : les équations de Friedmann	187
4.4. Le "problème" de la constante cosmologique	190

Chapitre 11 : Théories quantiques relativistes 195

1. Equation de Schrödinger	195
2. Equation de Klein-Gordon	196
3. Equation de Pauli	197
4. Equations de Weyl	198
4.1. Recherche des équations pour des particules sans masse de spin $\frac{1}{2}$	198
4.2. Equation covariante	199
4.2.1. Spineurs et matrices de Dirac	199
4.2.2. Énergies négatives	199
4.2.3. Forme covariante de l'équation de Weyl	200
5. Equation de Dirac	201
5.1. Equation covariante pour des fermions massifs de spin $\frac{1}{2}$	201
5.2. Equation de Dirac covariante sous champ	202
5.3. Une formulation alternative	203
6. Synthèse	204

Eléments de bibliographie 207

1. Incontournables...	207
-------------------------------	-----

2. D'excellents compléments	207
3. Des textes plus élémentaires	208
4. Pour prolonger la réflexion...	209
5. Aspects historiques - épistémologiques	210
6. Des articles pédagogiques sur des sujets ponctuels	211

Préambule

Ce cours est à peu près du niveau de la licence de physique, avec parfois quelques notions un peu plus élaborées peut-être (certainement). Le cours donné en licence repose sur les chapitres 1 à 9. Les deux chapitres suivants (toujours inachevés) sont ajoutés pour leur intérêt physique et pour me faire plaisir.

Le premier chapitre est une introduction qui discute les notions classique et relativiste de l'espace-temps et qui présente le programme de la relativité dès lors qu'on en a admis les postulats : refonte de la cinématique (chapitre 2) et de la dynamique (chapitre 4), reformulation de l'électromagnétisme (chapitre 5). Entre-temps, au chapitre 3 sont données quelques notions de calcul tensoriel qui permettent d'écrire des expressions conformes aux principes de relativité de manière à ce que la covariance soit évidente (on dit manifeste). Le chapitre 6 donne une formulation alternative basée sur le calcul variationnel. Les trois chapitres 7 à 9 ont pour objet d'inclure la gravitation. Le chapitre 7 discute du principe d'équivalence, le chapitre 8 donne quelques notions de géométrie non euclidienne en partant de l'exemple de la géométrie des surfaces (Gauss) et en admettant la validité plus générale de certains résultats qui sont établis ici dans des cas particuliers. Le chapitre 9 établit la forme de la métrique en présence d'un champ de gravitation statique à symétrie sphérique (Schwarzschild).

Dans le chapitre 10 on présente quelques notions de cosmologie et le chapitre 11 a pour but d'établir les équations qui régissent la physique quantique et satisfont à la covariance relativiste (de manière rapide et sans en discuter les conséquences).

La bibliographie est présentée à la fin, mais on peut d'ores et déjà indiquer les ouvrages les plus utilisés (parfois abondamment) dans ce cours :

L.D. Landau et E. Lifshitz, *Théorie des champs*, 3ème édition, Editions MIR, Moscou 1970.

I.R. Kenyon, *General relativity*, Oxford University Press, New York 1990.

M.V. Berry, *Principles of cosmology and gravitation*, Institute of Physics Publishing, Bristol 1989.

M. Lambert, *Relativité restreinte et électromagnétisme*, Ellipses, Paris 2000.

Notes de cours de David Sénéchal,
en particulier *Mécanique I* et *Ondes électromagnétiques* que l'on peut trouver sur
<http://www.physique.usherb.ca/~dsenech/>

Par ailleurs, s'il y a une référence incontournable dans le domaine (même si je l'ai assez peu utilisée), c'est probablement

S. Weinberg, *Gravitation and Cosmology*, Wiley, New-York 1972.

Finalement, ce texte est disponible en ligne à l'adresse

http://www.lpm.u-nancy.fr/webperso/berche.b/0-REL_tot.ps

Au cas où cette adresse ne soit plus fonctionnelle, on peut toujours tenter une version .pdf ou une version compressée .ps.gz.

Chapitre 1

Le principe de relativité

Dans ce chapitre sont présentées les conceptions classique et relativiste de la structure de l'espace-temps. De nombreux textes de présentation de la relativité discutent évidemment les aspects développés ici. Je tiens à mentionner, même si ce n'est pas disponible facilement sous forme écrite, un exposé remarquable de Ph. Lombard (Nancy 1) donné aux "Archives Henri Poincaré" (Nancy 2, <http://www.univ-nancy2.fr/ACERHP/philscien/pspage.html>). On consultera plus aisément par exemple le chapitre sur la relativité dans le livre de Gruber.

1. Perspective historique

1.1. Du principe d'inertie au principe de relativité

Le principe de relativité est lié de très près au principe d'inertie. On peut y associer trois noms :

- i) Galilée (1564-1642) qui a entre autres travaux scientifiques fondé la cinématique, étudié la chute des corps, la composition des mouvements et par là ébauché le principe d'inertie, sans pour autant pouvoir lui donner corps car la notion de causalité n'est pas encore claire à l'époque de la Renaissance.
- ii) Newton (1642-1727), le père de la conception d'un espace et d'un temps absolus. Il cherche à établir des liens de causalité (notion de force), ce qui l'amène à fonder la dynamique : les lois générales du mouvement exprimées sous forme différentielle (déterministe). Le principe d'inertie occupe une place centrale (dans un certain référentiel absolu, un corps qui n'est soumis à

aucune force persévère indéfiniment dans son état de mouvement rectiligne et uniforme) et permet la définition de référentiels inertiels (ceux dans lesquels le principe d'inertie vaut). Le rôle privilégié joué par ces référentiels inertiels dans la forme des lois de la dynamique conduit à l'invariance galiléenne ou la première formulation du principe de relativité, à savoir l'invariance des lois de la dynamique par changement de référentiel inertiel. On parle également de relativité galiléenne (expression forgée par Poincaré bien plus tard) et de référentiels galiléens. Newton élabore également une théorie de la gravitation (attraction universelle).

- iii) Einstein (1879-1955), dont les contributions à la physique sont célèbres. Il étend le principe de relativité à l'ensemble des lois de la physique. Cela le conduit à la théorie de la relativité restreinte (special relativity) qui nécessite une refonte révolutionnaire des concepts d'espace et de temps, une révision de la cinématique puis une nouvelle formulation de la dynamique. L'électromagnétisme n'est pas fondamentalement touché par cette nouvelle théorie, mais reformulé de manière extrêmement compacte. Le rôle particulier joué par la gravitation amène Einstein à proposer une extension du principe de relativité aux référentiels accélérés, c'est la relativité générale.

1.2. Dynamique et gravitation newtoniennes

Dans les *Principia mathematica*⁽¹⁾, Newton pose les fondements des lois de la dynamique et de celles de la gravitation. Ce sont deux corpus formels distincts, mais qui une fois combinés expliquent les mouvements par exemple des planètes.

1.2.1. Les lois du mouvement

L'espace et le temps ont un caractère absolu (c'est-à-dire indépendant de tout observateur). La causalité est exprimée par le concept de force : la force est définie comme la cause de toute déviation au principe d'inertie.

- i) Il existe un référentiel absolu dans lequel un corps qui n'est soumis à aucune force suit un mouvement de translation rectiligne et uniforme. C'est le principe d'inertie. Il existe une classe de référentiels équivalents, appelés référentiels inertiels, dans lesquels le principe d'inertie est vérifié.
- ii) Le principe fondamental de la dynamique exprime que le *changement de mouvement* $m\Delta v$ est égal à la *force motrice* $F\Delta t$. C'est une loi différentielle maintenant écrite sous la forme

$$\frac{d\mathbf{p}}{dt} = \mathbf{F}, \quad (1.1)$$

où $\mathbf{p} = m\mathbf{v}$ est la quantité de mouvement.

La forme différentielle permet une intégration numérique qui met clairement en évidence le déterminisme : la cinématique définit la vitesse

$$v_x(t) = \frac{dx(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{x(t + \Delta t) - x(t)}{\Delta t},$$

⁽¹⁾I. Newton, *De Philosophiae Naturalis Principia Mathematica* (Christian Bourgois Éditeur, Paris 1985).

soit $x(t + \Delta t) = x(t) + \Delta t v_x(t) + O(\Delta t^2)$, et la loi du mouvement

$$\frac{dp_x(t)}{dt} = F_x(t)$$

conduit à $p_x(t + \Delta t) = p_x(t) + \Delta t F_x(t) + O(\Delta t^2)$. Ainsi, connaissant les conditions initiales x_0 et v_{x0} , il suffit d'itérer pour connaître $x(t)$ et $v_x(t)$ à tout instant avec une précision arbitraire en choisissant Δt assez petit.

Traisons l'exemple du rappel élastique à deux dimensions, $\mathbf{F} = -k\mathbf{r}$. On prend $k = 1$, $m = 1$, $\Delta t = 0.2$ et les conditions initiales $x_0 = 1$, $y_0 = 0$, $v_{x0} = 0$ et $v_{y0} = 0.5$. Par itérations successives, on obtient le tableau suivant :

Tableau 1.1 Intégration numérique des équations de Newton.

itération	x	y	v_x	v_y	Δx	Δy	Δv_x	Δv_y
1	1.000	0.000	0.000	0.500	0.000	0.100	-0.200	0.000
2	1.000	0.100	-0.200	0.500	-0.040	0.100	-0.200	-0.020
3	0.960	0.200	-0.400	0.480	-0.080	0.096	-0.192	-0.040
4	0.880	0.296	-0.592	0.440	-0.118	0.088	-0.176	-0.059
5	0.762	0.384	-0.768	0.381	-0.154	0.076	-0.152	-0.077
...								
calcul exact								
	1.000	0.000	0.000	0.500				
	0.960	0.140	-0.280	0.480				
	0.880	0.237	-0.474	0.440				
	0.762	0.324	-0.648	0.381				
	...							

Le calcul différentiel remplace évidemment avantageusement l'intégration numérique et donne directement

$$\begin{aligned} x(t) &= \cos t \\ y(t) &= 0.5 \sin t \\ x^2 + (y/0.5)^2 &= 1, \end{aligned}$$

l'équation de la trajectoire est donc une ellipse.

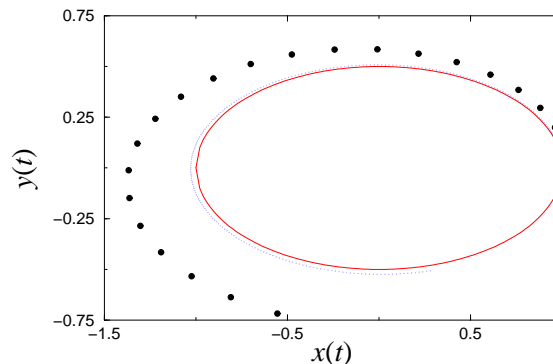


Figure 1.1 Comparaison de la solution analytique avec l'intégration numérique pour un pas grossier $\Delta t = 0.2$ (cercles pleins) et pour un pas fin $\Delta t = 0.02$ (pointillés, confondus à cette échelle avec la solution analytique).

1.2.2. Gravitation

Disposant du principe fondamental de la dynamique, Newton a également trouvé la loi de force responsable du mouvement elliptique ⁽²⁾ observé des planètes. Dans le cas plus simple d'une trajectoire circulaire, on montre facilement que la force est radiale. On a en effet

$$m\mathbf{v}(t + dt) - m\mathbf{v}(t) = dt\mathbf{F}(t),$$

or $\mathbf{v}(t) = v(t)\hat{\mathbf{u}}_\tau(t)$ et $\mathbf{v}(t + dt) = (v(t) + dv)\hat{\mathbf{u}}_\tau(t + dt)$ ($\hat{\mathbf{u}}_\tau(t)$ est le vecteur unitaire tangent à l'instant t), la différence $m\mathbf{v}(t + dt) - m\mathbf{v}(t)$ est donc radiale, dirigée vers le centre de la trajectoire, soit $\mathbf{F} = -F\hat{\mathbf{u}}_r$. Par ailleurs les lois de Kepler (la loi des aires notamment) exigent que le module de la force varie en $1/r^2$:

$$\mathbf{F}(\mathbf{r}) = -G\frac{mm'}{r^2}\hat{\mathbf{u}}_r. \quad (1.2)$$

1.2.3. Invariance galiléenne et principe d'inertie

On appelle invariance galiléenne le fait que l'on ne puisse pas déterminer si un référentiel inertiel est en translation uniforme par rapport à un autre au moyen d'une expérience de mécanique. C'est un fait d'expérience (c'est le cas, rapporté par Galilée dans les *Dialogues*⁽³⁾, d'un objet tombant depuis la vigie au pied du mât d'un bateau en mouvement calme), c'est également une conséquence des lois de la dynamique et de la conception newtonienne de l'espace et du temps. L'appellation "principe de relativité galiléenne" est due à Poincaré, on en donne une formulation précise : *Les lois de la mécanique prennent la même forme mathématique dans tous les référentiels inertiels, i.e. dans lesquels le principe d'inertie est valable.*

On peut établir, dans le formalisme newtonien, l'invariance galiléenne en trois étapes essentielles.

- i) Transformation de Galilée : selon Newton, "le temps est absolu, homogène et s'écoule uniformément", ce qui veut dire qu'entre deux référentiels \mathcal{R} et \mathcal{R}' supposés synchronisés et coïncidants à l'origine des temps $t = t' = 0$, on a la relation très importante $t = t'$. Si de plus \mathcal{R}' est en mouvement à vitesse \mathbf{u} constante parallèlement à l'axe commun $Ox = Ox'$ par rapport à \mathcal{R} , on a $x' = x - ut$, $y' = y$ et $z' = z$. La transformation

$$\begin{aligned} \mathbf{r}' &= \mathbf{r} - \mathbf{u}t \\ t' &= t \end{aligned} \quad (1.3)$$

constitue la transformation de Galilée.

- ii) Principe d'inertie : le principe d'inertie étant supposé vérifié dans le référentiel \mathcal{R} , la transformation de Galilée étend ce principe à tout référentiel \mathcal{R}' en translation uniforme. En effet, si $\mathbf{r}(t) = \mathbf{v}t$ avec \mathbf{v} constante, alors dans \mathcal{R}' $\mathbf{r}'(t) = \mathbf{r}(t) - \mathbf{u}t$ entraîne que $\mathbf{v}'(t) = d\mathbf{r}'/dt' = \mathbf{v} - \mathbf{u}$ est aussi

⁽²⁾Si l'on néglige les perturbations dues aux autres corps et les corrections relativistes, le mouvement relatif de deux corps en interaction gravitationnelle est elliptique. Nous verrons que l'explication des écarts très faibles à la trajectoire elliptique n'a pu être fournie que par la relativité générale.

⁽³⁾G. Galilée, *Dialogues et Lettres choisies*, Discours sur deux Sciences Nouvelles, trad. P.H. Michel (Hermann, Paris 1966).

constante. Il existe donc toute une classe de référentiels équivalents. Il reste à identifier l'un d'eux. Newton postule l'existence d'un référentiel absolu dans lequel le principe d'inertie est valable, ainsi que la relation fondamentale de la dynamique qui, comme on l'a déjà vu, mesure l'écart au principe d'inertie. La vitesse de la lumière y prend également une certaine valeur. En bonne approximation, un référentiel centré sur le Soleil et dont les trois axes pointent vers des étoiles distantes identifiées (leur mouvement relativement au Soleil est donc imperceptible) est inertiel. C'est le référentiel héliocentrique.

- iii) Invariance galiléenne : c'est une généralisation consistant à étendre, par transformation de Galilée, toutes les lois de la mécanique aux référentiels inertiels. Si dans \mathcal{R} on a $\frac{d\mathbf{p}}{dt} = \mathbf{F}$, $\mathbf{p} = m\mathbf{v}$, $\mathbf{v} = d\mathbf{r}/dt$, alors dans \mathcal{R}' on aura $\mathbf{p}' = m\mathbf{v}'$ avec $\mathbf{v}' = \mathbf{v} - \mathbf{u}$, soit

$$\frac{d\mathbf{p}'}{dt} = m \frac{d\mathbf{v}}{dt} - m \underbrace{\frac{d\mathbf{u}}{dt}}_0 = \frac{d\mathbf{p}}{dt},$$

ou encore

$$\frac{d\mathbf{p}'}{dt} = \mathbf{F}', \quad \text{avec} \quad \mathbf{F}' = \mathbf{F}.$$

Cette relation exprime l'invariance de la forme de la relation fondamentale de la dynamique dans les deux référentiels et du même coup la loi de transformation des forces (les forces actives sont inchangées par changement de référentiel galiléen). On constate de plus que si \mathcal{R}' n'est pas un référentiel galiléen, de nouvelles "forces" (appelées inertielles) apparaissent :

$$\frac{d\mathbf{p}'}{dt} = \mathbf{F}' - m \frac{d\mathbf{u}}{dt}.$$

2. Structure métrique de l'espace-temps newtonien

Au risque de présenter quelques redondances avec le paragraphe précédent, nous allons maintenant discuter plus en détail les notions d'espace et de temps chez Newton, avant de les mettre à mal avec la relativité.

2.1. L'espace physique

Il est continu, homogène, isotrope et tridimensionnel. Mathématiquement, l'espace physique est isomorphe à \mathbb{R}^3 muni de sa structure affine (droite, parallélisme) et de sa structure métrique (distance, angle, théorème de Pythagore). Cette conception repose sur les notions expérimentales d'alignement, (arête d'un solide), de parallélisme. Bien plus tard le théorème de Noether donnera de plus un contenu physique à l'homogénéité et l'isotropie de l'espace, propriétés intimement liées aux lois de conservation de la quantité de mouvement et du moment cinétique.

2.2. Le temps

Il est uniforme et orienté. Mathématiquement le temps est représenté par \mathbb{R} , droite ordonnée, ce qui définit la chronologie (passé, futur) et permet de donner une structure causale. L'uniformité sous-entend l'invariance par translation dans le temps et est associée à la conservation de l'énergie.

2.3. Espace-temps classique

L'espace-temps classique est perçu comme un contenant, le théâtre dans lequel se produisent des événements. Il préexiste aux objets physiques.

$\mathbb{R}^3 \times \mathbb{R}$ est constitué de feuillet, les strates spatiales euclidiennes successives. Toute droite orientée coupant les strates $3d$ est un axe des temps possible, mais $\mathbb{R}^3 \times \mathbb{R}$ ne possède pas de structure métrique car il n'existe classiquement aucune vitesse fondamentale v_0 qui permette de relier de façon univoque la coordonnée temporelle aux coordonnées spatiales. Les axes des temps permis sont en fait identifiés aux trajectoires de particules libres dans le référentiel absolu.

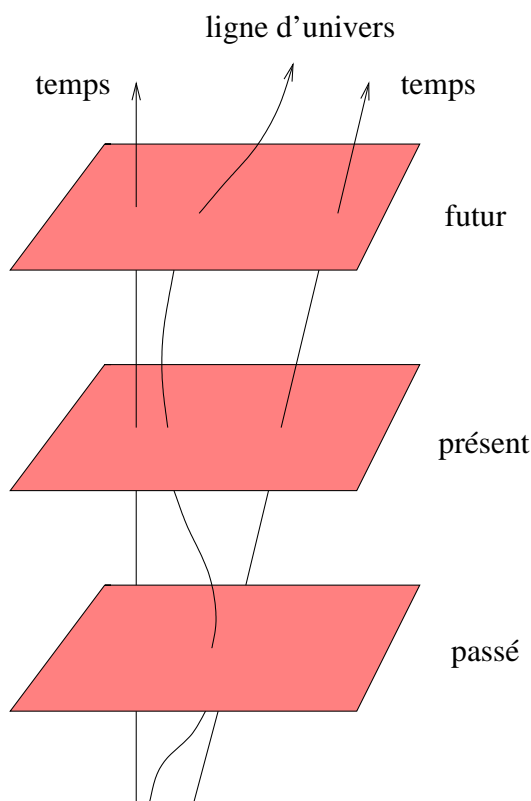


Figure 1.2 Structure de l'espace-temps newtonien en succession de strates spatiales euclidiennes. La métrique est tridimensionnelle dans chaque strate spatiale et il y a simultanéité de tous les événements au sein d'une strate spatiale.

Il y a en revanche deux invariants distincts : quel que soit le référentiel dans lequel on se place pour les mesurer, la distance $|\Delta \mathbf{r}|$ et la durée Δt entre deux événements sont inchangées.

La structure est causale dans le sens où le passé peut influencer le présent ou l'avenir, et aucune trajectoire ne peut remonter le temps. En revanche n'importe quel événement du passé pourrait avoir influencé le présent car aucune vitesse limite n'empêche la propagation de l'information. Cela provient de ce que la physique newtonienne postule l'existence d'interactions instantanées à distance, d'où la simultanéité à l'intérieur des strates spatiales. Une conséquence qui sera mise à mal avec la relativité est que l'univers global est accessible à l'observation.

2.4. Référentiels inertiels et gravitation

2.4.1. Gravitation et forces d'inertie

Nous n'avons pas encore précisé complètement la notion de référentiel. Au sens d'Einstein, le référentiel est défini au moyen d'un corps pratiquement rigide (i.e. un corps rigide en pratique). Cela consiste à se placer dans une zone de l'espace close, sans fenêtres pour observer un éventuel extérieur (ce serait faire des mesures) et à ne considérer que les champs de forces dont les sources (charges, masses) sont situées à l'intérieur. L'objectif est de pouvoir faire des mesures en maîtrisant totalement l'appareillage susceptible d'intervenir dans les expériences. De ce point de vue, lorsque les sources des champs sont à l'extérieur, inaccessibles à l'expérimentateur, celui-ci en est réduit à des conjectures. La gravitation par exemple n'est pas contrôlée par l'expérimentateur qui ne peut pas décider, par les expériences qu'il mène, s'il est au voisinage d'une source gravitationnelle, cause de son poids, ou si ce sont des effets d'inertie dus à l'état d'accélération éventuelle de son référentiel qui le maintiennent en contact avec le sol. Les deux points de vue sont localement équivalents et ce résultat est connu sous le nom de principe d'équivalence faible.

Détaillons cette correspondance, ce qui nous amènera à repréciser la nature des référentiels inertiels. En l'absence totale de pesanteur, un certain référentiel \mathcal{R} inertiel est au repos ou en translation uniforme à vitesse constante de sorte que la force, par exemple électrique, agissant sur une charge q de masse m produit une accélération,

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}.$$

Dans \mathcal{R}' , en mouvement à vitesse quelconque $\mathbf{u}(t)$ par rapport à \mathcal{R} , cette même relation s'écrit

$$\mathbf{F} - m \frac{d\mathbf{u}}{dt} = \frac{d\mathbf{p}'}{dt}, \text{ avec } \mathbf{p}' = \mathbf{p} - m\mathbf{u},$$

de sorte que les effets non inertiels (dus aux forces dites d'inertie, ce qui peut être source de confusion) sont visibles : même en l'absence de force \mathbf{F} , la masse voit sa quantité de mouvement varier sous l'effet du terme $-m \frac{d\mathbf{u}}{dt}$. Considérons maintenant la même charge soumise au même champ de force \mathbf{F} dans le voisinage d'un champ gravitationnel et cherchons un référentiel inertiel. Le référentiel immobile (par rapport à la source de gravitation) \mathcal{R}_g cette fois ne convient plus, car la relation fondamentale de la dynamique s'écrit

$$\mathbf{F} + m\mathbf{g} = \frac{d\mathbf{p}}{dt},$$

de sorte que même si l'on supprime \mathbf{F} , $\frac{d\mathbf{p}}{dt}$ est non nul, ce qui ne satisfait pas au principe d'inertie. On rappelle que la source de \mathbf{g} étant externe, la force $m\mathbf{g}$ est omniprésente mais l'expérimentateur n'a aucun moyen certain de le savoir et il pourrait aussi bien attribuer la variation observée de \mathbf{p} au terme $-m \frac{d\mathbf{u}}{dt}$ du référentiel accéléré précédent. Ce référentiel \mathcal{R}_g est donc plutôt à rapprocher du référentiel accéléré sans gravitation \mathcal{R}' . Si l'on considère en revanche le référentiel en chute

libre \mathcal{R}'_g dans le champ de pesanteur, les deux termes $m\mathbf{g}$ et $-m\frac{d\mathbf{u}}{dt}$ apparaissent simultanément,

$$\mathbf{F} + m\mathbf{g} - m\frac{d\mathbf{u}}{dt} = \frac{d\mathbf{p}'}{dt},$$

de sorte que si l'on choisit $\mathbf{u}(t) = \mathbf{g}t$ ils se compensent et en l'absence de force \mathbf{F} , la masse m suit une trajectoire rectiligne à vitesse uniforme *dans le référentiel en chute libre*. En présence d'un champ de gravitation, les référentiels inertiels sont les référentiels en chute libre. Inversement, on peut confondre un référentiel non inertiel en présence d'un champ de gravitation \mathbf{g} avec un référentiel accéléré avec une accélération $\frac{d\mathbf{u}}{dt} = -\mathbf{g}$.

2.4.2. Caractère local d'un référentiel inertiel en présence d'un champ de gravitation

Les référentiels en chute libre sont inertiels en bonne approximation, mais sont en fait seulement localement inertiels. En effet deux objets qui ne seraient soumis à aucune force devraient rester immobiles l'un par rapport à l'autre et effectivement, en vertu du principe d'équivalence, ils tombent tous deux en chute libre avec une accélération indépendante de leur masse, mais qui varie toutefois légèrement d'un point à un point voisin en raison du caractère non uniforme du champ gravitationnel. C'est ce qu'on appelle les effets de marée ; ils produisent une légère déviation au principe d'inertie et l'équivalence entre gravitation et accélération n'est que locale.

Par exemple deux corps ponctuels distants de $l = 25$ m lâchés d'une altitude $h = 250$ m se rapprocheront, au cours de leur chute pour atteindre le sol, d'une quantité $(R_T + h)\theta - R_T\theta$, soit avec $\theta \simeq l/(R_T + h)$, d'une quantité égale à $h\theta \simeq 0.98$ mm. Si les deux corps étaient placés sur une même verticale, à des altitudes $h \pm l/2$ avec les mêmes valeurs numériques, on aurait cette fois à résoudre l'équation

$$\ddot{z} = -GM_T/(R_T + z)^2 \simeq -\frac{GM_T}{R_T^2}(1 - 2z/R_T)$$

qui admet les solutions

$$z(t) = \frac{1}{2}R_T + (h \pm l/2 - R_T/2) \cosh \omega t$$

avec les valeurs

$$\omega = \sqrt{2g/R_T}$$

$$\Delta t = \sqrt{2g/h}.$$

On trouve que les deux corps se rapprochent alors de $z_+(\Delta t) - z_-(\Delta t) - l \simeq 1.95$ mm. On voit sur ces exemples que tant que l'on s'intéresse à des expériences locales, les effets de marée jouent peu, mais ils doivent être pris en considération pour une description générale valable à toute échelle.

3. Invariance de la vitesse de la lumière dans les référentiels inertiels

3.1. Expérience de Michelson et Morley

Les premières mesures de la vitesse de la lumière sont dues à Roemer à partir de l'observation des durées d'occultation des satellites de Jupiter (elles donnaient déjà des valeurs supérieures à 200 000 km/s). Des mesures plus précises ont été obtenues au cours du XIX^{ème} siècle, notamment par Fizeau. Selon la conception de l'époque, la lumière, comme l'ensemble des ondes électromagnétiques et à l'image des ondes acoustiques par exemple, se propage par l'intermédiaire des vibrations d'un milieu appelé éther. A cette époque cependant commencent à émerger des difficultés car on doit attribuer à l'éther, le milieu de propagation des ondes électromagnétiques, des propriétés contradictoires. Par exemple la lumière étant une onde transverse, l'analogie avec les ondes élastiques requiert alors que l'éther possède un module de cisaillement non nul et donc s'apparente à un solide alors que comme son nom l'indique il est très discret et semble impossible à détecter. C'est dans ce contexte que Michelson, puis Michelson et Morley vont mener une série d'expériences entre 1881 et 1887 cherchant à mettre en évidence le mouvement de la Terre par rapport à l'éther. Ils utilisent pour cela un interféromètre de Michelson dans lequel les interférences sont obtenues en superposant les faisceaux lumineux issus des réflexions sur deux bras perpendiculaires qui sont supposés être entraînés dans le mouvement orbital de la Terre de manières différentes par rapport à l'éther.

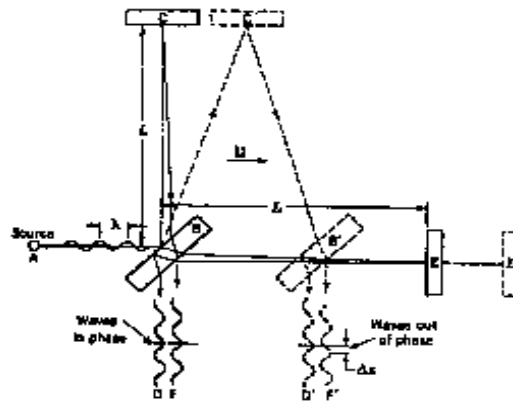


Figure 1.3 Schéma de l'expérience de Michelson et Morley : une source monochromatique émet en A un rayonnement de longueur d'onde λ qui arrive en B sur une séparatrice. Une partie du faisceau est réfléchi vers un miroir C distant de L , dans la direction perpendiculaire au mouvement de la Terre par rapport à l'éther (à vitesse u supposée uniforme sur la durée de l'expérience), l'autre partie du faisceau est transmise vers un miroir E dans la direction du mouvement de la Terre. Après réflexion sur les miroirs C et E , les deux faisceaux se recombinent et produisent une figure d'interférences (en D et F) si les deux trajets ont une différence de phase. Les traits pleins correspondent au référentiel terrestre et les pointillés à l'éther.

La lumière se propage à la vitesse c par rapport à l'éther. On suppose que la Terre, dans son mouvement orbital autour du Soleil (à $u \simeq 30$ km/s) se déplace par rapport à l'éther, lui-même supposé immobile par exemple par rapport au référentiel

héliocentrique qui a déjà un rôle très particulier dans la physique newtonienne et, avec son caractère absolu, peut très bien constituer le milieu dans lequel l'éther est au repos en bonne approximation⁽⁴⁾. Raisonnons dans le référentiel où l'éther est au repos. Pour les deux faisceaux, les trajets de la source à la séparatrice sont identiques. Il suffit d'étudier les bras de réflexion aller-retour vers les miroirs. Pour le bras perpendiculaire $BC'B''$ (C' repère la position du miroir au moment où le faisceau lui parvient et B'' celle de la séparatrice après réflexion sur le miroir lorsque le faisceau lui parvient à son tour, un peu plus tard) la longueur totale du trajet effectué à la vitesse c , $c\tau_{BC'B''}$ vaut deux fois la longueur géométrique du trajet aller, $2\sqrt{L^2 + |\mathbf{u}|^2\tau_{BC'B''}^2}/4$. Par élimination, on arrive à

$$\tau_{BC'B''} = \frac{2L/c}{\sqrt{1 - |\mathbf{u}|^2/c^2}}.$$

Pour le second bras la situation est différente car le miroir E fuit d'abord le faisceau à l'aller alors qu'au retour la séparatrice vient à la rencontre du faisceau. Pendant la durée $\tau_{BE'}$ de l'aller, le miroir E s'est déplacé en E' d'une distance de $|\mathbf{u}|\tau_{BE'}$ et la lumière pendant ce temps parcourt la distance $c\tau_{BE'} = L + |\mathbf{u}|\tau_{BE'}$, soit

$$\tau_{BE'} = \frac{L/c}{1 - |\mathbf{u}|/c}.$$

Pour le trajet retour un raisonnement analogue conduit à $c\tau_{E'B''} = L - |\mathbf{u}|\tau_{E'B''}$, soit

$$\tau_{E'B''} = \frac{L/c}{1 + |\mathbf{u}|/c}.$$

Les deux contributions additionnées donnent au total pour le bras parallèle

$$\tau_{BE'B''} = \frac{2L/c}{1 - |\mathbf{u}|^2/c^2},$$

c'est-à-dire que les deux faisceaux doivent arriver dans le détecteur avec un décalage de

$$\Delta\tau = \frac{2L}{c} \left(\frac{1}{1 - |\mathbf{u}|^2/c^2} - \frac{1}{\sqrt{1 - |\mathbf{u}|^2/c^2}} \right)$$

produisant une figure d'interférences mesurable en principe. Si l'ensemble du dispositif est ensuite tourné de 90° les rôles sont inversés et la différence de marche $\delta = c\Delta\tau$ joue en sens contraire, de sorte qu'entre les deux situations la figure d'interférences doit se déplacer de $n = 2\delta/\lambda$. Compte-tenu de ce que $|\mathbf{u}| \ll c$, on parvient à

$$n \simeq \frac{2L}{\lambda} \frac{|\mathbf{u}|^2}{c^2}$$

⁽⁴⁾ Il n'est pas capital ici que l'éther soit au repos dans le référentiel héliocentrique, mais comme ce dernier est inertiel en très bonne approximation, il joue déjà un rôle singulier comme référentiel absolu.

qui est de l'ordre de 0.05 pour $L \simeq 1$ m et $\lambda \simeq 4000$ Å.

Les résultats des expériences successives n'ont conduit en fait à aucune variation mesurable de la position de la figure d'interférences.

Tableau 1.2 Résultats des expériences de Michelson et Morley.

		L	$2L \mathbf{u} ^2/\lambda c^2$	$n_{obs.}$
1881	Michelson et Morley	120 cm	0.04	0.01
1887	Michelson et Morley	1100 cm	0.40	0.005
1930	Goos	2100 cm	0.75	0.001

Il n'y a plus dès lors que deux conclusions possibles, soit la Terre est immobile par rapport à l'éther, ou encore l'éther est emporté par la Terre dans son mouvement autour du Soleil (et alors pourquoi une telle situation spécifique de la Terre), soit tout simplement l'éther n'existe pas.

Cette dernière position s'est finalement imposée lentement. Les équations de Maxwell d'ailleurs n'exigent pas de milieu de propagation pour les ondes électromagnétiques, puisque ce sont les variations spatio-temporelles du champ électrique qui génèrent celles du champ magnétique et réciproquement. On peut toutefois noter que très tôt des voix se sont élevées contre la nécessité d'un éther, comme par exemple le physicien danois Ludwig Valentin Lorenz en 1867⁽⁵⁾ :

L'hypothèse d'un éther est déraisonnable parce que c'est un milieu sans substance qui n'a été imaginé que parce que la lumière est conçue de la même manière que le son. Ce doit donc être un milieu d'excessivement grande élasticité et faible densité pour expliquer la très grande vitesse de la lumière. Ce n'est certainement pas scientifique d'inventer une nouvelle substance quand son existence n'est pas révélée de meilleure manière.

3.2. Effet Doppler classique

Une autre difficulté provient de ce que les phénomènes électromagnétiques ne respectent pas l'invariance galiléenne. On peut mettre en évidence ce phénomène en comparant deux situations a priori équivalentes, mais où l'effet Doppler donne finalement des résultats différents selon que c'est la source ou l'observateur qui est en mouvement.

Considérons une source électromagnétique émettant des pulses à intervalles de temps réguliers T_S (ce qui définit la période) et se déplaçant à vitesse \mathbf{u} par rapport à un observateur immobile. On supposera que l'observateur est à grande distance $r \gg |\mathbf{u}|T_S$ et dans une direction inclinée d'un angle θ par rapport à la direction de la vitesse. On se limite ici à un calcul au premier ordre en $|\mathbf{u}|/c$. Bien que la dissymétrie entre les deux situations envisagées (source ou observateur en mouvement) n'apparaisse pas à cet ordre dans l'expression de la fréquence mesurée par l'observateur, il est manifeste qu'une différence apparaîtrait aux ordres supérieurs (les deux expressions obtenues selon que c'est la source ou l'observateur qui est en mouvement sont différentes).

On peut considérer deux instants successifs $t = 0$ et T_S où la source, aux positions S_1 et S_2 (avec $S_1S_2 = |\mathbf{u}|T_S$), émet des pulses :

$$\begin{aligned} t = 0 & : \quad \text{émission d'un premier pulse,} & \text{réception à} & \quad t = r_1/c \\ t = T_S & : \quad \text{émission d'un second pulse,} & \text{réception à} & \quad t = T_S + r_2/c \end{aligned}$$

⁽⁵⁾Cité dans J.D. Jackson and L.B. Okun, *Historical roots of gauge invariance*, <http://xxx.lpthe.jussieu.fr/hep-ph/0012061>.

La période mesurée par l'observateur vaut donc $T_O = T_S + \frac{r_2 - r_1}{c}$ et dans les conditions indiquées précédemment, il vient $r_2 - r_1 \simeq -S_1 S_2 \cos \theta = -|\mathbf{u}| T_S \cos \theta$, soit

$$T_O \simeq T_S \left(1 - \frac{|\mathbf{u}|}{c} \cos \theta \right),$$

ou

$$\nu_O \simeq \nu_S \left(1 + \frac{|\mathbf{u}|}{c} \cos \theta \right).$$

C'est l'expression classique du décalage Doppler entre fréquence à l'émission et fréquence à la réception d'un signal électromagnétique.

La situation symétrique a priori consiste à supposer cette fois la source immobile, mais l'observateur en mouvement, dans les mêmes conditions d'approximation. On a cette fois réception du signal par un observateur occupant deux positions successives, O_1 et O_2 , donc à nouveau

$$\begin{aligned} t = 0 & : \text{émission d'un premier pulse,} & \text{réception à} & t = r_1/c \\ t = T_S & : \text{émission d'un second pulse,} & \text{réception à} & t = T_S + r_2/c, \end{aligned}$$

mais avec $O_1 O_2 = |\mathbf{u}| T_O$ cette fois, soit $r_2 - r_1 \simeq -O_1 O_2 \cos \theta = -|\mathbf{u}| T_O \cos \theta$, soit

$$T_O \simeq T_S \left(1 + \frac{|\mathbf{u}|}{c} \cos \theta \right)^{-1},$$

ou

$$\nu_O \simeq \nu_S \left(1 + \frac{|\mathbf{u}|}{c} \cos \theta \right).$$

On constate que les expressions sont identiques au premier ordre en $|\mathbf{u}|/c$ mais diffèrent au-delà, ce qui signifie que l'invariance galiléenne n'est pas vérifiée pour ce type de phénomène.

3.3. Statut des équations de Maxwell, programme de la théorie de la relativité

Le résultat négatif des expériences de Michelson et Morley, conduisant à l'inutilité de l'éther, semble alors indiquer que la vitesse de propagation des ondes électromagnétiques (et de la lumière en particulier) est la même quel que soit l'état de mouvement relatif de la source et de l'observateur (en se limitant ici au cas de mouvements uniformes). C'est donc une constante invariante par changement de référentiel inertiel.

On vient de voir par ailleurs que les phénomènes liés à l'électromagnétisme (par le cas de l'effet Doppler) ne respectent pas l'invariance galiléenne par changement de référentiel inertiel. On peut rendre manifeste cette non-invariance galiléenne des lois de l'électromagnétisme en s'intéressant aux équations de Maxwell et en montrant qu'elles ne sont pas invariantes dans la transformation de Galilée, ce que l'on admettra ici. On se trouve donc dans une situation inacceptable, puisque la valeur de la vitesse de la lumière est déterminée par les équations de Maxwell. On peut en effet combiner

(dans le vide) $\text{rot } \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$ et $\text{rot } \mathbf{B} = \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}$ pour obtenir l'équation des ondes $\text{rot rot } \mathbf{E} = -\Delta \mathbf{E} = -\varepsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}$, ou

$$\square \mathbf{E} \equiv \vec{\nabla}^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \vec{0}$$

avec $c = (\varepsilon_0 \mu_0)^{-1/2}$.

Cela soulève donc quelques difficultés que l'on ne peut lever qu'au prix de modifications théoriques (et conceptuelles) radicales.

- i) Le fait que c soit manifestement conséquence des équations de Maxwell et qu'elle soit constante par changement de référentiel inertiel milite en faveur de la validité des équations de Maxwell dans tous les référentiels inertiels, c'est-à-dire en faveur de leur invariance galiléenne (mais pas par transformation de Galilée⁽⁶⁾).*
- ii) Si c'est le cas, il s'agit donc de trouver une nouvelle transformation des coordonnées d'espace et de temps, qui préserve la forme des équations de Maxwell et qui remplace la transformation de Galilée.*
- iii) Cela doit vraisemblablement entraîner une révision des lois de la dynamique, puisque d'une part l'invariance galiléenne de la mécanique est un fait d'expérience, et que d'autre part les lois de la dynamique newtoniennes sont invariantes par transformation de Galilée, elles ne le sont probablement pas par la nouvelle transformation qui reste à trouver.*
- iv) Enfin il est très probable que la cinématique elle-même soit à reconstruire, car la transformation de Galilée est simplement l'expression des transformations de coordonnées par changement de référentiel.*
- v) Sortant du cadre de ce cours (cf. chapitre 11), il faudrait ensuite reformuler la théorie quantique, puisque l'équation de Schrödinger n'est pas invariante relativiste.*

Ces différents points constituent le programme de la relativité. Dans l'édifice classique, on ne conservera que l'électromagnétisme.

3.4. Conséquences immédiates de l'invariance de c

Nous allons ici montrer que la conception newtonienne d'un temps absolu doit en effet être remise en question dès lors que l'on admet l'invariance de la valeur de la vitesse de propagation des ondes électromagnétiques c (par changement de référentiel inertiel).

Établissons tout d'abord un résultat indispensable, l'invariance des longueurs transverses. On considère une expérience de mécanique très simple dans laquelle deux règles identiques parallèles sont mises en mouvement relativement l'une à l'autre dans la direction perpendiculaire à leur longueur (notée L_{\perp}). Supposons que la règle 1 possède deux pointes susceptibles de s'encaster dans deux marques correspondantes sur la règle 2. La valeur de L_{\perp} est déterminée par la distance entre ces pointes (ou ces marques). Considérons la règle 1 en mouvement vers la règle 2 immobile. Lors du choc, les deux pointes (L'_{\perp}) viennent produire de nouvelles marques. Si les nouvelles marques sont à l'extérieur des anciennes (L_{\perp}), on en conclura que la longueur transverse en mouvement est plus grande que la même longueur au repos, $L'_{\perp} > L_{\perp}$. Mais dans ce cas on peut envisager l'expérience

⁽⁶⁾ On conserve le terme d'invariance galiléenne pour l'invariance par changement de référentiel inertiel, même si c'est une autre transformation mathématique qui entre en jeu.

du point de vue de la règle 1 qui, immobile, voit arriver sur elle la règle 2 en mouvement et donc telle que les marques originales (L_{\perp}) soient plus éloignées que les deux pointes (L'_{\perp}), c'est-à-dire que l'invariance galiléenne impose deux conclusions contradictoires qu'on ne peut concilier qu'en contredisant les prémices. On en déduit donc que les longueurs transverses sont invariantes par changement de référentiel inertiel (ou non inertiel d'ailleurs, tant que c'est transverse par rapport au mouvement) $L_{\perp} = L'_{\perp}$.

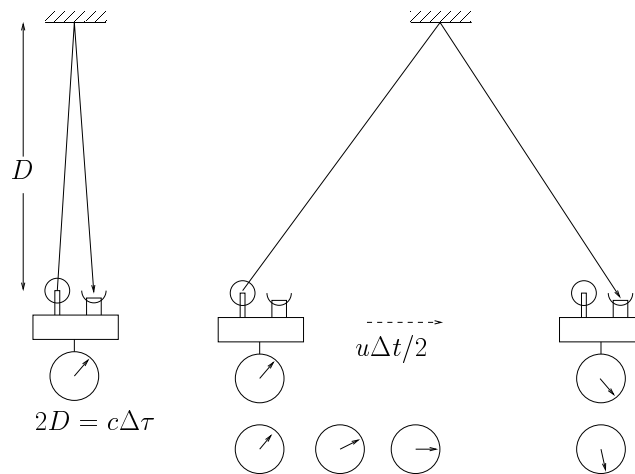


Figure 1.4 Schéma de l'expérience considérée dans le texte. On distingue l'indication de l'horloge embarquée (temps propre) de celles des horloges au sol, synchronisées.

On considère maintenant une expérience d'électromagnétisme : un pulse lumineux (bleu) est émis à $t' = 0$, parcourt une distance L_{\perp} , se réfléchit sur un miroir pour être absorbé par une cellule au temps $t' = \Delta\tau = 2L_{\perp}/c$ (les deux instants sont mesurés par un même observateur). Supposons que l'ensemble expérience-observateur (constituant le référentiel \mathcal{R}') soit solidaire d'un dispositif qui se déplace perpendiculairement au pulse à la vitesse \mathbf{u} par rapport à un second référentiel \mathcal{R} . Dans \mathcal{R} , de nombreux observateurs synchronisés sont placés à différents points de l'espace. Pour les synchroniser, on peut supposer qu'ils sont distants de 1 m les uns des autres. Le premier envoie un signal lumineux (rouge⁽⁷⁾) à $t = 0$. Le second a préalablement réglé sa montre à $t = 1/c$ seconde, le second à $t = 2/c$ seconde, etc. Et chacun de ces observateurs fait démarrer sa montre au passage du signal rouge émis par le premier. Cette technique de synchronisation ne pose pas de problème car tous ces observateurs sont immobiles les uns par rapport aux autres. Dans \mathcal{R} , l'émission et l'absorption du pulse bleu ne se produisent pas au même endroit. L'émission a lieu auprès de l'observateur O_1 , à $t = 0$. A l'instant où la lumière se réfléchit sur le miroir,

⁽⁷⁾ On utilise ici volontairement deux couleurs pour ne pas confondre le second signal, qui ne sert qu'à synchroniser les horloges dans \mathcal{R} et le premier qui constitue l'expérience proprement dite.

le dispositif s'est déplacé transversalement de $|\mathbf{u}|\Delta t/2$ si Δt est la durée totale aller-retour dans \mathcal{R} , la lumière a donc parcouru une distance $\sqrt{L_{\perp}^2 + |\mathbf{u}|^2(\Delta t/2)^2}$ qui vaut également par définition (avec c invariante) $c\Delta t/2$. En remplaçant $L_{\perp} = c\Delta\tau/2$ on a donc entre les durées mesurées dans les deux référentiels l'expression

$$\Delta t = \frac{\Delta\tau}{\sqrt{1 - |\mathbf{u}|^2/c^2}}.$$

On a utilisé ici la notation τ pour le temps mesuré par l'observateur solidaire de l'expérience (le référentiel propre).

Les deux horloges, en mouvement uniforme l'une par rapport à l'autre (ou aux autres !), ne mesurent donc pas la même durée entre deux événements donnés. C'est donc une remise en question du temps absolu de Newton et cela invalide la transformation de Galilée.

4. Le principe de relativité d'Einstein

4.1. Les postulats de la relativité restreinte

Compte-tenu des conclusions auxquelles nous sommes parvenus au § 3, les postulats d'Einstein, qui fondent la théorie de la relativité, peuvent s'exprimer de la manière suivante :

- i) *La vitesse de la lumière a la même valeur dans tous les référentiels inertiels (elle est en particulier indépendante de l'état de mouvement uniforme de la source).*
- ii) *Les lois de la physique (et pas simplement de la mécanique) ont la même forme dans tous les référentiels inertiels. On dira qu'il y a invariance galiléenne des lois de la physique, sans pour autant entendre invariance par transformation de Galilée, mais invariance par changement de référentiel galiléen ⁽⁸⁾.*

On peut noter comme conséquence immédiate que les valeurs numériques des constantes sans dimension (par exemple la constante de structure fine de l'électromagnétisme, $\alpha = e^2/\hbar c$) doivent être les mêmes dans tous les référentiels inertiels.

Revenons sur le programme défini précédemment. Il découle de ces postulats qu'il faut établir la nouvelle transformation des coordonnées spatio-temporelles assurant l'invariance des équations de Maxwell (on a déjà vu quelle forme prenait le passage de t à τ) et qu'il faut revoir les lois de la dynamique, mais que la dynamique newtonienne doit constituer la limite non relativiste de cette nouvelle dynamique, comme la transformation de Galilée doit être retrouvée dans la limite $c \rightarrow \infty$. Cette nouvelle transformation est en fait connue en 1905, c'est la transformation de Lorentz, établie indépendamment par Lorentz et Fitzgerald pour rendre compte des résultats négatifs de l'expérience de Michelson et Morley. Notons enfin que la valeur de c ayant maintenant un caractère universel, cela va permettre d'établir

⁽⁸⁾ Pour éviter toute confusion, on parle parfois d'invariance relativiste des lois de la physique, avec toute l'ambiguïté de comporter dans une même expression deux termes antinomiques comme invariance relativiste !

une métrique quadridimensionnelle fondée sur un nouvel invariant appelé intervalle. Cela signifie que la relativité restreinte fournit un nouveau contenu géométrique à l'espace-temps qui sera maintenant isomorphe à \mathbb{R}^4 . De manière naturelle, on sera donc amené à une formulation quadridimensionnelle des lois de la physique, facilitée par l'usage du calcul tensoriel.

4.2. Intervalle quadridimensionnel

4.2.1. La notion d'intervalle

On considère deux événements : émission d'un photon au point \mathbf{r}_1 à l'instant t_1 , puis absorption en \mathbf{r}_2 à t_2 . Ces coordonnées sont précisées dans le référentiel \mathcal{R} . La distance parcourue par le photon peut s'écrire $c\Delta t = |\Delta\mathbf{r}|$ avec $\Delta t = t_2 - t_1$ et $|\Delta\mathbf{r}| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$. On peut encore écrire cela sous la forme

$$c^2(\Delta t)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2.$$

Dans un second référentiel \mathcal{R}' en mouvement par rapport à \mathcal{R} , mais également inertiel, on a de même

$$c^2(\Delta t')^2 = (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2,$$

de sorte que l'on peut écrire

$$c^2(\Delta t)^2 - |\Delta\mathbf{r}|^2 = c^2(\Delta t')^2 - |\Delta\mathbf{r}'|^2$$

même si les durées Δt et $\Delta t'$ et les distances $|\Delta\mathbf{r}|$ et $|\Delta\mathbf{r}'|$ diffèrent.

De manière très générale, on appelle intervalle entre deux événements la quantité

$$(\Delta s)^2 = c^2(\Delta t)^2 - |\Delta\mathbf{r}|^2.$$

On constate que pour la lumière $\Delta s = 0$ et que c'est vrai dans tous les référentiels inertiels. Par ailleurs, l'existence d'une transformation non triviale entre t et t' comme on l'a vu au §3 entraîne l'existence d'une transformation analogue entre coordonnées spatiales. L'intervalle définit la métrique quadridimensionnelle de l'espace-temps de la relativité. On parle de métrique pseudo-euclidienne, ou métrique de Minkowski (1907). C'est l'équivalent du théorème de Pythagore en géométrie euclidienne. Pour des événements infiniment voisins, on définit l'intervalle infinitésimal

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$$

dans \mathbb{R}^4 au lieu de $\mathbb{R}^3 \times \mathbb{R}$ classiquement.

On introduit parfois une coordonnée temporelle imaginaire (c'est la notation initialement introduite par Minkowski) $\theta = ict$ pour retrouver (à un signe global près) une métrique euclidienne,

$$ds^2 = -(d\theta^2 + dx^2 + dy^2 + dz^2),$$

de sorte que $-ds^2$ apparaisse comme la distance entre deux points voisins de $\mathbb{R}^3 \times ic\mathbb{R}$. Nous n'utiliserons que rarement cette notation dans ce cours.

4.2.2. Invariance de l'intervalle

L'invariance de l'intervalle est une propriété essentielle, déjà pressentie par le fait que dans tous les référentiels inertiels $ds^2 = 0$ entre événements liés par un signal lumineux. Cette propriété d'invariance va jouer un rôle déterminant dans la formulation de la théorie. On en propose ici une démonstration simple, due à Landau et Lifshitz. Soient \mathcal{R} et \mathcal{R}' deux référentiels inertiels dans lesquels on note ds et ds' les valeurs de l'intervalle entre deux événements donnés. On cherche une transformation de l'un à l'autre :

$$ds^2 = f(ds'^2) = f(0) + ds'^2 f'(0) + \dots$$

Pour la lumière, on a déjà établi que $ds^2 = ds'^2 = 0$, ce qui impose que $f(0) = 0$. En se limitant au premier ordre⁽⁹⁾, l'uniformité et l'isotropie de l'espace exigent que la transformation (qui ne peut dépendre que de l'état de mouvement relatif des deux référentiels) ne dépende que de $|\mathbf{u}|$, soit

$$ds^2 = a(|\mathbf{u}|) ds'^2$$

où $a(|\mathbf{u}|) = f'(0)$. On considère maintenant un troisième référentiel inertiel, de vitesse \mathbf{v} par rapport à \mathcal{R} et \mathbf{w} par rapport à \mathcal{R}' , soit $\mathbf{w} = \mathbf{v} - \mathbf{u}$ et $|\mathbf{w}| = |\mathbf{u}|^2 + |\mathbf{v}|^2 - 2|\mathbf{u}||\mathbf{v}|\cos\theta$. On a donc $ds^2 = a(|\mathbf{u}|) ds'^2 = a(|\mathbf{v}|) ds''^2$ et $ds'^2 = a(|\mathbf{w}|) ds''^2$, soit encore $ds^2 = a(|\mathbf{u}|)a(|\mathbf{w}|) ds''^2$ ce qui impose

$$a(|\mathbf{w}|) = a(|\mathbf{v}|)/a(|\mathbf{u}|).$$

Comme par ailleurs $a(|\mathbf{w}|)$ dépend de θ l'angle entre les vitesses relatives, mais que ni $a(|\mathbf{u}|)$ ni $a(|\mathbf{v}|)$ n'en dépendent, on doit nécessairement avoir $a(|\mathbf{w}|) = \text{const} = a(|\mathbf{u}|) = a(|\mathbf{v}|)$ et du coup $a(|\mathbf{w}|) = 1$, soit

$$ds'^2 = ds^2.$$

L'intervalle entre deux événements est bien invariant par changement de référentiel inertiel. On constate ainsi que les postulats de la relativité restreinte induisent une métrique quadridimensionnelle, et que les lois de la physique se traduisent par des invariants dans cette géométrie. Il n'existe donc aucun référentiel privilégié absolu, ce qui rend caduque l'hypothèse de l'éther. En revanche, parmi l'ensemble de tous les référentiels, ceux qui sont inertiels jouent un rôle privilégié, car ce sont ceux dans lesquels les lois de la physique prennent la forme la plus simple. C'était déjà le cas des lois de la dynamique avec l'invariance galiléenne, mais on verra que cela va maintenant s'étendre à toutes les lois de la physique.

4.3. Structure causale de l'espace-temps de Minkowski

Soient deux événements dans \mathcal{R} et l'intervalle entre eux,

$$\Delta s^2 = c^2 \Delta t^2 - |\Delta \mathbf{r}|^2.$$

Il existe un référentiel \mathcal{R}' tel que les deux événements aient lieu au même endroit ($\Delta \mathbf{r}' = \vec{0}$) si

$$\Delta s^2 = c^2 (\Delta t')^2 > 0.$$

⁽⁹⁾ On peut toujours choisir ds assez petit pour que cela soit licite.

Le référentiel \mathcal{R}' n'existe que si $\Delta s^2 > 0$ (ou si Δs est réel) et on dit que *l'intervalle Δs est du genre temps*. Cela signifie que la distance $\Delta \mathbf{r}$ mesurée dans \mathcal{R} ou dans tout autre référentiel inertiel est plus courte que la distance $c\Delta t$ parcourue pendant le même temps par la lumière dans ce même référentiel. *Deux événements séparés par un intervalle du genre temps peuvent donc être reliés par un signal lumineux ou, de manière plus générale, il est possible qu'il existe une relation de cause à effet entre ces deux événements.* Si \mathcal{R}' est "solidaire" du système en évolution, il s'agit du référentiel propre et l'on note $\Delta s^2 = c^2 \Delta \tau^2$ où $\Delta \tau$ est le temps propre indiqué par une horloge "attachée" au système. Ceci se généralise, même si le système évolue à vitesse variable et ne constitue donc pas un référentiel inertiel. Pour des intervalles élémentaires on peut toujours définir un *référentiel inertiel localement tangent* au référentiel propre dans lequel on a alors

$$ds^2 = c^2 d\tau^2 = c^2 dt^2 - |d\mathbf{r}|^2$$

où

$$d\tau = dt \sqrt{1 - |\mathbf{v}|^2/c^2}, \quad \frac{|\mathbf{v}|^2}{c^2} = \frac{1}{c^2} \left(\frac{d\mathbf{r}}{dt} \right)^2.$$

A tout instant il est possible de définir un référentiel inertiel localement tangent, même si celui-ci change le long de la ligne d'univers du système étudié. Cela permet de suivre la trajectoire depuis tout autre référentiel inertiel \mathcal{R} et l'expression précédente pour des grandeurs infinitésimales se généralise aux variations finies,

$$\Delta \tau = \int_1^2 dt \sqrt{1 - |\mathbf{v}|^2(t)/c^2}$$

où $\mathbf{v}(t)$ est la vitesse mesurée dans \mathcal{R} . Un observateur solidaire de \mathcal{R} mesure entre les événements 1 et 2 une durée $\Delta t = \int_1^2 dt$ différente de $\Delta \tau$. Comme $\sqrt{1 - |\mathbf{v}|^2(t)/c^2} < 1$ on a

$$\Delta \tau < \Delta t.$$

Le temps propre, indiqué par une horloge solidaire du mouvement du système étudié, s'écoule plus lentement que le temps mesuré par une autre horloge de référence au repos. C'est un *effet géométrique de perspective* lié à l'invariance de l'intervalle. De même que la longueur d'une règle dans l'espace euclidien est plus grande que n'importe laquelle de ses projections dans n'importe quel repère, le temps propre dans l'espace de Minkowski est plus court que n'importe laquelle de ses "projections" dans n'importe quel référentiel inertiel.

Notons que si $\Delta s^2 < 0$, alors il est impossible de trouver \mathcal{R}' pour annuler $\Delta \mathbf{r}'$ et l'on a l'inégalité $|\Delta \mathbf{r}|^2 > c^2 \Delta t^2$ dans \mathcal{R} comme dans tout autre référentiel inertiel. Aucun signal, même se propageant à la vitesse de la lumière, ne peut joindre les deux événements considérés qui ne sont donc liés par *aucune relation de causalité* et qui n'ont donc même *aucune relation d'antériorité* l'un par rapport à l'autre. *L'intervalle Δs est du genre espace*, Δs est imaginaire pur.

Le lieu des points où $\Delta s = 0$ définit le *cône de lumière* et sépare les deux types d'intervalles. Seul un signal se propageant à la vitesse c (par exemple la lumière) peut joindre deux événements du *genre lumière* séparés par un intervalle nul. Le cône de lumière est défini par référence à un événement particulier, appelons-le

ici et maintenant. Les événements situés à l'intérieur du cône de lumière sont soit dans le passé absolu, soit dans le futur absolu de l'événement de référence. Ils définissent les événements qui sont potentiellement en relation causale avec *ici et maintenant.* En dehors du cône de lumière, les événements sont dans l'ailleurs absolu, ils n'entretiennent aucune relation de causalité ni d'antériorité avec *ici et maintenant.*

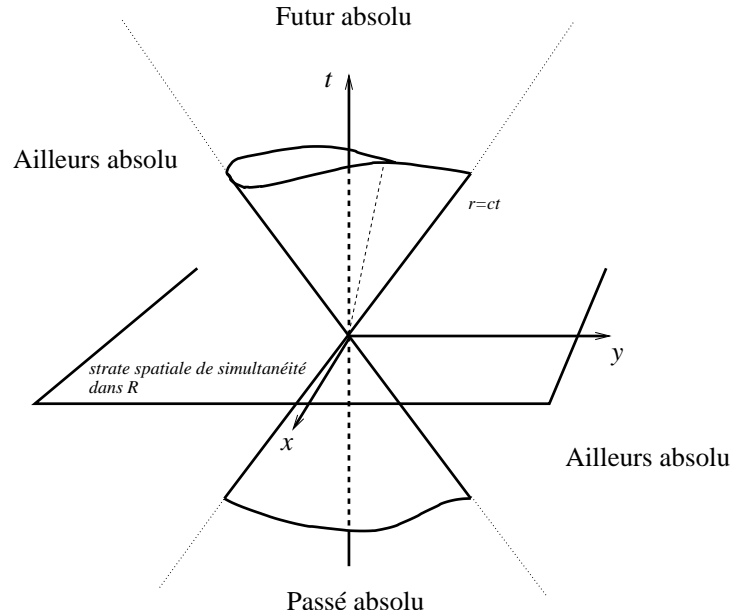


Figure 1.5 Structure de l'espace-temps minkowskien : Le cône de lumière d'un événement de référence (invariant par changement de référentiel inertiel) définit le lieu des autres événements qui peuvent avoir une relation de causalité avec lui, dans le passé ou le futur absolu. En dehors de ce cône de lumière il n'existe aucune relation causale possible. N'importe quel axe à l'intérieur du cône de lumière définit un axe des temps et la strate spatiale de simultanéité qui lui est associée.

Le cône de lumière est le même pour tous les référentiels inertiels ($ds = 0$). Le choix d'une strate spatiale de simultanéité perpendiculaire à l'axe du cône représentant l'axe du temps est conventionnel et ne vaut que dans un référentiel particulier. Une conséquence importante de l'ailleurs absolu est l'inaccessibilité de l'univers global : pour que deux événements aient interagi, ou pour qu'ils interagissent dans le futur, il faut que leurs cônes de lumière se chevauchent partiellement à un moment ou à un autre.

Chapitre 2

Transformation de Lorentz

On aborde la cinématique relativiste par la recherche d'une loi de transformation des coordonnées d'espace-temps par changement de référentiel inertiel qui soit compatible avec le principe de relativité énoncé précédemment, c'est-à-dire la transformation de Lorentz. Par extension on parlera ultérieurement de transformation de Lorentz pour d'autres quantités physiques obéissant à la même loi de transformation, par exemple la transformation de l'énergie-impulsion.

1. Transformation de Lorentz des coordonnées d'espace-temps

1.1. Etablissement de la transformation de Lorentz

On considère deux référentiels, \mathcal{R} et \mathcal{R}' , le second étant animé d'une vitesse uniforme $\mathbf{u} = u\hat{\mathbf{u}}_x$ par rapport au premier⁽¹⁰⁾. On cherche une transformation

$$\begin{aligned}x'_i &= \Xi_i(x, y, z, t), & x_i &= x, y, z \\t' &= \Theta(x, y, z, t)\end{aligned}$$

⁽¹⁰⁾ On utilisera en général la notation \mathbf{u} pour la vitesse relative de deux référentiels et \mathbf{v} pour la vitesse d'une particule test. La notation $\hat{\mathbf{u}}_\alpha$ représente le vecteur unitaire dans la direction générée par la variable α .

compatible avec le principe de relativité, c'est-à-dire telle que c soit identique dans \mathcal{R} et \mathcal{R}' .

- i) La transformation cherchée doit être linéaire, sinon un corps au repos dans \mathcal{R} ne serait pas en mouvement uniforme dans \mathcal{R}' en violation du principe d'inertie. Par exemple si l'on avait $x' = x - ut + at^2$, $t' = t$, $v'_x = v_x - u + 2at$ où le dernier terme pose problème. Les fonctions Ξ_i et Θ sont donc linéaires en leurs variables.
- ii) On note \mathbf{r}_\perp les coordonnées perpendiculaires à \mathbf{u} . La seule transformation acceptable des coordonnées transverses est de la forme $\mathbf{r}'_\perp = K(u)\mathbf{r}_\perp$ avec $K(0) = 1$ et $K(-u) = K(u)$ car elle ne doit pas dépendre de la direction (ce sont les coordonnées transverses) et elle doit redonner l'identité pour une vitesse nulle. On a aussi $\mathbf{r}_\perp = K(-u)\mathbf{r}'_\perp$, soit $\mathbf{r}_\perp = K(-u)K(u)\mathbf{r}_\perp$ ou encore

$$K(-u)K(u) = K^2(u) = 1.$$

Ces contraintes entraînent finalement

$$K(u) = 1$$

comme on l'avait anticipé au premier chapitre. Les coordonnées transverses à la vitesse relative ne sont pas affectées par le changement de référentiel.

- iii) On cherche finalement

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} a(u) & b(u) \\ d(u) & e(u) \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}$$

où les coefficients a, b, d et e ne dépendent que de u . La vitesse v d'un objet dans \mathcal{R} se transforme comme

$$v' = \frac{dx'}{dt'} = \frac{a(u)dx + b(u)dt}{d(u)dx + e(u)dt} = \frac{a(u)v + b(u)}{d(u)v + e(u)}.$$

En particulier, l'origine de \mathcal{R}' se déplace à la vitesse u par rapport à \mathcal{R} , soit

$$0 = \frac{a(u)u + b(u)}{d(u)u + e(u)}$$

qui impose $b(u) = -a(u)u$. Par ailleurs un objet au repos dans \mathcal{R} a une vitesse $v' = -u$ mesurée dans \mathcal{R}' , soit

$$-u = b(u)/e(u)$$

ou $b(u) = -e(u)u$. On a ainsi l'égalité $a(u) = e(u)$. Il reste à exprimer l'invariance de c :

$$c = \frac{a(u)c - b(u)u}{d(u)c + a(u)},$$

d'où $d(u) = -a(u)u/c^2$. La transformation cherchée est de la forme

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = a(u) \begin{pmatrix} 1 & -u \\ -u/c^2 & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} = [A(u)] \begin{pmatrix} x \\ t \end{pmatrix}.$$

Les contraintes $a(0) = 1$ et $a(-u) = a(u)$ entraînent

$$\begin{pmatrix} x \\ t \end{pmatrix} = [A(-u)] \begin{pmatrix} x' \\ t' \end{pmatrix} = [A(-u)][A(u)] \begin{pmatrix} x \\ t \end{pmatrix},$$

soit $[A(-u)][A(u)] = [1]$ qui donne sous forme développée

$$a(-u)a(u) \begin{pmatrix} 1 & u \\ u/c^2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -u \\ -u/c^2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

On en déduit l'expression du dernier paramètre

$$a(u) \equiv \gamma = \frac{1}{\sqrt{1 - u^2/c^2}}.$$

La transformation cherchée s'écrit

$$\begin{cases} x' = \gamma(x - ut) \\ y' = y \\ z' = z \\ t' = \gamma(t - ux/c^2). \end{cases}$$

Elle constitue la transformation de Lorentz. On pose en général $\beta = u/c$, soit

$$\begin{cases} x' = \gamma(x - \beta ct) \\ y' = y \\ z' = z \\ ct' = \gamma(ct - \beta x). \end{cases}$$

On note souvent $\gamma_{\mathbf{u}}$ et $\beta_{\mathbf{u}}$, notamment si plusieurs vitesses interviennent et pas seulement la vitesse relative des deux référentiels, ce qui est le cas le plus courant.

On retrouve bien entendu la transformation de Galilée des coordonnées dans la limite où c tend vers l'infini. C'est indispensable pour que la nouvelle cinématique redonne la cinématique traditionnelle à la limite non relativiste. On constate également que la transformation du temps établie au chapitre précédent est incluse dans la transformation de Lorentz.

1.2. Transformation de Lorentz pour une vitesse de direction quelconque

On peut aussi écrire vectoriellement, dans le cas d'une vitesse relative arbitraire, $\beta_{\mathbf{u}} = \frac{\mathbf{u}}{c}$, $\mathbf{r}'_{\parallel} = \gamma_{\mathbf{u}}\mathbf{r}_{\parallel} - \beta_{\mathbf{u}}\gamma_{\mathbf{u}}ct$, $\mathbf{r}'_{\perp} = \mathbf{r}_{\perp}$. Or $\mathbf{r}_{\parallel} = (\mathbf{r} \cdot \beta_{\mathbf{u}})\beta_{\mathbf{u}}/|\beta_{\mathbf{u}}|^2$ et $\mathbf{r}_{\perp} = \mathbf{r} - \mathbf{r}_{\parallel}$ d'où

$$\begin{aligned} \mathbf{r}' &= \mathbf{r}'_{\parallel} + \mathbf{r}'_{\perp} \\ &= \gamma_{\mathbf{u}}\mathbf{r}_{\parallel} - \beta_{\mathbf{u}}\gamma_{\mathbf{u}}ct + \mathbf{r}_{\perp} \\ &= \mathbf{r} + (\gamma_{\mathbf{u}} - 1)\mathbf{r}_{\parallel} - \beta_{\mathbf{u}}\gamma_{\mathbf{u}}ct \\ &= \mathbf{r} + (\gamma_{\mathbf{u}} - 1)\frac{(\mathbf{r} \cdot \beta_{\mathbf{u}})\beta_{\mathbf{u}}}{|\beta_{\mathbf{u}}|^2} - \beta_{\mathbf{u}}\gamma_{\mathbf{u}}ct \\ ct' &= \gamma_{\mathbf{u}}(ct - \beta_{\mathbf{u}} \cdot \mathbf{r}). \end{aligned}$$

1.3. Loi de composition des vitesses

On a également automatiquement la loi de composition des vitesses qui découle de ce qui précède. Avec $a(u) = \gamma$, $b(u) = -\gamma u$, $d(u) = -\gamma u/c^2$ et $e(u) = \gamma$, il vient

$$v'_x = \frac{\gamma(v_x - u)}{-\gamma uv_x/c^2 + \gamma} = \frac{v_x - u}{1 - uv_x/c^2}.$$

C'est la loi de composition des vitesses parallèles (à la vitesse relative de déplacement des deux référentiels). On obtient la transformation inverse par changement de signe, $u \rightarrow -u$:

$$v_x = \frac{v'_x + u}{1 + uv'_x/c^2}.$$

De cette relation on déduit que même si u et v'_x sont arbitrairement proches de c , la vitesse v_x observée dans \mathcal{R} ne peut pas dépasser c . La vitesse de la lumière est donc une vitesse limite. C'est en fait la vitesse de propagation de particules non massives dont la lumière (avec les photons) est l'exemple le plus connu. A priori, les gravitons devraient également se déplacer à cette vitesse.

On peut aussi calculer

$$\frac{v'_y}{c} = \frac{1}{c} \frac{dy'}{dt'} = \frac{dy}{\gamma(cdt - \beta dx)} = \frac{v_y}{c} \frac{\sqrt{1 - u^2/c^2}}{1 - uv_x/c^2},$$

et finalement ces expressions se rassemblent sous une forme vectorielle. La loi de transformation (ou de composition des vitesses) devient

$$\begin{aligned} \mathbf{v}' &= \frac{d\mathbf{r}'_{\parallel}}{dt'} + \frac{d\mathbf{r}'_{\perp}}{dt'} \\ &= \frac{\gamma_{\mathbf{u}} d\mathbf{r}_{\parallel} - \boldsymbol{\beta}_{\mathbf{u}} \gamma_{\mathbf{u}} c dt}{\frac{\gamma_{\mathbf{u}}}{c} (cdt - \boldsymbol{\beta}_{\mathbf{u}} d\mathbf{r})} + \frac{d\mathbf{r}_{\perp}}{\frac{\gamma_{\mathbf{u}}}{c} (cdt - \boldsymbol{\beta}_{\mathbf{u}} d\mathbf{r})} \\ &= \frac{\mathbf{v}_{\parallel} - \boldsymbol{\beta}_{\mathbf{u}} c + \mathbf{v}_{\perp} / \gamma_{\mathbf{u}}}{1 - \boldsymbol{\beta}_{\mathbf{u}} \mathbf{v} / c} \\ \mathbf{v}'_{\parallel} &= \frac{\mathbf{v}_{\parallel} - \boldsymbol{\beta}_{\mathbf{u}} c}{1 - \boldsymbol{\beta}_{\mathbf{u}} \mathbf{v} / c} \\ \mathbf{v}'_{\perp} &= \frac{\sqrt{1 - \boldsymbol{\beta}_{\mathbf{u}}^2}}{1 - \boldsymbol{\beta}_{\mathbf{u}} \mathbf{v} / c} \mathbf{v}_{\perp}. \end{aligned}$$

On peut en donner une forme un peu plus explicite en développant le numérateur, $\mathbf{v}_{\parallel} - \boldsymbol{\beta}_{\mathbf{u}} c + \mathbf{v}_{\perp} / \gamma_{\mathbf{u}} = v \cos \theta \hat{\mathbf{u}}_{\parallel} + \sqrt{1 - |\mathbf{u}|^2/c^2} v \sin \theta \hat{\mathbf{u}}_{\perp} - \frac{|\mathbf{u}|}{c} c \hat{\mathbf{u}}_{\parallel}$, dont le carré donne $v^2 \cos^2 \theta - 2uv \cos \theta + u^2 + v^2 \sin^2 \theta - u^2 v^2 \sin^2 \theta / c^2$, que l'on peut comparer à une expression plus compacte, $(\mathbf{v} - \boldsymbol{\beta}_{\mathbf{u}} c)^2 - (\mathbf{v} \wedge \boldsymbol{\beta}_{\mathbf{u}})^2 = v^2 - 2uv \cos \theta + u^2 - u^2 v^2 \sin^2 \theta / c^2$, identique. Pour une direction arbitraire, on a donc

$$\mathbf{v}' = \frac{\sqrt{(\mathbf{v} - \boldsymbol{\beta}_{\mathbf{u}} c)^2 - (\mathbf{v} \wedge \boldsymbol{\beta}_{\mathbf{u}})^2}}{1 - \boldsymbol{\beta}_{\mathbf{u}} \mathbf{v} / c}.$$

Là encore bien entendu, on note que la limite $c \rightarrow 0$ redonne la loi de composition traditionnelle des vitesses issue de la transformation de Galilée. On rend explicite les corrections relativistes en développant

$$\gamma = (1 - \beta^2)^{-1/2} = 1 + \frac{1}{2}\beta^2 + \frac{3}{8}\beta^4 + \dots$$

1.4. Rapidité et rotation

La représentation graphique d'un événement se fait en spécifiant (t, \mathbf{r}) ou (ct, \mathbf{r}) . On simplifie en ne considérant qu'une seule coordonnée spatiale. Un changement de référentiel inertiel apparaît alors comme une rotation dans l'espace de Minkowski. La transformation de Lorentz est encore appelée transformation du groupe de Lorentz⁽¹¹⁾. On pose

$$\begin{aligned} |\mathbf{u}|/c &= \beta = \tanh \eta \\ (1 - \beta^2)^{-1/2} &= \cosh \eta = \gamma \\ \beta(1 - \beta^2)^{-1/2} &= \sinh \eta = \beta\gamma. \end{aligned}$$

Le paramètre η est la rapidité. Il vient alors

$$\begin{aligned} x' &= x \cosh \eta - ct \sinh \eta \\ ct' &= ct \cosh \eta - x \sinh \eta, \end{aligned}$$

c'est-à-dire une transformation de coordonnées analogue à une rotation, mais avec des fonctions hyperboliques. On peut noter que

$$\begin{aligned} \cosh \eta &= \cos(i\eta) \\ \sinh \eta &= -i \sin(i\eta) \end{aligned}$$

et la transformation de Lorentz devient

$$\begin{aligned} x' &= x \cos i\eta + ict \sin i\eta \\ ict' &= -x \sin i\eta + ict \cos i\eta \end{aligned}$$

ou, sous forme matricielle,

$$\begin{pmatrix} x' \\ ict' \end{pmatrix} = \begin{pmatrix} \cos i\eta & \sin i\eta \\ -\sin i\eta & \cos i\eta \end{pmatrix} \begin{pmatrix} x \\ ict \end{pmatrix}.$$

Il s'agit formellement d'une rotation d'angle imaginaire $i\eta$ dans le plan complexe (x, ict) ⁽¹²⁾. La rapidité η n'est pas périodique, elle varie de $-\infty$ à $+\infty$. On adoptera plutôt la notation

$$\begin{pmatrix} x' \\ ct' \end{pmatrix} = \begin{pmatrix} \cosh \eta & -\sinh \eta \\ -\sinh \eta & \cosh \eta \end{pmatrix} \begin{pmatrix} x \\ ct \end{pmatrix}.$$

⁽¹¹⁾ L'appellation provient des travaux de Poincaré qui a mis en évidence la structure de groupe.

⁽¹²⁾ On doit cette représentation complexe à Minkowski. Comme on l'a déjà mentionné, nous n'adopterons pas ce point de vue ici.

On déduit des propriétés des fonctions hyperboliques que deux transformations de Lorentz consécutives de rapidités η_1 et η_2 suivant un même axe sont équivalentes à une transformation unique de rapidité $\eta_1 + \eta_2$. Graphiquement, la rotation apparaît comme sur la figure ci-dessous.

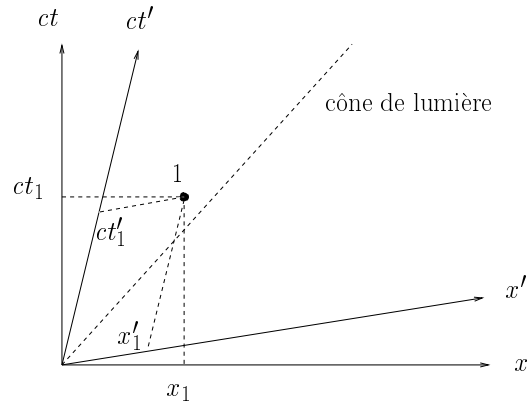


Figure 2.1 Représentation plane d'une "rotation de Lorentz". Le point 1 représente un événement quelconque.

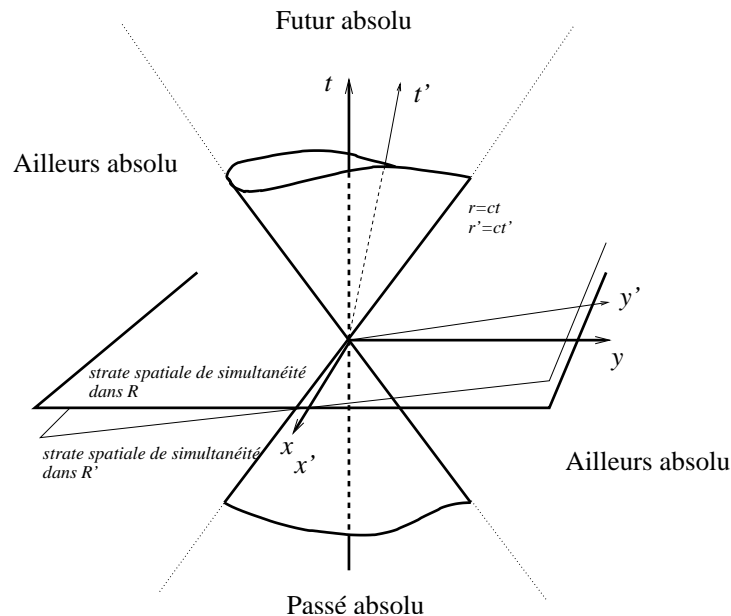


Figure 2.2 Changement de référentiel inertiel dans l'espace-temps minkowskien : n'importe quel axe à l'intérieur du cône de lumière définit un axe des temps et la strate spatiale de simultanéité qui lui est associée. Un changement d'axe des temps (représenté ici en traits plus fins) modifie la strate de simultanéité. Le changement de référentiel est analogue à une *rotation hyperbolique*, ici autour de l'axe des x invariant.

2. Quelques conséquences de la transformation de Lorentz

2.1. Intervalle invariant

Soient deux événements (ct_1, \mathbf{r}_1) et (ct_2, \mathbf{r}_2) et l'intervalle entre eux mesuré dans \mathcal{R} , $s^2 = c^2t^2 - |\mathbf{r}|^2$, où $t = t_2 - t_1$ et $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. Dans \mathcal{R}' on a immédiatement

$$\begin{aligned} s'^2 &= c^2t'^2 - |\mathbf{r}'|^2 \\ &= |\mathbf{r}|^2 \sinh^2 \eta + c^2t^2 \cosh^2 \eta - |\mathbf{r}|^2 \cosh^2 \eta - c^2t^2 \sinh^2 \eta \\ &= (c^2t^2 - |\mathbf{r}|^2)(\cosh^2 \eta - \sinh^2 \eta) = s^2 \end{aligned}$$

qui assure en effet l'invariance de l'intervalle.

2.2. Contraction des longueurs

La relativité prévoit que les objets en mouvement sont contractés dans le sens du mouvement. On a déjà mentionné que les longueurs transverses sont inchangées. Pour ce qui est des longueurs longitudinales (c'est-à-dire parallèlement à la vitesse), on obtient la longueur $L = x_2 - x_1$ de l'objet (une règle par exemple) par les coordonnées des extrémités mesurées dans \mathcal{R} (immobile) au même instant t . Aux deux événements (x_1, ct) et (x_2, ct) dans \mathcal{R} correspondent dans un référentiel \mathcal{R}' (solidaire de la règle, animée par rapport à \mathcal{R} d'une vitesse \mathbf{u}) les événements

$$\begin{aligned} (x'_1, ct'_1) &= \gamma(x_1 - \beta ct, ct - \beta x_1) \\ (x'_2, ct'_2) &= \gamma(x_2 - \beta ct, ct - \beta x_2), \end{aligned}$$

soit

$$L_0 = x'_2 - x'_1 = \gamma(x_2 - x_1) = \gamma L.$$

Comme $\gamma > 1$, la longueur L mesurée dans \mathcal{R} est plus petite que la longueur propre L_0 :

$$L = L_0 \sqrt{1 - |\mathbf{u}|^2/c^2}.$$

On parle de la contraction des longueurs en mouvement. On peut noter également que les mesures de position, simultanées dans \mathcal{R} , ne le sont pas dans \mathcal{R}' . Il est amusant aussi de mentionner l'existence d'un paradoxe consistant à faire passer une règle en mouvement à travers un trou faiblement incliné plus court que la longueur propre de la règle⁽¹³⁾.

⁽¹³⁾ On pourra consulter à ce propos le livre de Rougé.

2.3. Dilatation du temps

Une autre conséquence notoire de la transformation de Lorentz est qu'une horloge en mouvement retarde (s'écoule plus lentement) par rapport à une horloge immobile ⁽¹⁴⁾. On l'a déjà évoqué en exprimant le temps propre en toute généralité par

$$\Delta\tau = \int_1^2 \underbrace{dt}_{> d\tau} \underbrace{\sqrt{1 - |\mathbf{v}(t)|^2/c^2}}_{< 1}.$$

C'est encore plus évident à partir de l'intervalle,

$$\begin{aligned} ds^2 &= c^2 d\tau^2 \\ &= c^2 dt^2 - |d\mathbf{r}|^2, \end{aligned}$$

donc $d\tau < dt$ et par extension la même inégalité vaut entre durées finies.

Reconsidérons toutefois un cas d'école, comme la règle du paragraphe précédent. L'horloge est immobile en x' dans \mathcal{R}' et sa période (le temps écoulé entre deux "tics" successifs de l'horloge) définit l'unité de temps, $T_0 = t'_2 - t'_1$. On passe au référentiel \mathcal{R} , qui voit l'horloge se déplacer à vitesse \mathbf{v} ,

$$\begin{aligned} (x_1, ct_1) &= \gamma(x' + \beta ct'_1, ct'_1 + \beta x') \\ (x_2, ct_2) &= \gamma(x' + \beta ct'_2, ct'_2 - \beta x'), \end{aligned}$$

d'où une période mesurée dans \mathcal{R} qui vaut

$$\begin{aligned} T &= t_2 - t_1 \\ &= \frac{\gamma}{c}(ct'_2 + \beta x' - ct'_1 - \beta x') \\ &= \gamma T_0, \end{aligned}$$

soit encore

$$T = \frac{T_0}{\sqrt{1 - |\mathbf{u}|^2/c^2}}.$$

On dit qu'il y a dilatation du temps (il s'agit peut-être de la dilatation de l'unité de mesure du temps, mais je ne suis pas certain de la signification précise de ce terme) dans le référentiel propre de l'horloge, ce qui signifie que le temps "s'écoule" plus lentement dans un référentiel en mouvement par rapport à un autre référentiel. Serait-ce là l'origine d'une expression du type "tout est relatif", passée dans le vocabulaire courant pour caractériser la théorie de la relativité ?

⁽¹⁴⁾ Une autre façon de l'exprimer consiste à dire que par un effet géométrique de perspective dû à la métrique de l'espace-temps, le temps propre est plus court que le temps mesuré par tout autre observateur que celui qui est solidaire du mouvement étudié.

3. Vérifications expérimentales

3.1. L'expérience du Mont Washington

Les muons sont des particules chargées dont la masse est environ 200 fois celle de l'électron et qui se désintègrent spontanément en

$$\mu^- \rightarrow e^- + \nu_\mu + \bar{\nu}_e,$$

avec une période, mesurée dans leur référentiel propre, valant $T_0 = 1.53 \mu s$. Ces muons sont très abondants dans le rayonnement cosmique. Dans l'expérience de Rossi et Hall (1941), reprise par Frisch et Smith (1963), un compteur réglé pour détecter les muons de vitesse moyenne $0.992 c$ et situé à 1910 mètres d'altitude au sommet du Mont Washington dans le New-Hampshire enregistre 563 ± 10 muons par heure. Un détecteur identique au premier situé à 3 mètres d'altitude au niveau de la mer enregistre 408 ± 9 événements par heure. La durée de parcours dans le référentiel terrestre est de $\Delta t = 1907/0.992c = 6.408 \mu s$, soit près de 4 fois la période de désintégration. S'il y avait en moyenne 563 muons par heure au sommet, en l'absence d'effet relativiste, on s'attendrait à en observer environ $563/2^4$ au niveau de la mer, soit une trentaine. L'explication réside dans la dilatation du temps car dans le référentiel propre des muons il s'est écoulé beaucoup moins de temps,

$$t'_2 - t'_1 = \gamma^{-1}(t_2 - t_1) = \sqrt{1 - 0.992^2} \times 6.408 = 0.809 \mu s.$$

Ainsi, ayant mesuré $N_0 = 563$ muons par heure au sommet, on en attend

$$N = N_0 e^{-(t'_2 - t'_1) \ln 2 / T_0} = 563 \underbrace{e^{-0.322}}_{1 - 0.322 + 0.322^2 / 2 + \dots} \simeq 410$$

au niveau de la mer, en très bon accord avec la valeur mesurée. On peut bien entendu faire un calcul plus précis et un calcul d'erreur. Ce type d'expérience est réalisé quotidiennement dans les accélérateurs de particules et fournit une confirmation évidente de la dilatation du temps.

3.2. L'effet Doppler relativiste

Reprenons le problème de l'effet Doppler longitudinal avec une onde électromagnétique, problème pour lequel nous avons obtenu une dissymétrie dans le traitement non relativiste. En conservant les notations précédentes, nous avons établi une relation entre période émise, T_S , et période mesurée, T_O , sous la forme

$$T_O = T_S + \frac{r_2 - r_1}{c}.$$

Si la source est en mouvement, cela conduit à $T_O = T_S - \frac{u}{c} T_S$, alors que si c'est l'observateur qui est en mouvement, on obtient $T_O = T_S - \frac{u}{c} T_O$ (on se limite au cas où source et observateur se rapprochent).

Dans un traitement relativiste, ces deux relations restent correctes, mais si la source se déplace, dans la première expression, T_S et T_O sont mesurés dans le

référentiel fixe de l'observateur, et si l'on ramène la période mesurée, T_O à la période propre émise, $\tau_S = \gamma^{-1}T_S$, on obtient

$$T_O = \gamma\tau_S\left(1 - \frac{u}{c}\right) = \frac{1 - u/c}{\sqrt{1 - u^2/c^2}}\tau_S = \sqrt{\frac{1 - u/c}{1 + u/c}}\tau_S.$$

Si c'est l'observateur qui est en mouvement, on a cette fois $T_O = \gamma\tau_O$ et si l'on compare la période émise à celle mesurée par l'observateur dans son propre référentiel, on trouve

$$\tau_O = \gamma^{-1}T_O = \sqrt{1 - u^2/c^2}(1 + u/c)^{-1}T_S = \sqrt{\frac{1 - u/c}{1 + u/c}}T_S$$

c'est-à-dire que les deux situations sont maintenant traitées de manière symétrique, en accord avec le fait expérimental.

4. Calcul k

4.1. Principe de calcul k

On doit le calcul k à une conférence de vulgarisation prononcée par Hermann Bondi ⁽¹⁵⁾. C'est une manière simple et élégante de présenter les conséquences cinématiques de la relativité.

On suppose qu'un observateur immobile dans \mathcal{R} (on l'appellera **FIX**) utilise une technique de radar (émission, réflexion sur un miroir solidaire de \mathcal{R}' , détection) pour déterminer la position d'un observateur dans \mathcal{R}' , mobile par rapport à \mathcal{R} (on l'appellera **MOB**). Pour cela, **FIX** émet des pulses lumineux à intervalles de temps réguliers, t_0 . **FIX** trace un diagramme t, x qui indique le mouvement de **MOB**. On suppose pour simplifier dans un premier temps que **MOB** est animé d'un mouvement à vitesse constante u suivant Ox . Le point important est l'existence d'une relation linéaire entre les durées mesurées par **FIX** et **MOB**. Les horloges des deux personnages sont supposées synchronisées à l'origine, lorsque **FIX** et **MOB** étaient au même point. La linéarité de la relation entre durées impose que le pulse émis à l'instant $t_1 = t_0$ à l'horloge de **FIX** est reçu par **MOB** à $t'_a = kt_0$ (temps indiqué par son horloge locale), puis après réflexion, il revient à **FIX** à l'instant $t_2 = kt'_a = k(kt_0)$ (avec le même k par symétrie, puisqu'on pourrait considérer que **FIX** reçoit un signal émis par **MOB** à t'_a). Il est très important de noter que les instants t_1 , t'_a et t_2 ne sont pas mesurés par le même observateur (les primes permettent de les distinguer), mais surtout, ne se réfèrent pas aux mêmes événements (d'où la distinction t_1 puis t'_a). On a ici la succession des trois événements

- i. événement 1, émission d'un signal par **FIX**.
- ii. événement a , réflexion du signal par **MOB**.
- iii. événement 2, réception du signal par **FIX**.

⁽¹⁵⁾ Hermann Bondi, *Relativity and common sense* (Dover, New-York 1964).

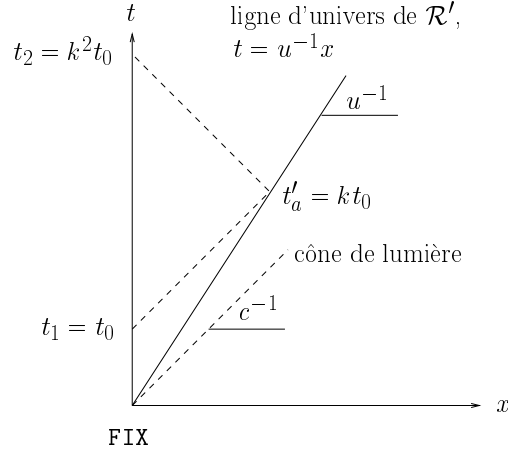


Figure 2.3 Principe du calcul k . Les temps indiqués sont mesurés sur les horloges locales et se réfèrent à trois événements distincts.

Le problème est pour ainsi dire terminé. Il suffit maintenant d'exprimer k . Pour **FIX**, l'aller-retour a duré $k^2 t_0 - t_0$, soit pour l'aller $t_{1a} = \frac{1}{2}(k^2 - 1)t_0$. Pendant ce temps, **MOB** a parcouru une distance ct_{1a} (toujours mesurée par **FIX**), soit $\frac{1}{2}c(k^2 - 1)t_0$. On peut maintenant exprimer la date de l'événement a . A l'horloge de **FIX**, la réflexion du signal par **MOB** a lieu après une durée égale à t_0 + la durée de l'aller, c'est-à-dire $t_a = t_0 + \frac{1}{2}(k^2 - 1)t_0 = \frac{1}{2}(k^2 + 1)t_0$. La vitesse u de **MOB** par rapport à **FIX** vaut donc $u = ct_{1a}/(t_0 + t_{1a})$, soit

$$u = c \frac{k^2 - 1}{k^2 + 1}, \quad k = \sqrt{\frac{1 + u/c}{1 - u/c}}.$$

On peut vérifier par exemple que ce résultat est en accord avec la "dilatation" du temps. Par exemple le signal est réfléchi par **MOB** à l'instant $t_a = \frac{1}{2}(k^2 + 1)t_0$ à l'horloge de **FIX**. L'horloge de **MOB** doit donc indiquer pour ce même événement $t'_a = \gamma^{-1}t_a$ qui donne en développant

$$\frac{1}{2} \sqrt{1 - u^2/c^2} \left(\frac{1 + u/c}{1 - u/c} + 1 \right) t_0 = \sqrt{\frac{1 + u/c}{1 - u/c}} t_0$$

qui vaut bien kt_0 .

4.2. Effet Doppler

On retrouve facilement l'effet Doppler longitudinal ($\theta = 0$ ou π) en l'occurrence ici lorsque l'observateur s'éloigne de la source dans la direction d'émission de l'onde. La fréquence d'émission vaut

$$\nu_{\text{ém.}} = \frac{1}{t_0}$$

et celle observée (dans le référentiel propre de l'observateur)

$$\nu'_{\text{obs.}} = \frac{1}{kt_0} = \frac{1}{k} \nu'_{\text{ém.}} = \sqrt{\frac{1 - u/c}{1 + u/c}} \nu_{\text{ém.}}$$

On peut encore écrire

$$\nu'_{\text{obs.}} = \frac{1 - u/c}{\sqrt{1 - u^2/c^2}} \nu_{\text{ém.}}$$

C'est le décalage Doppler relativiste vers la rouge, utilisé pour les mesures de vitesse de récession des étoiles et des galaxies.

Tableau 2.1 Comparaison entre les expressions classique et relativiste pour l'effet Doppler.

u/c	$\nu'_{\text{obs.}}/\nu_{\text{ém.}}$		
	classique	relativiste	δ
1 %	0.99	0.99005	5.10^{-5}
10 %	0.90	0.904	0.44 %
20 %	0.80	0.82	2.4 %
80 %	0.20	0.33	39 %
90 %	0.10	0.23	56 %
99 %	0.01	0.07	85 %

La comparaison entre les expressions classique et relativiste est donnée dans le tableau ci-joint. Si l'on veut rendre évidente la différence, on peut exprimer

$$\begin{aligned} \text{classique :} \quad & \nu'_{\text{obs.}} = \nu_{\text{ém.}} \left(1 - \frac{u}{c}\right) \\ \text{relativiste :} \quad & \nu'_{\text{obs.}} \simeq \nu_{\text{ém.}} \left(1 - \frac{u}{c} + \frac{1}{2} \frac{u^2}{c^2} + O\left(\frac{u^3}{c^3}\right)\right) \end{aligned}$$

et le terme correctif peut devenir dominant si la vitesse u cesse d'être petite devant c . On quantifie la différence par exemple par

$$\delta = \frac{(\nu'_{\text{obs.}}/\nu_{\text{ém.}})_{\text{rel.}} - (\nu'_{\text{obs.}}/\nu_{\text{ém.}})_{\text{class.}}}{(\nu'_{\text{obs.}}/\nu_{\text{ém.}})_{\text{rel.}}} = 1 - \gamma^{-1}.$$

4.3. Paradoxe des jumeaux

C'est un paradoxe (lorsqu'il est incomplètement posé) proposé par Langevin. Deux jumeaux sont sur Terre. L'un d'eux, MOB, embarque dans une fusée et effectue un long voyage à vitesse constante, puis revient sur Terre. A son retour, il se dit plus jeune que son jumeau FIX, resté sur Terre, puisque son temps propre s'est écoulé plus lentement. Inversement, FIX, le jumeau sédentaire, peut considérer que c'est lui, avec la Terre, qui a effectué un long voyage à vitesse $-u$, (puis est revenu à vitesse u) par rapport à MOB, resté dans une fusée immobile. FIX en conclut donc qu'il est le plus jeune. Ces deux conclusions ne sont compatibles que si les deux jumeaux ont en

fait vieillir au même rythme, mais alors la fameuse “dilatation” du temps est remise en cause.

Le paradoxe tient en fait dans la dissymétrie des référentiels. Volontairement on ne l’a pas précisé dans l’énoncé initial, mais **FIX** et la Terre sont supposés en bonne approximation constituer un référentiel inertiel (on néglige bien entendu tous les aspects liés à la gravitation, au mouvement non inertiel de la Terre, autant de choses que l’on peut éliminer dans une expérience de pensée). Dans ce cas, **MOB** a nécessairement dû être soumis à une accélération au départ, puis une décélération et une nouvelle accélération pour entamer son retour et une dernière décélération à l’arrivée. **MOB** n’est donc en aucun cas dans un référentiel inertiel. Et de fait, c’est bien lui, en raison de ces phases d’accélération, qui a vu son rythme cardiaque évoluer plus lentement que **FIX**.

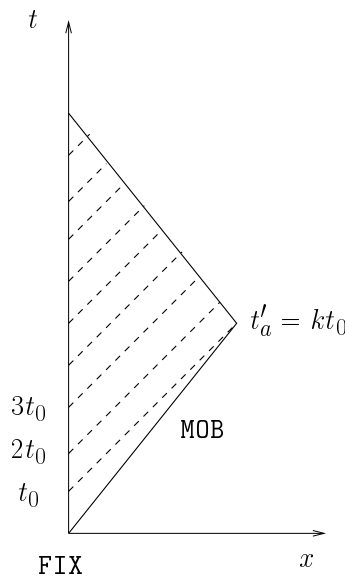


Figure 2.4 Paradoxe des jumeaux : **FIX** envoie des pulses lumineux à intervalles de temps réguliers, séparés de t_0 . **MOB** reçoit ces pulses, ce qui permet de comparer les horloges internes des deux jumeaux.

On s’en convainc aisément par le calcul k . **FIX** envoie des pulses lumineux à intervalles de temps réguliers, séparés de t_0 , et pour permettre de comparer les horloges internes des deux jumeaux, **MOB** reçoit ces pulses. On suppose pour simplifier le calcul, que le premier pulse, émis à t_0 par **FIX**, est reçu par **MOB** lorsqu’il est au point le plus éloigné de son voyage. Pour **MOB**, l’aller dure alors kt_0 , soit une durée totale de voyage aller-retour de $2kt_0$. Pour **FIX** en revanche, il s’écoule un temps t_0 avant l’émission du premier pulse, qui ne lui reviendra après réflexion qu’au bout du temps $k(kt_0)$. La lumière aura donc effectué l’aller-retour, depuis son instant d’émission, en un temps $k^2t_0 - t_0$, soit encore pour l’aller seulement $\frac{1}{2}(k^2 - 1)t_0$. Pour **FIX**, **MOB** est à son point le plus éloigné au bout d’une durée $\frac{1}{2}(k^2 - 1)t_0 + t_0 = \frac{1}{2}(k^2 + 1)t_0$. On en déduit pour chaque jumeau la durée totale du voyage de **MOB** :

$$\text{MOB : } \quad 2kt_0,$$

$$\text{FIX : } \quad (k^2 + 1)t_0.$$

Par exemple, si $u/c = 0.80$, $k = 3$ et on aura par exemple avec $t_0 = 1$ an,

$$\text{MOB : } \quad 6 \text{ ans,}$$

$$\text{FIX : } \quad 10 \text{ ans.}$$

On peut noter qu'il est sans importance sur le principe du calcul de considérer un mouvement plus réaliste, c'est-à-dire avec des phases d'accélération-décélération non instantanées.

5. Généralisation à des transformations non linéaires

Des transformations non linéaires des coordonnées spatio-temporelles ont été proposées, notamment par V. Fock ⁽¹⁶⁾. On peut introduire pour cela une échelle de longueur universelle, l , et définir la transformation

$$ct' = \frac{\gamma(ct - \beta x)}{1 + (\gamma - 1)ct/l - \gamma\beta x/l},$$

$$x' = \frac{\gamma(x - \beta ct)}{1 + (\gamma - 1)ct/l - \gamma\beta x/l},$$

$$y' = \frac{y}{1 + (\gamma - 1)ct/l - \gamma\beta x/l},$$

$$z' = \frac{z}{1 + (\gamma - 1)ct/l - \gamma\beta x/l}.$$

Cette transformation redonne la transformation de Lorentz dans la limite des courtes distances par rapport à l'échelle l . Pour donner un sens à cette longueur caractéristique, il est naturel d'introduire la longueur de Planck, que l'on peut former par une combinaison des constantes fondamentales \hbar , c et G , soit

$$l_P = (G\hbar/c^3)^{1/2}.$$

La valeur de l_P est extrêmement faible, $l_P \simeq 10^{-35}$ m, plus petite que toute échelle de longueur actuellement accessible et la transformation ci-dessus diffère de la transformation de Lorentz aux longueurs plus grandes que l_P , ce qui ne constitue pas la bonne limite. Par ailleurs, rappelons que la non linéarité ne permet plus de satisfaire au principe d'inertie dans tous les référentiels en translation uniforme par rapport à l'un d'entre eux où ce principe serait satisfait. Ces inconvénients sont très défavorables à de telles transformations, mais nous verrons au chapitre 4 qu'elles ont aussi certains avantages.

⁽¹⁶⁾ V. Fock, *The theory of Space-Time and Gravitation*, Pergamon, New-York, 1964.

Chapitre 3

Formalisme tensoriel

La relativité est une théorie qui, par certains aspects, est assez surprenante et pour laquelle les raisonnements intuitifs sont parfois sans valeur, voire mis en défaut. Il est très utile de développer un formalisme mathématique efficace sur lequel on puisse se reposer pour en faire une théorie efficace (certes assez abstraite)⁽¹⁷⁾. Cette théorie est celle des tenseurs. Les tenseurs ont été introduits par les mathématiciens italiens Ricci et Levi-Civita. La nouvelle branche des mathématiques ainsi créée s'appelait alors le calcul différentiel absolu. Dans ce chapitre on introduit les tenseurs et on présente les rudiments de base du calcul tensoriel⁽¹⁸⁾. Pour aborder la relativité générale, quelques notions plus avancées sont indispensables et ne sont pas présentées ici.

1. Quadrivecteurs et tenseurs

1.1. Quadrivecteurs contravariants et covariants

En relativité, la structure de l'espace-temps suggère l'introduction de quadrivecteurs ou 4-vecteurs, A^μ ayant une composante temporelle A^0 et trois composantes spatiales A^i . On adopte à partir de maintenant la convention selon laquelle les indices grecs peuvent prendre 4 valeurs, $\mu = 0, 1, 2, 3$, la valeur 0 étant toujours réservée aux composantes temporelles, et les indices latins peuvent prendre 3

⁽¹⁷⁾ cf M. Longair, *Theoretical concepts in physics*, Cambridge University Press, Cambridge, 2003.

⁽¹⁸⁾ On pourra consulter L.D. Landau et E. Lifshitz, *Théorie des champs*, 3ème édition, Editions MIR, Moscou 1970, §6 ou les notes de cours de David Sénéchal, *Ondes électromagnétiques*, §4.2.

valeurs, $i = 1, 2, 3$ et sont réservés aux composantes spatiales exclusivement⁽¹⁹⁾. On peut écrire indifféremment un quadrivecteur comme

$$A^\mu \equiv (A^0, A^i) \equiv (A^0, A^1, A^2, A^3). \quad (3.1)$$

La notation en exposant est réservée au label des composantes. Si l'on veut former le carré ou une autre puissance d'une quantité, on utilise des parenthèses, par exemple $(A^i)^2$, sauf s'il n'y a aucun risque de confusion, comme dans $x^2 + y^2 + z^2$. Lorsque c'est nécessaire on peut également utiliser la notation vectorielle habituelle pour rassembler les composantes spatiales,

$$A^\mu = (A^0, \mathbf{A}).$$

Il se peut qu'il soit nécessaire de distinguer le vecteur ordinaire \mathbf{A} des composantes spatiales du quadrivecteur, qui dans ce cas sont notées \mathbf{A} . Ce type de distinction sera utile au chapitre suivant, par exemple pour distinguer les composantes spatiales $\mathbf{v} = \frac{d\mathbf{r}}{d\tau}$ de la quadrivitesse en fonction de la vitesse ordinaire $\mathbf{v} = \frac{d\mathbf{r}}{dt}$, $\mathbf{v} = \gamma_{\mathbf{v}} \mathbf{v}$.

Par définition⁽²⁰⁾, un quadrivecteur *contravariant* est l'ensemble de 4 scalaires A^μ (la position de l'indice en haut est très importante) obéissant à la loi de transformation de Lorentz par changement de référentiel inertiel :

$$\begin{aligned} A'^0 &= \gamma(A^0 - \beta A^1) \\ A'^1 &= \gamma(A^1 - \beta A^0) \\ A'^2 &= A^2 \\ A'^3 &= A^3 \end{aligned}$$

avec $\beta = |\mathbf{u}|/c$ et par convention, si rien n'indique le contraire, la vitesse (constante) \mathbf{u} du référentiel primé par rapport au référentiel non primé est portée par l'axe 1,

⁽¹⁹⁾ Les notations ici sont assez diverses. Certains auteurs (comme Landau et Lifshitz) réservent les lettres latines plutôt que grecques pour les indices spatio-temporels, d'autres font courir les indices de 1 à 4, la quatrième composante étant temporelle, elle est parfois complexe (comme chez Pauli qui note $x^\mu \equiv (x^1, x^2, x^3, x^4) = (x, y, z, ict)$ là où nous écrivons $x^\mu \equiv (x^0, x^1, x^2, x^3) = (ct, x, y, z)$). Cela implique un choix de métrique qui détermine le signe de la norme invariante (ce que l'on appelle la signature). C'est également question de goût. Par exemple, pour Pauli qui préfère utiliser la notation initiale de Minkowski, on a $x_\mu x^\mu = x^2 + y^2 + z^2 - c^2 t^2$ (en réalité la distinction entre composantes contravariantes et covariantes n'est pas indispensable à ce niveau avec cette métrique) là où nous nous conformons plutôt à la convention de Landau et Lifshitz dans les éditions récentes (à partir de la 3ème édition française, conforme à la 5ème édition russe, les auteurs sont passés de la métrique de Minkowski à celle que nous employons ici, en passe de devenir universellement admise comme l'écrit Lifshitz dans la préface), avec $x_\mu x^\mu = c^2 t^2 - x^2 - y^2 - z^2$. Notons que dans la 1ère édition, intitulée *Théorie du champ*, Landau et Lifshitz utilisaient, comme Pauli, la notation complexe $x^\mu \equiv (x^1, x^2, x^3, x^4) = (x, y, z, ict)$. Les inconditionnels érudits (dont j'aimerais faire partie, mais n'étant pas russophone je me limite au titre...) pourront se reporter à Л.Д. Ландау, Е.М. Лифшиц, Теория Поля, Физматлит, Москва 2003. Weinberg utilise les lettres grecques, mais réserve celles du début de l'alphabet $\alpha, \beta, \gamma, \dots$ au cas de référentiels inertiels et celles du milieu, λ, μ, ν, \dots aux référentiels arbitraires.

⁽²⁰⁾ Les lecteurs un peu rigoureux peuvent consulter W. Appel, *Mathématiques pour la physique et les physiciens !*, H & K éditions, Paris 2002, chap.15. Cet ouvrage est d'ailleurs conseillé aussi pour ses autres chapitres...

$\mathbf{u} = (u, 0, 0) = u \hat{\mathbf{u}}_1$. On dit que les A^μ se transforment comme les composantes d'un vecteur à 4 dimensions dans une rotation de Lorentz, ce que l'on écrit de la façon suivante⁽²¹⁾ :

$$A'^\mu = \Lambda^\mu{}_\nu A^\nu. \quad (3.2)$$

Cette écriture est très importante ; la convention d'Einstein de sommation sur les indices répétés en position haute et basse est sous-entendue. Après sommation sur cet indice (ν ici), la quantité au second membre ne dépend plus de cet indice, il est dit muet. L'écriture ci-dessus résume donc 4 équations (une pour chacune des 4 valeurs possibles pour l'indice μ) et chacune de ces équations est une somme de 4 termes, par exemple

$$A'^2 = \Lambda^2{}_0 A^0 + \Lambda^2{}_1 A^1 + \Lambda^2{}_2 A^2 + \Lambda^2{}_3 A^3.$$

C'est donc équivalent à une notation matricielle. La matrice de rotation de Lorentz prend la forme :

$$[\Lambda^\mu{}_\nu] = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (3.3)$$

Un exemple de quadrivecteur contravariant est donné par

$$\begin{aligned} x^\mu &= (x^0, x^i) = (x^0, \mathbf{r}) \\ &= (ct, \mathbf{r}) = (ct, x, y, z), \end{aligned}$$

ou, pour des variations infinitésimales de coordonnées,

$$dx^\mu = (dx^0, dx^i) = (cdt, d\mathbf{r}) = (cdt, dx, dy, dz).$$

La partie spatiale du 4-vecteur, \mathbf{r} , coïncide dans ce cas avec le vecteur position ordinaire \mathbf{r} .

L'opération de sommation sur un indice muet est appelée contraction. Un exemple particulièrement important est obtenu lorsque l'on exprime la variation d'une quantité scalaire Q lors d'un déplacement élémentaire dans l'espace-temps :

$$\delta Q = \frac{\partial Q}{\partial x^\mu} \delta x^\mu.$$

La quantité au premier membre étant scalaire, celle du second membre doit l'être également, il s'agit donc d'une contraction sur l'indice muet μ qui exige que l'on note la dérivée avec un indice en bas,

$$\frac{\partial Q}{\partial x^\mu} \equiv \partial_\mu Q.$$

⁽²¹⁾ Les positions des indices (haut et bas, avant et arrière) sont importantes.

On écrit donc

$$\delta Q = \partial_\mu Q \delta x^\mu$$

où la convention d'Einstein rend la sommation explicite. La dérivée quadridimensionnelle ou quadri-gradient (l'opérateur gradient auquel on a ajouté la dérivée temporelle)

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} \equiv \left(\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right) \quad (3.4)$$

se transforme différemment d'un quadrivecteur contravariant. On peut en effet faire apparaître explicitement un changement de référentiel inertiel :

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \underbrace{\frac{\partial x'^\nu}{\partial x^\mu}}_{\Lambda^\nu{}_\mu} \frac{\partial}{\partial x'^\nu} = \Lambda^\nu{}_\mu \partial'_\nu$$

où l'on a utilisé la définition

$$x'^\nu = \Lambda^\nu{}_\mu x^\mu$$

pour identifier $\frac{\partial x'^\nu}{\partial x^\mu} = \Lambda^\nu{}_\mu$. On a donc

$$\partial_\mu = \Lambda^\nu{}_\mu \partial'_\nu$$

et par inversion, en faisant $\mathbf{u} \rightarrow -\mathbf{u}$, on a

$$\partial'_\mu = (\Lambda^{-1})^\nu{}_\mu \partial_\nu \quad (3.5)$$

avec la matrice de transformation inverse de Lorentz (qui est d'ailleurs également une transformation de Lorentz)

$$[(\Lambda^{-1})^\nu{}_\mu] = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.6)$$

et toujours avec la définition $\beta = |\mathbf{u}|/c$. Le quadrivecteur ∂_μ se transforme avec la matrice de transformation inverse, $[\Lambda^{-1}]$, contrairement aux quadrivecteurs contravariants qui se transforment avec la matrice directe $[\Lambda]$. On dit que ∂_μ est un quadrivecteur *covariant*. Par extension, tous les quadrivecteurs pour lesquels l'indice de composante est noté en bas sont des quadrivecteurs covariants. La quadri-dérivée d'une quantité scalaire par rapport aux composantes contravariantes des coordonnées, $\partial_\mu Q$, définit donc un quadrivecteur covariant. C'est un résultat très général. On note également

$$\partial_\mu Q = Q_{,\mu}. \quad (3.7)$$

Soit A^μ un 4-vecteur contravariant, $A^\mu = (A^0, A^1, A^2, A^3)$. Par définition, il se transforme avec la matrice directe $[\Lambda]$. Le 4-vecteur covariant A_μ qui lui est associé, $A_\mu = (A_0, A_1, A_2, A_3)$, devant se transformer avec la matrice inverse, on a donc

$$\begin{aligned} A'^0 &= \gamma(A^0 - \beta A^1), & A'_0 &= \gamma(A_0 + \beta A_1), \\ A'^1 &= \gamma(A^1 - \beta A^0), & A'_1 &= \gamma(A_1 + \beta A_0), \\ A'^2 &= A^2, & A'_2 &= A_2, \\ A'^3 &= A^3, & A'_3 &= A_3. \end{aligned}$$

Les objets A^μ et A_μ représentent la même grandeur physique, leurs composantes sont donc liées et pour être conformes aux transformations ci-dessus, les composantes temporelles doivent être identiques alors que les composantes spatiales doivent être changées de signe ⁽²²⁾ :

$$A_0 = A^0, \quad A_1 = -A^1, \quad A_2 = -A^2, \quad A_3 = -A^3.$$

Pour éviter toute confusion, on choisit généralement, lorsqu'on explicite les composantes, de travailler soit avec les composantes contravariantes, soit avec les composantes covariantes. On choisit ici d'utiliser le plus souvent possible les composantes contravariantes, on écrira par exemple

$$A^\mu = (A^0, \mathbf{A}), \quad A_\mu = (A^0, -\mathbf{A}). \quad (3.8)$$

Des exemples de quadrivecteurs covariants sont donnés par

$$\begin{aligned} x_\mu &= (ct, -\mathbf{r}) = (ct, -x, -y, -z), \\ dx_\mu &= (cdt, -d\mathbf{r}) = (cdt, -dx, -dy, -dz). \end{aligned}$$

On notera la particularité de la quadri-dérivée pour ce qui concerne le signe de la partie spatiale.

$$\begin{aligned} \partial_\mu &\equiv \frac{\partial}{\partial x^\mu} \equiv \left(\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right), \\ \partial^\mu &\equiv \frac{\partial}{\partial x_\mu} \equiv \left(\frac{1}{c} \frac{\partial}{\partial t}, -\nabla \right). \end{aligned}$$

On définit le carré (on parle parfois de norme) d'un quadrivecteur par la quantité

$$\begin{aligned} A_\mu A^\mu &\equiv A^\mu A_\mu = A_0 A^0 + A_1 A^1 + A_2 A^2 + A_3 A^3 \\ &= (A_0)^2 - (A_1)^2 - (A_2)^2 - (A_3)^2 \\ &= (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2 \\ &= (A^0)^2 - |\mathbf{A}|^2. \end{aligned} \quad (3.9)$$

⁽²²⁾ On se limite ici à des systèmes de coordonnées cartésiennes. Cette propriété n'est plus exacte dans un système de coordonnées arbitraire, mais alors la relation $A_\mu = g_{\mu\nu} A^\nu$ que l'on verra un peu plus loin reste vraie.

Cette quantité est une *contraction* sur l'indice muet μ , elle ne dépend donc plus d'aucun indice et constitue un scalaire invariant (ce ne serait pas le cas de combinaisons scalaires du type $(A^0)^2 + (A^1)^2 + (A^2)^2 + (A^3)^2$ par exemple). Par extension, on définit le produit invariant des quadrivecteurs ⁽²³⁾ A^μ et B^μ par (on ne précise plus cette fois les différentes formes équivalentes)

$$A_\mu B^\mu = A^0 B^0 - \mathbf{AB}. \quad (3.10)$$

On vérifie bien qu'il s'agit d'un scalaire invariant de Lorentz (ou scalaire de Lorentz) en écrivant le produit dans le référentiel primé :

$$\begin{aligned} A'_\mu B'^\mu &= A'^0 B'^0 - \mathbf{A}'\mathbf{B}' \\ &= \gamma^2(A^0 - \beta A^1)(B^0 - \beta B^1) - \gamma^2(A^1 - \beta A^0)(B^1 - \beta B^0) - A^2 B^2 - A^3 B^3 \\ &= A^0 B^0 - \mathbf{AB} \\ &= A_\mu B^\mu. \end{aligned}$$

1.2. Le tenseur métrique

A partir du quadrivecteur contravariant A^μ , on définit le tenseur métrique $g_{\mu\nu}$ qui fait passer au quadrivecteur covariant associé

$$A_\mu = g_{\mu\nu} A^\nu \quad (3.11)$$

par sommation sur l'indice ν . On précisera au paragraphe suivant la définition d'un tenseur, on se contente pour le moment de la dénomination de tenseur métrique sans davantage de précision. On voit que $g_{\mu\nu}$ doit être la matrice diagonale qui assure que $A_0 = A^0$ et $A_i = -A^i$, soit

$$[g_{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (3.12)$$

Dans une expression telle que $A_\mu = g_{\mu\nu} A^\nu$, le second membre dépend a priori de trois indices, mais il s'agit en fait d'une contraction, puisque l'un des indices (ν en l'occurrence) est répété en positions haute et basse, il disparaît donc après sommation et la quantité ainsi définie ne dépend que de l'indice restant, μ , ce qui en fait un quadrivecteur. On notera que l'indice non répété dans le membre de droite, μ , occupe la même position basse dans les deux membres de l'expression.

On peut également effectuer la transformation inverse

$$A^\mu = g^{\mu\nu} A_\nu, \quad (3.13)$$

⁽²³⁾ Encore une fois, l'égalité ci-dessous vaut pour des systèmes de coordonnées orthogonales, mais l'expression $A_\mu B^\mu$ est bien un invariant dans tout système de coordonnées.

$$[g^{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (3.14)$$

L'égalité $g_{\mu\nu} = g^{\mu\nu}$ est très particulière. Elle n'est possible qu'en coordonnées cartésiennes et parce que $g_{\mu\nu}$ est diagonal.

Le carré invariant d'un quadrivecteur peut maintenant s'écrire à l'aide uniquement de la forme contravariante et du tenseur métrique (ou à l'aide de la forme covariante uniquement) :

$$A_\mu A^\mu = g_{\mu\nu} A^\mu A^\nu = g^{\mu\nu} A_\mu A_\nu. \quad (3.15)$$

L'introduction du tenseur métrique vient au départ de ce qu'il définit la métrique par l'intervalle infinitésimal

$$\begin{aligned} ds^2 &= dx_\mu dx^\mu \\ &= g_{\mu\nu} dx^\mu dx^\nu \\ &= c^2 dt^2 - dx^2 - dy^2 - dz^2, \end{aligned}$$

soit $\text{sign}(g_{\mu\nu}) = (1, -1, -1, -1)$ où sign (pour signature) signifie que l'on ne donne que les éléments diagonaux (en toute rigueur, leur signe seulement). Cette écriture est très intéressante, car elle rend manifeste l'invariance de l'intervalle : comme ds^2 représente un carré par l'intermédiaire d'une contraction, c'est automatiquement un scalaire invariant de Lorentz.

On peut noter que la forme du tenseur métrique dépend du système de coordonnées. En coordonnées sphériques, on aura $dx^\mu = (cdt, dr, d\theta, d\varphi)$, et $ds^2 = c^2 dt^2 - dr^2 - r^2 d\theta^2 - r^2 \sin^2 \theta d\varphi^2 = g_{\mu\nu} dx^\mu dx^\nu$, soit

$$[g_{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2 \theta \end{pmatrix}. \quad (3.16)$$

Dans ce cas, $g^{\mu\nu}$ et $g_{\mu\nu}$ ne sont plus représentés par la même matrice. En effet, comme $dx_\mu = g_{\mu\nu} dx^\nu$ et $dx^\mu = g^{\mu\nu} dx_\nu$, l'intervalle s'écrit

$$ds^2 = dx_\mu dx^\mu = g_{\mu\nu} g^{\mu\nu} \underbrace{dx_\nu dx^\nu}_{ds^2},$$

soit $g_{\mu\nu} g^{\mu\nu} = 1$ d'où l'on déduit

$$g^{\mu\nu} = \frac{1}{g} G^{\mu\nu}$$

où g est le déterminant de $g_{\mu\nu}$ et $G^{\mu\nu}$ la matrice de ses cofacteurs ($g_{\mu\nu}$ étant symétrique, il n'est pas nécessaire de transposer la matrice des cofacteurs). On a donc

$$[g^{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -\frac{1}{r^2} & 0 \\ 0 & 0 & 0 & -\frac{1}{r^2 \sin^2 \theta} \end{pmatrix}.$$

1.3. Tenseurs

On définit un tenseur de rang deux par l'introduction de deux indices. Un tenseur deux fois contravariant est un objet $A^{\mu\nu}$ qui obéit, par changement de référentiel inertiel, à la transformation de Lorentz généralisée aux deux indices :

$$A'^{\mu\nu} = \Lambda^\mu_\sigma \Lambda^\nu_\tau A^{\sigma\tau}. \quad (3.17)$$

Pour un tenseur deux fois covariant on a de même

$$A'_{\mu\nu} = (\Lambda^{-1})^\sigma_\mu (\Lambda^{-1})^\tau_\nu A_{\sigma\tau}. \quad (3.18)$$

Pour un tenseur mixte de rang deux,

$$A'^\mu_\nu = \Lambda^\mu_\sigma (\Lambda^{-1})^\tau_\nu A^\sigma_\tau. \quad (3.19)$$

On note toujours les mêmes règles en calcul tensoriel : les indices contractés disparaissent et les indices non sommés conservent leur position des deux côtés de l'égalité. Un quadrivecteur est un tenseur de rang 1. On peut former des tenseurs de rang arbitraire par produits de tenseurs de rang plus faible :

$$\begin{aligned} A^{\mu\nu} &= a^\mu b^\nu, \\ B^\mu_\nu &= a^\mu c_\nu, \\ C^\mu_{\nu\rho} &= a^\mu b_\nu d_\rho, \\ D_\mu^{\nu\rho} &= b_\mu A^{\nu\rho}, \dots \end{aligned}$$

En ce qui concerne le passage des composantes covariantes ou mixtes aux composantes contravariantes par exemple, on utilise à nouveau le tenseur métrique. Par exemple de $g_{\nu\mu} A^{\mu\sigma} = A_\nu^\sigma$ on déduit $A_0^1 = g_{0\mu} A^{\mu 1} = A^{01}$ et de $g_{\nu\mu} A^{\sigma\nu} = A^\sigma_\mu$ on déduit $A^0_1 = g_{\nu 1} A^{0\nu} = -A^{01}$, ce qui justifie la nécessité de distinguer en général A^μ_ν et A_ν^μ . De manière générale, l'élévation ou l'abaissement d'un indice temporel laisse la valeur de la composante inchangée alors que l'élévation ou l'abaissement d'un indice spatial en change le signe. Voici quelques exemples :

$$\begin{aligned} A_{00} &= A^{00}, \\ A_{01} &= -A^{01}, \\ A_{11} &= A^{11}, \\ A_0^0 &= A^{00}, \\ A_0^1 &= A^{01}, \\ A_1^0 &= -A^{01}, \\ A_1^1 &= -A^{11} \dots \end{aligned}$$

On peut résumer ces transformations en notation matricielle, en introduisant une notation simplifiée,

$$[A_{\mu\nu}] = \left(\begin{array}{c|c} \boxed{A_{00}} & \boxed{A_{0j}} \\ \hline \boxed{A_{i0}} & \boxed{A_{ij}} \end{array} \right) = \begin{pmatrix} A_{00} & A_{0j} \\ A_{i0} & A_{ij} \end{pmatrix},$$

$$[A_{\mu\nu}] = \begin{pmatrix} A_{00} & A_{0j} \\ A_{i0} & A_{ij} \end{pmatrix} = \begin{pmatrix} A^{00} & -A^{0j} \\ -A^{i0} & A^{ij} \end{pmatrix}, \quad (3.20)$$

$$[A^\mu{}_\nu] = \begin{pmatrix} A^0_0 & A^0_j \\ A^i_0 & A^i_j \end{pmatrix} = \begin{pmatrix} A^{00} & -A^{0j} \\ A^{i0} & -A^{ij} \end{pmatrix}, \quad (3.21)$$

$$[A_\mu{}^\nu] = \begin{pmatrix} A_0^0 & A_0^j \\ A_i^0 & A_i^j \end{pmatrix} = \begin{pmatrix} A^{00} & A^{0j} \\ -A^{i0} & -A^{ij} \end{pmatrix}. \quad (3.22)$$

L'exemple des coordonnées comme quadrivecteur est particulièrement important car il permet d'introduire la matrice de changement de coordonnées de manière simple. Si l'on note x^μ et x'^μ les coordonnées respectives dans les deux référentiels, il existe nécessairement une transformation faisant passer des unes aux autres, $x'^\mu = x'^\mu(x^0, x^1, x^2, x^3)$, soit

$$\begin{aligned} dx'^\mu &= \frac{\partial x'^\mu}{\partial x^\nu} dx^\nu, \\ &= \frac{\partial x'^\mu}{\partial x^0} dx^0 + \frac{\partial x'^\mu}{\partial x^1} dx^1 + \frac{\partial x'^\mu}{\partial x^2} dx^2 + \frac{\partial x'^\mu}{\partial x^3} dx^3, \end{aligned}$$

ce qui permet d'identifier la transformation

$$\Lambda^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu}, \quad (3.23)$$

puisque dx'^μ est un quadrivecteur contravariant qui se transforme par définition avec la matrice $[\Lambda^\mu{}_\nu]$. On peut noter que les positions respectives en haut et en bas des indices μ et ν sont conformes aux prescriptions concernant la nature covariante de la quadri-dérivée. Par inversion de la relation de transformation des coordonnées, on peut écrire

$$dx^\mu = \frac{\partial x^\mu}{\partial x'^\nu} dx'^\nu,$$

ce qui identifie la matrice inverse

$$(\Lambda^{-1})^\mu{}_\nu = \frac{\partial x^\mu}{\partial x'^\nu}. \quad (3.24)$$

On a bien

$$\frac{\partial x'^{\mu}}{\partial x^{\nu}} \frac{\partial x^{\nu}}{\partial x'^{\mu}} = \Lambda^{\mu}_{\nu} (\Lambda^{-1})^{\nu}_{\mu} = 1$$

et c'est bien cette matrice inverse qui assure la transformation des coordonnées covariantes, puisque

$$\begin{aligned} ds^2 &= dx_{\mu} dx^{\mu} \\ &= dx_{\mu} \underbrace{(\Lambda^{-1})^{\mu}_{\nu} dx'^{\nu}}_{dx'^{\mu}} \\ &= \underbrace{dx'_{\nu} (\Lambda^{-1})^{\mu}_{\nu} dx_{\mu}}_{dx'_{\nu}} dx'^{\nu}. \end{aligned}$$

Par ailleurs, d'après les règles régissant l'élévation et l'abaissement des indices, on établit la relation

$$[\Lambda_{\nu}^{\mu}] = \begin{pmatrix} \Lambda_0^0 & \Lambda_0^j \\ \Lambda_i^0 & \Lambda_i^j \end{pmatrix} = \begin{pmatrix} \Lambda_0^0 & -\Lambda_j^0 \\ -\Lambda_0^i & \Lambda_j^i \end{pmatrix},$$

ce qui prouve, en utilisant les formes des matrices $[\Lambda^{\mu}_{\nu}]$ et $[(\Lambda^{-1})^{\nu}_{\mu}]$, que ⁽²⁴⁾

$$(\Lambda^{-1})^{\mu}_{\nu} \equiv \Lambda_{\nu}^{\mu}. \quad (3.25)$$

On note que la distinction entre T^{μ}_{ν} et T_{ν}^{μ} n'est importante que si le tenseur est non diagonal.

Par conséquent, dans les transformations du référentiel initial au référentiel primé, chaque indice contravariant fait apparaître un facteur Λ^*_{\ast} et chaque indice covariant fait intervenir un terme Λ_{\ast}^* et il apparaît autant de facteurs dans la transformation que le rang du tenseur. Les exemples précédents se généralisent aisément, par exemple

$$A'^{\mu}_{\nu\rho} = \Lambda^{\mu}_{\sigma} \Lambda_{\nu}^{\tau} \Lambda_{\rho}^{\lambda} A^{\sigma}_{\tau\lambda}.$$

À propos de $g_{\mu\nu}$, nous avons employé à plusieurs reprises le terme de tenseur métrique. Pour prouver qu'il s'agit bien d'un tenseur, il suffit de vérifier qu'il se

⁽²⁴⁾ Il suffit de noter le changement $\beta \rightarrow -\beta$ qui donne

$$[(\Lambda^{-1})^{\mu}_{\nu}] = \begin{pmatrix} (\Lambda^{-1})^0_0 & (\Lambda^{-1})^0_j \\ (\Lambda^{-1})^i_0 & (\Lambda^{-1})^i_j \end{pmatrix} = \begin{pmatrix} \Lambda^0_0 & -\Lambda^0_j \\ -\Lambda^i_0 & \Lambda^i_j \end{pmatrix}.$$

transforme convenablement par changement de référentiel. Pour cela, on exprime à nouveau l'invariance de l'intervalle,

$$\begin{aligned} ds^2 &= g_{\mu\nu} dx^\mu dx^\nu \\ &= g_{\mu\nu} \frac{\partial x^\mu}{\partial x'^\sigma} \frac{\partial x^\nu}{\partial x'^\tau} dx'^\sigma dx'^\tau \\ &= g'_{\sigma\tau} dx'^\sigma dx'^\tau, \end{aligned}$$

soit

$$\begin{aligned} g'_{\sigma\tau} &= \frac{\partial x^\mu}{\partial x'^\sigma} \frac{\partial x^\nu}{\partial x'^\tau} g_{\mu\nu} \\ &= \Lambda_\sigma^\mu \Lambda_\tau^\nu g_{\mu\nu}, \end{aligned}$$

ce qui est bien la transformation d'un tenseur de rang deux, deux fois covariant. Ce point est particulièrement important, car il existe d'autres symboles obéissant aux mêmes règles concernant les indices, les changements de signes relatifs à leurs modifications de position, mais qui ne sont pas des tenseurs car ils n'obéissent pas à la loi de transformation convenable par changement de référentiel inertiel. Par exemple Λ^μ_ν est relatif à un changement de référentiel mais pas à une observable dans un référentiel donné, ce n'est pas un tenseur. Un autre tenseur important est le tenseur de Kronecker. Il apparaît par exemple lorsque l'on établit l'invariance du produit scalaire de deux quadrivecteurs,

$$\begin{aligned} A'_\mu B'^\mu &= \Lambda_\mu^\sigma \underbrace{\Lambda_\tau^\mu}_{\delta^\sigma_\tau} A_\sigma B^\tau \\ &= A_\sigma B^\sigma. \end{aligned}$$

δ^σ_τ est le tenseur unité dont les composantes diagonales valent 1 et les composantes non diagonales sont nulles. Il joue un rôle particulier vis-à-vis du tenseur métrique. En effet, par abaissement du premier indice, on obtient un tenseur deux fois covariant qui n'est autre que $g_{\mu\nu}$ et par élévation du second indice, on obtient un tenseur deux fois contravariant qui coïncide avec $g^{\mu\nu}$:

$$\begin{aligned} \delta_{\mu\nu} &= g_{\mu\nu}, \\ \delta^{\mu\nu} &= g^{\mu\nu}. \end{aligned}$$

On note aussi qu'il n'est pas utile de distinguer les formes δ^μ_ν et δ_ν^μ , plus simplement notées δ^μ_ν , car le tenseur de Kronecker est diagonal.

2. Exemples

2.1. Temps propre

Le temps propre d'un objet en mouvement est le temps tel qu'il s'écoule dans le référentiel propre, solidaire de l'objet. On peut l'exprimer sous diverses formes,

$$d\tau = \frac{1}{c} \sqrt{ds^2} = \frac{1}{c} \sqrt{dx_\mu dx^\mu} = dt \sqrt{1 - |\mathbf{v}|^2/c^2} = \frac{1}{\gamma_{\mathbf{v}}} dt,$$

où \mathbf{v} , la vitesse de l'objet par rapport à un référentiel inertiel dans lequel le temps est noté t , peut éventuellement dépendre du temps (le référentiel propre peut subir une accélération sans pour autant que la relativité restreinte cesse d'être efficace). Le temps propre infinitésimal s'exprime par une contraction, c'est donc un scalaire invariant (tenseur de rang zéro si l'on veut). Par intégration le long de la ligne d'univers on obtient la durée propre entre deux événements,

$$\tau(ab) = \int_a^b dt \sqrt{1 - |\mathbf{v}(t)|^2/c^2}.$$

Cette relation est à comprendre au sens où l'on passe du référentiel d'un observateur (qui mesure t) à toute une succession de référentiels inertiels tangents à la ligne d'univers de la particule test, chacun étant supposé correctement synchronisé à l'instant coïncidant avec le référentiel de l'observateur.

2.2. Quadrivitesse

La vitesse ordinaire est définie en référence à un temps absolu,

$$\mathbf{v} = \frac{d\mathbf{r}}{dt}.$$

On définit par analogie la 4-vitesse ou quadrivitesse en référence au seul temps privilégié⁽²⁵⁾, le temps propre, et l'on a

$$v^\mu = \frac{dx^\mu}{d\tau}. \quad (3.26)$$

Comme $d\tau$ est une différentielle invariante, v^μ est un quadrivecteur contravariant. En fonction de la vitesse ordinaire, les composantes v^0 et v^i de la quadrivitesse s'écrivent

$$v^\mu = \left(c \frac{dt}{d\tau}, \frac{d\mathbf{r}}{d\tau} \right),$$

⁽²⁵⁾ Certains auteurs définissent la quadrivitesse par rapport à l'intervalle, $v^\mu = \frac{dx^\mu}{ds}$, cela simplifie certains facteurs c , mais au prix des dimensions car $v_\mu v^\mu$ par exemple n'est plus égal à c^2 , mais vaut 1 (sans dimension). C'est le choix de Landau et Lifshitz par exemple, L.D. Landau et E. Lifshitz, *Théorie des champs*, Editions MIR, Moscou 1970.

soit, avec $d\tau = \sqrt{1 - |\mathbf{v}|^2/c^2} dt$ ou $dt/d\tau = \gamma_{\mathbf{v}}$,

$$\begin{aligned} v^\mu &= (v^0, \mathbf{v}) \\ &= (c\gamma_{\mathbf{v}}, \gamma_{\mathbf{v}}\mathbf{v}) = \left(\frac{c}{\sqrt{1 - |\mathbf{v}|^2/c^2}}, \frac{\mathbf{v}}{\sqrt{1 - |\mathbf{v}|^2/c^2}} \right). \end{aligned} \quad (3.27)$$

On constate que l'apparition du facteur $\gamma_{\mathbf{v}}$ devant la vitesse ordinaire pour les composantes spatiales de la quadrivitesse, v^i , est un effet cinématique dû au passage du temps propre au temps de l'observateur et que ce facteur n'a rien à voir avec un aspect dynamique ou inertiel lié à la masse. Dans ce cours, on utilisera toujours m pour la masse au repos et nous éviterons l'appellation masse en mouvement pour $\gamma_{\mathbf{v}}m$, comme on le voit parfois.

Disposant de la quadrivitesse, on peut former plusieurs invariants par contraction. Par exemple la norme invariante

$$v_\mu v^\mu = \frac{dx_\mu}{d\tau} \frac{dx^\mu}{d\tau} = \frac{ds^2}{d\tau^2} = c^2, \quad (3.28)$$

ou la combinaison

$$v_\mu dx^\mu = \frac{dx_\mu}{d\tau} dx^\mu = \frac{ds^2}{d\tau} = c^2 d\tau$$

qui permet de redéfinir le temps propre

$$d\tau = \frac{1}{c^2} v_\mu dx^\mu.$$

On peut retrouver la loi de transformation des composantes ordinaires de la vitesse par application de la transformation des composantes d'un quadrivecteur contravariant : $v'^\mu = \Lambda^\mu_\nu v^\nu$ donne

$$\begin{pmatrix} v'^0 \\ v'^1 \\ v'^2 \\ v'^3 \end{pmatrix} = \gamma_{\mathbf{u}} \begin{pmatrix} 1 & -\beta_{\mathbf{u}} & 0 & 0 \\ -\beta_{\mathbf{u}} & 1 & 0 & 0 \\ 0 & 0 & 1/\gamma_{\mathbf{u}} & 0 \\ 0 & 0 & 0 & 1/\gamma_{\mathbf{u}} \end{pmatrix} \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix},$$

\mathbf{u} étant la vitesse du référentiel primé par rapport au référentiel initial ⁽²⁶⁾, portée comme d'habitude par l'axe 1. En faisant le rapport entre la deuxième et la première équation, on a

$$\frac{v'^1}{v'^0} = \frac{\gamma_{\mathbf{v}'} v'_x}{\gamma_{\mathbf{v}'} c} = \frac{-\beta_{\mathbf{u}} v^0 + v^1}{v^0 - \beta_{\mathbf{u}} v^1}, \quad \text{soit} \quad v'_x = \frac{v_x - u}{1 - uv_x/c^2},$$

ce qui établit la transformation des vitesses longitudinales. Le rapport entre troisième et première donne de même la transformation des vitesses transverses,

$$\frac{v'^2}{v'^0} = \frac{\gamma_{\mathbf{v}'} v'_y}{\gamma_{\mathbf{v}'} c} = \frac{1/\gamma_{\mathbf{u}} v^2}{v^0 - \beta_{\mathbf{u}} v^1}, \quad \text{soit} \quad v'_y = \frac{v_y \sqrt{1 - u^2/c^2}}{1 - uv_x/c^2}.$$

⁽²⁶⁾ Pour éviter toute confusion entre les facteurs relativistes $\gamma_{\mathbf{u}}$ et $\beta_{\mathbf{u}}$ relatifs au changement de référentiel et $\gamma_{\mathbf{v}}$ et $\beta_{\mathbf{v}}$ relatifs au mouvement de la particule vue depuis \mathcal{R} , Jackson introduit pour les premiers la notation Γ et B .

2.3. Quadriaccélération

Par généralisation, on définit la quadriaccélération ou 4-accélération,

$$a^\mu = \frac{dv^\mu}{d\tau} = \frac{d^2x^\mu}{d\tau^2} = \gamma_{\mathbf{v}} \frac{dv^\mu}{dt}. \quad (3.29)$$

Les composantes explicites s'expriment sous la forme

$$\begin{aligned} a^0 &= \gamma_{\mathbf{v}} \frac{dv^0}{dt} \\ &= \gamma_{\mathbf{v}} \frac{d}{dt}(c\gamma_{\mathbf{v}}) \\ &= \gamma_{\mathbf{v}} c \left[-\frac{1}{2} \left(1 - \frac{|\mathbf{v}|^2}{c^2}\right)^{3/2} \left(-2 \frac{\mathbf{v}}{c^2} \frac{d\mathbf{v}}{dt}\right) \right] \\ &= \frac{\mathbf{u}\mathbf{a}}{c(1 - |\mathbf{v}|^2/c^2)^2}, \end{aligned}$$

pour la composante temporelle et, pour les composantes spatiales, un calcul analogue conduit à

$$\mathbf{a} = \frac{\mathbf{a}}{1 - |\mathbf{v}|^2/c^2} + \frac{(\mathbf{u}\mathbf{a})\mathbf{u}}{c(1 - |\mathbf{v}|^2/c^2)^2}.$$

De nouveau on peut construire des invariants scalaires, comme

$$v_\mu a^\mu = v_\mu \frac{dv^\mu}{d\tau} = \frac{1}{2} \frac{d}{d\tau}(v_\mu v^\mu) = 0,$$

4-vitesse et 4-accélération sont perpendiculaires dans tous les référentiels inertiels, ou encore

$$a_\mu a^\mu = -\gamma_{\mathbf{v}}^2 (|\mathbf{a}|^2 + (\gamma_{\mathbf{v}}/c)^2 (\mathbf{u}\mathbf{a})^2),$$

expression qui conserve la même valeur numérique dans tous les référentiels inertiels, même si les valeurs de \mathbf{v} et \mathbf{a} , elles, peuvent changer (uniquement \mathbf{v} en fait).

2.4. Quadricourant

On a déjà évoqué que l'électromagnétisme est compatible avec la relativité restreinte. Ses lois ont donc déjà une forme covariante, même si celle-ci n'est pas forcément manifeste. La loi de conservation de la charge électrique s'écrit

$$\nabla \mathbf{j} + \frac{\partial \rho}{\partial t} = 0 \quad (3.30)$$

et suggère l'introduction d'un quadrivecteur densité de courant

$$j^\mu = (j^0, j^i) = (c\rho, \mathbf{j}), \quad (\mathbf{j} = \mathbf{j}) \quad (3.31)$$

car on a alors une contraction

$$\partial_\mu j^\mu = \frac{1}{c} \frac{\partial}{\partial t}(c\rho) + \nabla \cdot \mathbf{j}$$

qui assure que la conservation de la charge électrique

$$\partial_\mu j^\mu = 0 \quad (3.32)$$

est satisfaite dans tous les référentiels inertiels dès qu'elle l'est dans l'un d'entre eux. On dispose ainsi automatiquement de la loi de transformation du quadricourant :

$$j'^\mu = \Lambda^\mu{}_\nu j^\nu.$$

A titre d'illustration, on considère un fil rectiligne (de section S) parcouru par un courant d'intensité $I = \mathbf{j} \cdot \mathbf{S} = n|q_e||\mathbf{v}|S$. Les porteurs de charges libres (électrons de conduction), de densité de charge $\rho_- = -n|q_e|$ sont animés d'un mouvement d'ensemble à la vitesse $\mathbf{v} = -|\mathbf{v}|\hat{\mathbf{u}}_z$ et les ions (noyaux entourés du reste des électrons) sont en densité $\rho_+ = +n|q_e|$ et sont immobiles dans \mathcal{R} . Une charge $q = |q_e|$ est située à la distance r à l'extérieur du fil ; elle est animée dans \mathcal{R} de la même vitesse \mathbf{v} que les électrons de conduction. Dans le référentiel \mathcal{R} , le champ magnétique vaut

$$\mathbf{B} = \frac{\mu_0 |\mathbf{j}| S}{2\pi r} \hat{\mathbf{u}}_\varphi = \frac{\mu_0 n |q_e| |\mathbf{v}| S}{2\pi r} \hat{\mathbf{u}}_\varphi$$

ce qui donne lieu à une force de Lorentz sur la charge q en mouvement,

$$\mathbf{F} = q|\mathbf{v}|(-\hat{\mathbf{u}}_z) \wedge \mathbf{B} = \frac{\mu_0 n |q_e|^2 |\mathbf{v}|^2 S}{2\pi r} \hat{\mathbf{u}}_r.$$

On se place maintenant dans le référentiel \mathcal{R}' se déplaçant à vitesse \mathbf{v} des électrons de conduction. Le quadri-courant $j^\mu = (\rho c, \mathbf{j}) = (\rho c, j_z, j_x, j_y)$ (on choisit l'axe Oz comme axe de la transformation) se transforme par transformation de Lorentz suivant

$$\begin{pmatrix} \rho' c \\ j'_z \\ j'_x \\ j'_y \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \rho c \\ j_z \\ j_x \\ j_y \end{pmatrix} = \begin{pmatrix} \gamma\rho c - \beta\gamma j_z \\ -\beta\gamma\rho c + \gamma j_z \\ j_x \\ j_y \end{pmatrix}$$

La densité de charge associée aux ions, ρ_+ , se transforme alors suivant $\rho'_+ = \gamma\rho_+$ (il n'y a pas de courant associé aux ions dans \mathcal{R}), alors que pour les électrons on a la transformation inverse (c'est dans \mathcal{R}' que les électrons sont au repos), $\rho'_- = \gamma^{-1}\rho_-$. Dans le référentiel en mouvement, le conducteur apparaît donc chargé,

$$\rho' = \rho'_+ + \rho'_- = \beta^2 \gamma n |q_e|.$$

Cela génère un champ électrique (les longueurs transverses, telles que r ici, sont inchangées) tel que $2\pi r E' = \frac{1}{\epsilon_0} \rho' S$, soit

$$\mathbf{E}' = \frac{(|\mathbf{v}|^2/c^2) \gamma n |q_e| S \mu_0 c^2}{2\pi r} \hat{\mathbf{u}}_r$$

d'où résulte la force

$$\mathbf{F}' = \gamma \frac{\mu_0 n |q_e|^2 |\mathbf{v}|^2 S}{2\pi r} \hat{\mathbf{u}}_r.$$

Le champ magnétique est simplement donné par

$$\mathbf{B}' = \frac{\mu_0 |\mathbf{j}'| S}{2\pi r} \hat{\mathbf{u}}_\varphi = \gamma \frac{\mu_0 n |q_e| |\mathbf{v}| S}{2\pi r} \hat{\mathbf{u}}_\varphi,$$

puisque $j' = \gamma j$. Ces expressions sont conformes aux lois de transformation des champs ci-dessous (que l'on établira plus tard)

$$\begin{aligned} \mathbf{E}'_{\parallel} &= \mathbf{E}_{\parallel}, & \mathbf{E}'_{\perp} &= \gamma_{\mathbf{u}} (\mathbf{E}_{\perp} + \boldsymbol{\beta}_{\mathbf{u}} c \wedge \mathbf{B}_{\perp}) \\ \mathbf{B}'_{\parallel} &= \mathbf{B}_{\parallel}, & \mathbf{B}'_{\perp} &= \gamma_{\mathbf{u}} (\mathbf{B}_{\perp} - (\boldsymbol{\beta}_{\mathbf{u}}/c) \wedge \mathbf{E}_{\perp}). \end{aligned}$$

et elles imposent de plus que $F' = \gamma F$ c'est-à-dire que \mathbf{F} se transforme comme un quadri-vecteur.

3. Intérêt du formalisme tensoriel en relativité

L'intérêt essentiel de la notation tensorielle, outre sa forme extrêmement compacte, est qu'elle assure automatiquement qu'une expression obéit aux bonnes lois de transformation par changement de référentiel inertiel : si une loi physique s'exprime par une égalité entre deux tenseurs dans un référentiel donné,

$$G^{\mu\nu} = T^{\mu\nu}$$

par exemple, dans le référentiel primé elle prend la même forme entre les quantités primées correspondantes,

$$G'^{\mu\nu} = T'^{\mu\nu},$$

avec bien sûr $G'^{\mu\nu} = \Lambda^\mu_\sigma \Lambda^\nu_\tau G^{\sigma\tau}$ et $T'^{\mu\nu} = \Lambda^\mu_\sigma \Lambda^\nu_\tau T^{\sigma\tau}$. Une telle expression qui conserve la même forme est dite covariante (sous-entendu par changement de référentiel inertiel) et le formalisme consistant à écrire les lois physiques sous forme tensorielle est dit manifestement covariant (car sous cette forme, la covariance est manifeste). Un autre avantage du formalisme tensoriel est qu'il permet souvent une généralisation relativiste des lois classiques. A l'instar de l'analyse dimensionnelle, il produit en effet des contraintes fortes, puisqu'une égalité tensorielle ne peut intervenir qu'entre tenseurs de même rang et de même nature, ce qui limite en général la forme d'une loi physique, aux constantes sans dimension près. Par exemple le mouvement d'une charge dans un champ électrique est décrit classiquement par la relation fondamentale de la dynamique

$$m \frac{d^2 \mathbf{r}}{dt^2} = q \mathbf{E}.$$

On verra que le champ électrique intervient, sous la forme E_x/c par exemple, dans les composantes d'un tenseur de rang deux, le tenseur champ électromagnétique

$F_{\mu\nu}$. On généralise logiquement le premier membre de la relation fondamentale de la dynamique par l'expression $m \frac{d^2 x^\mu}{d\tau^2}$, tenseur contravariant de rang 1. Ecrire alors au second membre $qF^{\mu\nu}$ pose deux difficultés. D'une part du point de vue dimensionnel (c'est E^i/c qui intervient dans $F^{\mu\nu}$), on pourrait alors écrire $qF^{\mu\nu}c$, mais il reste la seconde difficulté, car cette expression est un tenseur de rang 2. Pour en faire un tenseur de rang 1, il faut assurer la contraction d'un des indices contravariants et le plus simple est de remplacer c par v_ν , soit

$$m \frac{d^2 x^\mu}{d\tau^2} = qF^{\mu\nu} v_\nu.$$

Cette dernière expression est manifestement covariante, elle est correcte du point de vue dimensionnel, elle a donc toutes les chances d'être exacte, à une constante sans dimension près éventuellement. Notons que si le système de coordonnées est absolument général, il faut passer à une extension de la dérivée, la dérivée covariante et l'expression prend une forme analogue, $m \frac{D^2 x^\mu}{D\tau^2} = qF^{\mu\nu} v_\nu$ où le sens de cette dérivée covariante sera évoqué au chapitre 9.

Chapitre 4

Dynamique relativiste

Ce chapitre est assez largement inspiré, pour ce qui concerne les collisions en particulier, des ouvrages de Landau et Lifshitz et de Rougé. La partie traitant du problème de Kepler classique, notamment de l'influence des autres planètes sur le mouvement de Mercure, est tirée de Landau et Lifshitz, *Mécanique* et de B. Davies, *Elementary theory of perihelion precession*, Am. J. Phys. **51** (1983) 909.

1. Quantité de mouvement et énergie

1.1. Préambule

La généralisation relativiste de la quantité de mouvement et de l'énergie n'est pas une question triviale. Le théorème de Noether précise que pour une théorie définie par l'action

$$\int L dt,$$

l'homogénéité de l'espace entraîne la conservation de la quantité

$$\frac{\partial L}{\partial \mathbf{v}}$$

qui peut donc être définie comme la quantité de mouvement \mathbf{p} . De même l'homogénéité du temps implique la conservation de

$$\mathbf{v} \frac{\partial L}{\partial \mathbf{v}} - L$$

que l'on identifie à l'énergie. Ces résultats sont bien connus en mécanique analytique. En relativité, le principe de moindre action repose, comme on le verra plus tard, sur l'action

$$\begin{aligned} S &= -mc \int ds \\ &= -mc^2 \int d\tau \\ &= -mc^2 \int \sqrt{1 - \frac{|\mathbf{v}|^2}{c^2}} dt, \end{aligned} \tag{4.1}$$

c'est-à-dire sur le lagrangien

$$L = -mc^2 \sqrt{1 - \frac{|\mathbf{v}|^2}{c^2}}. \tag{4.2}$$

Il en découle immédiatement que

$$\begin{aligned} \mathbf{p} &= \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1 - \frac{|\mathbf{v}|^2}{c^2}}} \\ E &= \mathbf{v} \frac{\partial L}{\partial \mathbf{v}} - L = \frac{m|\mathbf{v}|^2}{\sqrt{1 - \frac{|\mathbf{v}|^2}{c^2}}} + mc^2 \sqrt{1 - \frac{|\mathbf{v}|^2}{c^2}} \\ &= \frac{mc^2}{\sqrt{1 - \frac{|\mathbf{v}|^2}{c^2}}}. \end{aligned} \tag{4.3}$$

La mécanique analytique fournit donc une définition sans ambiguïté de \mathbf{p} et de E qui diffèrent des expressions classiques. Nous allons justifier ces relations au moyen du formalisme covariant.

1.2. Quadrivecteur énergie-impulsion

Intuitivement il est naturel de remplacer l'expression classique $\mathbf{p} = m\mathbf{v} = m \frac{d\mathbf{r}}{dt}$ par une relation dans laquelle la trajectoire est paramétrée par le temps propre, $\mathbf{p} = m \frac{d\mathbf{r}}{d\tau}$, pour les composantes spatiales p^i d'un 4-vecteur p^μ :

$$\mathbf{p} = m \frac{d\mathbf{r}}{dt} \frac{dt}{d\tau} = \frac{m\mathbf{v}}{\sqrt{1 - |\mathbf{v}|^2/c^2}}$$

qui coïncide avec la définition naturelle $p^i = mv^i$ faisant intervenir les composantes spatiales de la 4-vitesse. Il suffit alors de généraliser à la composante temporelle, $p^0 = mv^0$, soit

$$p^0 \equiv \frac{E}{c} = \frac{mc}{\sqrt{1 - |\mathbf{v}|^2/c^2}},$$

où E/c est introduit pour des raisons dimensionnelles. Le quadrivecteur énergie-impulsion est donc défini par

$$p^\mu = mv^\mu = (\gamma_{\mathbf{v}} mc, \gamma_{\mathbf{v}} m\mathbf{v}). \quad (4.4)$$

C'est bien un quadrivecteur, ou tenseur de rang 1, contravariant car il est proportionnel à v^μ qui en est également un.

On peut noter que

$$\begin{aligned} \gamma_{\mathbf{v}} &= \left(1 - \frac{|\mathbf{v}|^2}{c^2}\right)^{-1/2} \\ &= 1 + \frac{1}{2} \frac{|\mathbf{v}|^2}{c^2} + \frac{3}{8} \frac{|\mathbf{v}|^4}{c^4} + \dots, \end{aligned}$$

soit

$$\begin{aligned} E = p^0 c &= mc^2 \left(1 + \frac{1}{2} \frac{|\mathbf{v}|^2}{c^2} + \frac{3}{8} \frac{|\mathbf{v}|^4}{c^4} + \dots\right) \\ &= mc^2 + \frac{1}{2} m |\mathbf{v}|^2 + \frac{3}{8} m \frac{|\mathbf{v}|^4}{c^2} + \dots \end{aligned}$$

où les termes successifs représentent l'énergie au repos pour mc^2 , l'énergie cinétique classique pour $\frac{1}{2} m |\mathbf{v}|^2$ et la première correction relativiste pour $\frac{3}{8} m \frac{|\mathbf{v}|^4}{c^2} = \frac{1}{2} m |\mathbf{v}|^2 (3|\mathbf{v}|^2/4c^2)$.

Compte-tenu de ce qui précède, l'énergie cinétique relativiste s'exprime par l'énergie à laquelle on soustrait l'énergie au repos,

$$K = E - mc^2 = mc^2(\gamma_{\mathbf{v}} - 1).$$

Pour la quantité de mouvement, un développement limité à faible vitesse donne

$$\mathbf{p} = m\mathbf{v} \left(1 + \frac{1}{2} \frac{|\mathbf{v}|^2}{c^2} + \dots\right).$$

L'apparition du concept d'énergie au repos, mc^2 , est l'un des résultats célèbres d'Einstein qui a ouvert de nouveaux champs d'activité, notamment avec la physique nucléaire. Une autre relation particulièrement importante en dynamique relativiste est fournie par le carré invariant du quadrivecteur énergie-impulsion :

$$\begin{aligned} p_\mu p^\mu &= m^2 v_\mu v^\mu \\ &\equiv m^2 c^2 \\ &= (p^0)^2 - (p^1)^2 - (p^2)^2 - (p^3)^2 \\ &= E^2/c^2 - |\mathbf{p}|^2. \end{aligned} \quad (4.5)$$

C'est un invariant souvent écrit sous la forme

$$E^2 = |\mathbf{p}|^2 c^2 + m^2 c^4. \quad (4.6)$$

Cette relation (en l'absence d'énergie potentielle) est notamment utile dans l'étude des collisions relativistes. On peut remarquer que l'on a aussi

$$E = mc^2 + K$$

qui conduit à une relation entre énergie cinétique et quantité de mouvement qui remplace l'expression classique $K = |\mathbf{p}|^2/2m$,

$$K = \sqrt{|\mathbf{p}|^2 c^2 + m^2 c^4} - mc^2.$$

Si $mc^2 \gg |\mathbf{p}|^2/2m$, on retrouve l'expression classique.

Sachant par ailleurs que p^μ est un quadrivecteur contravariant, on a automatiquement sa loi de transformation,

$$p'^\mu = \Lambda^\mu{}_\nu p^\nu, \quad (4.7)$$

soit

$$\begin{aligned} p'^0 &= \gamma_{\mathbf{u}}(p^0 - \beta_{\mathbf{u}} p^1), \\ p'^1 &= \gamma_{\mathbf{u}}(p^1 - \beta_{\mathbf{u}} p^0), \\ p'^2 &= p^2, \\ p'^3 &= p^3, \end{aligned} \quad (4.8)$$

pour un mouvement relatif à vitesse \mathbf{u} entre les deux référentiels parallèlement à l'axe 1, ce qui donne en particulier

$$\begin{aligned} p'_x &= \frac{p_x - |\mathbf{u}|E/c^2}{\sqrt{1 - |\mathbf{u}|^2/c^2}}, \\ E' &= \frac{E - |\mathbf{u}|p_x}{\sqrt{1 - |\mathbf{u}|^2/c^2}}. \end{aligned} \quad (4.9)$$

1.3. Digression : masse au repos, masse en mouvement

A la suite d'Einstein, certains auteurs introduisent la définition d'une masse au repos ⁽²⁷⁾, c'est le m que nous utilisons, ainsi qu'une masse en mouvement différente censée simplifier certaines expressions en les rendant similaires aux expressions newtoniennes correspondantes, et caractérisant l'augmentation d'inertie d'un objet en mouvement,

$$m(\mathbf{v}) = \frac{m}{\sqrt{1 - |\mathbf{v}|^2/c^2}}.$$

⁽²⁷⁾ Elle est souvent notée m_0 , mais cela n'a rien à voir avec un indice covariant, d'ailleurs la masse est un scalaire.

Cela permet de conserver pour la quantité de mouvement l'expression classique $\mathbf{p} = m(\mathbf{v})\mathbf{v}$ au lieu de $\mathbf{p} = \gamma_{\mathbf{v}}m\mathbf{v}$. Ce n'est cependant pas rigoureux car comme on l'a déjà observé l'apparition du facteur γ est due à un effet cinématique et non pas dynamique ⁽²⁸⁾. C'est particulièrement clair lorsque l'on décompose

$$\mathbf{p} = m \frac{d\mathbf{r}}{d\tau} = m\gamma_{\mathbf{v}} \frac{d\mathbf{r}}{dt} = \gamma_{\mathbf{v}}m\mathbf{v},$$

où le facteur relativiste $\gamma_{\mathbf{v}}$ provient du passage du temps propre τ à celui de l'observateur, t , il est donc porté par la vitesse et non pas par la masse. De plus, entre les composantes spatiales des quadrivecteurs p^μ et v^μ , on a bien de toute manière la relation $\mathbf{p} = m\mathbf{v}$. L'inutilité du concept de masse en mouvement est d'autant plus criante qu'elle ne simplifie en rien l'expression de l'énergie cinétique, car

$$\frac{1}{2}m(\mathbf{v})|\mathbf{v}|^2 = \frac{1}{2}\gamma_{\mathbf{v}}m|\mathbf{v}|^2 \simeq \frac{1}{2}m|\mathbf{v}|^2 + \frac{1}{4}m\frac{|\mathbf{v}|^4}{c^2} + \dots$$

ne coïncide pas avec

$$K = (\gamma_{\mathbf{v}} - 1)mc^2 \simeq \frac{1}{2}m|\mathbf{v}|^2 + \frac{3}{8}m\frac{|\mathbf{v}|^4}{c^2} + \dots$$

L'intérêt en revanche de $m(\mathbf{v})$ est que cette quantité est liée à l'énergie totale, $E = m(\mathbf{v})c^2$ et que c'est donc une quantité qui se conserve lors des collisions.

1.4. Particules de masse nulle

La relation $E^2 = |\mathbf{p}|^2c^2 + m^2c^4$ permet d'envisager des particules de masse nulle. On a dans ce cas

$$E = |\mathbf{p}|c. \tag{4.10}$$

Si l'on utilise par ailleurs

$$E = \frac{mc^2}{\sqrt{1 - |\mathbf{v}|^2/c^2}}$$

avec $m = 0$, on voit qu'il faut au moins une forme indéterminée pour s'accorder à la valeur finie $|\mathbf{p}|c$, ce qui exige qu'une particule de masse nulle se déplace à la vitesse c . En vertu de la loi de composition des vitesses qui laisse la valeur de c invariante, on en déduit qu'une particule de masse nulle est animée d'une vitesse c , quel que soit le référentiel inertiel d'où elle est considérée. C'est le cas pour le photon, pour lequel la borne supérieure expérimentale est de

$$m_\gamma \leq 4.10^{-51}\text{kg}$$

⁽²⁸⁾ Voir par exemple C. Lanczos, *The variational principles of mechanics*, Dover, New York 1986, p. 318.

et c'était le cas pensait-on récemment des trois neutrinos, pour lesquels les bornes sont moins sévères (pour comparaison, $m_e = 0.511$ MeV),

$$m_{\nu_e} \leq 15 \text{ eV},$$

$$m_{\nu_\mu} \leq 0.17 \text{ MeV},$$

$$m_{\nu_\tau} \leq 24 \text{ MeV},$$

avant que l'on ne découvre les oscillations de masse des neutrinos électronique et muonique. C'est encore le cas prévu par la théorie pour le graviton.

Par ailleurs, la correspondance entre variables dynamiques et variables ondulatoires en mécanique quantique

$$E = \hbar\omega, \quad \mathbf{p} = \hbar\mathbf{k}$$

suggère l'introduction d'un quadrivecteur d'onde

$$k^\mu \equiv (\omega/c, \mathbf{k})$$

tel que

$$k'^0 = \gamma_{\mathbf{u}}(k^0 - \beta_{\mathbf{u}}k^1)$$

$$k'^1 = \gamma_{\mathbf{u}}(k^1 - \beta_{\mathbf{u}}k^0)$$

$$k'^2 = k^2$$

$$k'^3 = k^3.$$

La relation de dispersion $E = |\mathbf{p}|c$ se traduit en $\omega = c|\mathbf{k}|$.

On peut noter que la loi de transformation redonne l'expression relativiste de l'effet Doppler. Prenons l'exemple de l'effet Doppler longitudinal, où le référentiel \mathcal{R}' s'éloigne de \mathcal{R} à vitesse \mathbf{u} suivant l'axe 1. On a

$$\omega'/c = \frac{\omega/c - k_x|\mathbf{u}|/c}{\sqrt{1 - |\mathbf{u}|^2/c^2}},$$

soit sous la forme habituelle, $\omega' = \omega \sqrt{\frac{1 - |\mathbf{u}|/c}{1 + |\mathbf{u}|/c}}$.

2. Cinématique des réactions entre particules

2.1. Terminologie

On parle de cinématique des réactions entre particules car on se place ici dans une description 'macroscopique' des collisions (élastiques, auquel cas les particules conservent leur intégrité avant et après l'interaction, mais aussi inélastiques, il peut alors apparaître de nouvelles particules dans l'état final, ou certaines des particules incidentes peuvent par exemple se désexciter). Dans tous les cas, on ignore les détails qui relèvent de la dynamique et on ne cherche pas à préciser les détails de la trajectoire sous l'effet de l'interaction. Les transferts d'énergie et d'impulsion sont simplement déduits des lois de conservation de ces quantités, mais en aucun cas de la résolution d'une équation dynamique. Cela implique le plus souvent que l'état initial soit supposé entièrement déterminé et que dans l'état final, l'identité des particules et les angles d'éjection soient connus.

2.2. Le problème de la conservation de la masse

Considérons la collision dans \mathcal{R} de deux corps de masses m_1 et m_2 conduisant à un état final constitué de n particules de masses m'_i , $i = 1, \dots, n$. La conservation de la quantité de mouvement classique s'écrit

$$m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2 = \sum_i m'_i \mathbf{v}'_i.$$

Cette équation est également vraie dans un autre référentiel galiléen \mathcal{R}' se déplaçant à la vitesse \mathbf{u} par rapport à \mathcal{R} ,

$$m_1(\mathbf{v}_1 - \mathbf{u}) + m_2(\mathbf{v}_2 - \mathbf{u}) = \sum_i m'_i(\mathbf{v}'_i - \mathbf{u}).$$

Il en découle automatiquement la conservation de la masse totale

$$m_1 + m_2 = \sum_i m'_i.$$

Cette équation n'est plus vérifiée en relativité ou la quantité qui se conserve serait en revanche

$$\gamma_1 m_1 + \gamma_2 m_2 = \sum_i \gamma_i m'_i.$$

La relativité amène ainsi à considérer des échanges de masse en énergie.

2.3. Notion de masse d'un système composé

On considère n particules de masses m_i et animées de vitesses \mathbf{v}_i . En omettant l'énergie potentielle d'interaction, l'impulsion et l'énergie totale valent

$$\mathbf{p}_{\text{tot}} = \sum_i \frac{m_i \mathbf{v}_i}{\sqrt{1 - |\mathbf{v}_i|^2/c^2}},$$

$$E_{\text{tot}} = \sum_i \frac{m_i c^2}{\sqrt{1 - |\mathbf{v}_i|^2/c^2}}.$$

On définit le référentiel du centre d'impulsion (où les quantités sont notées avec des primes) par $\mathbf{p}'_{\text{tot}} = 0$. Il se déplace à vitesse \mathbf{u} suivant un axe choisi comme axe 1 par rapport au référentiel initial (dit référentiel "laboratoire"). On a

$$p'_{\text{tot}}{}^1 = \gamma_{\mathbf{u}}(p_{\text{tot}}^1 - \beta_{\mathbf{u}} E_{\text{tot}}/c) = 0$$

$$\beta_{\mathbf{u}} = \frac{|\mathbf{u}|}{c} = cp_{\text{tot}}^1/E_{\text{tot}}.$$

Dans le référentiel du centre d'impulsion, l'énergie non potentielle vaut donc

$$\begin{aligned} E'_{\text{tot}} &= p'_{\text{tot}}{}^0 c = \gamma_{\mathbf{u}} (E_{\text{tot}} - c \beta_{\mathbf{u}} \underbrace{p_{\text{tot}}^1}_{\frac{\beta_{\mathbf{u}} E_{\text{tot}}}{c}}) \\ &= \gamma_{\mathbf{u}} E_{\text{tot}} (1 - \beta_{\mathbf{u}}^2) \\ &= E_{\text{tot}} \sqrt{1 - |\mathbf{u}|^2/c^2}, \end{aligned}$$

ou encore

$$E_{\text{tot}} = \gamma_{\mathbf{u}} E'_{\text{tot}}.$$

On définit la masse M du système composé comme s'il était immobile dans le référentiel du centre d'impulsion, soit

$$E'_{\text{tot}} = M c^2$$

ce qui exige une expression qui légitime cette définition dans le référentiel initial

$$E_{\text{tot}} = \frac{M c^2}{\sqrt{1 - |\mathbf{u}|^2/c^2}}.$$

Si les particules n'ont aucune énergie potentielle mutuelle, on a donc (ici estimé dans \mathcal{R}')

$$M = \sum_i \frac{m_i}{\sqrt{1 - |\mathbf{v}'_i|^2/c^2}},$$

où \mathbf{v}'_i est la vitesse de la particule i dans le référentiel du centre d'impulsion.

Si au contraire les particules interagissent, la masse totale contient des termes provenant de cette énergie mutuelle. On peut envisager deux cas extrêmes :

- Dans un gaz chaud raréfié, l'énergie totale des particules est essentiellement cinétique car elles n'interagissent presque pas, et on a des facteurs $\gamma_{\mathbf{v}'_i}$ grands, de sorte que $M > \sum_i m_i$.
- Dans un système où l'énergie potentielle de cohésion est très élevée et fortement négative (un noyau par exemple), la masse totale est inférieure à la masse des constituants. On peut en effet écrire

$$\begin{aligned} E_{\text{tot}} &= \sum_i \frac{m_i c^2}{\sqrt{1 - |\mathbf{v}_i|^2/c^2}} + \frac{1}{2} \sum_{ij} V(\mathbf{r}_{ij}), \\ E'_{\text{tot}} &= \sum_i \frac{m_i c^2}{\sqrt{1 - |\mathbf{v}'_i|^2/c^2}} + \frac{1}{2} \sum_{ij} V(\mathbf{r}'_{ij}) \\ &= M c^2 \end{aligned}$$

soit

$$M c^2 = \sum_i \frac{m_i c^2}{\sqrt{1 - |\mathbf{v}'_i|^2/c^2}} + \frac{1}{2} \sum_{ij} V(\mathbf{r}'_{ij})$$

où le dernier terme est très négatif.

2.4. Désintégrations de particules

On considère dans ce qui suit des collisions entre particules. L'état initial et l'état final sont considérés bien avant et après l'interaction et celle-ci est supposée à courte portée de sorte que dans l'état initial comme dans l'état final, l'énergie totale des particules comprend exclusivement l'énergie au repos et l'énergie cinétique, à l'exclusion de toute forme d'énergie potentielle. La désintégration de particules est un processus manifestement inélastique car l'intégrité des particules n'est pas conservée. C'est le cas le plus simple d'interaction que l'on puisse envisager, puisque dans le référentiel propre de la particule initiale, celle-ci étant au repos, son énergie totale, avant l'interaction, est l'énergie au repos.

On considère la désintégration d'une particule a de masse m_a initialement au repos dans \mathcal{R} en deux particules 1 et 2 de masses m_1 et m_2 , d'énergies E_1 et E_2 et de quantités de mouvement \mathbf{p}_1 et \mathbf{p}_2 . La conservation de l'énergie-impulsion p^μ peut s'écrire à partir de $p_2^\mu = p_a^\mu - p_1^\mu$ ou de $p_1^\mu = p_a^\mu - p_2^\mu$. A cette relation de conservation (la première par exemple) est associé un invariant

$$p_{2\mu}p_2^\mu = p_{a\mu}p_a^\mu - 2p_{a\mu}p_1^\mu + p_{1\mu}p_1^\mu.$$

L'intérêt manifeste de l'invariant sous cette forme est qu'il apparaît comme une contraction. C'est donc un scalaire de Lorentz dont la valeur est la même dans tous les référentiels inertiels. En particulier, exprimée dans le référentiel du centre de masse \mathcal{R} , cette égalité donne de manière explicite $p_a^\mu = (E_a/c, \mathbf{0})$, $p_1^\mu = (E_1/c, \mathbf{p}_1)$ et $p_2^\mu = (E_2/c, \mathbf{p}_2)$, soit

$$m_2^2c^2 = m_a^2c^2 - 2E_aE_1/c^2 + m_1^2c^2,$$

d'où l'on déduit (la seconde équation s'obtient par la symétrie évidente du problème)

$$E_1 = \frac{m_a^2c^2 + m_1^2c^2 - m_2^2c^2}{2m_a},$$

$$E_2 = \frac{m_a^2c^2 - m_1^2c^2 + m_2^2c^2}{2m_a}.$$

On déduit l'impulsion commune par

$$|\mathbf{p}_1| = |\mathbf{p}_2| = \frac{1}{c}\sqrt{E_1^2 - m_1^2c^4}$$

$$= \frac{1}{2m_a}c\sqrt{\lambda(m_a^2, m_1^2, m_2^2)}$$

avec $\lambda(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2yz - 2zx$.

On peut traiter une application à la désintégration $\Lambda^0 \rightarrow p + \pi^-$. On donne $m_{\Lambda^0} = 1.116 \text{ GeV}/c^2$, $m_p = 0.938 \text{ GeV}/c^2$ et $m_{\pi^-} = 0.140 \text{ GeV}/c^2$. Les énergies du proton et du méson π produits valent $E_p = 0.943 \text{ GeV}$ et $E_{\pi^-} = 0.173 \text{ GeV}$.

2.5. Collisions élastiques

On considère une collision élastique entre deux corps de masses m_1 et m_2 . Les particules conservent leur intégrité. L'énergie et l'impulsion sont conservées, soit en termes de quadri-vecteur, $p_1^\mu + p_2^\mu = q_1^\mu + q_2^\mu$ où $p_i^\mu = (E_i/c, \mathbf{p}_i)$ est la quadri-impulsion initiale de la particule i et $q_f^\mu = (\mathcal{E}_f/c, \mathbf{q}_f)$ est celle de la particule f après la collision. Cette loi de conservation est réécrite sous la forme

$$q_2^\mu = p_1^\mu + p_2^\mu - q_1^\mu$$

et exploitée en exprimant la contraction associée,

$$q_{2\mu}q_2^\mu = p_{1\mu}p_1^\mu + p_{2\mu}p_2^\mu + q_{1\mu}q_1^\mu + 2p_{1\mu}p_2^\mu - 2p_{1\mu}q_1^\mu - 2p_{2\mu}q_1^\mu,$$

soit, en utilisant $p_{i\mu}p_i^\mu = q_{i\mu}q_i^\mu = m_i^2c^2$,

$$m_1^2c^2 + p_{1\mu}p_2^\mu - p_{1\mu}q_1^\mu - p_{2\mu}q_1^\mu = 0.$$

De même, l'égalité

$$q_1^\mu = p_1^\mu + p_2^\mu - q_2^\mu$$

fournit une expression analogue déduite par permutation des q_1 et q_2 :

$$m_2^2c^2 + p_{1\mu}p_2^\mu - p_{1\mu}q_2^\mu - p_{2\mu}q_2^\mu = 0.$$

Dans le cas d'une collision sur cible immobile, on a dans le référentiel du laboratoire par exemple $p_2^\mu = (m_2c, \vec{0})$. On déduit des deux expressions précédentes

$$m_1^2c^2 + E_1m_2 - (E_1\mathcal{E}_1/c^2 - |\mathbf{p}_1||\mathbf{q}_1| \cos \theta_1) - m_2\mathcal{E}_1 = 0$$

$$m_2^2c^2 + E_1m_2 - (E_1\mathcal{E}_2/c^2 - |\mathbf{p}_1||\mathbf{q}_2| \cos \theta_2) - m_2\mathcal{E}_2 = 0,$$

(les angles d'émission des particules après collision sont mesurés depuis la trajectoire incidente de la particule 1) soit encore les angles en fonction des énergies,

$$\cos \theta_1 = \frac{\mathcal{E}_1(E_1/c^2 + m_2) - E_1m_2 - m_1^2c^2}{|\mathbf{p}_1||\mathbf{q}_1|}$$

$$\cos \theta_2 = \frac{\mathcal{E}_2(E_1/c^2 + m_2) - E_1m_2 - m_2^2c^2}{|\mathbf{p}_1||\mathbf{q}_2|}.$$

On préfère en général déduire les énergies des produits à partir des valeurs mesurées des angles d'émission. Par exemple pour la particule 2 on commence par exprimer l'angle comme

$$\cos \theta_2 = \frac{(E_1 + m_2c^2)(\mathcal{E}_2 - m_2c^2)}{|\mathbf{p}_1||\mathbf{q}_2|c^2},$$

que l'on élève au carré, puis on utilise pour le dénominateur $|\mathbf{p}_1|^2 c^2 = E_1^2 - m_1^2 c^4$ et $|\mathbf{q}_2|^2 c^2 = \mathcal{E}_2^2 - m_2^2 c^4$, soit

$$\cos^2 \theta_2 = \frac{(E_1 + m_2 c^2)^2 (\mathcal{E}_2 - m_2 c^2)}{(E_1 - m_1 c^2)^2 (\mathcal{E}_2 + m_2 c^2)},$$

et par élimination on en déduit

$$\mathcal{E}_2 = m_2 c^2 \frac{(E_1 + m_2 c^2)^2 + (E_1^2 - m_1^2 c^4) \cos^2 \theta_2}{(E_1 + m_2 c^2)^2 - (E_1^2 - m_1^2 c^4) \cos^2 \theta_2}.$$

Pour la seconde particule il suffit d'appliquer la conservation sous forme traditionnelle,

$$\mathcal{E}_1 = E_1 + E_2 - \mathcal{E}_2.$$

On peut appliquer ce formalisme par exemple à l'effet Compton ⁽²⁹⁾ (diffusion élastique d'un photon de haute énergie sur un électron, supposé libre). Il est préférable de revenir à la première équation

$$m_1^2 c^2 + E_1 m_2 - (E_1 \mathcal{E}_1 / c^2 - |\mathbf{p}_1| |\mathbf{q}_1| \cos \theta_1) - m_2 \mathcal{E}_1 = 0$$

en choisissant le photon comme particule 1, l'électron étant la cible immobile 2. On en déduit

$$|\mathbf{p}_1| m_2 c - |\mathbf{p}_1| |\mathbf{q}_1| + |\mathbf{p}_1| |\mathbf{q}_1| \cos \theta_1 - |\mathbf{q}_1| m_2 c = 0$$

soit

$$|\mathbf{q}_1| = \frac{|\mathbf{p}_1| m_2 c}{|\mathbf{p}_1| (1 - \cos \theta_1) + m_2 c},$$

que l'on écrit plus souvent sous la forme

$$\frac{1}{|\mathbf{q}_1|} - \frac{1}{|\mathbf{p}_1|} = \frac{1}{m_2 c} (1 - \cos \theta_1),$$

avec la relation de de Broglie $|\mathbf{p}_1| = h/\lambda_i$ pour les photons incidents et pour le faisceau diffusé, $|\mathbf{q}_1| = h/\lambda_d$, soit $\lambda_d - \lambda_i = \frac{h}{m_2 c} (1 - \cos \theta_1)$.

⁽²⁹⁾ M.H. Shamos *Great experiments in physics*, Dover New-York 1987, p.348.

2.6. Collisions inélastiques et seuil de réaction

Lors des collisions inélastiques, la nature des particules impliquées est modifiée après l'interaction. La désintégration en est un cas particulier. Il peut s'agir également des mêmes particules, mais dont l'état interne a été modifié, par exemple absorption d'un photon par excitation électronique d'un atome. Cela n'est possible évidemment que pour des particules ayant une structure interne. En général, dans une collision inélastique, une partie de l'énergie des particules incidentes est transformée en énergie de liaison pour produire de nouvelles particules. La notion de seuil en énergie, vis-à-vis de tel état final, est donc primordiale.

Considérons le cas de deux particules 1 et 2 dans l'état initial qui produisent n particules dans l'état final. On note $p_i^\mu = (E_i/c, \mathbf{p}_i)$ et $q_f^\mu = (\mathcal{E}_f/c, \mathbf{q}_f)$ les quadri-impulsions dans l'état initial et dans l'état final respectivement. La conservation de l'énergie-impulsion s'écrit

$$\begin{aligned} P^\mu &= Q^\mu, \\ P^\mu &= \sum_i p_i^\mu = \left(\sum_i E_i/c, \sum_i \mathbf{p}_i \right), \\ Q^\mu &= \sum_f q_f^\mu = \left(\sum_f \mathcal{E}_f/c, \sum_f \mathbf{q}_f \right), \end{aligned}$$

et l'invariant associé est noté

$$c^2 Q_\mu Q^\mu = M^2 c^4 = S^2,$$

où S est appelé énergie seuil de la réaction. En effet, si l'on se place dans le référentiel \mathcal{R}' du centre d'impulsion où $Q'^\mu = \sum_f q_f'^\mu = (\sum_f \mathcal{E}'_f/c, \vec{0})$, on a

$$\sum_f \mathcal{E}'_f = \sqrt{c^2 Q'_\mu Q'^\mu} = S$$

qui représente l'énergie minimale des produits de la collision, car

$$\sum_f \mathcal{E}'_f = \sum_f \sqrt{|\mathbf{q}'_f|^2 c^2 + m_f^2 c^4} \geq \sum_f m_f c^2.$$

En mettant à profit la conservation $P^\mu = Q^\mu$, la réaction envisagée n'est donc permise que si

$$S = \sqrt{c^2 P_\mu P^\mu} \geq \sum_f m_f c^2.$$

Dans une collision traditionnelle, on envoie un faisceau incident, $p_1^\mu = (E_1/c, \mathbf{p}_1)$, sur une cible immobile, $p_2^\mu = (m_2 c, \vec{0})$. Dans ce référentiel on a donc

$$\begin{aligned} P_\mu P^\mu &= (E_1/c + m_2 c)^2 - |\mathbf{p}_1|^2 \\ &= E_1^2/c^2 + 2E_1 m_2 + m_2^2 c^2 - |\mathbf{p}_1|^2 \\ &= m_1^2 c^2 + m_2^2 c^2 + 2E_1 m_2 \end{aligned}$$

et la condition de seuil pour un état final donné s'écrit

$$S = c\sqrt{m_1^2c^2 + m_2^2c^2 + 2E_1m_2} \geq \sum_f m_f c^2.$$

On transforme en général cette expression pour qu'elle apparaisse sous la forme d'une condition sur l'énergie des particules du faisceau incident,

$$E_1 \geq \frac{\left(\sum_f m_f c\right)^2 - m_1^2c^2 - m_2^2c^2}{2m_2}.$$

Dans le cas d'un collisionneur en revanche, les impulsions des deux particules incidentes sont opposées (le référentiel du laboratoire se confond alors avec le référentiel du centre d'impulsion). On a cette fois

$$\begin{aligned} P'_\mu P'^\mu &= (E'_1 + E'_2)^2/c^2 - |\mathbf{p}'_1 + \mathbf{p}'_2|^2 \\ &= [(E'_1 + E'_2)/c]^2 \end{aligned}$$

et la réaction est permise si

$$S = E'_1 + E'_2 \geq \sum_f m_f c^2.$$

Le collisionneur est beaucoup plus efficace pour accélérer les particules à haute énergie. On peut donner une illustration dans le cas des expériences de Carlo Rubbia (production des bosons d'échange de l'interaction faible, W^\pm et Z^0 , au CERN). Ces expériences furent réalisées à l'aide de deux faisceaux d'antiprotons de 270 GeV, ce qui conduit à $E'_1 + E'_2 = 540$ GeV. Pour atteindre la même efficacité dans une expérience sur cible immobile, il aurait fallu fournir au faisceau incident l'énergie E_1 telle que

$$\sqrt{m_1^2c^4 + m_2^2c^4 + 2E_1m_2c^2} = 540 \text{ GeV},$$

soit, avec $mc^2 \simeq 0.938$ GeV, $E_1 \simeq 160\,000$ GeV !

Lorsque les conditions d'énergie supérieure ou égale au seuil sont satisfaites, on peut ensuite déterminer les énergies et impulsions des produits de la réaction. Considérons par exemple la collision inélastique suivante :

$$1 + 2 \rightarrow 3 + 4,$$

étudiée dans le référentiel \mathcal{R}' du centre d'impulsion. On a

$$\begin{aligned} P'^\mu &= ((E'_1 + E'_2)/c, \vec{0}), \\ Q'^\mu &= ((\mathcal{E}'_3 + \mathcal{E}'_4)/c, \mathbf{q}'_3 + \mathbf{q}'_4) \end{aligned}$$

avec $Q'_\mu Q'^\mu = S^2/c^2$ et $P'_\mu P'^\mu = Q'_\mu Q'^\mu$ soit $E'_1 + E'_2 = S$. Pour déterminer \mathcal{E}'_3 , on procède comme dans le cas d'une désintégration en formant $q'^\mu_4 = P'^\mu - q'^\mu_3$ dont l'invariant donne

$$\begin{aligned} q'^\mu_{4\mu} q'^\mu_4 &= m_4^2 c^2 \\ &= (E'_1 + E'_2)^2/c^2 + m_3^2 c^2 - 2(E'_1 + E'_2)/c \times \mathcal{E}'_3/c \\ &= S^2/c^2 + m_3^2 c^2 - 2S\mathcal{E}'_3/c^2. \end{aligned}$$

On en déduit les expressions cherchées (avec les permutations d'indices nécessaires)

$$\begin{aligned} \mathcal{E}'_3 &= \frac{S^2/c^2 + m_3^2 c^2 - m_4^2 c^2}{2S/c^2} \\ \mathcal{E}'_4 &= \frac{S^2/c^2 - m_3^2 c^2 + m_4^2 c^2}{2S/c^2} \\ E'_1 &= \frac{S^2/c^2 + m_1^2 c^2 - m_2^2 c^2}{2S/c^2} \\ E'_2 &= \frac{S^2/c^2 - m_1^2 c^2 + m_2^2 c^2}{2S/c^2} \end{aligned}$$

et les impulsions s'en déduisent immédiatement,

$$\begin{aligned} |\mathbf{p}'_1| = |\mathbf{p}'_2| &= \lambda^{1/2}(S^2/c^4, m_1^2, m_2^2)c^3/2S \\ |\mathbf{q}'_3| = |\mathbf{q}'_4| &= \lambda^{1/2}(S^2/c^4, m_3^2, m_4^2)c^3/2S \end{aligned}$$

avec la même fonction λ que dans le cas des désintégrations.

3. Forces conservatives et équations du mouvement

3.1. Notion de force

3.1.1. Expression de la force et de la relation fondamentale de la dynamique

Pour que les notions de force conservative et d'énergie potentielle conservent un sens en relativité, il faut que l'expression $E_{\text{tot}} = E + U(\mathbf{r})$ soit toujours valable, où E comprend l'énergie de masse et l'énergie cinétique, $E = mc^2 + K$ ou encore $E^2 = |\mathbf{p}|^2 c^2 + m^2 c^4$. On en déduit

$$2E \frac{dE}{dt} = 2c^2 \mathbf{p} \frac{d\mathbf{p}}{dt}$$

soit, reporté dans

$$\begin{aligned}\frac{dE_{\text{tot}}}{dt} &= \frac{dE}{dt} + \frac{dU(\mathbf{r})}{dt} \\ &= \frac{c^2}{E} \mathbf{p} \frac{d\mathbf{p}}{dt} + \vec{\nabla}U(\mathbf{r}) \frac{d\mathbf{r}}{dt}.\end{aligned}$$

On a utilisé le fait que l'énergie potentielle est indépendante du temps, $dU(\mathbf{r}, t) = \frac{\partial U(\mathbf{r}, t)}{\partial t} dt + \vec{\nabla}U(\mathbf{r}, t) d\mathbf{r}$ se simplifie donc en ne laissant que la partie convective. On peut alors poursuivre le calcul,

$$\begin{aligned}\frac{dE_{\text{tot}}}{dt} &= \frac{c^2}{\gamma_{\mathbf{v}} mc^2} \gamma_{\mathbf{v}} m \mathbf{v} \frac{d\mathbf{p}}{dt} + \vec{\nabla}U(\mathbf{r}) \mathbf{v} \\ &= \mathbf{v} \left(\frac{d\mathbf{p}}{dt} - \mathbf{F} \right).\end{aligned}$$

La contrainte de conservation de l'énergie conduit alors à la définition de la force en relativité qui est une modification simple de la loi dynamique de Newton,

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt} \frac{m\mathbf{v}}{\sqrt{1 - |\mathbf{v}|^2/c^2}},$$

où \mathbf{v} est la vitesse ordinaire de la particule.

En effectuant le calcul de la dérivée, il vient

$$\mathbf{F} = \frac{m\mathbf{a}}{\sqrt{1 - |\mathbf{v}|^2/c^2}} + \frac{m\mathbf{v}(\mathbf{a}\mathbf{v})/c^2}{(1 - |\mathbf{v}|^2/c^2)^{3/2}}$$

qui redonne bien entendu l'expression newtonienne dans la limite $c \rightarrow \infty$.

3.1.2. Quadri-force

On peut rassembler en un seul concept la puissance et la force, on définit pour cela

$$\Phi^\mu = \frac{dp^\mu}{d\tau} = \gamma_{\mathbf{v}} \frac{dp^\mu}{dt}.$$

En notant par ailleurs que $p^\mu = (\gamma_{\mathbf{v}} mc, \gamma_{\mathbf{v}} m\mathbf{v}) = (E/c, \mathbf{p})$, il vient pour la composante temporelle

$$\Phi^0 = \gamma_{\mathbf{v}} \frac{d}{dt} (\gamma_{\mathbf{v}} mc) = \gamma_{\mathbf{v}} \frac{d}{dt} \left(\frac{E}{c} \right) = \frac{1}{c} \gamma_{\mathbf{v}} \frac{dE}{dt}$$

ce qui signifie que $\frac{dE}{dt} = c\gamma_{\mathbf{v}}^{-1}\Phi^0 = c\sqrt{1 - \frac{|\mathbf{v}|^2}{c^2}}\Phi^0$ représente la puissance fournie par la force $\mathbf{F} = \frac{d\mathbf{p}}{dt}$ dans le référentiel de l'observateur.

Pour les composantes spatiales on a de même directement

$$\Phi^i = \gamma_{\mathbf{v}} \frac{d\mathbf{p}}{dt} = \frac{1}{\sqrt{1 - \frac{|\mathbf{v}|^2}{c^2}}} \mathbf{F},$$

ainsi que les lois de transformation par changement de référentiel inertiel.

3.2. Mouvement d'une charge dans un champ magnétique uniforme

On considère le mouvement d'une charge q dans un champ magnétique constant et uniforme $\mathbf{B} = B\hat{\mathbf{u}}_z$. En anticipant sur la chapitre suivant, on admet que la force de Lorentz conserve la forme habituelle

$$\mathbf{F} = q\mathbf{v} \wedge \mathbf{B}.$$

On a alors

$$\frac{d\mathbf{p}}{dt} = q\mathbf{v} \wedge \mathbf{B}$$

qui impose $\frac{d\mathbf{p}}{dt} \perp \mathbf{v}$, soit $\frac{d\mathbf{p}}{dt} \perp \mathbf{p}$ et par conséquent $d(\mathbf{p}^2) = 2\mathbf{p}d\mathbf{p} = 0$. La norme $|\mathbf{p}|$ de \mathbf{p} est donc conservée. Cela entraîne que $\mathbf{v}/\sqrt{1 - \frac{|\mathbf{v}|^2}{c^2}}$ est également conservé, donc

$$\frac{|\mathbf{v}|^2}{1 - |\mathbf{v}|^2/c^2} = \text{const},$$

$$|\mathbf{v}|^2 = \text{const} \times (1 - |\mathbf{v}|^2/c^2),$$

$$|\mathbf{v}|^2(1 + \text{const}/c^2) = \text{const},$$

$$|\mathbf{v}|^2 = \text{const},$$

donc finalement $d(\mathbf{v}^2) = 2\mathbf{v}d\mathbf{v} = 0$, ce qui implique comme dans le cas non relativiste que $\mathbf{a} = \frac{d\mathbf{v}}{dt} \perp \mathbf{v}$ et que la trajectoire de la particule soit hélicoïdale.

On peut maintenant s'intéresser à l'équation de la trajectoire qui se simplifie grâce à l'observation ci-dessus,

$$\gamma_{\mathbf{v}} m \mathbf{a} = q\mathbf{v} \wedge \mathbf{B}$$

que l'on peut projeter,

$$m\gamma_{\mathbf{v}} \begin{pmatrix} \frac{dv_x}{dt} \\ \frac{dv_y}{dt} \\ \frac{dv_z}{dt} \end{pmatrix} = q \begin{vmatrix} \hat{\mathbf{u}}_x & v_x & 0 \\ \hat{\mathbf{u}}_y & v_y & 0 \\ \hat{\mathbf{u}}_z & v_z & B \end{vmatrix} = qB \begin{pmatrix} v_y \\ -v_x \\ 0 \end{pmatrix}.$$

En développant, on obtient

$$\begin{cases} \ddot{v}_x + \left(\frac{qB}{\gamma_{\mathbf{v}}m}\right)^2 v_x = 0 \\ \ddot{v}_y + \left(\frac{qB}{\gamma_{\mathbf{v}}m}\right)^2 v_y = 0 \\ \frac{dv_z}{dt} = 0 \end{cases}$$

ce qui donne finalement l'équation d'un cercle dans la plan perpendiculaire au champ, mais avec une pulsation cyclotron modifiée par rapport à la situation non relativiste,

$$\omega_c = \frac{qB}{m\gamma_{\mathbf{v}}} = \omega_c^{\text{class.}} \sqrt{1 - |\mathbf{v}|^2/c^2}.$$

C'est en raison de cette dépendance en vitesse de la pulsation que le principe simplifié du cyclotron ne fonctionne plus lorsque la vitesse devient trop élevée. Ce phénomène est relié au problème de la contraction des longueurs pour l'observateur immobile. Dans le référentiel propre de la charge en mouvement circulaire, la longueur propre dans la direction tangentielle au mouvement, $d\lambda$, est plus grande que la longueur dl mesurée par un observateur immobile, $d\lambda = \gamma_{\mathbf{v}} dl$, de sorte que le périmètre d'un cercle de rayon R , $\rho = 2\pi R / \sqrt{1 - |\mathbf{v}|^2/c^2}$ est vu plus grand dans le référentiel propre que la valeur $2\pi R$ mesurée par FIX (bien entendu, la longueur transverse, R , est invariable). On conserve toutefois une relation analogue à la relation classique pour le rayon de la trajectoire,

$$R = \frac{|\mathbf{v}|}{\omega_c} = \frac{\gamma_{\mathbf{v}} m |\mathbf{v}|}{qB} = \frac{|\mathbf{p}|}{qB}.$$

La mesure de R donne accès à \mathbf{p} et par suite, à E , \mathbf{v} ,... validant les mesures de type chambres à bulles.

3.3. Problème de Kepler

Il est instructif, avant de considérer le problème de Kepler en relativité restreinte, de repréciser en quoi ce problème a joué un rôle central tout d'abord en dynamique newtonienne, puis dans les observations astronomiques, notamment dans le cas de la planète Mercure. Nous allons tout d'abord reprendre le problème initial puis ses perturbations en dynamique classique, avant d'aborder la solution relativiste, ce qui mettra en évidence l'incapacité de la théorie à ce niveau à rendre compte quantitativement des observations.

3.3.1. Les états liés du potentiel newtonien

Lorsque l'énergie totale E est négative, on sait que le mouvement est borné (cercle ou ellipse). Rappelons sommairement la solution du problème de Kepler en dynamique newtonienne. Le moment cinétique donne $\dot{r} = \frac{|\boldsymbol{\sigma}|}{mr^2} \frac{dr}{d\phi}$ que l'on utilise dans l'expression de la conservation de l'énergie pour obtenir

$$\frac{1}{r^2} \left(\frac{dr}{d\phi} \right)^2 + 1 = \frac{2mr^2}{|\boldsymbol{\sigma}|^2} \left(E + \frac{GM_{\odot}m}{r} \right).$$

Le changement de variable $u = \frac{1}{r}$ permet de réécrire cette équation sous la forme

$$\left(\frac{du}{d\phi} \right)^2 + u^2 = \frac{2m}{|\boldsymbol{\sigma}|^2} (E + GM_{\odot}mu).$$

En dérivant par rapport à ϕ , on obtient l'équation du mouvement

$$\frac{d^2u}{d\phi^2} + u = \frac{GM_{\odot}m^2}{|\boldsymbol{\sigma}|^2}. \quad (4.11)$$

Une solution particulière de l'équation complète est donnée par $u_0 = \frac{GM_{\odot}m^2}{|\boldsymbol{\sigma}|^2}$ et la solution générale de l'équation sans second membre est de la forme $u_1(\phi) =$

$A \cos(\phi - \psi)$. En faisant le choix de conditions initiales telles que le facteur de phase ψ soit nul, on obtient la solution complète

$$u(\phi) = \frac{GM_{\odot} m^2}{|\boldsymbol{\sigma}|^2} (1 + e \cos \phi) \quad (4.12)$$

où l'excentricité de l'ellipse est identifiée en réinjectant cette solution dans l'équation du mouvement :

$$\frac{1}{r} = \frac{1}{p} (1 + e \cos \phi), \quad \frac{1}{p} = \frac{GM_{\odot} m^2}{|\boldsymbol{\sigma}|^2}, \quad e = \sqrt{1 + \frac{2E|\boldsymbol{\sigma}|^2}{G^2 M_{\odot}^2 m^3}}. \quad (4.13)$$

Le mouvement est bien indépendant de la masse m de la planète (masse test) car $E \sim m$, $|\boldsymbol{\sigma}| \sim m$, donc e et p ne dépendent pas de m .

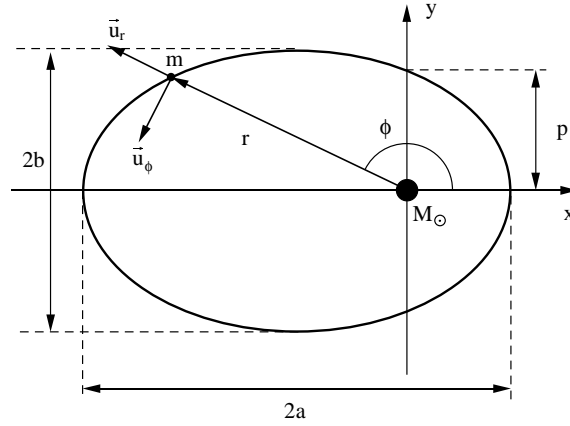


Figure 4.1 Trajectoire elliptique d'une planète de masse m dans le champ de gravitation du Soleil de masse M_{\odot} . Si le champ est rigoureusement newtonien, la trajectoire est une ellipse fermée.

Les caractéristiques géométriques de l'ellipse (demi grand-axe et demi petit-axe) valent :

$$a = \frac{p}{1 - e^2} \quad \text{et} \quad b = \frac{p}{\sqrt{1 - e^2}} \quad (4.14)$$

et la période T du mouvement se déduit simplement de la loi des aires :

$$|\boldsymbol{\sigma}| = mr^2 \dot{\phi} = 2m \times \frac{1}{2} r^2 \frac{d\phi}{dt} = 2m \frac{dS}{dt}$$

qui donne, en intégrant sur une période :

$$|\boldsymbol{\sigma}|T = 2m\pi ab, \quad T = \frac{2\pi mp^2}{(1 - e^2)^{3/2}}. \quad (4.15)$$

3.3.2. Ecart au comportement newtonien, déplacement du périhélie des planètes

On considère ici la trajectoire d'une masse m liée dans le champ de gravitation du Soleil. On suppose qu'au champ purement newtonien $-\frac{GM_\odot m}{r}$ s'ajoute une perturbation centrale $\delta U(r)$. Celle-ci modifie, comme on va le voir, la nature de la trajectoire, mais le mouvement reste plan. Cette fois, $U(r) = -\frac{GM_\odot m}{r} + \delta U(r)$, d'où l'on déduit de nouveau

$$\frac{d\phi}{dr} = \frac{|\boldsymbol{\sigma}|}{mr^2} \left(\frac{2}{m} \left(E - U(r) - \frac{|\boldsymbol{\sigma}|^2}{2mr^2} \right) \right)^{-1/2} \quad (4.16)$$

que l'on peut réexprimer sous une forme plus pratique ⁽³⁰⁾ pour les calculs

$$\frac{d\phi}{dr} = -m \frac{\partial}{\partial |\boldsymbol{\sigma}|} \left(\frac{2}{m} \left(E - U(r) - \frac{|\boldsymbol{\sigma}|^2}{2mr^2} \right) \right)^{1/2}.$$

Lorsque r varie de r_{\min} à r_{\max} puis à r_{\min} , le rayon vecteur a tourné de $\Delta\phi_0$ tel que

$$\Delta\phi_0 = 2 \int_{r_{\min}}^{r_{\max}} \frac{d\phi}{dr} dr = -2m \frac{\partial}{\partial |\boldsymbol{\sigma}|} \int_{r_{\min}}^{r_{\max}} \left(\frac{2}{m} \left(E - U(r) - \frac{|\boldsymbol{\sigma}|^2}{2mr^2} \right) \right)^{1/2} dr.$$

Si $U(r)$ est purement newtonien, $\Delta\phi_0 = 2\pi$ car le mouvement est parfaitement elliptique. Si $U(r) = -\frac{GM_\odot m}{r} + \delta U(r)$, on a alors une *rosette de Sommerfeld* et le périhélie se déplace de $\Delta\phi$ (à 2π près). Pour calculer ce déplacement en présence de perturbations centrales, on développe l'intégrand I au premier ordre en $\delta U(r)$ ⁽³¹⁾ :

$$I = \sqrt{\frac{2}{m} \left(E + \frac{GM_\odot m}{r} - \frac{|\boldsymbol{\sigma}|^2}{2mr^2} \right)} - \frac{1}{2} \frac{2}{m} \delta U(r) \left[\frac{2}{m} \left(E + \frac{GM_\odot m}{r} - \frac{|\boldsymbol{\sigma}|^2}{2mr^2} \right) \right]^{-1/2} + \mathcal{O}(\delta U^2(r)).$$

Le terme d'ordre 0 donne une contribution 2π (problème newtonien non perturbé) et le terme d'ordre 1 s'écrit simplement

$$\Delta\phi_0 \simeq 2m \frac{\partial}{\partial |\boldsymbol{\sigma}|} \frac{1}{m} \int_0^\pi \frac{mr^2}{|\boldsymbol{\sigma}|} \delta U(r) d\phi$$

en faisant usage de la solution au voisinage de la solution non perturbée.

On obtient finalement le déplacement du périhélie :

$$\Delta\phi_0 = \frac{\partial}{\partial |\boldsymbol{\sigma}|} \left(\frac{2m}{|\boldsymbol{\sigma}|} \int_0^\pi r^2 \delta U(r) d\phi \right). \quad (4.17)$$

Si la perturbation est de la forme $\delta U(r) = \frac{\beta}{r^2}$, on obtient $\Delta\phi_0 = -\frac{2\pi\beta}{GM_\odot mp}$, et si elle est de la forme $\delta U(r) = \frac{\gamma}{r^3}$, on trouve

$$\Delta\phi_0 = \frac{\partial}{\partial |\boldsymbol{\sigma}|} \frac{2m}{|\boldsymbol{\sigma}|} \int_0^\pi \frac{\gamma}{p} (1 + e \cos \phi) d\phi = -\frac{6\pi\gamma}{GM_\odot mp^2}$$

au voisinage de la solution non perturbée.

⁽³⁰⁾ On remarque que $-m \frac{\partial}{\partial \sigma} \left(-\frac{2}{m} \frac{\sigma^2}{2mr^2} \right)^{+1/2} = \frac{\sigma}{2mr^2}$.

⁽³¹⁾ L.D. Landau et E. Lifshitz, Mécanique (Editions Mir, Moscou 1981), p. 74

3.3.3. Les observations expérimentales dans le cas de Mercure

Dès 1850, Urbain Jean Joseph Le Verrier (1811-1877) montrait sur la base d'observations astronomiques que la théorie de Newton n'était pas totalement satisfaisante. Le mouvement de la planète Mercure en particulier présentait une anomalie inexplicée. Si sa trajectoire était bien plane, elle s'effectuait le long d'une ellipse tournant lentement dans son plan. En tenant compte de la précession générale des équinoxes ainsi que de l'influence des autres planètes sur le mouvement de Mercure, Le Verrier estimait qu'il restait une avance résiduelle du périhélie de Mercure de 38'' d'arc par siècle. Pour expliquer cette anomalie, diverses hypothèses (parfois osées !) ont été avancées⁽³²⁾ :

Pour donner à moindre frais une explication des désaccords, les astronomes se sont efforcés de trouver une interprétation qui laisse inchangée la loi de Newton. Les moyens les plus simples consistaient à supposer l'existence d'un corps passé, jusque-là, inaperçu ou d'une propriété encore insoupçonnée. (...) C'est ainsi qu'on imagine successivement la présence d'une petite planète intramercurelle, puis d'un anneau de petites planètes. On suppose encore la possibilité d'un aplatissement de la couronne solaire et enfin la présence de matière diffuse, en forme de lentille biconvexe, qui s'étendrait, avec une densité décroissante, au-delà de l'orbite de la terre.

Ces hypothèses ont toutes été infirmées par l'expérience ou rejetées (dans le cas de la lentille de matière) parce que totalement arbitraires. L'aplatissement éventuel du Soleil en revanche conduit à un moment quadropolaire J_2 difficile à évaluer. Néanmoins, les observations récentes de la forme du Soleil montrent que cet effet contribue très peu à l'écart au résultat newtonien⁽³³⁾. On en est alors arrivé à remettre en cause la loi d'attraction universelle de Newton⁽³⁴⁾ :

Si l'agrément avec l'expérience vient à faire défaut, on peut essayer, en renonçant à la simplicité, de modifier très légèrement cette loi. (...) En supposant que la loi d'attraction est représentée par un monôme de la forme

$$F(r) = -K \frac{MM'}{r^n},$$

Hall avait montré qu'il existe une avance des périhélies si $n > 2$, et que la valeur $n = 2.000\ 000\ 16$ suffit pour énoncer une prédiction correcte de l'anomalie excentrique de Mercure.

De nouveau, une telle modification n'apporte rien, car si elle résoud le problème dans le cas de Mercure, elle est mise en défaut pour les autres planètes. Il fallait donc conclure à une discordance entre théorie et expérience. Ce désaccord ne devait être expliqué qu'en 1915 par un calcul en perturbation en Relativité Générale réalisé par Einstein.

L'avance observée du périhélie dans le cas de Mercure est de l'ordre de 5600'' d'arc par siècle. Des deux effets invoqués par Le Verrier, celui de la précession générale contribue pour environ 5000'', les perturbations cumulées dues aux autres planètes apportent une correction supplémentaire de l'ordre de 10% et l'avance résiduelle inexplicée est finalement inférieure à 1% de la valeur observée.

- **La précession générale des équinoxes**

La contribution la plus importante à l'avance du périhélie de Mercure, observée depuis la Terre, est due au fait que les observations ne sont pas faites dans un référentiel immobile par rapport au Soleil. La Terre n'est pas tout à fait sphérique,

⁽³²⁾ M.A. Tonnelat, Histoire du Principe de Relativité (Flammarion, Paris 1971), pp. 336-339

⁽³³⁾ Voir les résultats présentés dans le tableau récapitulatif des différentes contributions.

⁽³⁴⁾ M.A. Tonnelat, op. cit., p. 338.

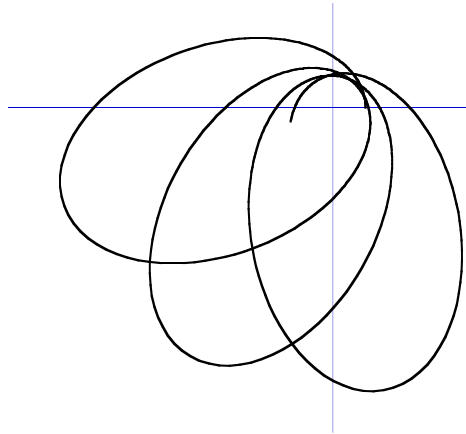


Figure 4.2 Avance du périhélie dans le cas d'une trajectoire elliptique. Sur cet exemple, l'avance est de l'ordre de 35 degrés par révolution.

elle présente un léger aplatissement aux pôles et un renflement au niveau de l'équateur. De l'action combinée du Soleil et de la Lune (l'effet de la Lune est dominant), il résulte un couple de torsion non constant, dépendant des positions relatives des trois astres, dont l'effet est de provoquer la précession de l'axe de rotation diurne de la Terre avec une période de l'ordre de 26 000 ans. Ce mouvement de précession a été observé pour la première fois par Hipparque en 125 BC, et a été expliqué par Newton. L'étoile polaire n'indique par conséquent pas tout à fait la direction du Nord. A l'époque d'Hipparque, elle était dans une direction inclinée à 12° du pôle, elle est aujourd'hui à 1° et sera à 46° dans 13 000 ans. Un simple calcul d'ordre de grandeur montre qu'il en résulte une avance apparente dans les observations astronomiques de $\frac{360 \times 3600}{260} \simeq 5000''$ d'arc par siècle.

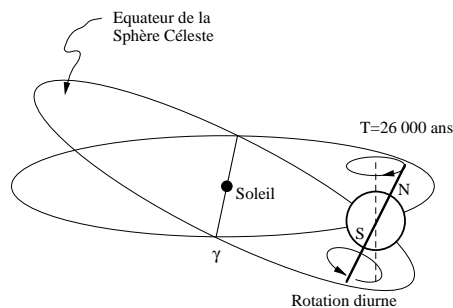


Figure 4.3 La précession générale des équinoxes (point γ). La direction des équinoxes est donnée par l'intersection du plan de l'écliptique avec le plan équatorial de la sphère céleste (parallèle à l'équateur terrestre). Comme l'axe de rotation diurne de la Terre précesse, sous les influences combinées du Soleil et de la Lune, avec une période d'environ 26 000 ans, le plan équatorial n'a pas une orientation constante.

Planète	M	r_{\max}	r_{\min}	\bar{r}	T	e	α
Mercure	$3.303 \cdot 10^{23}$	69.7	45.9	57.9	0.240 85	0.205 6	7.00°
Vénus	$48.70 \cdot 10^{23}$	109	107.4	108.2	0.615 21	0.006 8	3.39°
Terre	$59.75 \cdot 10^{23}$	152.1	147.1	149.6	1.000 04	0.016 7	–
Mars	$6.418 \cdot 10^{23}$	249.1	206.7	227.9	1.880 89	0.093 4	1.85°
Jupiter	$189.9 \cdot 10^{25}$	815.7	740.9	778.3	11.862 23	0.048 5	1.30°
Saturne	$56.86 \cdot 10^{25}$	1 507	1 347	1 427.0	29.455 77	0.055 6	2.49°
Uranus	$8.66 \cdot 10^{25}$	3 004	2 735	2 869.6	84.013 9	0.047 2	0.77°
Neptune	$10.30 \cdot 10^{25}$	4 537	4 456	4 496.6	164.793	0.008 6	1.77°
Pluton	$\simeq 1 \cdot 10^{22}$	7 375	4 425	5 900.0	247.7	0.250	17.2°

Tableau 4.1 Données numériques relatives aux planètes du système solaire. Les masses (M) sont indiquées en kg, les distances (aphélie, périhélie et rayon moyen de l'orbite) en millions de km, la période de révolution (T) en années tropiques et l'inclinaison de l'orbite (α) est repérée par rapport au plan de l'écliptique, e est l'excentricité.

• Les perturbations dues à la présence des autres planètes

Le tableau 4.1 donne les valeurs numériques⁽³⁵⁾ relatives aux orbites et masses des planètes du système solaire. Ces données permettront de déterminer l'influence des planètes sur le mouvement de Mercure.

L'effet, sur le mouvement de Mercure, d'une planète externe (de masse M_P) est évalué en assimilant la planète en moyenne à un anneau de matière (de rayon R_P , le rayon moyen de l'orbite de la planète P). Le rayon de la planète étant négligeable devant celui de son orbite, on remplace l'anneau par un fil circulaire de densité de masse $\lambda = M_P/2\pi R_P$ ⁽³⁶⁾. Les plans des orbites des différentes planètes sont pratiquement confondus, il est alors évident que, par symétrie, cette distribution de masse conduit, en tout point situé à l'intérieur et dans le plan de l'orbite de la planète P , à une force de gravitation centrale (par rapport au Soleil) et dirigée vers l'extérieur.

Etant données deux positions diamétralement opposées P_1 et P_2 de la planète P sur son orbite circulaire, soit S le Soleil et M la position de Mercure, la contribution du champ de gravitation $d\mathbf{G}$, projetée sur l'axe SM s'écrit :

$$|d\mathbf{G}| = G\lambda R_p d\theta \left(\frac{\cos \phi_1}{r_1^2} - \frac{\cos \phi_2}{r_2^2} \right).$$

Dans l'approximation où $R_{\text{Mercure}} \ll R_p$ ⁽³⁷⁾, on peut considérer que $\phi_1 \simeq \phi_2 \simeq \theta$ et remplacer $dm = \lambda R_p d\theta$ par $\lambda r_1 d\phi_1$ ou $\lambda r_2 d\phi_2$. On obtient ainsi simplement

$$|d\mathbf{G}| = G\lambda \cos \theta \left(\frac{1}{r_1} - \frac{1}{r_2} \right) d\theta.$$

⁽³⁵⁾ Atlas Universalis de l'Astronomie, pp. 50-52

⁽³⁶⁾ M.P. Price and W.F. Rush, Nonrelativist contribution to Mercury's precession, *Am. J. Phys.* **47** (1979) 531

⁽³⁷⁾ Hypothèse peu satisfaisante, par exemple pour Vénus ou la Terre le rapport R_{Mercure}/R_P est de l'ordre de 0.5 et 0.4. Pour les planètes plus extérieures, l'approximation est meilleure. Celle-ci conduira néanmoins à des résultats tout à fait acceptables, même dans le cas des planètes internes.

Avec

$$\begin{aligned} r_1 &= -r \cos \theta + \sqrt{r^2 \cos^2 \theta - (r^2 - R_P^2)} \\ r_2 &= +r \cos \theta + \sqrt{r^2 \cos^2 \theta - (r^2 - R_P^2)} \end{aligned}$$

on déduit finalement la contribution au champ de gravitation de la planète P :

$$\mathbf{G}(r) = + \frac{GM_P}{2R_P} \frac{r}{R_P^2 - r^2} \mathbf{u}_r. \quad (4.18)$$

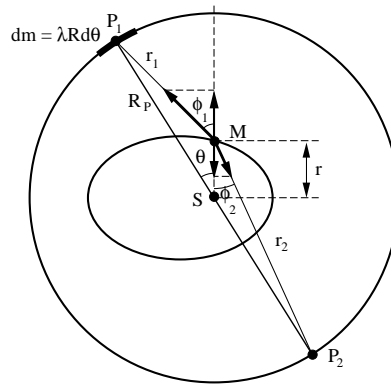


Figure 4.4 Influence d'une planète extérieure, assimilée à un anneau de matière, sur le mouvement de Mercure. Les orbites des différentes planètes du système solaire (à l'exception de Pluton, dont l'influence est négligeable), sont à peu près coplanaires.

L'influence des planètes externes sur le mouvement de Mercure s'évalue alors en perturbations de manière analogue au calcul précédent. En présence d'une force centrale perturbatrice par rapport à l'attraction newtonienne :

$$F(r) = - \frac{GM_{\odot} m}{r^2} (1 + g(r)),$$

le déplacement du périhélie peut s'écrire sous la forme⁽³⁸⁾ :

$$\Delta \phi_0 \simeq \frac{2}{e} \int_0^{\pi} g(r) \cos \phi \, d\phi. \quad (4.19)$$

En remplaçant $r(\phi)$ par la solution non perturbée, on arrive à une intégrale délicate. On utilise alors une approximation discrète sous la forme :

$$\Delta \phi_0 \simeq \frac{\pi}{Ne} \sum_{i=1}^{2N} g(r_i) \cos \phi_i, \quad \phi_i = \left(i - \frac{1}{2} \right) \pi / 2N$$

⁽³⁸⁾ B. Davies, Elementary theory of perihelion precession, *Am. J. Phys.* **51** (1983) 909

pour évaluer numériquement l'intégrale. La correction relative à la force de gravitation due à la planète P s'écrit :

$$g_P(r) = -\frac{M_P}{2M_\odot} \frac{(r/R_P)^3}{1 - (r/R_P)^2}, \quad (4.20)$$

elle est représentée sur la figure (1.6). L'allure de la courbe est suffisamment régulière pour que l'on puisse espérer une convergence numérique rapide de l'intégrale discrète. Les résultats sont rassemblés dans le tableau (4.2) (la masse du Soleil est de $M_\odot = 1.98 \cdot 10^{30}$ kg) :

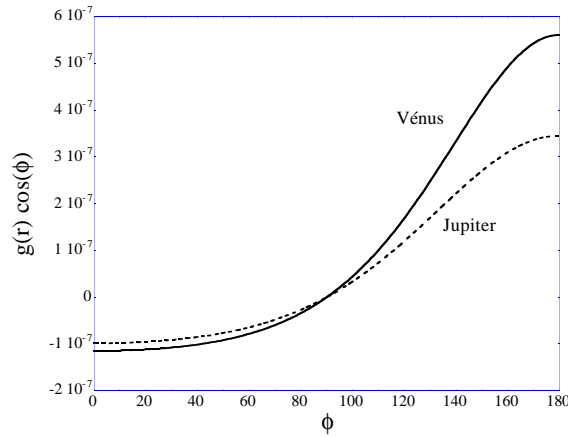


Figure 4.5 Correction à la force de gravitation newtonienne. Les perturbations les plus importantes dues à Vénus et à Jupiter sur le mouvement de Mercure sont représentées.

Planète	M_P/M_\odot	p/R_P	$N = 1$	$N = 2$	$N \geq 3$
Vénus	$2.460 \cdot 10^{-6}$	0.512	254.09	273.25	273.30
Terre	$3.018 \cdot 10^{-6}$	0.370	86.99	91.48	91.49
Mars	$3.241 \cdot 10^{-7}$	0.243	2.25	2.35	2.35
Jupiter	$9.591 \cdot 10^{-4}$	$7.114 \cdot 10^{-2}$	151.02	156.75	156.75
Saturne	$2.872 \cdot 10^{-4}$	$3.880 \cdot 10^{-2}$	7.29	7.56	7.56
Uranus	$4.374 \cdot 10^{-5}$	$1.929 \cdot 10^{-2}$	0.14	0.14	0.14
Neptune	$5.202 \cdot 10^{-5}$	$1.231 \cdot 10^{-2}$	0.04	0.04	0.04
Pluton	$\simeq 5.051 \cdot 10^{-9}$	$0.938 \cdot 10^{-2}$	0.00	0.00	0.00

Tableau 4.2 Influence des planètes externes sur le mouvement de Mercure. Le tableau donne les contributions, planète par planète, à l'avance du périhélie de Mercure. Les résultats, exprimés en secondes d'arc par siècle, convergent très rapidement (dès $N = 3$ pour avoir deux décimales).

Cause	Avance du périhélie de Mercure
Mercure (???)	0.025 ± 0.00
Vénus	277.856 ± 0.68
Terre	90.038 ± 0.08
Mars	2.536 ± 0.00
Jupiter	153.584 ± 0.00
Saturne	7.302 ± 0.01
Uranus	0.141 ± 0.00
Neptune	0.042 ± 0.00
Applatissage du Soleil	0.010 ± 0.02
Précession générale	5025.645 ± 0.50
Somme	5557.18 ± 0.85
Avance observée	5599.74 ± 0.41
Différence	42.56 ± 0.94

Tableau 4.3 Contribution des planètes externes et de l'applatissage du Soleil à l'avance du périhélie de Mercure. Les résultats sont tirés de l'article de Clemence de 1947, qui présente aussi les résultats concernant l'avance du périhélie de la Terre.

La somme des contributions des différentes planètes obtenue par ce calcul approché est de $531.63''$ d'arc par siècle, valeur extrêmement proche du résultat de Clemence⁽³⁹⁾ de $531.54''$ (tableau 4.3) considéré comme le travail de référence dans ce domaine⁽⁴⁰⁾.

3.3.4. La rosette de Sommerfeld et l'avance du périhélie de Mercure en dynamique relativiste

On se propose ici de traiter le problème de Kepler en relativité restreinte. Le calcul n'est pas exact, puisque l'on considère une interaction à distance instantanée. On a

$$\mathbf{p} = \gamma m \mathbf{v}, \quad \gamma = \frac{1}{\sqrt{1 - \frac{|\mathbf{v}|^2}{c^2}}}$$

et l'équation du mouvement pour une force centrale prend la forme

$$\frac{d\mathbf{p}}{dt} = \mathbf{F}(r) = -\frac{dU(r)}{dr} \mathbf{u}_r,$$

ce qui impose, comme dans le cas non relativiste, un mouvement plan. En appliquant $\mathbf{r} \wedge \dots$ à l'équation ci-dessus, on montre en effet que le moment cinétique $\boldsymbol{\sigma} = \mathbf{r} \wedge \mathbf{p}$ reste constant. La norme $|\boldsymbol{\sigma}|$ constitue une première constante du mouvement :

$$|\boldsymbol{\sigma}| = \gamma m r^2 \dot{\phi}. \quad (4.21)$$

⁽³⁹⁾ G.M. Clemence, The Relativity Effect in Planetary Motions, *Rev. Mod. Phys.* **19** (1947) 361, G.M. Clemence, Relativity Effect in Planetary Motions, *Proc. Am. Phil. Soc.* **93** (1949) 532

⁽⁴⁰⁾ C.W. Misner, K.S. Thorne and J.A. Wheeler, *Gravitation* (Freeman, San Francisco 1973), pp. 1112-1113

Une seconde constante du mouvement est fournie par l'énergie totale E . Rappelons qu'en l'absence d'énergie potentielle, l'énergie s'écrit $E = \gamma mc^2 = mc^2 + K$, ce qui permet d'exprimer l'énergie cinétique sous la forme⁽⁴¹⁾ $K = mc^2(\gamma - 1)$. En présence d'énergie potentielle on a donc

$$E = mc^2(\gamma - 1) + U(r). \quad (4.22)$$

Cette équation donne

$$\gamma^2 = \left(1 + \frac{E - U(r)}{mc^2} \right)^2,$$

et de la définition de γ (avec l'expression de la vitesse en coordonnées polaires), on déduit aussi

$$\gamma^2 = \left[1 - \frac{1}{c^2} \left(\frac{dr}{dt} \right)^2 - \frac{r^2}{c^2} \left(\frac{d\phi}{dt} \right)^2 \right]^{-1}.$$

En éliminant $\dot{\phi}$ grâce à l'expression du moment cinétique, on a donc deux expressions de γ^2 :

$$\gamma^2 = 1 + \frac{|\boldsymbol{\sigma}|^2}{m^2 c^2 r^4} \left(\frac{dr}{d\phi} \right)^2 + \frac{|\boldsymbol{\sigma}|^2}{m^2 c^2 r^2} = \left(1 + \frac{E - U(r)}{mc^2} \right)^2.$$

Le changement de variable usuel $u = \frac{1}{r}$ donne

$$1 + \frac{|\boldsymbol{\sigma}|^2}{m^2 c^2} \left[\left(\frac{du}{d\phi} \right)^2 + u^2 \right] = \left(1 + \frac{E - U(u)}{mc^2} \right)^2.$$

que l'on dérive par rapport à ϕ pour obtenir l'équation du mouvement du problème de Kepler en Relativité Restreinte :

$$\frac{d^2 u}{d\phi^2} + \left(1 - \frac{G^2 M_\odot^2 m^2}{|\boldsymbol{\sigma}|^2 c^2} \right) u = \frac{GM_\odot m^2}{|\boldsymbol{\sigma}|^2} \left(1 + \frac{E}{mc^2} \right). \quad (4.23)$$

Cette équation ne diffère qualitativement de l'équation non relativiste que par le coefficient ω inférieur à 1, qui caractérise la vitesse angulaire du périhélie :

$$\omega^2 = 1 - \frac{G^2 M_\odot^2 m^2}{|\boldsymbol{\sigma}|^2 c^2},$$

et les solutions sont de la forme

$$\frac{1}{r} = \frac{1}{\bar{p}} (1 + \bar{e} \cos \omega \phi) \quad (4.24)$$

⁽⁴¹⁾ A partir de $E^2 = m^2 c^4 + p^2 c^2 = (mc^2 + K)^2$, on obtient le développement habituel $K = \frac{p^2}{2m} - \frac{1}{8} \frac{p^4}{m^3 c^2} + \dots$

où le paramètre \bar{p} et l'excentricité \bar{e} sont proches des valeurs classiques.

Si le premier périhélie est à l'origine $\phi_1 = 0$, après une révolution, le périhélie suivant se trouve à l'angle ϕ_2 tel que $\omega\phi_2 = 2\pi$, soit

$$\phi_2 \simeq 2\pi \left(1 + \frac{G^2 M_\odot^2 m^2}{2|\sigma|^2 c^2} \right).$$

On en déduit l'avance du périhélie par révolution :

$$\Delta\phi_0 = \frac{\pi G^2 M_\odot^2 m^2}{|\sigma|^2 c^2}.$$

En utilisant les valeurs classiques des paramètres de l'ellipse, $|\sigma|^2 = m^2 GM_\odot p$, $p = a(1 - e^2)$, on obtient par révolution la valeur

$$\Delta\phi_0 = \frac{\pi GM_\odot}{a(1 - e^2)c^2}. \quad (4.25)$$

En utilisant les valeurs numériques, on obtient $7.13''$ d'arc par siècle. Cette valeur est manifestement insuffisante pour expliquer la différence entre valeur observée et valeur calculée, mais elle montre bien qu'il s'agit d'une avance du périhélie.

4. Retour aux transformations non linéaires

La cohérence entre gravitation et théories quantiques reste un défi de la physique actuelle. Plusieurs théories alternatives de la gravité quantique existent (notamment la célèbre théorie des cordes) et toutes considèrent qu'à des échelles inférieures à la longueur de Planck $l_P = (G\hbar/c^3)^{1/2}$, la relativité générale doit céder la place à la gravité quantique, c'est-à-dire que des effets purement quantiques de la gravitation doivent se manifester à cette échelle. En termes d'énergies, les effets quantiques apparaissent aux énergies supérieures à l'énergie de Planck, $E_P = \hbar c/l_P = (\hbar c^5/G)^{1/2}$. La difficulté provient de ce que l'énergie n'est pas un invariant relativiste, ce qui signifie que si les effets quantiques se manifestent dans un référentiel inertiel donné, il n'en est pas nécessairement de même dans d'autres référentiels inertiels, ce qui est bien entendu regrettable. On préférerait une théorie qui prédisse le recours à la gravité quantique d'une manière indépendante du référentiel auquel est lié l'observateur.

Différentes alternatives à la transformation de Lorentz ont été envisagées, dont celle de Fock évoquée au chapitre 2. Récemment, une extension conduisant à une échelle d'énergie invariante a été proposée par Magueijo et Smolin ⁽⁴²⁾. On considère la transformation non linéaire

$$\begin{aligned} p^{0'} &= \frac{\gamma(p^0 - \beta p^1)}{1 + (\gamma - 1)p^0/l - \gamma\beta p^1/l}, \\ p^{1'} &= \frac{\gamma(p^1 - \beta p^0)}{1 + (\gamma - 1)p^0/l - \gamma\beta p^1/l}, \\ p^{2'} &= \frac{p^2}{1 + (\gamma - 1)p^0/l - \gamma\beta p^1/l}, \\ p^{3x'} &= \frac{p^3}{1 + (\gamma - 1)p^0/l - \gamma\beta p^1/l}. \end{aligned}$$

⁽⁴²⁾ J. Magueijo and L. Smolin, Phys. Rev. Lett. **88**, 190403 (2002).

Elle redonne la transformation de Lorentz aux énergies inférieures à l'énergie de Planck $E_P = \hbar c/l_P$ qui est un invariant de cette transformation. L'invariant associé au carré s'écrit maintenant

$$p_\mu p^\mu = \frac{g_{\mu\nu} p^\mu p^\nu}{(1 - l_P p^0/\hbar)^2} = m^2 c^2.$$

On a vu au chapitre 2 que cette transformation présente une difficulté lorsqu'elle est appliquée aux coordonnées spatio-temporelles, car elle redonne cette fois la transformation de Lorentz dans la mauvaise limite des distances courtes devant l_P et non pas des grandes distances.

Chapitre 5

Formulation covariante de l'électromagnétisme et de l'électrodynamique

L'électromagnétisme et les équations de Maxwell satisfont au principe de relativité. Il est toutefois très utile d'en donner une formulation covariante pour rendre cette propriété manifeste. On peut consulter de nombreux ouvrages sur cet aspect, notamment Jackson ou Rohrlich.

1. Quadrivecteur potentiel électromagnétique

1.1. Equations de propagation des ondes

On rappelle la forme habituelle des équations de Maxwell dans le vide,

$$\begin{aligned}
 \operatorname{div} \mathbf{E} &= \rho / \varepsilon_0 \\
 \operatorname{div} \mathbf{B} &= 0 \\
 \operatorname{rot} \mathbf{E} &= -\partial_t \mathbf{B} \\
 \operatorname{rot} \mathbf{B} &= \mu_0 \mathbf{j} + \varepsilon_0 \mu_0 \partial_t \mathbf{E}
 \end{aligned}
 \tag{5.1}$$

et la relation entre champs et potentiels,

$$\begin{aligned}
 \mathbf{E} &= -\vec{\nabla} \phi - \partial_t \mathbf{A} \\
 \mathbf{B} &= \operatorname{rot} \mathbf{A}.
 \end{aligned}
 \tag{5.2}$$

La quatrième équation de Maxwell, écrite en fonction des potentiels, donne

$$\underbrace{\text{rot rot } \mathbf{A}}_{\vec{\nabla}(\text{div } \mathbf{A}) - \vec{\nabla}^2 \mathbf{A}} = \mu_0 \mathbf{j} + \varepsilon_0 \mu_0 (-\partial_t \vec{\nabla} \phi - \partial_t^2 \mathbf{A})$$

soit

$$\vec{\nabla}^2 \mathbf{A} - \frac{1}{c^2} \partial_t^2 \mathbf{A} = -\frac{1}{\varepsilon_0 c^2} \mathbf{j} + \vec{\nabla}(\text{div } \mathbf{A} + \frac{1}{c^2} \partial_t \phi).$$

De même, en travaillant sur la première des équations de Maxwell, on obtient

$$\text{div}(-\vec{\nabla} \phi - \partial_t \mathbf{A}) = \rho / \varepsilon_0,$$

soit

$$\vec{\nabla}^2 \phi + \partial_t \text{div } \mathbf{A} = -\rho / \varepsilon_0.$$

Si l'on impose la jauge de Lorenz-Lorentz ⁽⁴³⁾

$$\text{div } \mathbf{A} + \frac{1}{c^2} \partial_t \phi = 0, \quad (5.3)$$

on obtient les équations de propagation des potentiels,

$$\begin{aligned} \vec{\nabla}^2 \mathbf{A} - \frac{1}{c^2} \partial_t^2 \mathbf{A} &= -\frac{1}{\varepsilon_0 c^2} \mathbf{j} \\ \vec{\nabla}^2 \phi - \frac{1}{c^2} \partial_t^2 \phi &= -\rho / \varepsilon_0. \end{aligned} \quad (5.4)$$

1.2. Jauge de Lorenz et équations de propagation en notation covariante

La forme même de la jauge de Lorenz suggère l'introduction d'un quadripotentiel

$$A^\mu = (\phi/c, \mathbf{A}) \quad (5.5)$$

de sorte que comme l'équation de conservation du quadricourant $j^\mu = (\rho c, \mathbf{j})$, elle apparaisse comme une contraction invariante,

$$\partial_\mu A^\mu = \frac{1}{c} \frac{\partial}{\partial t} \left(\frac{\phi}{c} \right) + \text{div } \mathbf{A} = 0. \quad (5.6)$$

⁽⁴³⁾ Cette relation, très importante puisqu'il s'agit d'une jauge covariante, est due au danois Ludvig Valentin Lorenz, mais elle fut popularisée par un quasi-homonyme, le physicien hollandais incontournable, Hendrik Antoon Lorentz, auquel l'usage en a attribué la paternité.

Les équations de propagation s'obtiennent de même de manière covariante au moyen de l'opérateur d'Alembertien,

$$\begin{aligned}\partial_\mu &= \left(\frac{1}{c}\partial_t, \vec{\nabla}\right) \\ \partial^\mu &= \left(\frac{1}{c}\partial_t, -\vec{\nabla}\right) \\ \partial_\mu \partial^\mu &= \left(\frac{1}{c^2}\partial_t^2, -\vec{\nabla}^2\right)\end{aligned}$$

soit

$$\begin{aligned}\partial_\mu \partial^\mu A^0 &= \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \left(\frac{\phi}{c}\right) - \vec{\nabla}^2 \left(\frac{\phi}{c}\right) = \frac{1}{\varepsilon_0 c} \rho = \frac{1}{\varepsilon_0 c^2} j^0 \\ \partial_\mu \partial^\mu \mathbf{A} &= \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{A} - \vec{\nabla}^2 \mathbf{A} = \frac{1}{\varepsilon_0 c^2} \mathbf{j}\end{aligned}$$

ce que l'on combine en une seule équation,

$$\partial_\mu \partial^\mu A^\nu = \mu_0 j^\nu, \quad (5.7)$$

avec $\mu_0 = 1/\varepsilon_0 c^2$.

On peut remarquer que les équations de propagation font intervenir $\mu_0 j^\nu$ au second membre, c'est-à-dire un tenseur contravariant de rang 1. Elles font par ailleurs appel à des dérivées secondes, donc des termes de la forme $\partial_\mu \partial^\nu A^\sigma$, mais il est nécessaire de faire apparaître une contraction pour obtenir un tenseur de rang 1,

$$\begin{aligned}\partial_\mu \partial^\mu &= \text{Tr}(\partial_\mu \partial^\mu) \\ &= \partial_0 \partial^0 + \partial_1 \partial^1 + \partial_2 \partial^2 + \partial_3 \partial^3 \\ &= (\partial_0)^2 - (\partial_1)^2 - (\partial_2)^2 - (\partial_3)^2 \\ &= \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2.\end{aligned} \quad (5.8)$$

et finalement la seule forme possible est bien $\partial_\mu \partial^\mu A^\nu = \mu_0 j^\nu$.

1.3. Transformation des potentiels

Il découle automatiquement de leur définition que les composantes du quadri-potentiel se transforment par changement de référentiel inertiel conformément à

$$A'^\mu = \Lambda^\mu{}_\nu A^\nu, \quad (5.9)$$

les lois des transformation de ϕ et \mathbf{A} s'ensuivent, $\phi'/c = \gamma\phi/c - \beta\gamma A_x$, $A'_x = -\beta\gamma\phi/c + \gamma A_x$, $A'_y = A_y$ et $A'_z = A_z$.

2. Tenseur champ électromagnétique et équations de Maxwell

2.1. Définition du tenseur champ électromagnétique

On cherche à définir un tenseur représentant le champ électromagnétique (qu'on appelle souvent tenseur de Faraday). On sait que A^μ est un 4-vecteur. Par ailleurs la relation $\mathbf{B} = \text{rot } \mathbf{A}$ contient des termes de la forme

$$B_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}$$

qui suggèrent de définir un tenseur de rang 2 antisymétrique

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \quad (5.10)$$

On a trivialement des zéros sur la diagonale, $F^{00} = F^{ii} = 0$ et $F^{\mu\nu} = -F^{\nu\mu}$ ou encore $F_{\mu\nu} = -F_{\nu\mu}$. Les composantes non nulles s'expriment explicitement en fonction des champs \mathbf{E} et \mathbf{B} .

$$\begin{aligned} F^{0i} &= \partial^0 A^i - \partial^i A^0 \\ &= \partial_0 A^i + \partial_i A^0 \\ &= \frac{1}{c} \frac{\partial}{\partial t} A^i + \frac{\partial}{\partial x^i} \left(\frac{\phi}{c} \right) \\ &\equiv -\frac{1}{c} E^i, \end{aligned}$$

composante ordinaire de \mathbf{E} le long de l'axe i .

$$\begin{aligned} F^{ij} &= \partial^i A^j - \partial^j A^i \\ &= -\partial_i A^j + \partial_j A^i \\ &= -\frac{\partial}{\partial x^i} A^j + \frac{\partial}{\partial x^j} A^i \\ &\equiv \pm B^k \quad \text{et permutations circulaires,} \end{aligned}$$

composante ordinaire de \mathbf{B} (avec permutations circulaires sur les trois indices i, j, k). Finalement on obtient le tenseur champ électromagnétique deux fois contravariant :

$$\begin{aligned} F^{\mu\nu} &= \partial^\mu A^\nu - \partial^\nu A^\mu = \frac{\partial A^\nu}{\partial x_\mu} - \frac{\partial A^\mu}{\partial x_\nu}. \\ [F^{\mu\nu}] &= \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}. \end{aligned} \quad (5.11)$$

Le tenseur deux fois covariant

$$F_{\mu\nu} = g_{\mu\sigma} g_{\nu\tau} F^{\sigma\tau}$$

s'en déduit en conservant le signe de $F_{00} = F^{00}$ et des composantes purement spatiales, $F_{ij} = F^{ij}$ et en changeant le signe des composantes mixtes $F_{0j} = -F^{0j}$:

$$[F_{\mu\nu}] = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix}. \quad (5.12)$$

On définit également un tenseur champ électromagnétique dual. Celui-ci est formé au moyen du tenseur de Levi-Civita

$$\epsilon^{\mu\nu\sigma\tau} = \begin{cases} +1 & \text{si } \mu, \nu, \sigma, \tau = 0, 1, 2, 3 \text{ et permutations paires} \\ -1 & \text{si permutations impaires} \\ 0 & \text{si deux indices ou plus sont égaux} \end{cases}$$

En particulier on a $\epsilon_{\mu\nu\sigma\tau} = -\epsilon^{\mu\nu\sigma\tau}$. On définit le tenseur dual $\bar{F}^{\mu\nu}$ par la contraction

$$\bar{F}^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\sigma\tau} F_{\sigma\tau}.$$

Ses composantes s'obtiennent à partir de celles de $F^{\mu\nu}$ en changeant \mathbf{E}/c en \mathbf{B} et \mathbf{B} en $-\mathbf{E}/c$:

$$[\bar{F}^{\mu\nu}] = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z/c & -E_y/c \\ B_y & -E_z/c & 0 & E_x/c \\ B_z & E_y/c & -E_x/c & 0. \end{pmatrix} \quad (5.13)$$

2.2. Changement de référentiel inertiel

Utilisant les règles de transformation des tenseurs, on obtient aisément les transformations des champs électromagnétiques par changement de référentiel inertiel. $F^{\mu\nu}$ étant un tenseur de rang deux, deux fois contravariant, il se transforme par définition selon

$$F'^{\mu\nu} = \Lambda^\mu_\sigma \Lambda^\nu_\tau F^{\sigma\tau}.$$

A titre d'exemple, établissons la loi de transformation de la composante $F^{10} = E_x/c$.

$$F'^{10} = \Lambda^1_\sigma \Lambda^0_\tau F^{\sigma\tau},$$

expression qui constitue une somme double (sur les deux indices σ et τ). Elle comprend 16 termes parmi lesquels seuls les termes impliquant Λ^1_0 , Λ^1_1 , Λ^0_0 et Λ^0_1 sont non nuls. Il reste

$$\begin{aligned} F'^{10} &= \Lambda^1_0 \Lambda^0_0 \underbrace{F^{00}}_0 + \Lambda^1_0 \Lambda^0_1 F^{01} + \Lambda^1_1 \Lambda^0_0 F^{10} + \Lambda^1_1 \Lambda^0_1 \underbrace{F^{11}}_0 \\ &= (-\beta_{\mathbf{u}} \gamma_{\mathbf{u}}) \times (-\beta_{\mathbf{u}} \gamma_{\mathbf{u}}) \times (-E_x/c) + \gamma_{\mathbf{u}} \times \gamma_{\mathbf{u}} \times (E_x/c), \end{aligned}$$

soit

$$E'_x/c = \gamma_{\mathbf{u}}^2 (1 - \beta_{\mathbf{u}}^2) E_x/c \equiv E_x/c.$$

Etablissons de même la transformation de $F^{20} = E_y/c$:

$$\begin{aligned} F'^{20} &= \Lambda^2_{\sigma} \Lambda^0_{\tau} F^{\sigma\tau} \\ &= \Lambda^2_2 \Lambda^0_0 F^{20} + \Lambda^2_2 \Lambda^0_1 F^{21} \\ &= 1 \times \gamma_{\mathbf{u}} \times (E_y/c) + 1 \times (-\beta_{\mathbf{u}} \gamma_{\mathbf{u}}) \times B_z, \end{aligned}$$

ou encore

$$E'_y/c = \gamma_{\mathbf{u}} (E_y/c - \beta_{\mathbf{u}} B_z).$$

On constate qu'en relativité, les composantes de \mathbf{E} et \mathbf{B} se mêlent en général, ce qui justifie leur regroupement au sein d'une entité unique. Si l'on développe les relations de transformation, on obtient :

$$\begin{aligned} E'_x &= E_x \\ E'_y &= \gamma_{\mathbf{u}} (E_y - \beta_{\mathbf{u}} c B_z) \\ E'_z &= \gamma_{\mathbf{u}} (E_z + \beta_{\mathbf{u}} c B_y) \\ B'_x &= B_x \\ B'_y &= \gamma_{\mathbf{u}} (B_y + (\beta_{\mathbf{u}}/c) E_z) \\ B'_z &= \gamma_{\mathbf{u}} (B_z - (\beta_{\mathbf{u}}/c) E_y) \end{aligned} \tag{5.14}$$

et sous forme condensée on écrit :

$$\begin{aligned} \mathbf{E}'_{\parallel} &= \mathbf{E}_{\parallel}, \\ \mathbf{B}'_{\parallel} &= \mathbf{B}_{\parallel}, \\ \mathbf{E}'_{\perp} &= \gamma_{\mathbf{u}} (\mathbf{E}_{\perp} + \beta_{\mathbf{u}} c \wedge \mathbf{B}_{\perp}), \\ \mathbf{B}'_{\perp} &= \gamma_{\mathbf{u}} (\mathbf{B}_{\perp} - (\beta_{\mathbf{u}}/c) \wedge \mathbf{E}_{\perp}). \end{aligned} \tag{5.15}$$

En réalisant des contractions sur tous les indices, on obtient les invariants du champ électromagnétique :

$$-\frac{1}{2} F^{\mu\nu} F_{\mu\nu} = \mathbf{E}^2/c^2 - \mathbf{B}^2, \quad -\frac{1}{4} \bar{F}^{\mu\nu} F_{\mu\nu} = \mathbf{E} \cdot \mathbf{B}/c. \tag{5.16}$$

Pour établir l'expression de ces invariants il faut développer les doubles sommes sur les indices.

$$\begin{aligned}
 -\frac{1}{2}F^{\mu\nu}F_{\mu\nu} &= -\frac{1}{2}\left(\underbrace{F^{00}F_{00}}_{(F^{00})^2} + \underbrace{\sum_i \sum_j F^{ij}F_{ij}}_{\sum_i \sum_j (F^{ij})^2} \right. \\
 &\quad \left. + \underbrace{\sum_i F^{0i}F_{0i} + \sum_i F^{i0}F_{i0}}_{-2\sum_i (F^{0i})^2} \right) \\
 &= -\frac{1}{2}(0 + 2(B_x^2 + B_y^2 + B_z^2) - 2(E_x^2 + E_y^2 + E_z^2)/c^2) \\
 &= \mathbf{E}^2/c^2 - \mathbf{B}^2.
 \end{aligned}$$

De même, l'expression développée du second invariant donne

$$\begin{aligned}
 \frac{1}{4}\bar{F}^{\mu\nu}F_{\mu\nu} &= \frac{1}{4}\left(\underbrace{\bar{F}^{00}F_{00}}_0 + \underbrace{\sum_i \sum_j \bar{F}^{ij}F_{ij}}_{2(-B_z(E_z/c) + B_y(-E_y/c) - B_x(E_x/c))} \right. \\
 &\quad \left. + \underbrace{\sum_i \bar{F}^{0i}F_{0i} + \sum_i \bar{F}^{i0}F_{i0}}_{-2(B_z E_z/c + B_y E_y/c - B_x E_x/c)} \right) \\
 &= \frac{1}{4}(-4)\mathbf{EB}/c.
 \end{aligned}$$

De ces deux invariants se déduisent quelques propriétés immédiates.

- i) Par exemple si l'on a l'égalité $|\mathbf{B}| = |\mathbf{E}|/c$ dans un référentiel inertiel, il en est de même dans tous les référentiels inertiels.
- ii) S'il existe un référentiel inertiel dans lequel \mathbf{B} est nul et \mathbf{E} prend une valeur \mathbf{E}_0 , alors dans aucun référentiel inertiel $|\mathbf{E}|$ ne pourra être inférieur à $|\mathbf{E}_0|$.
- iii) Si \mathbf{E} et \mathbf{B} sont perpendiculaires dans un référentiel inertiel, alors ils le sont dans tous les référentiels inertiels.

2.3. Equations de Maxwell covariantes

Construisons les contractions du tenseur champ électromagnétique, par exemple $\partial_\mu F^{\mu\nu}$. Pour $\nu = 0$, on a

$$\begin{aligned}
 \partial_\mu F^{\mu 0} &= \partial_0 F^{00} + \partial_1 F^{10} + \partial_2 F^{20} + \partial_3 F^{30} \\
 &= 0 + \frac{\partial}{\partial x}(E_x/c) + \frac{\partial}{\partial y}(E_y/c) + \frac{\partial}{\partial z}(E_z/c) \\
 &= \frac{1}{c} \operatorname{div} \mathbf{E} = \frac{1}{\varepsilon_0 c^2} \rho c \\
 &= \mu_0 j^0.
 \end{aligned}$$

Pour $\nu = 1$, on a de même

$$\begin{aligned}\partial_\mu F^{\mu 1} &= \partial_0 F^{01} + \partial_1 F^{11} + \partial_2 F^{21} + \partial_3 F^{31} \\ &= \frac{1}{c} \frac{\partial}{\partial t} (-E_x/c) + 0 + \frac{\partial}{\partial y} B_z + \frac{\partial}{\partial z} (-B_y) \\ &= -\frac{1}{c^2} \frac{\partial E_x}{\partial t} + (\mathbf{rot} \mathbf{B}) \mathbf{u}_x \\ &= \mu_0 j^1,\end{aligned}$$

et les deux dernières composantes spatiales s'en déduisent par permutation. On a donc

$$\partial_\mu F^{\mu\nu} = \mu_0 j^\nu, \quad (5.17)$$

expression qui regroupe les deux équations de Maxwell avec sources,

$$\begin{aligned}\operatorname{div} \mathbf{E} &= \frac{\rho}{\varepsilon_0} \\ \mathbf{rot} \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} &= \mu_0 \mathbf{j}.\end{aligned} \quad (5.18)$$

Le second couple d'équations de Maxwell s'obtient plus aisément après l'introduction du tenseur champ électromagnétique dual. Examinons la contraction $\partial_\mu \bar{F}^{\mu\nu}$. Pour $\nu = 0$, on a

$$\begin{aligned}\partial_\mu \bar{F}^{\mu 0} &= \partial_0 \bar{F}^{00} + \partial_1 \bar{F}^{10} + \partial_2 \bar{F}^{20} + \partial_3 \bar{F}^{30} \\ &= 0 + \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} \\ &= \operatorname{div} \mathbf{B}.\end{aligned}$$

Pour $\nu = 1$, on a de même

$$\begin{aligned}\partial_\mu \bar{F}^{\mu 1} &= \partial_0 \bar{F}^{01} + \partial_1 \bar{F}^{11} + \partial_2 \bar{F}^{21} + \partial_3 \bar{F}^{31} \\ &= \frac{1}{c} \frac{\partial}{\partial t} (-B_x) + 0 + \frac{\partial}{\partial y} (-E_z/c) + \frac{\partial}{\partial z} (E_y/c) \\ &= -\frac{1}{c} \frac{\partial B_x}{\partial t} - \frac{1}{c} (\mathbf{rot} \mathbf{E}) \mathbf{u}_x.\end{aligned}$$

On en déduit l'expression unifiée du second couple d'équations de Maxwell, les équations sans second membre,

$$\partial_\mu \bar{F}^{\mu\nu} = 0. \quad (5.19)$$

Les équations de Maxwell manifestement covariantes s'écrivent donc :

$$\begin{aligned}\partial_\mu F^{\mu\nu} &= \mu_0 j^\nu, \\ \partial_\mu \bar{F}^{\mu\nu} &= \partial_\mu F^{\nu\lambda} + \partial_\nu F^{\lambda\mu} + \partial_\lambda F^{\mu\nu} = 0.\end{aligned} \quad (5.20)$$

2.4. Force de Lorentz

Comme on l'a suggéré au chapitre 3, la notation covariante est un guide pour l'expression des lois physiques en relativité. Reprenons l'exemple du mouvement d'une charge q dans un champ électromagnétique. On sait maintenant que ce champ s'exprime sous la forme d'un tenseur de rang 2, $F^{\mu\nu}$ et que ses composantes spatio-temporelles F^{0i} font intervenir E_{x_i}/c alors que les composantes purement spatiales sont liées à celles de \mathbf{B} . Il est donc logique de généraliser la relation fondamentale de la dynamique

$$m \frac{d^2 \mathbf{r}}{dt^2} = q(\mathbf{E} + \mathbf{v} \wedge \mathbf{B})$$

à une expression covariante,

$$m \frac{d^2 x^\mu}{d\tau^2} = q F^{\mu\nu} v_\nu.$$

Pour interpréter les composantes de cette équation, on peut l'écrire sous forme matricielle et passer à la limite classique $c \rightarrow \infty$.

$$m \frac{d}{d\tau} \begin{pmatrix} \gamma_{\mathbf{v}} c \\ \gamma_{\mathbf{v}} v_x \\ \gamma_{\mathbf{v}} v_y \\ \gamma_{\mathbf{v}} v_z \end{pmatrix} = q \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} \gamma_{\mathbf{v}} c \\ -\gamma_{\mathbf{v}} v_x \\ -\gamma_{\mathbf{v}} v_y \\ -\gamma_{\mathbf{v}} v_z \end{pmatrix}$$

La première équation s'écrit encore

$$mc \frac{d\gamma_{\mathbf{v}}}{d\tau} = \frac{q}{c} \gamma_{\mathbf{v}} \mathbf{E} \cdot \mathbf{v}$$

où il apparaît la puissance classique

$$q\mathbf{E} \cdot \mathbf{v} = \mathcal{P} = -\vec{\nabla}U(\mathbf{r}) \cdot \mathbf{v} = -\frac{dU(\mathbf{r})}{dt}$$

(la partie magnétique ne travaille pas). En développant la dérivée du premier membre on obtient

$$\frac{m}{1 - |\mathbf{v}|^2/c^2} \mathbf{v} \frac{d\mathbf{v}}{d\tau} = q\mathbf{E} \cdot \mathbf{v},$$

soit, à la limite $c \rightarrow \infty$,

$$m\mathbf{v} \frac{d\mathbf{v}}{dt} = \frac{d}{dt} \left(\frac{1}{2} m |\mathbf{v}|^2 \right) = \mathcal{P} = -\frac{dU(\mathbf{r})}{dt}.$$

Le contenu énergétique de la première équation est ainsi lié à la définition de la puissance et au théorème de l'énergie cinétique. Pour ce qui est des équations spatiales, on peut par exemple considérer le cas de la première qui donne

$$m \frac{d(\gamma_{\mathbf{v}} v_x)}{d\tau} = q\gamma_{\mathbf{v}} (E_x + B_z v_y - B_y v_z)$$

où l'on voit clairement apparaître la force de Lorentz au second membre.

3. Applications

3.1. Champ d'une charge ponctuelle en mouvement arbitraire

3.1.1. Quadripotentiel A^μ

On considère une charge q animée d'un mouvement arbitraire par rapport à un observateur situé en \mathbf{r}_0 . À l'instant t_0 (mesuré dans son propre référentiel), le champ électromagnétique ressenti est celui de la charge lorsqu'elle était à la position retardée $\mathbf{r}_q(t_r)$ telle que

$$|\mathbf{r}_q(t_r) - \mathbf{r}_0| = c(t_0 - t_r)$$

en raison de la vitesse finie de propagation des interactions électromagnétiques. On introduit pour la suite la notation $\mathbf{R}(t_r) = \mathbf{r}_q(t_r) - \mathbf{r}_0$ et la notation quadridimensionnelle associée, $R_\sigma(t_r) = (c(t_0 - t_r), -\mathbf{R}(t_r))$. Ce quadrivecteur est par définition du genre lumière, puisque

$$R_\sigma(t_r)R^\sigma(t_r) = c^2(t_0 - t_r)^2 - |\mathbf{R}(t_r)|^2 = 0.$$

Dans la notation tensorielle, on appelle x_0^μ la quadriposition de l'observateur et $z^\mu(t)$ celle de la charge.

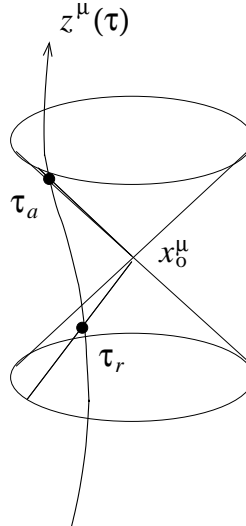


Figure 5.1 Cône de lumière pour un observateur en x_0^μ . La ligne représente la ligne d'univers $z^\mu(\tau)$ d'une charge q en mouvement arbitraire qui intercepte le cône de lumière à l'instant retardé τ_r (mesuré dans son référentiel propre) puis à nouveau à l'instant avancé τ_a .

Pour trouver la forme du quadripotentiel $A^\mu(x_0^\nu)$ au point \mathbf{r}_0, t_0 dans le référentiel lié à l'observateur, on utilise la covariance des expressions tensorielles, c'est-à-dire que l'on se place tout d'abord dans le référentiel propre de la charge (le temps propre dans ce référentiel \mathcal{R}' qui accompagne la charge dans son mouvement

est noté τ et les quantités exprimées dans ce référentiel sont notées avec des primes) où le quadripotential prend une expression simplifiée, puisque seule la partie temporelle est non nulle. On exprime cette partie temporelle sous forme manifestement covariante (c'est-à-dire comme un tenseur contravariant de rang 1) et l'expression correspondante dans le référentiel original en découle immédiatement.

Dans le référentiel propre de la charge, le quadripotential se réduit à la partie électrostatique, soit

$$\phi'(\mathbf{r}'_0, \tau_0) = \frac{q}{4\pi\varepsilon_0} \frac{1}{|\mathbf{R}'(\tau_r)|}.$$

Comme on a dans \mathcal{R}' , tout comme dans \mathcal{R} d'ailleurs,

$$R'_{\sigma}(\tau_r)R'^{\sigma}(\tau_r) = c^2(\tau_0 - \tau_r)^2 - |\mathbf{R}'(\tau_r)|^2 = 0,$$

où $R'^{\sigma}(\tau_r) = (c(\tau_0 - \tau_r), \mathbf{R}'(\tau_r))$, on peut écrire $|\mathbf{R}'(\tau_r)| = c(\tau_0 - \tau_r)$ et

$$\frac{1}{c}\phi'(\mathbf{r}'_0, \tau_0) = \frac{q}{4\pi\varepsilon_0 c^2(\tau_0 - \tau_r)}. \quad (5.21)$$

Dans ce même référentiel, on note que

$$\mathbf{A}'(\mathbf{r}'_0, \tau_0) = \vec{0}. \quad (5.22)$$

Le point essentiel est que dans le référentiel \mathcal{R}' , le quadripotential n'a qu'une partie temporelle non nulle, il doit donc être proportionnel à la quadrivitesse qui, dans ce même référentiel, prend la forme simplifiée $v'^{\mu}(\tau) = (c, \vec{0})$. C'est de plus compatible avec l'exigence d'obtenir pour A'^{μ} un tenseur contravariant de rang 1. Il faut alors faire également apparaître la quadrivitesse au dénominateur pour des raisons dimensionnelles, mais dans une contraction pour obtenir un scalaire au dénominateur. On forme ainsi la quantité

$$\frac{v'^{\mu}(\tau_r)}{v'_{\sigma}(\tau_r)R'^{\sigma}(\tau_r)} = \left(\frac{c}{c^2(\tau_0 - \tau_r)}, \vec{0} \right)$$

qui permet d'écrire

$$A'^{\mu}(x'^{\nu}_0) = \frac{q}{4\pi\varepsilon_0 c} \frac{v'^{\mu}(\tau_r)}{v'_{\sigma}(\tau_r)R'^{\sigma}(\tau_r)} = \frac{q}{4\pi\varepsilon_0 c} \frac{[v'^{\mu}]_{ret.}}{[v'_{\sigma}R'^{\sigma}]_{ret.}},$$

où la notation fréquemment employée $[\dots]_{ret.}$ signifie que la quantité entre crochets est évaluée au temps retardé.

L'intérêt de cette expression est évident : il s'agit d'une forme tensorielle, donc covariante par changement de référentiel inertiel, elle se généralise immédiatement à tout autre référentiel inertiel, en particulier \mathcal{R} .

$$A^{\mu}(x^{\nu}_0) = \frac{q}{4\pi\varepsilon_0 c} \frac{v^{\mu}(t_r)}{v_{\sigma}(t_r)R^{\sigma}(t_r)} = \frac{q}{4\pi\varepsilon_0 c} \frac{[v^{\mu}]_{ret.}}{[v_{\sigma}R^{\sigma}]_{ret.}}. \quad (5.23)$$

En exprimant dans \mathcal{R} les composantes de v^μ et le produit invariant $v_\sigma R^\sigma$, (il n'y a pas d'ambiguïté ici, on notera donc γ pour $\gamma_{\mathbf{v}}$)

$$\begin{aligned} v^\sigma(t_r) &= (\gamma c, \gamma \mathbf{v}(t_r)) \\ R^\sigma(t_r) &= (c(t_0 - t_r), \mathbf{R}(t_r)) \\ v_\sigma(t_r) R^\sigma(t_r) &= \gamma c^2(t_0 - t_r) - \gamma \mathbf{v}(t_r) \mathbf{R}(t_r) \\ &= \gamma(|\mathbf{R}(t_r)|c - \mathbf{R}(t_r) \mathbf{v}(t_r)). \end{aligned}$$

on obtient la forme explicite du potentiel électrostatique et du potentiel vecteur dans le référentiel de l'observateur,

$$\begin{aligned} A^\mu(\mathbf{r}_0, t_0) &= \frac{q}{4\pi\epsilon_0 c} \left(\frac{c}{|\mathbf{R}(t_r)|c - \mathbf{R}(t_r) \mathbf{v}(t_r)}, \frac{\mathbf{v}(t_r)}{|\mathbf{R}(t_r)|c - \mathbf{R}(t_r) \mathbf{v}(t_r)} \right) \\ \phi(\mathbf{r}_0, t_0) &= \frac{q}{4\pi\epsilon_0 |\mathbf{R}(t_r)|} (1 - \mathbf{R}(t_r) \mathbf{v}(t_r) / |\mathbf{R}(t_r)|c)^{-1} \\ \mathbf{A}(\mathbf{r}_0, t_0) &= \frac{q \mathbf{v}(t_r)}{4\pi\epsilon_0 c^2 |\mathbf{R}(t_r)|} (1 - \mathbf{R}(t_r) \mathbf{v}(t_r) / |\mathbf{R}(t_r)|c)^{-1}. \end{aligned} \quad (5.24)$$

On introduit souvent la notation abrégée $\boldsymbol{\beta}(t) = \mathbf{v}(t)/c$ et le vecteur unitaire dans la direction charge-observateur, $\hat{\mathbf{u}}_R(t) = \mathbf{R}(t)/|\mathbf{R}(t)|$, ce qui conduit à

$$\begin{aligned} \phi(\mathbf{r}_0, t_0) &= \frac{1}{4\pi\epsilon_0} \frac{q}{|\mathbf{R}(t_r)|} (1 - \boldsymbol{\beta}(t_r) \hat{\mathbf{u}}_R(t_r))^{-1} \\ \mathbf{A}(\mathbf{r}_0, t_0) &= \frac{\mu_0}{4\pi} \frac{q \mathbf{v}(t_r)}{|\mathbf{R}(t_r)|} (1 - \boldsymbol{\beta}(t_r) \hat{\mathbf{u}}_R(t_r))^{-1}. \end{aligned} \quad (5.25)$$

Ces potentiels sont connus sous le nom de potentiels de Liénard-Wiechert⁽⁴⁴⁾.

3.1.2. Composantes ordinaires des champs \mathbf{E} et \mathbf{B} de la charge en mouvement arbitraire

On peut à partir de ces potentiels obtenir l'expression explicite des champs par un calcul fastidieux dont on résume ci-dessous les étapes essentielles. Le champ électrique au point 0 est donné par

$$\mathbf{E}(\mathbf{r}_0, t_0) = -\frac{\partial \mathbf{A}(\mathbf{r}_0, t_0)}{\partial t_0} - \vec{\nabla} \phi(\mathbf{r}_0, t_0),$$

On a tout d'abord besoin d'exprimer $\frac{\partial t_r}{\partial t_0}$ et $\vec{\nabla} t_r$, car

$$\begin{aligned} \frac{\partial \mathbf{A}(\mathbf{r}_0, t_0)}{\partial t_0} &= \frac{\partial \mathbf{A}(\mathbf{r}_0, t_0)}{\partial t_r} \frac{\partial t_r}{\partial t_0}, \\ \vec{\nabla} \phi(\mathbf{r}_0, t_0) &= \vec{\nabla} \phi(\mathbf{r}_0, t_0) \Big|_{t_r} + \frac{\partial \phi(\mathbf{r}_0, t_0)}{\partial t_r} \vec{\nabla} t_r, \end{aligned}$$

⁽⁴⁴⁾ Ces potentiels ont été obtenus indépendamment avant l'élaboration de la relativité restreinte par Liénard en 1898 et par Wiechert en 1900.

où ... $|_{t_r}$ signifie que le gradient est calculé à t_r constant.

i) Calculons tout d'abord $\frac{\partial t_r}{\partial t_0}$. On a

$$\frac{\partial}{\partial t_r}(\mathbf{R}(t_r))^2 = 2\mathbf{R}(t_r)\frac{\partial\mathbf{R}(t_r)}{\partial t_r} = -2\mathbf{R}(t_r)\mathbf{v}(t_r)$$

qui vaut aussi

$$\frac{\partial}{\partial t_r}|\mathbf{R}(t_r)|^2 = 2|\mathbf{R}(t_r)|\frac{\partial|\mathbf{R}(t_r)|}{\partial t_r}$$

d'où l'on déduit

$$\frac{\partial|\mathbf{R}(t_r)|}{\partial t_r} = -\frac{\mathbf{R}(t_r)}{|\mathbf{R}(t_r)|}\mathbf{v}(t_r).$$

Par ailleurs, $t_0 = t_r + |\mathbf{R}(t_r)|/c$, soit

$$\begin{aligned}\frac{\partial t_0}{\partial t_r} &= 1 + \frac{1}{c}\frac{\partial|\mathbf{R}(t_r)|}{\partial t_r} \\ &= 1 - \frac{\mathbf{R}(t_r)}{|\mathbf{R}(t_r)|}\frac{\mathbf{v}(t_r)}{c} \\ &= 1 - \beta(t_r)\hat{\mathbf{u}}_R(t_r),\end{aligned}$$

d'où

$$\frac{\partial t_r}{\partial t_0} = (1 - \beta(t_r)\hat{\mathbf{u}}_R(t_r))^{-1} \quad (5.26)$$

ii) Pour calculer $\vec{\nabla}t_r$ il est plus commode d'utiliser les avantages de la notation tensorielle. En effet, si l'on exprime $\frac{1}{c}\frac{\partial t_r}{\partial t_0}$ comme la composante temporelle d'un quadrivecteur, le gradient cherché sera donné par les composantes spatiales de ce quadrivecteur.

$$\begin{aligned}\frac{1}{c}\frac{\partial t_r}{\partial t_0} &= \frac{|\mathbf{R}(t_r)|}{|\mathbf{R}(t_r)|c - \mathbf{R}(t_r)\mathbf{v}(t_r)} \\ &= \frac{\gamma_{\mathbf{v}}|\mathbf{R}(t_r)|}{v_{\sigma}(t_r)R^{\sigma}(t_r)} \\ &= \frac{\gamma_{\mathbf{v}}c(t_0 - t_r)}{v_{\sigma}(t_r)R^{\sigma}(t_r)},\end{aligned}$$

de sorte qu'on identifie le quadrivecteur covariant

$$\partial_{\mu}t_r = \frac{\partial t_r}{\partial x^{\mu}} = \left(\frac{1}{c}\frac{\partial t_r}{\partial t_0}, \vec{\nabla}t_r\right) = \gamma_{\mathbf{v}}\frac{R_{\mu}(t_r)}{v_{\sigma}(t_r)R^{\sigma}(t_r)}.$$

Il en résulte que

$$\begin{aligned}\vec{\nabla}t_r &= \gamma_{\mathbf{v}} \frac{-\mathbf{R}(t_r)}{v_{\sigma}(t_r)R^{\sigma}(t_r)} \\ &= -\frac{\mathbf{R}(t_r)}{|\mathbf{R}(t_r)|_c - \mathbf{R}(t_r)\mathbf{v}(t_r)}.\end{aligned}\quad (5.27)$$

iii) On peut maintenant calculer la dérivée du potentiel vecteur par rapport à l'instant retardé t_r . Pour alléger l'écriture, on ne précise pas dans les quelques équations qui suivent la dépendance temporelle des quantités écrites au second membre, mais elles sont toutes supposées être évaluées à l'instant retardé t_r .

$$\begin{aligned}\frac{\partial}{\partial t_r}\mathbf{A}(\mathbf{r}_0, t_0) &= \frac{\mu_0 q}{4\pi} \left[\left(\frac{\mathbf{a}}{|\mathbf{R}|} + \mathbf{v} \left(-\frac{1}{|\mathbf{R}|^2} \frac{\partial |\mathbf{R}|}{\partial t_r} \right) \right) (1 - \beta \hat{\mathbf{u}}_R)^{-1} \right. \\ &\quad \left. - \frac{\mathbf{v}}{|\mathbf{R}|} (1 - \beta \hat{\mathbf{u}}_R)^{-2} \frac{\partial}{\partial t_r} \left(-\frac{\mathbf{R}\mathbf{v}}{|\mathbf{R}|_c} \right) \right]_{ret.} \\ &= \frac{\mu_0 q}{4\pi} (1 - \beta \hat{\mathbf{u}}_R)^{-2} \left[\left(\frac{\mathbf{a}}{|\mathbf{R}|} - \frac{\mathbf{v}}{|\mathbf{R}|^2} \left(-\frac{\mathbf{R}\mathbf{v}}{|\mathbf{R}|} \right) \right) (1 - \beta \hat{\mathbf{u}}_R) \right. \\ &\quad \left. + \frac{\mathbf{v}}{|\mathbf{R}|_c} \left(\frac{-\mathbf{v}^2 + \mathbf{R}\mathbf{a}}{|\mathbf{R}|} - \frac{\mathbf{R}\mathbf{v}}{|\mathbf{R}|^2} \left(-\frac{\mathbf{R}\mathbf{v}}{|\mathbf{R}|} \right) \right) \right]_{ret.} \\ &= \frac{\mu_0 q}{4\pi} (1 - \beta \hat{\mathbf{u}}_R)^{-2} \left[\frac{\mathbf{a}}{|\mathbf{R}|} (1 - \beta \hat{\mathbf{u}}_R) + \frac{\mathbf{R}\mathbf{a}}{|\mathbf{R}|^2 c} \mathbf{v} \right. \\ &\quad \left. + \frac{\mathbf{R}\mathbf{v}}{|\mathbf{R}|^3} (1 - \beta \hat{\mathbf{u}}_R) \mathbf{v} - \frac{\mathbf{v}^2}{|\mathbf{R}|^2 c} \mathbf{v} + \frac{(\mathbf{R}\mathbf{v})^2}{|\mathbf{R}|^4 c} \mathbf{v} \right]_{ret.}.\end{aligned}$$

Après simplification des premier et troisième termes en partant de la fin, et multiplication par $\frac{\partial t_r}{\partial t_0}$, il vient finalement

$$\begin{aligned}-\frac{\partial}{\partial t_0}\mathbf{A}(\mathbf{r}_0, t_0) &= -\frac{\mu_0 q}{4\pi} (1 - \beta \hat{\mathbf{u}}_R)^{-3} \left[(1 - \beta \hat{\mathbf{u}}_R) \frac{\mathbf{a}}{|\mathbf{R}|} \right. \\ &\quad \left. + \left(\frac{\mathbf{a}}{|\mathbf{R}|} \hat{\mathbf{u}}_R \right) \beta + \left(\frac{c^2}{|\mathbf{R}|^2} \beta \hat{\mathbf{u}}_R \right) \beta - \frac{c^2 \beta^2}{|\mathbf{R}|^2} \beta \right]_{ret.}.\end{aligned}\quad (5.28)$$

On vérifie que tous les termes sont homogènes.

iv) Le calcul du gradient du potentiel scalaire est plus fastidieux. On a déjà calculé $\vec{\nabla}t_r = -\frac{1}{c}(1 - \beta \hat{\mathbf{u}}_R)^{-1} \hat{\mathbf{u}}_R$. Considérons ensuite

$$\begin{aligned}\frac{\partial \phi}{\partial t_r} &= \frac{q}{4\pi \varepsilon_0} \left[\left(-\frac{1}{|\mathbf{R}|^2} \frac{\partial |\mathbf{R}|}{\partial t_r} \right) (1 - \beta \hat{\mathbf{u}}_R)^{-1} \right. \\ &\quad \left. - (1 - \beta \hat{\mathbf{u}}_R)^{-2} \frac{1}{|\mathbf{R}|} \frac{\partial}{\partial t_r} \left(\frac{\mathbf{R}\mathbf{v}}{|\mathbf{R}|_c} \right) \right]_{ret.}\end{aligned}$$

Avec $\frac{\partial |\mathbf{R}|}{\partial t_r} = -\hat{\mathbf{u}}_R \mathbf{v}$ et $\frac{\partial}{\partial t_r} \left(\frac{\mathbf{R}\mathbf{v}}{|\mathbf{R}|c} \right) = \frac{-|\mathbf{v}|^2 + \mathbf{R}\mathbf{a}}{|\mathbf{R}|c} + \frac{\mathbf{R}\mathbf{v}}{|\mathbf{R}|^2} \frac{\mathbf{R}\mathbf{v}}{|\mathbf{R}|c}$, on obtient

$$\frac{\partial \phi}{\partial t_r} = \left[\frac{q}{4\pi\epsilon_0 |\mathbf{R}|^2} (1 - \beta \hat{\mathbf{u}}_R)^{-2} \left(-\beta \hat{\mathbf{u}}_R c + |\beta|^2 c - \frac{\mathbf{R}\mathbf{a}}{c} \right) \right]_{ret.}$$

d'où l'on déduit le premier terme

$$-\frac{\partial \phi}{\partial t_r} \vec{\nabla} t_r = \left[\frac{q}{4\pi\epsilon_0 |\mathbf{R}|^2 c} (1 - \beta \hat{\mathbf{u}}_R)^{-3} \left(-\beta \hat{\mathbf{u}}_R c + |\beta|^2 c - \frac{\mathbf{R}\mathbf{a}}{c} \right) \hat{\mathbf{u}}_R \right]_{ret.}.$$

On poursuit par le calcul du gradient à l'instant retardé

$$\begin{aligned} \vec{\nabla} \phi \Big|_{t_r} &= \frac{q}{4\pi\epsilon_0} \left[(1 - \beta \hat{\mathbf{u}}_R)^{-1} \vec{\nabla} \left(\frac{1}{|\mathbf{R}|} \right) \right. \\ &\quad \left. - (1 - \beta \hat{\mathbf{u}}_R)^{-2} \frac{1}{|\mathbf{R}|} \vec{\nabla} \left(\frac{\mathbf{R}\mathbf{v}}{|\mathbf{R}|c} \right) \right]_{ret.} \end{aligned}$$

Avec $\vec{\nabla} \left(\frac{\mathbf{R}\mathbf{v}}{|\mathbf{R}|c} \right) = \frac{\mathbf{v}}{|\mathbf{R}|c} - \mathbf{R}\mathbf{v} \frac{\mathbf{R}}{|\mathbf{R}|^3 c}$, il vient

$$-\vec{\nabla} \phi \Big|_{t_r} = \left[\frac{q}{4\pi\epsilon_0 |\mathbf{R}|^2} (1 - \beta \hat{\mathbf{u}}_R)^{-2} (\beta - \hat{\mathbf{u}}_R) \right]_{ret.}$$

et l'on combine les deux quantités pour obtenir

$$\begin{aligned} -\vec{\nabla} \phi(\mathbf{r}_0, t_0) &= \left[\frac{q}{4\pi\epsilon_0 |\mathbf{R}|^2} (1 - \beta \hat{\mathbf{u}}_R)^{-3} (\beta(1 - \beta \hat{\mathbf{u}}_R) \right. \\ &\quad \left. - \hat{\mathbf{u}}_R(1 - |\beta|^2) - \frac{\mathbf{R}\mathbf{a}}{c^2} \hat{\mathbf{u}}_R) \right]_{ret.}. \end{aligned} \quad (5.29)$$

v) Il reste simplement à exprimer le champ électrique

$$\begin{aligned} \mathbf{E}(\mathbf{r}_0, t_0) &= \frac{q}{4\pi\epsilon_0 [|\mathbf{R}|^2]_{ret.}} (1 - [\beta \hat{\mathbf{u}}_R]_{ret.})^{-3} \\ &\quad \times \left[(1 - |\beta|^2)(\hat{\mathbf{u}}_R - \beta) + \frac{|\mathbf{R}|}{c^2} (\mathbf{a}\hat{\mathbf{u}}_R(\hat{\mathbf{u}}_R - \beta) - \mathbf{a}(1 - \beta \hat{\mathbf{u}}_R)) \right]_{ret.}, \end{aligned}$$

et en utilisant $\mathbf{a}\hat{\mathbf{u}}_R = a_{\parallel}$ et $\mathbf{a} - (\mathbf{a}\hat{\mathbf{u}}_R)\hat{\mathbf{u}}_R = \mathbf{a}_{\perp}$ et de même pour les composantes de β on obtient

$$\begin{aligned} \mathbf{E}(\mathbf{r}_0, t_0) &= \frac{q}{4\pi\epsilon_0 [|\mathbf{R}|^2]_{ret.} (1 - [\beta \hat{\mathbf{u}}_R]_{ret.})^3} \\ &\quad \times \left[(1 - |\beta|^2)(\hat{\mathbf{u}}_R - \beta) - \frac{|\mathbf{R}|}{c^2} (\mathbf{a}_{\perp} - a_{\parallel}\beta - \beta_{\parallel}\mathbf{a}) \right]_{ret.}. \end{aligned} \quad (5.30)$$

Le champ total est constitué d'une partie statique et d'une partie proportionnelle à la vitesse retardée, toutes deux décroissant rapidement en $1/R^2$ et d'une partie faisant intervenir l'accélération, décroissant bien plus lentement en $1/R$ et qui constitue le champ de rayonnement.

iv) Le champ magnétique se déduit du champ électrique par

$$\mathbf{B}(\mathbf{r}_0, t_0) = \frac{[\hat{\mathbf{u}}_R]_{ret.}}{c} \wedge \mathbf{E}(\mathbf{r}_0, t_0). \quad (5.31)$$

v) Les expressions précédentes des champs \mathbf{E} et \mathbf{B} sont dues à Liénard et Wiechert. Il existe également une autre forme, due à Heaviside et Feynman⁽⁴⁵⁾,

$$\mathbf{E}(\mathbf{r}_0, t_0) = \frac{q}{4\pi\epsilon_0} \left(\frac{\hat{\mathbf{u}}_R}{|\mathbf{R}|^2} + \frac{|\mathbf{R}|}{c} \frac{d}{dt} \frac{\hat{\mathbf{u}}_R}{|\mathbf{R}|^2} + \frac{1}{c^2} \frac{d^2}{dt^2} \hat{\mathbf{u}}_R \right).$$

3.1.3. Cas limites

Partant des expressions générales, on peut considérer divers cas limites intéressants. Tout d'abord pour une particule immobile, on retrouve bien le champ coulombien habituel (heureusement !),

$$\mathbf{E}(\mathbf{r}_0, t_0) = \frac{q}{4\pi\epsilon_0 |\mathbf{R}|^2} \hat{\mathbf{u}}_R. \quad (5.32)$$

Pour une particule en mouvement uniforme, on obtient un champ décroissant en $1/|\mathbf{R}|^2$, de type coulombien,

$$\mathbf{E}(\mathbf{r}_0, t_0) = \frac{q}{4\pi\epsilon_0} \left[\frac{(1 - |\beta|^2)(\hat{\mathbf{u}}_R - \beta)}{|\mathbf{R}|^2(1 - \beta\hat{\mathbf{u}}_R)^3} \right]_{ret.}. \quad (5.33)$$

Il est remarquable de noter qu'en l'absence d'accélération (vitesse $\mathbf{v}(t_r) = \mathbf{v}$ uniforme), le champ électrique est centré sur la position actuelle que la particule occupe à l'instant t_0 . On a en effet

$$\hat{\mathbf{u}}_R(t_r) - \beta(t_r) = \frac{\mathbf{R}(t_r)}{c(t_0 - t_r)} - \frac{\mathbf{v}}{c},$$

or $\mathbf{R}(t_0) = \mathbf{R}(t_r) - (t_0 - t_r)\mathbf{v}$, d'où

$$\hat{\mathbf{u}}_R(t_r) - \beta(t_r) = \frac{\mathbf{R}(t_0)}{c(t_0 - t_r)}.$$

Il n'y a donc finalement aucun effet de retard et le champ est essentiellement coulombien et centré sur la position instantanée de la charge.

Si l'on considère maintenant une charge accélérée (\mathbf{a} non nul), mais ayant une faible vitesse de sorte que $\beta \ll 1$ soit négligeable, on obtient deux contributions dominantes,

$$\mathbf{E}(\mathbf{r}_0, t_0) = \frac{q}{4\pi\epsilon_0} \left[\frac{\hat{\mathbf{u}}_R}{|\mathbf{R}|^2} \right]_{ret.} - \frac{q}{4\pi\epsilon_0 c^2} \left[\frac{\mathbf{a}_\perp}{|\mathbf{R}|} \right]_{ret.}. \quad (5.34)$$

⁽⁴⁵⁾ L. Eyges, The classical electromagnetic field, Dover, New York 1972, p. 281 et R.P. Feynman, R.B. Leighton and M. Sands, The Feynman lectures on physics, vol. I chap. 28 et vol. II chap. 21, Addison-Wesley, Reading 1977.

3.1.4. Tenseur champ électromagnétique $F^{\mu\nu}$

Le formalisme tridimensionnel, on l'a vu au paragraphe précédent, est lourd à manipuler et les calculs sont finalement plus simples s'ils sont menés à l'aide du formalisme covariant⁽⁴⁶⁾. On calcule les dérivées à l'instant retardé

$$[\partial^\mu A^\nu]_{ret.} = \frac{\mu_0 q}{4\pi} \left[\frac{\partial}{\partial x_\mu} \left(\frac{v^\nu}{\rho} \right) \right]_{ret.},$$

où l'on a posé

$$\rho = v_\sigma R^\sigma / c.$$

On développe la dérivée

$$\begin{aligned} [\partial^\mu A^\nu]_{ret.} &= \frac{\mu_0 q}{4\pi} \left[\frac{1}{\rho} \frac{\partial v^\nu}{\partial x_\mu} + v^\nu \frac{\partial}{\partial x_\mu} \left(\frac{1}{\rho} \right) \right]_{ret.} \\ &= \frac{\mu_0 q}{4\pi} \left[\frac{1}{\rho} \partial^\mu v^\nu + v^\nu \partial^\mu \left(\frac{1}{\rho} \right) \right]_{ret.}, \end{aligned}$$

puis on utilise

$$\begin{aligned} \partial^\mu v^\nu &= \frac{\partial v^\nu}{\partial \tau} \frac{\partial \tau}{\partial x_\mu} = a^\nu \partial^\mu \tau \\ \partial^\mu \left(\frac{1}{\rho} \right) &= \frac{\partial}{\partial \rho} \left(\frac{1}{\rho} \right) \frac{\partial \rho}{\partial x_\mu} = -\frac{1}{\rho^2} \partial^\mu \rho, \end{aligned}$$

il vient ensuite

$$[\partial^\mu A^\nu]_{ret.} = \frac{\mu_0 q}{4\pi} \left([a^\nu \partial^\mu \tau]_{ret.} - \frac{1}{[\rho^2]_{ret.}} [\partial^\mu \rho]_{ret.} \right),$$

d'où l'on déduit le tenseur champ électromagnétique au point d'observation en fonction de la position retardée de la charge.

$$\begin{aligned} F^{\mu\nu}(x_0^\sigma) &= \frac{\mu_0 q}{4\pi} ([a^\nu \partial^\mu \tau]_{ret.} - [a^\mu \partial^\nu \tau]_{ret.} \\ &\quad - [\partial^\mu \rho]_{ret.}/[\rho^2]_{ret.} + [\partial^\nu \rho]_{ret.}/[\rho^2]_{ret.}). \end{aligned} \tag{5.35}$$

⁽⁴⁶⁾ On en trouve l'exposé détaillé dans F. Rohrlich, Classical charged particles, Addison-Wesley, Redwood 1965, section 4-8, p. 83.

3.2. Application au cas d'une charge ponctuelle en mouvement uniforme

On considère à nouveau dans cette section le cas d'une charge q animée d'un mouvement à vitesse constante \mathbf{v} dans un référentiel inertiel \mathcal{R} . On se propose de calculer les champs \mathbf{E} et \mathbf{B} par transformation de Lorentz cette fois. Supposons que la vitesse de la particule définisse l'axe Ox , $\mathbf{v} = v\mathbf{u}_x$. Dans le référentiel \mathcal{R}' , solidaire du mouvement de la charge et dont les axes coïncident à un instant donné avec ceux de \mathcal{R} , le champ est purement coulombien,

$$\mathbf{E}'(\mathbf{r}') = \frac{q}{4\pi\epsilon_0} \frac{\mathbf{r}'}{r'^3},$$

$$\mathbf{B}'(\mathbf{r}') = \vec{0}.$$

On sait par ailleurs passer des coordonnées exprimées dans \mathcal{R}' à celles dans \mathcal{R} par une simple transformation de Lorentz, puisque ces deux référentiels sont inertiels,

$$\begin{aligned} ct' &= \gamma_{\mathbf{v}}(ct - \beta_{\mathbf{v}}x) \\ x' &= \gamma_{\mathbf{v}}(-\beta_{\mathbf{v}}ct + x) \\ y' &= y \\ z' &= z \end{aligned}$$

de même que l'on sait transformer les composantes des champs,

$$\begin{aligned} E'_x &= E_x \\ E'_y &= \gamma_{\mathbf{u}}(E_y - \beta_{\mathbf{u}}cB_z) \\ E'_z &= \gamma_{\mathbf{u}}(E_z + \beta_{\mathbf{u}}cB_y) \\ B'_x &= B_x \\ B'_y &= \gamma_{\mathbf{u}}(B_y + (\beta_{\mathbf{u}}/c)E_z) \\ B'_z &= \gamma_{\mathbf{u}}(B_z - (\beta_{\mathbf{u}}/c)E_y). \end{aligned}$$

Calculons alors les composantes des champs dans \mathcal{R} à l'instant $t = 0$. Pour cela on exprime les transformations inverses,

$$\begin{aligned} E_x &= E'_x \\ E_y &= \gamma_{\mathbf{u}}(E'_y + \beta_{\mathbf{u}}cB'_z) \\ E_z &= \gamma_{\mathbf{u}}(E'_z - \beta_{\mathbf{u}}cB'_y) \end{aligned}$$

avec les composantes de \mathbf{B}' nulles, soit

$$\begin{aligned} E_x &= E'_x = \frac{q}{4\pi\epsilon_0} \frac{\gamma_{\mathbf{v}}x}{(\gamma_{\mathbf{v}}^2x^2 + y^2 + z^2)^{3/2}}, \\ E_y &= \gamma_{\mathbf{v}}E'_y = \gamma_{\mathbf{v}} \frac{q}{4\pi\epsilon_0} \frac{y}{(\gamma_{\mathbf{v}}^2x^2 + y^2 + z^2)^{3/2}}, \\ E_z &= \gamma_{\mathbf{v}}E'_z = \gamma_{\mathbf{v}} \frac{q}{4\pi\epsilon_0} \frac{z}{(\gamma_{\mathbf{v}}^2x^2 + y^2 + z^2)^{3/2}}. \end{aligned}$$

Si l'on note θ l'angle entre \mathbf{v} et la direction d'observation, $x = r \cos \theta$, $y^2 + z^2 = r^2 \sin^2 \theta$, soit

$$\gamma_{\mathbf{v}}^2 x^2 + y^2 + z^2 = \gamma_{\mathbf{v}}^2 r^2 (1 - \beta_{\mathbf{v}}^2 \sin^2 \theta)$$

et on peut écrire vectoriellement

$$\mathbf{E}(\mathbf{r}, 0) = \frac{q}{4\pi\epsilon_0} \frac{(1 - \beta_{\mathbf{v}}^2)\mathbf{r}}{r^3(1 - \beta^2 \sin^2 \theta)^{3/2}}. \quad (5.36)$$

On retrouve de manière directe le fait que le champ électrique est centré sur la position instantanée de la charge, sans effet de retard.

Le calcul du champ magnétique est plus simple si l'on utilise la forme vectorielle de la transformation des champs,

$$\mathbf{B}'_{\perp} = \gamma_{\mathbf{u}}(\mathbf{B}_{\perp} - (\boldsymbol{\beta}_{\mathbf{u}}/c) \wedge \mathbf{E}_{\perp}) = \vec{0},$$

soit

$$\mathbf{B}_{\perp} = (\boldsymbol{\beta}_{\mathbf{u}}/c) \wedge \mathbf{E}_{\perp} = \frac{1}{c^2} \mathbf{v} \wedge \mathbf{E}.$$

On a aussi $\mathbf{B}'_{\parallel} = \mathbf{B}_{\parallel} = \vec{0}$, d'où

$$\mathbf{B} = (\boldsymbol{\beta}_{\mathbf{u}}/c) \wedge \mathbf{E} = \frac{1}{c^2} \mathbf{v} \wedge \mathbf{E}. \quad (5.37)$$

3.3. Calcul approché du champ électrique rayonné

Une méthode simplifiée de calcul du champ électrique rayonné par la charge accélérée, due à Thomson ⁽⁴⁷⁾, est basée sur le résultat précédent, à savoir que dans la limite faiblement relativiste, le champ électrique est coulombien et centré sur la position instantanée de la charge. On considère alors la charge immobile à l'origine à $t = 0$. Elle subit une accélération \mathbf{a} constante pendant la durée infinitésimale Δt , acquiert ainsi une vitesse $\mathbf{v} = \mathbf{a}\Delta t$ à laquelle elle se déplace à vitesse constante jusqu'au temps t auquel on calcule le champ. On se place dans la limite $\Delta t \ll t$ et $|\mathbf{a}|\Delta t \ll c$.

À des distances inférieures à $r = ct$ (dans une *sphère d'information*), le champ est coulombien et centré sur la position actuelle de la charge comme on l'a dit. On note \mathbf{u}_{\parallel} et \mathbf{u}_{\perp} les vecteurs unitaires parallèlement et perpendiculairement à la direction d'observation (on introduit aussi les notations \mathbf{u}_r et \mathbf{u}_{θ}). On a ainsi

$$E_{\parallel} \simeq \frac{q}{4\pi\epsilon_0 r^2}.$$

Au-delà de la distance $c(t+\Delta t) \simeq ct$, le champ est également coulombien, mais centré sur la position originale de la particule car l'information qu'elle a été accélérée n'est pas encore parvenue dans cette région.

⁽⁴⁷⁾ M. Longair, *Theoretical concepts in physics*, Cambridge University Press, Cambridge 2003, p.306.

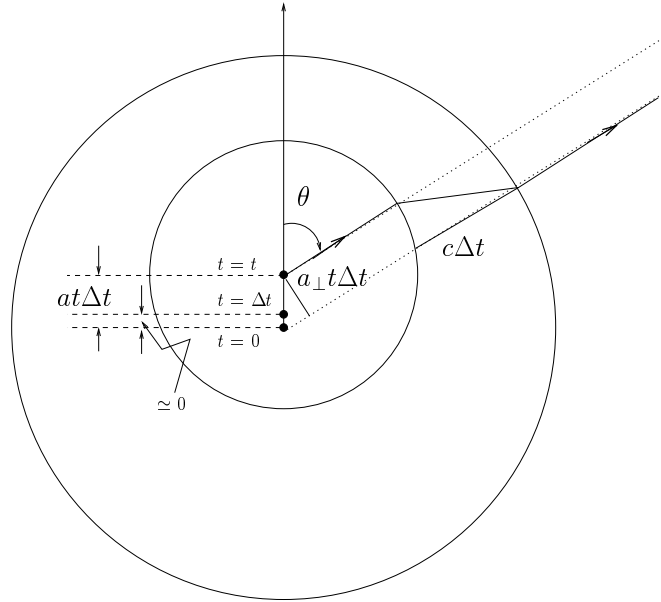


Figure 5.2 Charge immobile à $t = 0$, accélérée pendant un instant $\Delta t \ll t$, puis se déplaçant à vitesse constante jusqu'au temps t auquel on évalue le champ rayonné.

Une simple construction géométrique illustrée sur la figure ci-dessus conduit à la règle de proportionnalité

$$\frac{E_{\perp}}{E_{\parallel}} = \frac{a_{\perp} t \Delta t}{c \Delta t},$$

soit à l'ordre dominant

$$E_{\perp} = \frac{q}{4\pi\epsilon_0} \frac{a_{\perp}}{rc^2},$$

que l'on écrit en général en fonction de l'accélération retardée, puisqu'à l'instant t la charge n'est plus accélérée,

$$E_{\perp} = \frac{q}{4\pi\epsilon_0} \frac{a_{\perp}(t - r/c)}{rc^2}.$$

L'aspect rayonnement est illustré par la figure ci-dessous qui représente le développement temporel de la composante $a_{\perp} t \Delta t$ et sa dépendance angulaire (ici, Δt est totalement négligé).

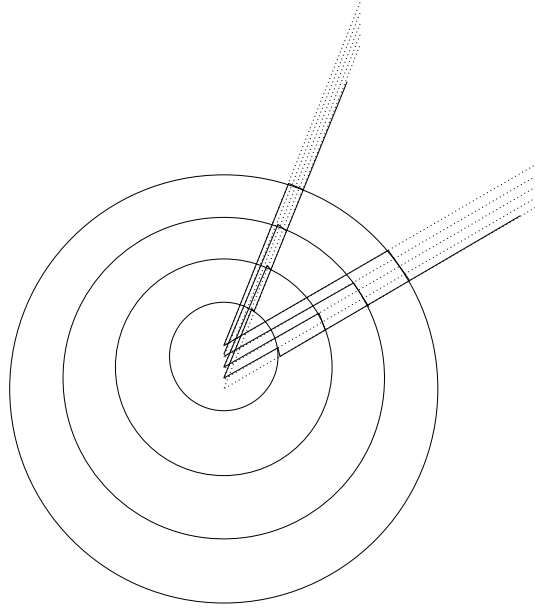


Figure 5.3 Dépendance temporelle et angulaire du champ rayonné.

Chapitre 6

Formulation lagrangienne de la relativité restreinte

Le principe de la méthode variationnelle⁽⁴⁸⁾ repose sur le choix d'une "famille de fonctions d'essais" supposée représenter une quantité physique à partir de laquelle on calcule une propriété déterminée. La minimisation de la valeur correspondante par rapport à un ou plusieurs paramètres variationnels fournit les conditions qui s'avèrent les plus voisines de la réalité physique parmi les fonctions d'essais choisies. La grandeur que l'on minimise est l'action S . Elle est définie comme une fonctionnelle du lagrangien. En mécanique usuelle, le lagrangien d'un système conservatif est donné par la différence entre énergie cinétique et énergie potentielle. Il n'a pas d'autre signification que de permettre de retrouver la dynamique du problème. Le lagrangien adapté à l'électromagnétisme (et donc équivalent aux équations de Maxwell) a été obtenu par Schwarzschild en 1903, avant la théorie de la relativité restreinte. Dans le cas d'une théorie des champs, comme l'électromagnétisme, on écrit le lagrangien comme l'intégrale spatiale d'une densité lagrangienne. En relativité, l'action étant un scalaire, elle doit se mettre sous la forme d'un invariant de Lorentz, c'est-à-dire qu'on doit pouvoir l'écrire à partir d'une contraction, ou plus précisément comme l'intégrale sur la ligne d'univers d'une contraction.

L'intérêt majeur de cette formulation réside dans son très grand degré de généralité, qui permet, dans un cadre unique, de traiter de nombreuses théories différentes.

⁽⁴⁸⁾ On pourra consulter L.D. Landau et E. Lifshitz, *Théorie des champs*, 3ème édition, Editions MIR, Moscou 1970, chap. II et III ; J.D. Jackson, *Classical electrodynamics*, Wiley, New York 1999 (3ème édition), §12.1 ; M. Lambert, *Relativité restreinte et électromagnétisme*, Ellipses, Paris 2000, chap. 8, 11 et 12, ou les notes de cours de David Sénéchal, *Ondes électromagnétiques*, §4.5.

1. Le principe de moindre action pour la particule libre

1.1. Le principe de moindre action en physique classique

On considère un système à n degrés de liberté $\alpha = 1, \dots, n$ décrit par les coordonnées $q_\alpha(t)$ et les vitesses $\dot{q}_\alpha(t)$. On cherche une équation régissant l'évolution du système d'un point $A(q_\alpha(t_a), t_a)$ à un point $B(q_\alpha(t_b), t_b)$. La dynamique du système est entièrement déterminée par une fonction L appelée lagrangien :

$$L(q_\alpha(t), \dot{q}_\alpha(t), t).$$

Le principe de moindre action d'Hamilton stipule que la trajectoire réelle du système est telle qu'une quantité $S[q_\alpha(t)]$, appelée action, est stationnaire :

$$S[q_\alpha(t)] = \int_{t_a}^{t_b} L(q_\alpha(t), \dot{q}_\alpha(t), t) dt, \quad (6.1)$$

$$\delta S[q_\alpha(t)] = 0.$$

C'est une fonctionnelle de $q_\alpha(t)$ et $\dot{q}_\alpha(t)$. Ce principe signifie que si l'on considère la trajectoire réelle notée $q_\alpha^0(t)$ entre A et B et une autre trajectoire joignant les points A et B et différant de la trajectoire réelle d'une quantité $\delta q_\alpha(t)$:

$$q_\alpha(t) = q_\alpha^0(t) + \delta q_\alpha(t),$$

alors la variation (fonctionnelle) de l'action est nulle au premier ordre en $\delta q_\alpha(t)$. On peut montrer qu'en fait la trajectoire réelle correspond à un minimum pour l'action ⁽⁴⁹⁾.

Par généralisation de la différentielle d'une fonction de plusieurs variables, $S(q_i, \dot{q}_i) = \sum_i L(q_i, \dot{q}_i)$,

$$dS(q_i, \dot{q}_i) = \sum_i \left(\frac{\partial L}{\partial q_i} dq_i + \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i \right),$$

la variation de S s'écrit

$$\delta S[q_\alpha(t)] = \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial q_\alpha} \delta q_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} \delta \dot{q}_\alpha \right) dt.$$

Les variables $q_\alpha(t)$ et $\dot{q}_\alpha(t)$ ne sont pas indépendantes, car entre t_a et t_b la connaissance de $q_\alpha(t)$ détermine $\dot{q}_\alpha(t)$. Pour éliminer $\delta \dot{q}_\alpha(t) = \frac{d}{dt}(\delta q_\alpha(t))$, on intègre par parties le second terme :

$$\int_{t_a}^{t_b} \frac{\partial L}{\partial \dot{q}_\alpha} \delta \dot{q}_\alpha dt = \left[\frac{\partial L}{\partial \dot{q}_\alpha} \delta q_\alpha(t) \right]_{t_a}^{t_b} - \int_{t_a}^{t_b} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) \delta q_\alpha dt.$$

⁽⁴⁹⁾ Corollaire : le calcul de l'action à partir d'une trajectoire différente conduit nécessairement à une borne supérieure pour S .

Le terme intégré est nul à cause des conditions aux bornes, $\delta q_\alpha(t_a) = \delta q_\alpha(t_b) = 0$, et il reste :

$$\delta S[q_\alpha(t)] = \int_{t_a}^{t_b} \left[\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) \right] \delta q_\alpha(t) dt.$$

Comme la variation $\delta q_\alpha(t)$ autour de la trajectoire réelle est arbitraire, pour assurer $\delta S = 0$, l'intégrand doit être nul, d'où l'on déduit les équations d'Euler – Lagrange, satisfaites par la trajectoire réelle $q_\alpha^0(t)$:

$$\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) = 0. \quad (6.2)$$

On obtient une équation différentielle semblable pour chaque degré de liberté. On emploie parfois une notation compacte,

$$\frac{\partial L}{\partial \mathbf{r}} - \frac{d}{dt} \left(\frac{\partial L}{\partial \mathbf{v}} \right) = \mathbf{0}. \quad (6.3)$$

Dans le cas d'une particule ponctuelle de masse m , d'énergie cinétique K , soumise à une énergie potentielle extérieure \mathcal{U} , le lagrangien s'écrit $L(\mathbf{r}, \mathbf{v}) = K(\mathbf{v}) - \mathcal{U}(\mathbf{r})$. On définit également l'impulsion

$$p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha}, \quad \mathbf{p} = \frac{\partial L}{\partial \mathbf{v}}, \quad (6.4)$$

et l'hamiltonien (égal à l'énergie totale pour un système conservatif)

$$\begin{aligned} H(p_\alpha(t), q_\alpha(t)) &= \sum_\alpha p_\alpha \dot{q}_\alpha - L(q_\alpha(t), \dot{q}_\alpha(t)) \\ &= \mathbf{p}\mathbf{v} - L(\mathbf{r}, \mathbf{v}). \end{aligned} \quad (6.5)$$

On réserve l'appellation d'hamiltonien lorsque l'énergie totale est exprimée en fonction de l'impulsion de la position (après élimination de la vitesse).

1.2. Le principe de moindre action pour la particule libre en relativité

1.2.1. L'action de la particule libre en dynamique relativiste

Partons de la particule libre en physique classique, $L(\mathbf{r}, \mathbf{v}) = K(\mathbf{v}) = \frac{1}{2}m|\mathbf{v}|^2$ où $\mathbf{v} = \frac{d\mathbf{r}}{dt}$, soit

$$S_{class.} = \frac{1}{2}m \int \mathbf{v} \frac{d\mathbf{r}}{dt} dt = \frac{1}{2}m \int \mathbf{v} d\mathbf{r}$$

et $\delta S_{class.} = 0$ s'écrit encore ⁽⁵⁰⁾

$$\delta \left(m \int \mathbf{v} \, d\mathbf{r} \right) = 0. \tag{6.6}$$

Or $\mathbf{v} \, d\mathbf{r}$ est relié à la partie spatiale du produit invariant ⁽⁵¹⁾ $v_\mu \, dx^\mu$, puisque

$$\begin{aligned} dx^\mu &= (c \, dt, \, d\mathbf{r}), \\ v_\mu &= \frac{dx_\mu}{d\tau} = \left(c \frac{dt}{d\tau}, \, -\frac{d\mathbf{r}}{d\tau} \right) = (c\gamma, \, -\gamma\mathbf{v}), \\ v_\mu \, dx^\mu &= \gamma c^2 \, dt - \gamma \mathbf{v} \, d\mathbf{r}. \end{aligned}$$

Comme $\gamma \rightarrow 1$ dans la limite non relativiste, on passe du cas classique au cas relativiste en faisant simplement

$$\delta \left(m \int \mathbf{v} \, d\mathbf{r} \right) = 0 \longrightarrow \delta \left(-m \int v_\mu \, dx^\mu \right) = 0. \tag{6.7}$$

On peut donner des formes équivalentes en notant que

$$\begin{aligned} -m v_\mu \, dx^\mu &= -m \frac{dx_\mu}{d\tau} \, dx^\mu \\ &= -m \frac{ds^2}{d\tau} \\ &= -mc \, d\tau \frac{ds}{d\tau} \\ &= -mc \, ds \end{aligned}$$

ce qui conduit à l'expression traditionnelle de l'action de la particule libre en relativité restreinte

$$S_{rel.} = -mc \int ds. \tag{6.8}$$

Elle est simplement donnée (au facteur numérique $-mc$ près) par la longueur de l'intervalle le long de la ligne d'univers de la particule libre. C'est une interprétation géométrique élégante où l'on voit que la partie cinétique de l'énergie est totalement intégrée dans la géométrie de l'espace-temps. Cette quantité par ailleurs est un scalaire de Lorentz, ce qui assure en particulier comme c'est un invariant que si $\delta S_{rel.} = 0$ dans un référentiel inertiel, ce sera également vrai dans tous les autres référentiels inertiels.

Tableau 6.1 Principe de moindre action pour la particule libre.

Dynamique newtonienne		Dynamique relativiste
$\delta \left(m \int \mathbf{v} \, d\mathbf{r} \right) = 0$	\rightarrow	$\delta \left(-m \int v_\mu \, dx^\mu \right) = 0$
		$\delta \left(-mc \int ds \right) = 0$

⁽⁵⁰⁾ On omet le facteur 1/2 que l'on retrouvera à partir de l'expression relativiste par le développement limité d'une racine carrée.

⁽⁵¹⁾ Le fait qu'il s'agisse d'une forme invariante par changement de référentiel inertiel est bien entendu capital pour assurer la validité de la forme de l'action dans tous ces référentiels équivalents.

1.2.2. L'action en notation tridimensionnelle

En notation tridimensionnelle explicite on développe

$$\begin{aligned} S_{rel.} &= -mc \int ds \\ &= -mc \int (c^2 dt^2 - d\mathbf{r}^2)^{1/2} \\ &= -mc^2 \int dt \sqrt{1 - |\mathbf{v}|^2/c^2}, \end{aligned}$$

ce qui conduit au lagrangien

$$\begin{aligned} L &= -mc^2 \sqrt{1 - |\mathbf{v}|^2/c^2} \\ &\simeq -mc^2 \left(1 - \frac{1}{2} \frac{|\mathbf{v}|^2}{c^2} + \dots \right) \\ &\simeq -mc^2 + \frac{1}{2} m |\mathbf{v}|^2 + \dots \end{aligned}$$

A la constante $-mc^2$ près ⁽⁵²⁾ on retrouve bien le lagrangien classique de la particule libre. On peut alors calculer l'impulsion (qui confirme la forme induite lors du chapitre sur la dynamique relativiste) :

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - |\mathbf{v}|^2/c^2}}, \quad (6.9)$$

et l'énergie totale

$$E = \frac{mc^2}{\sqrt{1 - |\mathbf{v}|^2/c^2}}. \quad (6.10)$$

Cette dernière quantité s'exprime également en fonction de l'impulsion, on parle alors de l'hamiltonien,

$$H = c\sqrt{|\mathbf{p}|^2 + m^2c^2}. \quad (6.11)$$

L'équation d'Euler-Lagrange ici donne simplement

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \mathbf{v}} \right) = \frac{d\mathbf{p}}{dt} = \mathbf{0},$$

comme attendu pour la particule libre.

⁽⁵²⁾ Cette constante ne change rien bien entendu à l'extrémisation de l'action.

1.2.3. Le principe de moindre action dans le formalisme de Minkowski

On revient à l'expression de l'action exprimée sous forme manifestement covariante,

$$\begin{aligned} S_{rel.} &= -m \int v_\mu dx^\mu \\ &= -mc \int ds \\ &= -mc \int (dx_\mu dx^\mu)^{1/2}, \end{aligned}$$

soit

$$\delta S_{rel.} = -mc \int \delta(dx_\mu dx^\mu)^{1/2},$$

avec

$$\begin{aligned} \delta(dx_\mu dx^\mu)^{1/2} &= \frac{1}{2}(dx_\mu dx^\mu)^{-1/2} \underbrace{[\delta(dx_\mu) dx^\mu + dx_\mu \delta(dx^\mu)]}_{2 dx^\mu \delta(dx_\mu)}, \\ &= \frac{dx^\mu}{ds} \delta(dx_\mu) \\ &= \frac{1}{c} v^\mu \delta(dx_\mu). \end{aligned}$$

On a de plus $\delta(dx_\mu) = d(\delta x_\mu)$, soit

$$\begin{aligned} \delta S_{rel.} &= -m \int_a^b v^\mu d(\delta x_\mu) \\ &= -m [v^\mu \delta x_\mu]_a^b + m \int_a^b dv^\mu \delta x_\mu. \end{aligned}$$

Ici, a et b repèrent les coordonnées d'espace-temps de deux événements entre lesquels on minimise l'action et le terme de bords s'annule. Il est naturel de paramétrer la ligne d'univers par le temps propre, soit

$$\delta S_{rel.} = m \int_a^b \frac{dv^\mu}{d\tau} d\tau \delta x_\mu = 0$$

et cette égalité étant valable pour une variation arbitraire de la trajectoire d'univers, on a

$$a^\mu = \frac{dv^\mu}{d\tau} = \frac{d^2 x^\mu}{d\tau^2} = 0, \quad (6.12)$$

c'est-à-dire une généralisation évidente de l'équation classique de la trajectoire de la particule libre.

On peut noter que s'il s'agit de photons, $d\tau$ n'est pas défini (puisque $ds = 0$), et dans ce cas on introduit un paramètre arbitraire λ qui évolue le long de la ligne d'univers et

$$\begin{aligned} \delta S_{rel.} &= m \int_a^b \frac{dv^\mu}{d\lambda} d\lambda \delta x_\mu = 0, \\ \frac{dv^\mu}{d\lambda} &= \frac{d^2 x^\mu}{d\lambda^2} = 0. \end{aligned} \quad (6.13)$$

2. Particule chargée sous champ

2.1. Action invariante

On doit tout d'abord former une action invariante de Lorentz. Pour cela, la seule caractéristique électromagnétique de la particule étant sa charge et la seule caractéristique du champ étant le quadripotential, on forme la quantité $qA_\mu dx^\mu$, soit

$$S_q = \int_a^b (-mc ds - qA_\mu dx^\mu)$$

où le signe est justifié par la limite non relativiste et où

$$A^\mu = (\phi/c, \mathbf{A}).$$

On écrit encore

$$S_q = \int_a^b (-mc(dx_\mu dx^\mu)^{1/2} - qA_\mu dx^\mu). \quad (6.14)$$

2.2. Calcul variationnel en notation tridimensionnelle

Sous forme explicite, on a

$$\begin{aligned} S_q &= \int_a^b \left(-mc^2 dt \sqrt{1 - |\mathbf{v}|^2/c^2} - q \frac{\phi}{c} c dt + q \mathbf{A} \cdot d\mathbf{r} \right) \\ &= \int_a^b \left(-mc^2 \sqrt{1 - |\mathbf{v}|^2/c^2} - q\phi + q \mathbf{A} \cdot \mathbf{v} \right) dt \end{aligned}$$

ce qui conduit au lagrangien relativiste

$$L = -mc^2 \sqrt{1 - |\mathbf{v}|^2/c^2} - q\phi + q \mathbf{A} \cdot \mathbf{v}. \quad (6.15)$$

Dans la limite non relativiste, il devient

$$L \simeq -mc^2 + \frac{1}{2}m|\mathbf{v}|^2 - q\phi + q \mathbf{A} \cdot \mathbf{v}.$$

L'impulsion

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1 - |\mathbf{v}|^2/c^2}} + q\mathbf{A} \quad (6.16)$$

s'exprime en fonction de la quantité de mouvement

$$\boldsymbol{\pi} = \frac{m\mathbf{v}}{\sqrt{1 - |\mathbf{v}|^2/c^2}} \quad (6.17)$$

par

$$\mathbf{p} = \boldsymbol{\pi} + q\mathbf{A}. \quad (6.18)$$

L'hamiltonien prend la forme

$$\begin{aligned} H &= \mathbf{v} \frac{\partial L}{\partial \mathbf{v}} - L \\ &= \mathbf{v}(\boldsymbol{\pi} + q\mathbf{A}) + mc^2 \sqrt{1 - |\mathbf{v}|^2/c^2} + q\phi - q\mathbf{A}\mathbf{v} \\ &= \frac{mc^2}{\sqrt{1 - |\mathbf{v}|^2/c^2}} + q\phi. \end{aligned} \quad (6.19)$$

De $\boldsymbol{\pi} = \mathbf{p} - q\mathbf{A}$ on déduit $\pi^\mu = p^\mu - qA^\mu$ dont le carré invariant donne $\pi_\mu \pi^\mu = mv_\mu mv^\mu = m^2 c^2$ et de $p^\mu - qA^\mu = ((H - q\phi)/c, \mathbf{p} - q\mathbf{A})$ on déduit de même $(p_\mu - qA_\mu)(p^\mu - qA^\mu) = (H - q\phi)^2/c^2 - |\mathbf{p} - q\mathbf{A}|^2$, soit finalement l'expression invariante

$$(H - q\phi)^2 = m^2 c^4 + (\mathbf{p} - q\mathbf{A})^2 c^2.$$

L'hamiltonien devient

$$H = \sqrt{m^2 c^4 + c^2 (\mathbf{p} - q\mathbf{A})^2} + q\phi. \quad (6.20)$$

Aux faibles vitesses on peut simplifier

$$\mathbf{p} - q\mathbf{A} \simeq m\mathbf{v},$$

soit

$$\begin{aligned} H &\simeq mc^2 \left(1 + \frac{\mathbf{v}^2}{2c^2} \right) + q\phi \\ &= mc^2 + \frac{1}{2} m\mathbf{v}^2 + q\phi \\ &= mc^2 + \frac{1}{2m} (\mathbf{p} - q\mathbf{A})^2 + q\phi \end{aligned}$$

ce qui constitue une expression non relativiste fréquemment utilisée.

Notons enfin la forme des équations d'Euler-Lagrange appliquées au lagrangien relativiste (6.15).

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \mathbf{v}} \right) = \frac{d\mathbf{p}}{dt} = \frac{d}{dt} (\boldsymbol{\pi} + q\mathbf{A}),$$

$$\frac{\partial L}{\partial \mathbf{r}} = q\vec{\nabla}(\mathbf{A}\mathbf{v}) - q\vec{\nabla}\phi$$

On utilise ensuite $\vec{\nabla}(\mathbf{A}\mathbf{v}) = (\mathbf{A}\vec{\nabla})\mathbf{v} + (\mathbf{v}\vec{\nabla})\mathbf{A} + \mathbf{v} \wedge (\vec{\nabla} \wedge \mathbf{A}) + \mathbf{A} \wedge (\vec{\nabla} \wedge \mathbf{v})$ pour écrire

$$\frac{\partial L}{\partial \mathbf{r}} = q \underbrace{(\mathbf{v}\vec{\nabla})\mathbf{A}}_{\frac{d\mathbf{A}}{dt} - \frac{\partial \mathbf{A}}{\partial t}} + q\mathbf{v} \wedge (\vec{\nabla} \wedge \mathbf{A}) - q\vec{\nabla}\phi$$

d'où découle l'équation fondamentale de la dynamique

$$\frac{d\boldsymbol{\pi}}{dt} = q \left(-q\vec{\nabla}\phi - \frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \wedge \mathbf{rot} \mathbf{A} \right)$$

qui montre que l'expression de la force de Lorentz est une conséquence de la relativité et également que c'est une relation exacte.

2.3. Calcul variationnel dans le formalisme de Minkowski

On exprime cette fois la variation de l'action sous la forme

$$\begin{aligned}\delta S_q &= \int_a^b (-mc\delta(dx_\mu dx^\mu)^{1/2} - q\delta(A_\mu dx^\mu)) \\ &= -m[v_\mu \delta x^\mu]_a^b + \int_a^b m dv_\mu \delta x^\mu - q \int_a^b (A_\mu d\delta x^\mu + \delta A_\mu dx^\mu)\end{aligned}$$

Or on a aussi

$$\int_a^b A_\mu d\delta x^\mu = [A_\mu \delta x^\mu]_a^b - \int_a^b dA_\mu \delta x^\mu$$

de sorte qu'il reste finalement

$$\begin{aligned}\delta S_q &= -m[v_\mu \delta x^\mu]_a^b - q[A_\mu \delta x^\mu]_a^b \\ &\quad + \int_a^b (m dv_\mu \delta x^\mu + q dA_\mu \delta x^\mu - q\delta A_\mu dx^\mu).\end{aligned}$$

Les termes intégrés disparaissent car les variations δx^μ sont nulles aux extrémités de la ligne d'univers. En utilisant de plus

$$\begin{aligned}\delta A_\mu &= \frac{\partial A_\mu}{\partial x^\nu} \delta x^\nu, \\ dA_\mu &= \frac{\partial A_\mu}{\partial x^\nu} dx^\nu,\end{aligned}$$

il vient

$$\begin{aligned}\delta S_q &= \int_a^b (m dv_\mu \delta x^\mu + q \frac{\partial A_\mu}{\partial x^\nu} dx^\nu \delta x^\mu - q \frac{\partial A_\mu}{\partial x^\nu} \delta x^\nu dx^\mu) \\ &= \int_a^b \left(m \frac{dv_\mu}{d\tau} - q \left(\frac{\partial A_\nu}{\partial x^\mu} v^\nu - \frac{\partial A_\mu}{\partial x^\nu} v^\nu \right) \right) d\tau \delta x^\mu,\end{aligned}$$

d'où les équations du mouvement sous forme tensorielle

$$m \frac{dv_\mu}{d\tau} = q(\partial_\mu A_\nu - \partial_\nu A_\mu)v^\nu \quad (6.21)$$

ou encore en faisant apparaître le tenseur champ électromagnétique,

$$\begin{aligned}F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}, \\ ma_\mu &= qF_{\mu\nu}v^\nu.\end{aligned} \quad (6.22)$$

2.4. Equations d'Euler-Lagrange covariantes

Le calcul précédent conduit automatiquement à des équations analogues aux équations classiques d'Euler-Lagrange. Il est utile d'écrire directement ces équations sous forme covariante. Partant de l'action

$$S = \int L_\mu dx^\mu = \int L d\lambda,$$

où $L = L_\mu dx^\mu / d\lambda$, le paramètre scalaire λ pouvant être le temps propre τ , celui mesuré par un observateur t , ou encore l'intervalle s . On pose ici $\dot{x}^\mu = \frac{dx^\mu}{d\lambda}$, même si le paramètre λ n'est pas nécessairement le temps ici ⁽⁵³⁾ et les équations d'Euler-Lagrange prennent la forme

$$\frac{d}{d\lambda} \left(\frac{\partial L}{\partial \dot{x}^\mu} \right) - \frac{\partial L}{\partial x^\mu} = 0.$$

Dans le cas d'une charge q , de masse m , en présence d'un champ électromagnétique, l'action prend la forme

$$S = \int (-mv_\mu dx^\mu - qA_\mu dx^\mu),$$

de sorte que le lagrangien est donné par

$$L = (-mv_\mu - qA_\mu)\dot{x}^\mu.$$

On en déduit l'impulsion covariante ⁽⁵⁴⁾,

$$p_\mu = -\frac{\partial L}{\partial \dot{x}^\mu} = mv_\mu + qA_\mu.$$

On peut noter que cette expression est conforme à ce que nous avons déjà obtenu préalablement, puisque

$$p_\mu = (E/c, -\mathbf{p}) = m(\gamma c, -\gamma \mathbf{v}) + q(\phi/c, -\mathbf{A}).$$

L'équation d'Euler-Lagrange ci-dessus peut s'écrire ⁽⁵⁵⁾

$$\dot{p}_\mu = -\frac{\partial L}{\partial x^\mu},$$

⁽⁵³⁾ Nous introduisons ici une notation non conventionnelle pour l'action $\int L_\mu dx^\mu$ qui permet d'identifier facilement $L = L_\mu \dot{x}^\mu$ mais n'a pas davantage d'intérêt.

⁽⁵⁴⁾ Le signe $-$ dans la définition de p_μ est nécessaire pour retrouver les expressions classiques dans la limite non relativiste. On peut en comprendre l'origine en notant qu'il apparaît forcément un signe négatif dans les composantes spatiales, soit de p_μ , soit, si l'on calcule plutôt $p^\mu = -\frac{\partial L}{\partial \dot{x}_\mu}$, dans celles de x_μ au dénominateur.

⁽⁵⁵⁾ Même remarque à propos du signe qui découle ici de la définition choisie pour p_μ .

où $\dot{p}_\mu = m\dot{v}_\mu + q\dot{A}_\mu$ s'exprime encore plus explicitement en notant que $dA_\mu = \frac{\partial A_\mu}{\partial x^\nu} dx^\nu$, soit $\dot{A}_\mu = \frac{\partial A_\mu}{\partial x^\nu} \dot{x}^\nu$ et finalement

$$\dot{p}_\mu = m\dot{v}_\mu + q\partial_\nu A_\mu \dot{x}^\nu.$$

Au second membre il vient

$$\frac{\partial L}{\partial x^\mu} = -q\partial_\mu A_\nu \dot{x}^\nu,$$

en mettant à profit un changement d'indice $\mu \rightarrow \nu$ dans la contraction, effectuée sur un indice muet, du dernier terme de L . L'équation du mouvement devient immédiatement

$$m\dot{v}_\mu = -q\partial_\nu A_\mu \dot{x}^\nu + \partial_\mu A_\nu \dot{x}^\nu = -qF_{\nu\mu} \dot{x}^\nu = qF_{\mu\nu} \dot{x}^\nu.$$

On a mis à profit l'antisymétrie du tenseur champ électromagnétique dans le dernier terme.

2.5. Invariance de jauge

La définition de l'action d'une charge q sous champ est fondée sur la nécessité d'avoir un scalaire de Lorentz. On peut s'interroger sur le caractère général de l'expression obtenue (rappelons qu'elle fournit la bonne limite classique, mais est-ce la seule possibilité ?) Si l'on modifie le quadripotential par la transformation (appelée transformation de jauge)

$$A_\mu \rightarrow A'_\mu + \partial_\mu \chi, \quad (6.23)$$

l'action devient

$$\begin{aligned} S_q \rightarrow S'_q &= \int_a^b (-mc(dx_\mu dx^\mu)^{1/2} - qA_\mu dx^\mu - q\partial_\mu \chi dx^\mu) \\ &= S_q - q \int_a^b \frac{\partial \chi}{\partial x^\mu} dx^\mu \\ &= S_q - q \int_a^b d\chi \\ &= S_q - q(\chi(b) - \chi(a)) \end{aligned}$$

et ne varie que d'une quantité globale indépendante du chemin suivi entre a et b ⁽⁵⁶⁾. La minimisation de S'_q est donc équivalente à celle de S_q et la physique est invariante par changement de jauge. En notation tridimensionnelle, ce changement de jauge s'écrit

$$A'_\mu = (\phi'/c, -\mathbf{A}') = (\phi/c, -\mathbf{A}) + \left(\frac{1}{c} \frac{\partial \chi}{\partial t}, \vec{\nabla} \chi \right)$$

soit

$$\begin{aligned} \phi &\rightarrow \phi' = \phi + \frac{\partial \chi}{\partial t} \\ \mathbf{A} &\rightarrow \mathbf{A}' = \mathbf{A} - \vec{\nabla} \chi, \end{aligned} \quad (6.24)$$

et il laisse la physique du problème inchangée.

⁽⁵⁶⁾ Il est indispensable que le quadri-potential soit modifié par un terme en $\partial_\mu \chi$ pour l'action fasse intervenir une différentielle.

2.6. Densité lagrangienne

De manière générale, l'action S , scalaire invariant, peut s'écrire comme une intégrale sur un 4-élément de volume,

$$S = \int d^4x \mathcal{L}(x_\mu, \dot{x}_\mu)$$

où $d^4x = d^3r dt$ est l'élément de volume de l'espace-temps et $\mathcal{L}(x_\mu, \dot{x}_\mu)$ est la densité lagrangienne. Le lagrangien est donné par l'intégrale spatiale de la densité lagrangienne,

$$L = \int d^3r \mathcal{L}(x_\mu, \dot{x}_\mu).$$

L'intérêt de la densité $\mathcal{L}(x_\mu, \dot{x}_\mu)$ est qu'il s'agit d'un scalaire invariant. En effet, d^4x est un invariant car dans $d^4x = d^3r dt = d^2S_\perp dl_\parallel dt$, les sections transverses sont invariables, alors que le temps et les longueurs longitudinales ont des lois de transformation qui se compensent parfaitement.

Considérons maintenant une charge élémentaire $dq = \rho d^3r$ dans un champ électromagnétique. L'action comprend un terme d'interaction charge-champ que nous connaissons,

$$\begin{aligned} S_{\text{int.}} &= - \int dq A_\mu dx^\mu \\ &= - \int dq(\phi - \mathbf{A}\mathbf{v}) dt \\ &= - \int \rho(\phi - \mathbf{A}\mathbf{v}) d^3r dt \\ &= - \int j_\mu A^\mu d^4x \\ &= \int \mathcal{L}_{\text{int.}} d^4x. \end{aligned}$$

Il y a également un terme décrivant le champ libre. Par analogie avec la dépendance en x_μ et \dot{x}_μ de la densité lagrangienne de la particule libre, ce terme doit faire intervenir a priori une densité lagrangienne de la forme $\mathcal{L}_{\text{chp}}(A_\mu, \partial_\nu A_\mu)$. Comme les équations de Maxwell sont linéaires, il faut au plus une dépendance quadratique du lagrangien en fonction des champs. On forme la quantité $F_{\mu\nu}F^{\mu\nu}$ et pour être compatible avec les équations de Maxwell, on anticipe le préfacteur, soit

$$\mathcal{L}_{\text{chp}} = -\frac{1}{4\mu_0} F_{\mu\nu}F^{\mu\nu}.$$

Au total on écrit l'action comme

$$S = - \int \left(\frac{1}{4\mu_0} F_{\mu\nu}F^{\mu\nu} + j^\mu A_\mu \right) d^4x. \quad (6.25)$$

Cela permet bien de retrouver des équations de Maxwell car

$$\delta(F_{\mu\nu}F^{\mu\nu}) = 2F^{\mu\nu}\delta F_{\mu\nu} = 4F^{\mu\nu}\partial_\mu\delta A_\nu,$$

soit, après intégration par parties et élimination des termes de bords,

$$\delta S = - \int d^4x [j^\mu\delta A_\mu - \frac{1}{4\mu_0}4\partial_\mu F^{\mu\nu}\delta A_\nu]$$

ou, en changeant les indices du premier terme,

$$\partial_\mu F^{\mu\nu} = \mu_0 j^\nu.$$

L'autre partie des équations de Maxwell (sans sources) provient de la définition même du tenseur de Faraday $F^{\mu\nu}$.

Il est utile d'écrire la densité lagrangienne sous forme tridimensionnelle. En l'absence de charges, la densité lagrangienne du champ électromagnétique s'écrit

$$\frac{1}{2}\varepsilon_0\mathbf{E}^2 - \frac{1}{2\mu_0}\mathbf{B}^2.$$

Le terme densité se justifie, puisque pour obtenir une grandeur extensive on intègre sur l'espace ⁽⁵⁷⁾. En présence de charges (immobiles ou en mouvement), la densité lagrangienne s'écrit (Schwarzschild)

$$\mathcal{L} = \frac{1}{2}\varepsilon_0\mathbf{E}^2(\mathbf{r}) - \frac{1}{2\mu_0}\mathbf{B}^2(\mathbf{r}) + \rho(\mathbf{r})(\mathbf{A}(\mathbf{r}) \cdot \mathbf{v} - \phi(\mathbf{r})) \quad (6.26)$$

où \mathbf{v} est la vitesse des porteurs de charge, $\rho(\mathbf{r})$ la densité de charges et $\mathbf{j}(\mathbf{r}) = \rho(\mathbf{r})\mathbf{v}$ la densité de courant. En fonction des potentiels, on a

$$\mathcal{L} = \frac{1}{2}\varepsilon_0 \left(-\vec{\nabla}\phi - \frac{\partial\mathbf{A}}{\partial t} \right)^2 - \frac{1}{2\mu_0}(\mathbf{rot}\mathbf{A})^2 + \rho(\mathbf{A} \cdot \mathbf{v} - \phi). \quad (6.27)$$

Si on note de manière générique ψ_j pour ϕ , A_x , A_y et A_z , les quatre champs scalaires, \mathcal{L} est une fonction de ces champs et de leurs dérivées spatiales et temporelles

$$\mathcal{L} \left(\psi_j, \dot{\psi}_j, \frac{\partial\psi_j}{\partial x}, \frac{\partial\psi_j}{\partial y}, \frac{\partial\psi_j}{\partial z} \right).$$

La nouveauté pour une théorie des champs, par rapport à un système ayant un nombre fini de degrés de libertés, réside dans l'apparition des dérivées spatiales dans les équations d'Euler-Lagrange

$$\frac{\partial\mathcal{L}}{\partial\psi_j} - \sum_{\alpha=x,y,z} \frac{\partial}{\partial\alpha} \left[\frac{\partial\mathcal{L}}{\partial(\frac{\partial\psi_j}{\partial\alpha})} \right] - \frac{d}{dt} \left(\frac{\partial\mathcal{L}}{\partial\dot{\psi}_j} \right) = 0. \quad (6.28)$$

Appliquées aux composantes du quadripotentiel, les équations d'Euler – Lagrange permettent de retrouver les équations de propagation pour le potentiel scalaire et les composantes du potentiel vecteur. Postuler la minimisation de l'action définie par la densité lagrangienne de Schwarzschild conduit donc aux mêmes équations de propagation que les équations de Maxwell. Cette formulation de l'électromagnétisme est rarement employée en raison de son caractère assez axiomatique. Elle présente toutefois un grand intérêt physique, puisque de nombreux domaines de la physique peuvent être décrits par des principes variationnels qui apparaissent ainsi comme un contexte unificateur.

⁽⁵⁷⁾ $\frac{1}{2}\varepsilon_0\mathbf{E}^2$ par exemple est la densité d'énergie électrique associée à l'existence d'un champ électrique.

2.7. Généralisation

On a vu que la forme de l'action avait un contenu facile à interpréter,

$$S = \int L_\mu dx^\mu,$$

où $L_\mu = -mv_\mu - qA_\mu$, le premier terme donnant la contribution de l'énergie cinétique et le second le couplage (q) à un potentiel externe (A_μ). On omet ici la contribution du champ seul, $-\frac{1}{4\mu_0} \int F_{\mu\nu} F^{\mu\nu} d^4x$. On peut a priori généraliser à des interactions arbitraires, décrites par exemple par des potentiels de rang 2 couplés par une constante g , soit $L_\mu = -mv_\mu - qA_\mu - gB_{\mu\nu}v^\nu, \dots$. Les équations d'Euler-Lagrange donnent ensuite une dynamique que l'on peut comparer à l'expérience si on cherche à lui donner un sens physique. On peut ainsi former à volonté des lagrangiens très généraux satisfaisant au principe de covariance relativiste et étudier ensuite la dynamique qu'ils génèrent.

Chapitre 7

Géométrie des surfaces courbes

Nous allons introduire dans ce chapitre quelques éléments de géométrie dans des espaces non euclidiens (c'est-à-dire pourvus d'une courbure intrinsèque). Pour cela, l'essentiel des calculs est mené dans le cas simplifié des surfaces sphériques et les résultats sont supposés valables plus généralement. On peut consulter Kenyon ou Hakim pour davantage d'informations.

1. Le cinquième postulat d'Euclide

Les fondements de la géométrie ont été établis par Euclide d'Alexandrie au quatrième siècle BC dans les *Eléments*, un ouvrage de treize volumes dont les six premiers sont consacrés à la géométrie plane. Euclide y présente environ 400 propositions mathématiques et théorèmes basés sur un ensemble de dix axiomes ou postulats supposés évidents, les axiomes d'Euclide :

- i) Etant donnés deux points, il existe une droite qui les joint.
- ii) Une droite peut être prolongée indéfiniment.
- iii) Un cercle peut toujours être construit si l'on se donne son centre et un point de la circonférence.
- iv) Tous les angles droits sont égaux.
- v) Etant donnés une droite et un point, il existe une unique droite passant par ce point et ne coupant pas la droite initiale.

complétés par les notions logiques,

- vi) Des choses égales à une même troisième sont égales entre elles.
- vii) Si des choses égales sont ajoutées à des choses égales, les sommes sont encore égales.

- viii) Si des choses égales sont retranchées à des choses égales, les différences sont encore égales.
- ix) Les choses qui coïncident sont égales.
- x) Le tout est plus grand que la partie.

Parmi les postulats d'Euclide, le cinquième, appelé postulat de la parallèle ou simplement postulat d'Euclide, a été très largement étudié au cours des siècles dans l'idée de le déduire logiquement des quatre premières propositions⁽⁵⁸⁾. Ces tentatives ont toutes été vouées à l'échec, mais de nombreuses formulations alternatives ont été proposées, notamment "dans un triangle, la somme des angles est égale à π " due à Legendre⁽⁵⁹⁾.

La remise en question du cinquième postulat a conduit au développement des géométries non euclidiennes⁽⁶⁰⁾, a priori moins naturelles, mais qui ont trouvé un cadre d'application inattendu dans la théorie de la Relativité Générale.

Carl Friedrich Gauss (1777-1855) est considéré comme le précurseur des géométries non euclidiennes et en particulier l'inventeur de la géométrie elliptique dans laquelle il n'y a pas de parallèles, mais où les autres postulats d'Euclide sont vrais. Une représentation de cette géométrie consiste à définir les points comme étant répartis sur la surface d'une sphère (ce sont les intersections des diamètres de la sphère avec sa surface), et les "lignes", pour généraliser les droites, (on dit maintenant géodésiques), comme les intersections de la surface de la sphère avec les plans contenant le centre de la sphère. Deux points définissent alors de façon unique une ligne et un point est toujours donné par deux lignes. Cependant, dans cette géométrie, si l'on se donne une ligne AB et un point P , il n'existe aucune ligne passant par P et ne coupant pas AB (figure 7.1).

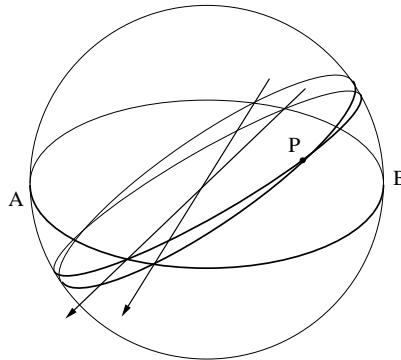


Figure 7.1 Géométrie de Gauss : le cinquième postulat d'Euclide n'est pas satisfait car en P on ne peut tracer aucune parallèle à AB .

⁽⁵⁸⁾ Tentatives par Ptolémée (?-168), Proclus (410-485), Nasir al din al Tusi (13^{ème} siècle), Levi ben Gerson (1288-1344), P. A. Cataldi (1548-1626), Giovanni Alfonso Borelli (1608-1679), Giordano Vitale (1633-1711), John Wallis (1616-1703), Gerolamo Saccheri (1667-1733), Johann Heinrich Lambert (1728-1777), Adrien-Marie Legendre (1752-1833).

⁽⁵⁹⁾ On trouve de nombreux détails historiques dans S. Weinberg, *Gravitation and Cosmology* (Wiley, New-York 1972), pp. 4-11.

⁽⁶⁰⁾ Article "Geometry", Mac Graw Hill Encyclopedia of Science and Technology, Article "Geometry", The New Encyclopaedia Britannica, Article "Géométrie", Encyclopédie Internationale des Sciences et Techniques.

Plus tard, Georg Friedrich Bernhard Riemann (1826-1866) a généralisé les travaux de Gauss à des surfaces courbes de dimension quelconque. Les points P sont spécifiés par des coordonnées x^i (comme latitude et longitude à la surface de la Terre), et le carré de la distance s^2 entre deux points est donné par une forme quadratique $s^2 = \int g_{ij} dx^i dx^j$, sommée sur les indices de 1 à n . Les fonctions g_{ij} sont les composantes du tenseur métrique et varient d'un point à un autre. Dans un changement de coordonnées, les g_{ij} sont transformées, mais les trajectoires géodésiques, ainsi qu'une combinaison des dérivées secondes des g_{ij} appelée courbure, sont indépendantes du choix des coordonnées et décrivent les propriétés intrinsèques de l'espace.

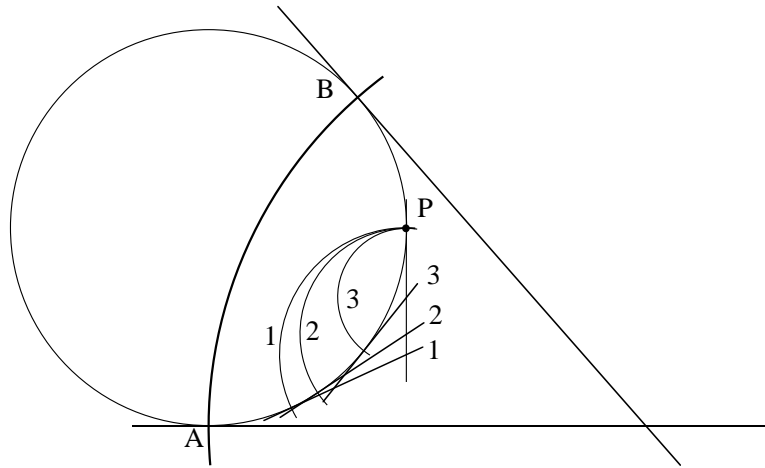


Figure 7.2 Géométrie de Bolyai et Lobachevski : étant donné une ligne et un point en dehors de cette ligne, on peut construire une infinité de lignes distinctes qui ne coupent pas la première.

Un autre type de géométrie non euclidienne (géométrie hyperbolique) a été développé par Janos Bolyai (1802-1860) et Nikolai Ivanovich Lobachevski (1793-1856) vers 1830. Le cinquième axiome d'Euclide n'y est pas non plus vérifié, car, en un point en dehors d'une ligne, on peut mener une infinité de lignes distinctes ne coupant pas la première ligne. On peut se représenter cette géométrie en définissant les lignes comme les arcs de cercle qui coupent à angle droit un grand cercle donné. Les points sont situés sur le grand cercle. Il est alors facile de voir que le cinquième postulat n'est pas satisfait (figure 7.2) : on peut en effet, en un point P extérieur à la ligne AB , construire une infinité d'arcs de cercle qui couperont perpendiculairement le grand cercle, sans pour autant couper AB . En un point P en dehors de la ligne AB , on peut construire des arcs de cercles tels que ceux qui sont représentés sur la figure 7.2.

2. Courbure des surfaces bidimensionnelles

2.1. Géodésiques et vols intercontinentaux

Considérons la surface bidimensionnelle d'une sphère de rayon R . Etant donnés deux points B et C , on cherche la plus courte distance s mesurée sur la sphère entre B et C . La courbe obtenue est une géodésique, notion qui généralise, pour une surface arbitraire, la notion de droite du plan.

La longueur d'une courbe de l'espace tridimensionnel euclidien est toujours supérieure ou égale à la longueur de toute projection plane de cette courbe. La courbe géodésique est donc nécessairement une courbe plane. Etant donnés B et C fixés, on choisit l'axe Oz sur la bissectrice de \widehat{BOC} et on compare les longueurs $s_1 = BC_1$ et $s_2 = BC_2$ des courbes planes du plan BOC et du plan perpendiculaire (figure 7.3).

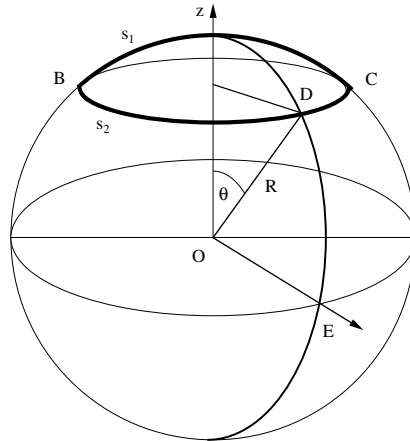


Figure 7.3 Géodésiques sur la sphère : comparaison entre les longueurs s_1 et s_2 de deux courbes planes sur la surface de la sphère. La courbe géodésique est donnée par la plus courte distance s sur la surface entre deux points. Dans le cas de la sphère, les géodésiques sont des arcs de grand cercle (s_1 ici).

On a $s_1 = 2R\theta$ alors que $s_2 = \pi R \sin \theta$. Comme $\pi \sin \theta \geq 2\theta$ sur l'intervalle $[0, \pi/2]$, $s_1 \leq s_2$ (il y a égalité en $\theta = 0$ et $\theta = \pi/2$). Les géodésiques de la sphère sont donc les arcs de grands cercles, trajets empruntés par les avions pour les vols intercontinentaux (en l'absence de toute autre considération telle que les vents en haute altitude, ...).

2.2. Le théorème de Pythagore et l'équation métrique

La notion de courbure locale d'une surface bidimensionnelle, généralisant celle de courbure d'une courbe plane, a été introduite par Leonhard Euler (1703-1783). Gauss a ensuite défini une grandeur intrinsèque pour caractériser la courbure d'une surface $2d$, il a introduit la notion de plan localement tangent et a mis en évidence la relation entre courbure et équation métrique (l'*excellent théorème* de Gauss). C'est son élève, Riemann, qui a généralisé ses travaux dans le cas des surfaces courbes de dimension quelconque. Nous allons ici établir quelques propriétés des surfaces courbes bidimensionnelles, et nous admettrons que la généralisation en dimension d est permise.

Considérons deux exemples de surfaces bidimensionnelles : la surface de la sphère et celle du cylindre. Etant donnés deux points B et C , on trace la courbe géodésique entre ces points.

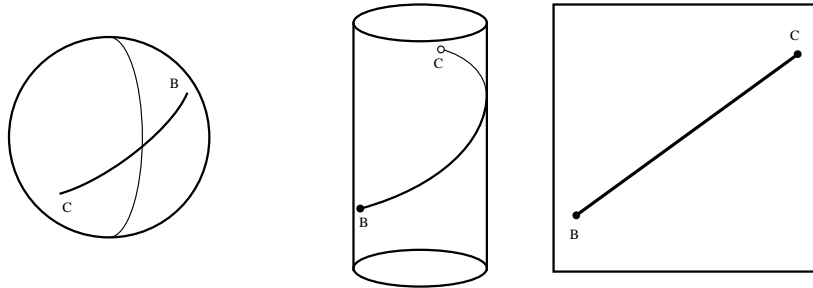


Figure 7.4 Géodésiques sur la sphère et le cylindre. Dans le second cas, après avoir coupé le cylindre parallèlement à son axe (sans traverser la géodésique afin d'éviter des problèmes de topologie), la surface peut être "aplatie" pour épouser rigoureusement un plan. La géodésique devient alors une ligne droite. La surface du cylindre est intrinsèquement plate. Il est impossible de réaliser une transformation semblable avec la sphère qui est intrinsèquement courbe.

Le cylindre peut être découpé parallèlement à son axe et déplié à plat. La géodésique apparaît ainsi comme une droite du plan. On dit que le cylindre est intrinsèquement plat (même si sa topologie diffère de celle du plan, il faut en particulier ici éviter que la coupure ne traverse la géodésique). Ce n'est évidemment pas le cas de la surface de la sphère.

Dans le cas de la surface cylindrique, on peut définir les coordonnées cartésiennes du plan : $B(y_1, z_1)$ et $C(y_2, z_2)$ permettant d'écrire la longueur s de la courbe (droite) BC sous la forme du théorème de Pythagore $s^2 = (y_2 - y_1)^2 + (z_2 - z_1)^2$. La métrique du plan est euclidienne et sous forme infinitésimale on obtient l'équation métrique euclidienne : $ds^2 = dy^2 + dz^2$. Sur le cylindre, le changement de variable $y = r\theta$ donne $s^2 = (r\theta_2 - r\theta_1)^2 + (z_2 - z_1)^2$ ou sous forme locale $ds^2 = r^2 d\theta^2 + dz^2$, en posant $dy = r d\theta$. La surface du cylindre peut ainsi être représentée par des coordonnées cartésiennes (figure (7.5)) analogues à celles du plan (aux problèmes de topologie près encore une fois), la métrique de la surface du cylindre étant euclidienne sous forme infinitésimale et sous forme globale.

On peut maintenant se poser le problème d'écrire l'analogue du théorème de Pythagore pour une surface sphérique. L'impossibilité de découper la sphère et

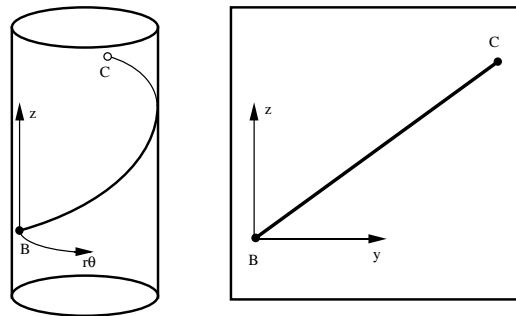


Figure 7.5 Cylindre et théorème de Pythagore. Les coordonnées z et $r\theta$ sur le cylindre permettent d'écrire une métrique euclidienne. Elles sont rigoureusement équivalentes aux coordonnées z et y du plan obtenu après avoir coupé le cylindre parallèlement à son axe.

de l'aplatir pour épouser un plan suggère des difficultés que d'autres propriétés connues révèlent d'ailleurs, on peut par exemple construire un triangle dont la somme des angles dépasse π (figure (7.6)⁽⁶¹⁾).

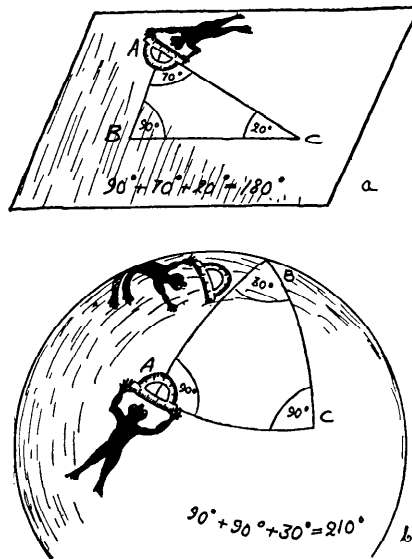


Figure 7.6 Tracé d'un triangle sur une sphère. La somme des angles dépasse 180° . Le postulat de la parallèle d'Euclide n'est pas vérifié.

Soient $M_1(\theta_1, \varphi_1)$ et $M_2(\theta_2, \varphi_2)$ deux points de la sphère de centre O (figure (7.7)). La longueur s de l'arc M_1M_2 mesuré sur la sphère vaut $s = R\beta$ où β est l'angle entre les vecteurs \mathbf{OM}_1 et \mathbf{OM}_2 . En décomposant \mathbf{OM}_1 et \mathbf{OM}_2 dans un

⁽⁶¹⁾ figure extraite de G. Gamow, *One, Two, Three... Infinity* (Dover, New-York 1988), p. 103.

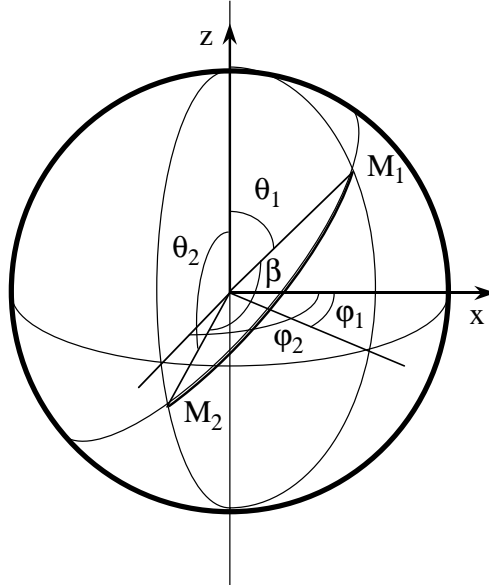


Figure 7.7 La surface de la sphère n'est pas euclidienne, en revanche le théorème de Pythagore peut s'écrire de façon infinitésimale. Cette surface est localement euclidienne.

repère cartésien de l'espace tridimensionnel euclidien, $\mathbf{OM}_1 = R \sin \theta_1 \cos \varphi_1 \hat{\mathbf{u}}_x + R \sin \theta_1 \sin \varphi_1 \hat{\mathbf{u}}_y + R \cos \theta_1 \hat{\mathbf{u}}_z$, on obtient simplement

$$\cos \beta = \frac{\mathbf{OM}_1 \cdot \mathbf{OM}_2}{R^2} = \cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2 \cos(\varphi_2 - \varphi_1),$$

soit

$$s^2 = R^2 \operatorname{Arccos}^2(\cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2 \cos(\varphi_2 - \varphi_1)).$$

Il est clair qu'il n'existe aucune transformation $\xi(\theta, \varphi)$, $\eta(\theta, \varphi)$ qui permette d'écrire s^2 comme $(\xi_2 - \xi_1)^2 + (\eta_2 - \eta_1)^2$. Si l'on considère un déplacement infinitésimal en revanche : $\theta_2 - \theta_1 = d\theta$, $\varphi_2 - \varphi_1 = d\varphi$, en développant au second ordre, on a

$$ds^2 = R^2 d\theta^2 + R^2 \sin^2 \theta d\varphi^2 = g_{\theta\theta} d\theta^2 + g_{\varphi\varphi} d\varphi^2 = d\xi^2 + d\eta^2,$$

en posant $d\xi = \sqrt{g_{\theta\theta}} d\theta$ et $d\eta = \sqrt{g_{\varphi\varphi}} d\varphi$. Les fonctions $g_{\theta\theta} = R^2$ et $g_{\varphi\varphi} = R^2 \sin^2 \theta$ sont les composantes du tenseur métrique, et l'équation métrique est localement euclidienne. Alors que θ et φ sont les coordonnées de Gauss, ξ et η sont les coordonnées du plan localement tangent. L'équation métrique indique que l'espace est courbe puisque ds^2 ne prend pas la même valeur pour une variation $d\theta$, $d\varphi$ donnée selon que l'on est au voisinage du pôle ou de l'équateur. Cependant, localement (c'est-à-dire dans une région de petite dimension devant le rayon de la sphère), les propriétés de la sphère peuvent être décrites par des coordonnées cartésiennes du plan tangent. C'est la propriété essentielle des espaces de Riemann que de pouvoir toujours définir un plan localement tangent (euclidien). Considérons

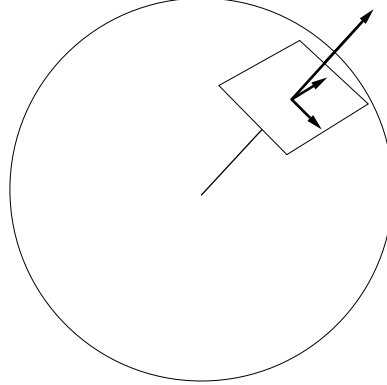


Figure 7.8 Plan localement tangent à une surface sphérique.

en effet une surface bidimensionnelle courbe quelconque et notons u et v les coordonnées de Gauss (dans la surface). L'équation métrique la plus générale est de la forme

$$ds^2 = g_{11} du^2 + 2g_{12} du dv + g_{22} dv^2. \quad (7.1)$$

Les g_{ij} sont des fonctions du point qui contiennent l'information sur la courbure et g_{12} est choisie positive. Au voisinage d'un point P on cherche à définir des coordonnées cartésiennes x et y telles que

$$\begin{aligned} du &= \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy, \\ dv &= \frac{\partial v}{\partial x} dx + \frac{\partial v}{\partial y} dy. \end{aligned}$$

L'équation métrique devient

$$ds^2 = g'_{11} dx^2 + 2g'_{12} dx dy + g'_{22} dy^2$$

avec

$$\begin{aligned} g'_{11} &= \left(\frac{\partial u}{\partial x}\right)^2 g_{11} + 2\frac{\partial u}{\partial x} \frac{\partial v}{\partial x} g_{12} + \left(\frac{\partial u}{\partial y}\right)^2 g_{22}, \\ g'_{12} &= \frac{\partial u}{\partial x} \frac{\partial u}{\partial y} g_{11} + \left(\frac{\partial u}{\partial x} \frac{\partial v}{\partial y} + \frac{\partial u}{\partial y} \frac{\partial v}{\partial x}\right) g_{12} + \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} g_{22}, \\ g'_{22} &= \left(\frac{\partial v}{\partial y}\right)^2 g_{11} + 2\frac{\partial u}{\partial y} \frac{\partial v}{\partial y} g_{12} + \left(\frac{\partial v}{\partial y}\right)^2 g_{22}. \end{aligned}$$

Le choix au point P des dérivées $\frac{\partial u}{\partial x} \dots$ et de leurs six dérivées premières indépendantes $\frac{\partial^2 u}{\partial x^2} \dots$ est arbitraire, ce qui laisse dix variables indépendantes. On peut imposer neuf contraintes pour choisir les fonctions g'_{ij} de sorte que $g'_{11} = g'_{22} = 1$, $g'_{12} = 0$ ainsi que $\frac{\partial g'_{11}}{\partial x} = \dots = 0$. Le dernier degré de liberté provient du choix de l'orientation des axes x et y dans le plan tangent où l'équation métrique a maintenant la forme euclidienne $ds^2 = dx^2 + dy^2$ (figure (7.8)).

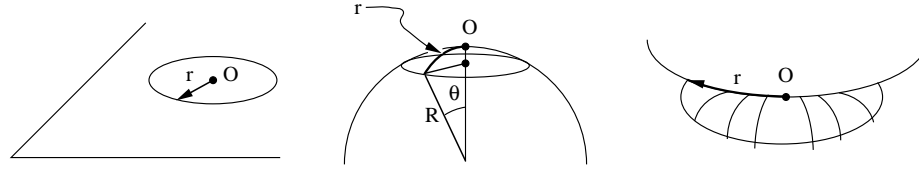


Figure 7.9 Cercle sur une surface bidimensionnelle courbe. Dans un espace plat, la circonférence du cercle vaut exactement $C_F = 2\pi r$, elle est inférieure dans un espace de courbure positive (surface de la sphère par exemple) $C_P < 2\pi r$ et supérieure dans le cas d'une surface de courbure négative (en selle de cheval par exemple) $C_N > 2\pi r$. Dans les trois cas, r est le rayon du cercle mesuré dans la surface. L'écart entre la circonférence C et $2\pi r$ donne accès à la courbure : $K = \frac{3}{\pi} \lim_{r \rightarrow 0} \frac{2\pi r - C}{r^3}$.

L'équation métrique originale s'écrit aussi

$$ds^2 = \left(\sqrt{g_{11}} du + \frac{g_{12}}{\sqrt{g_{11}}} dv \right)^2 + \left(g_{22} - \frac{g_{12}^2}{g_{11}} \right) dv^2,$$

d'où l'on déduit la relation entre les coordonnées de Gauss et les coordonnées euclidiennes du plan tangent :

$$dx = \sqrt{g_{11}} du + \frac{g_{12}}{\sqrt{g_{11}}} dv, \quad dy = \sqrt{g_{22} - \frac{g_{12}^2}{g_{11}}} dv.$$

Dans l'exemple de la sphère, $du = d\theta$, $dv = d\varphi$, $g_{11} = R^2$, $g_{12} = 0$ et $g_{22} = R^2 \sin^2 \theta$, d'où l'on déduit $dx = R d\theta$ et $dy = R \sin \theta d\varphi$.

Les deux surfaces que nous avons considérées nous permettent de définir un critère pour savoir si une surface est courbe. Etant donnée l'équation métrique en termes des variables de Gauss : $ds^2 = g_{uu} du^2 + 2g_{uv} du dv + g_{vv} dv^2$, s'il existe une transformation des coordonnées $x(u, v)$, $y(u, v)$ générale (au sens de non locale) telle que la métrique devienne euclidienne avec les nouvelles variables : $s^2 = x^2 + y^2$, l'espace est dit intrinsèquement plat. Dans le cas contraire, même s'il n'existe pas une telle transformation générale, il en existe toujours localement, $dx(du, dv)$, $dy(du, dv)$ et l'espace est courbe mais $ds^2 = dx^2 + dy^2$.

2.3. Mesure de la courbure d'une surface bidimensionnelle

La courbure gaussienne quantifie localement la courbure de toute surface bidimensionnelle. Cette notion, dont nous donnerons un sens précis un peu plus loin, peut être généralisée en toute dimension. Dans ce paragraphe, nous allons tout d'abord mettre en évidence quelques manifestations de la courbure d'une surface, nous établirons quelques résultats dans le cas des surfaces sphériques, et nous admettrons qu'ils ont une valeur générale.

- **Circonférence d'un cercle** : on trace un cercle dans une surface bidimensionnelle. Le rayon r est donné par la longueur de la courbe géodésique joignant le centre O du cercle à un point quelconque de la circonférence dans la surface (figure (7.9)).

Dans le cas d'un espace plat, la circonférence vaut exactement $\mathcal{C}_F = 2\pi r$ (l'indice F vaut pour "flat"). Sur la surface d'une sphère, cette circonférence est inférieure à l'expression $2\pi r$ puisque $\mathcal{C}_P = 2\pi R \sin \theta < 2\pi r$ (indice P pour positif). On dit que la surface de la sphère a une courbure positive. Enfin, dans le cas d'une surface en forme de selle de cheval, on a $\mathcal{C}_N > 2\pi r$, et la courbure est négative (d'où l'indice N). La différence $2\pi r - \mathcal{C}$ donne ainsi le signe de la courbure de la surface. On peut encore donner une expression plus précise en considérant le cas de la sphère. Dans la limite où $\theta = r/R$ est petit, on peut développer

$$\mathcal{C}_P = 2\pi R \frac{r}{R} \left(1 - \frac{1}{6} \frac{r^2}{R^2} + \dots \right)$$

soit

$$K = \frac{1}{R^2} = \frac{3}{\pi} \lim_{r \rightarrow 0} \frac{2\pi r - \mathcal{C}}{r^3}. \quad (7.2)$$

- Transport parallèle : on construit cette fois un circuit fermé sur la surface, et on déplace un vecteur de la surface le long du contour en lui gardant localement une orientation constante par rapport aux branches du circuit (figure (7.10)).

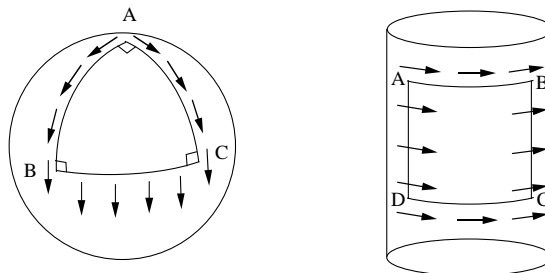


Figure 7.10 Transport parallèle : on trace un circuit fermé sur une surface et on translate de proche en proche le long du circuit un vecteur qui conserve une orientation constante par rapport aux contours du circuit. Dans le cas d'une surface courbe, le vecteur a tourné d'un angle α proportionnel à la courbure et à l'aire enfermée par le contour après le transport le long du circuit : $\alpha = K \times \mathcal{S}$

Il est clair qu'après cette opération sur le circuit fermé, dans le cas d'une surface plate (plan ou cylindre), le vecteur revient à sa position de départ. Dans le cas d'une surface courbe en revanche, le vecteur a tourné d'un angle α . On peut montrer dans le cas de la sphère, avec le circuit d'un huitième de sphère représenté sur la figure (7.10), que l'angle α est proportionnel à la courbure et à l'aire enfermée par le circuit : $\alpha = K \times \frac{1}{8} 4\pi R^2 = \frac{\pi}{2}$.

- Déviation géodésique : on considère deux géodésiques issues d'un même point O et la vitesse à laquelle elles se séparent. Après un déplacement d'une longueur s sur les géodésiques, les deux points courants se trouvent séparés d'une quantité η

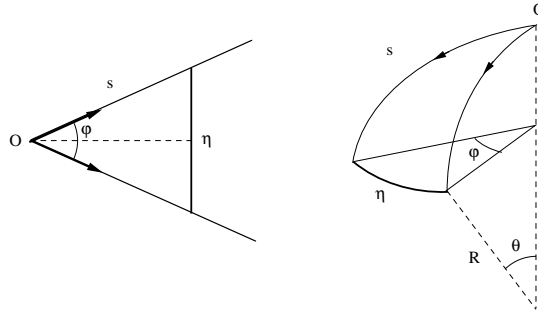


Figure 7.11 Déviation géodésique : on considère deux géodésiques issues d'un même point O . Après une distance s le long des géodésiques, les deux points courants sont distants d'une quantité η . Si la loi $\eta(s)$ est linéaire, l'espace est plat. L'équation différentielle qui régit le comportement de la fonction $\eta(s)$ fait intervenir la courbure : $\frac{d^2\eta}{ds^2} = -K\eta$.

(figure (7.11)). Dans le cas du plan, on a simplement $\eta = 2s \sin \frac{\varphi}{2} = C^{\text{te}} \times s$, soit une loi $\eta(s)$ linéaire, ce qui caractérise une surface sans courbure. Dans le cas de la sphère, on a $\eta = R\varphi \sin \theta$ et $\theta = s/R$, soit $\eta = R\varphi \sin \frac{s}{R}$. La déviation géodésique augmente donc moins vite sur une surface de courbure positive. On verrait de même qu'elle croît plus vite qu'une loi linéaire pour une surface de courbure négative. On peut relier la déviation géodésique à la courbure par une équation différentielle, la courbure mesurant l'écart au comportement linéaire :

$$\frac{d^2\eta}{ds^2} = -K\eta. \tag{7.3}$$

2.4. Courbure de Gauss et courbure moyenne

On cherche à trouver une grandeur locale intrinsèque pour caractériser la courbure d'une surface bidimensionnelle plongée dans l'espace $3d$. Par intrinsèque, on entend que cette grandeur doit dépendre exclusivement du point M auquel elle est calculée, mais pas de la direction choisie dans la surface, ni des propriétés de l'espace euclidien $3d$ dans lequel est plongée la surface.

Dans un premier temps, nous allons calculer la courbure d'une surface Σ en un point $M(x, y)$ de cette surface à partir de la notion de courbure d'une courbe plane (Euler). Ici, x, y et z sont les coordonnées dans l'espace euclidien. Soit \mathbf{n} , la normale en M à Σ , et soit π_t , le plan tangent en M . Etant donné un vecteur \mathbf{t} du plan tangent, le plan normal $\pi_n = (\mathbf{n}, \mathbf{t})$ coupe la surface Σ selon une courbe plane dont on peut calculer la courbure K_1 (figure (7.12)).

Bien entendu, si l'on choisit une autre direction \mathbf{t}' du plan tangent, le plan $\pi_{n'} = (\mathbf{n}, \mathbf{t}')$ définit une nouvelle courbe plane de courbure K_2 éventuellement différente. Ce n'est donc pas encore la grandeur intrinsèque cherchée, mais cette dernière s'exprimera à partir des courbures K_1 et K_2 .

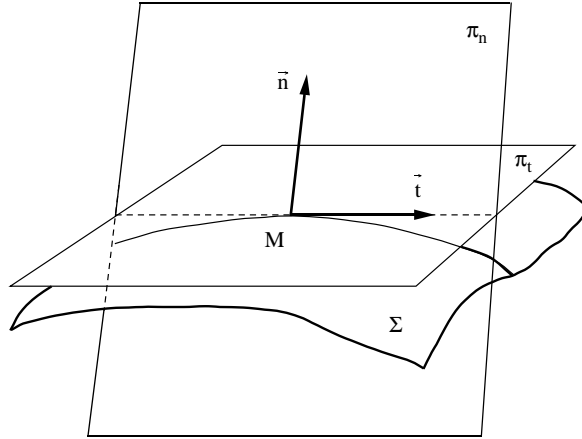


Figure 7.12 Courbure d'une surface bidimensionnelle Σ : en un point M de la surface, on trace les vecteurs normal et tangent \mathbf{n} et \mathbf{t} . L'intersection du plan (\mathbf{n}, \mathbf{t}) avec la surface Σ définit une courbe plane dont on peut calculer la courbure. Celle-ci ne caractérise cependant pas la surface au point M de façon intrinsèque, car la courbure obtenue dépend de la direction choisie pour le vecteur tangent.

A titre d'exemple, considérons la surface d'un cylindre de rayon r . On peut manifestement tracer sur la surface Σ deux types de géodésiques particulières (cercles de rayon r et droites verticales) de courbures $\frac{1}{r}$ et 0. On se place ici dans le cas d'une géodésique quelconque. Etant donné un point M , le vecteur normal \mathbf{n} et le plan tangent π_t . On considère dans le plan tangent un vecteur \mathbf{t} incliné d'un angle ϕ par rapport à l'horizontale (figure (7.13)).

L'intersection du plan π_n avec la surface Σ définit une ellipse d'équation

$$\frac{n^2}{a^2} + \frac{t^2}{b^2} = 1$$

dont le demi petit-axe a est donné par le rayon du cylindre r et le demi grand-axe dépend de l'angle ϕ : $b = \frac{r}{\cos \phi}$ (figure (7.14)).

Les points de l'ellipse, de coordonnées (n, t) , sont donnés par

$$n = \sqrt{r^2 - t^2 \cos^2 \phi}.$$

La dérivée première

$$\frac{\partial n}{\partial t} = -t \cos \phi (r^2 - t^2 \cos^2 \phi)^{-1/2}$$

est nulle en M et la dérivée seconde

$$\frac{\partial^2 n}{\partial t^2} = -\cos^2 \phi (r^2 - t^2 \cos^2 \phi)^{-1/2} - (t \cos^2 \phi)^2 (r^2 - t^2 \cos^2 \phi)^{-3/2}$$

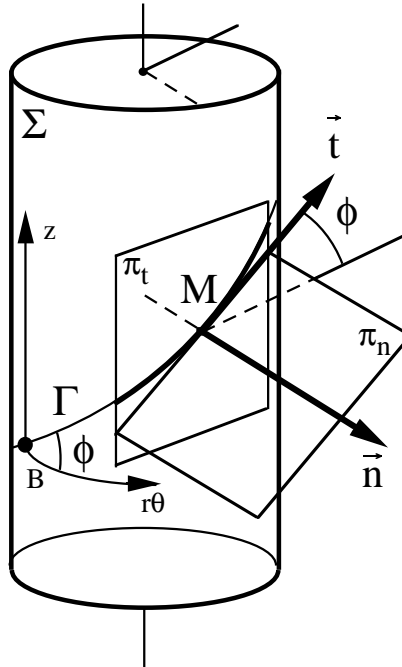


Figure 7.13 Courbure de la surface Σ d'un cylindre : au point M de Σ , on trace les vecteurs \mathbf{n} et \mathbf{t} . L'intersection du plan (\mathbf{n}, \mathbf{t}) avec la surface Σ définit une courbe plane de forme elliptique dont on peut calculer la courbure.

vaut $\frac{\cos^2 \phi}{r}$ en M , d'où l'on déduit la courbure en M :

$$K_\phi = \frac{1}{r} \cos^2 \phi.$$

On retrouve bien $K_{\phi=0} = \frac{1}{r}$ pour le cercle et $K_{\phi=\pi/2} = 0$ pour la droite verticale, et on constate effectivement que cette grandeur ne caractérise pas uniquement le point M puisqu'elle dépend de la direction choisie dans le plan tangent.

Afin de généraliser pour obtenir une courbure intrinsèque, on considère une surface Σ de \mathbb{R}^3 représentée au moins localement (c'est-à-dire dans le plan tangent) par une équation de type $n = f(x, x')$, où n est la coordonnée perpendiculaire et x et x' forment une base orthogonale du plan tangent. Au voisinage d'un point M de la surface de coordonnées (x_0, x'_0) , on peut effectuer un développement de Taylor :

$$\begin{aligned} n = n_0 + (x - x_0) \left(\frac{\partial f}{\partial x} \right)_M + (x' - x'_0) \left(\frac{\partial f}{\partial x'} \right)_M + (x - x_0)(x' - x'_0) \left(\frac{\partial^2 f}{\partial x \partial x'} \right)_M + \\ + \frac{1}{2}(x - x_0)^2 \left(\frac{\partial^2 f}{\partial x^2} \right)_M + \frac{1}{2}(x' - x'_0)^2 \left(\frac{\partial^2 f}{\partial x'^2} \right)_M + \dots \end{aligned}$$

où $n_0 = f(x_0, x'_0)$. On peut toujours choisir dans le plan tangent l'origine des axes x et x' de sorte que $x_0 = x'_0 = 0$. Par ailleurs les dérivées premières sont nulles. Il

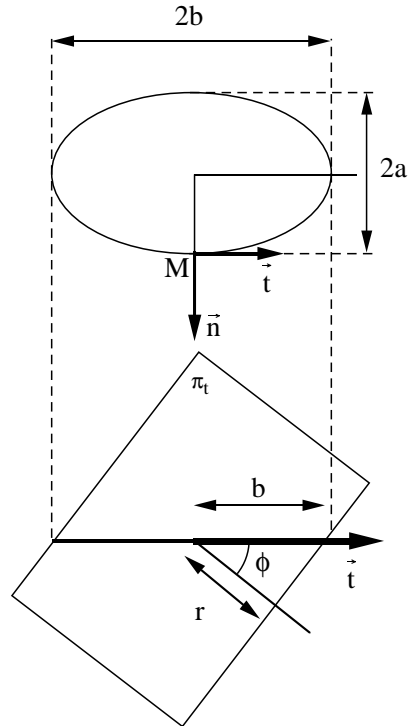


Figure 7.14 Courbure de l'ellipse au point M : l'équation de l'ellipse $\frac{n^2}{a^2} + \frac{t^2}{b^2} = 1$, où le demi petit-axe $a = r$ et le demi grand-axe $b = \frac{r}{\cos \phi}$ permettent de calculer la courbure de la courbe plane.

reste

$$n - n_0 = \frac{1}{2} f_{ij} x_i x_j + \mathcal{O}(x^3),$$

où la sommation de 1 à 2 sur les indices répétés est implicite et $x_1 = x$, $x_2 = x'$, et où

$$f_{ij} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_M.$$

L'anisotropie de la surface apparaît si $f_{ij} \neq C^{te} \times \delta_{ij}$. Soit maintenant $P(x, x')$ un point voisin de $M(0, 0)$. On peut définir une nouvelle base tournée (t, t') du plan tangent, $\mathbf{t} = \alpha_1 \mathbf{x} + \alpha_2 \mathbf{x}'$, telle que $\mathbf{MP} = |\mathbf{MP}| \mathbf{t} = t \mathbf{t}$. Les points M et P ont pour coordonnées dans la nouvelle base : $M(t = 0, t' = 0)$ et $P(t, t' = 0)$. On a donc $x = \alpha_1 t$ et $x' = \alpha_2 t$, de sorte que

$$n(t) = n_0 + \frac{1}{2} f_{ij} \alpha_i \alpha_j t^2,$$

ce qui permet de calculer les éléments K_{ij} d'une matrice de courbure :

$$\frac{\partial n}{\partial t} = f_{ij}\alpha_i\alpha_j t, \quad \text{nul en } M, \quad \frac{\partial^2 n}{\partial t^2} = f_{ij}\alpha_i\alpha_j, \quad \text{non nul en } M,$$

d'où la courbure dans la direction \mathbf{t} , $K_t = f_{ij}\alpha_i\alpha_j$ et la matrice de courbure

$$K_{ij} = f_{ij}\alpha_i\alpha_j \quad (7.4)$$

par application de la formule de courbure des courbes planes. K_{ij} dépend de la direction choisie dans le plan tangent, mais si l'on définit la matrice de courbure :

$$[K] = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial x'} \\ \frac{\partial^2 f}{\partial x' \partial x} & \frac{\partial^2 f}{\partial x'^2} \end{pmatrix}. \quad (7.5)$$

Il existe deux invariants qui définissent :

- la courbure de Gauss $\text{Det}(f_{ij}) = K_1 K_2 = K_G$,
- la courbure moyenne $\text{Tr}(f_{ij}) = K_1 + K_2 = 2K_M$, où K_1 et K_2 , appelées les courbures principales, sont les valeurs propres de $[K]$, les vecteurs propres définissant les directions principales. La courbure de Gauss caractérise de façon intrinsèque la surface au point M , elle est pour cette raison simplement appelée courbure de la surface, alors que la courbure moyenne est une caractéristique locale au point M en tant que surface plongée dans \mathbb{R}^3 et n'est donc pas une grandeur intrinsèque.

Reprenons l'exemple de la surface du cylindre (figure (7.15)).

L'équation de la surface du cylindre s'écrit $n = (r^2 - x^2)^{1/2}$ (n joue le rôle de y en cartésiennes). Dans la direction \mathbf{t} faisant avec l'horizontale un angle ϕ , on a $x = t \cos \phi - t' \sin \phi$ de sorte que l'équation dans cette nouvelle base devient

$$n(t, t') = (r^2 - (t \cos \phi - t' \sin \phi)^2)^{1/2}$$

La matrice de courbure, en orientant la normale vers l'intérieur, s'écrit

$$[K]_{t,t'} = [f_{ij}] = \begin{pmatrix} \frac{\cos^2 \phi}{r} & \frac{\cos \phi \sin \phi}{r} \\ \frac{\cos \phi \sin \phi}{r} & \frac{\sin^2 \phi}{r} \end{pmatrix}.$$

Les directions principales sont en fait les géodésiques particulières (grand cercle et droite verticale) et la matrice de courbure diagonalisée s'écrit

$$[K]_{x,z} = [f_{ij}] = \begin{pmatrix} r^{-1} & 0 \\ 0 & 0 \end{pmatrix},$$

d'où l'on déduit la courbure de Gauss $K_G = 0$ qui prouve que le cylindre a une surface intrinsèquement plate. La valeur non nulle $K_M = 1/2r$ de la courbure moyenne montre en revanche qu'un cylindre n'est cependant pas tout à fait un plan !

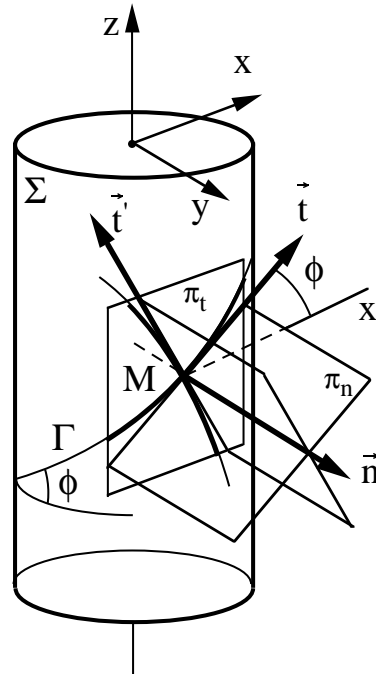


Figure 7.15 Courbure moyenne et courbure de Gauss dans le cas du cylindre : on calcule au point M les courbures des courbes planes contenues dans deux plans normaux perpendiculaires entre eux.

2.5. Courbure et équation métrique : l'excellent théorème de Gauss

Gauss a établi un théorème, connu sous le prestigieux nom d'excellent théorème de Gauss⁽⁶²⁾, qui relie la courbure au coefficient radial de la métrique d'un espace courbe. Nous allons démontrer ce théorème dans le cas particulier d'une surface sphérique et nous en admettrons la généralisation pour des surfaces arbitraires.

Afin de faire apparaître la composante radiale du tenseur métrique g_{rr} sur une surface sphérique, on choisit de travailler en coordonnées cylindriques (r, φ, z) (figure (7.16)).

Ecrivons dans un premier temps l'équation métrique sur la sphère de rayon R en coordonnées sphériques :

⁽⁶²⁾ "Theorema egregium" ou le très beau théorème. Voir par exemple I.R. Kenyon, *General relativity*, Oxford University Press, New York 1990, §3.4 ou M.V. Berry, *Principles of cosmology and gravitation*, Institute of Physics Publishing, Bristol 1989, Annexe B.

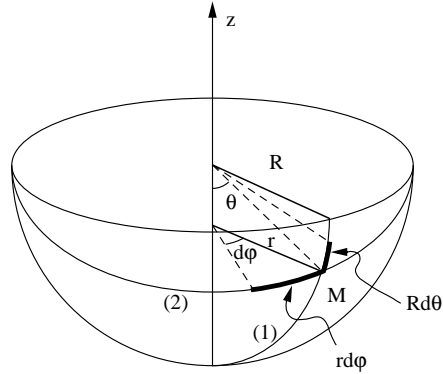


Figure 7.16 L'excellent théorème de Gauss établit une relation entre la composante radiale du tenseur métrique g_{rr} et la courbure de Gauss K de la surface.

$$ds = R d\theta \hat{\mathbf{u}}_\theta + R \sin \theta d\varphi \hat{\mathbf{u}}_\varphi,$$

soit

$$ds^2 = g_{\theta\theta} d\theta^2 + g_{\varphi\varphi} d\varphi^2 = R^2 d\theta^2 + R^2 \sin^2 \theta d\varphi^2. \quad (7.6)$$

En introduisant la coordonnée cylindrique $r = R \sin \theta$, $dr = R \cos \theta d\theta$, l'équation métrique devient $ds^2 = \frac{dr^2}{\cos^2 \theta} + r^2 d\varphi^2$ et en remplaçant $\cos^2 \theta$ par $1 - \frac{r}{R}$ il vient :

$$ds^2 = g_{rr} dr^2 + g_{\varphi\varphi} d\varphi^2 = \left(1 - \frac{r}{R}\right)^{-1} dr^2 + r^2 d\varphi^2. \quad (7.7)$$

Les deux géodésiques (1) et (2) représentées sur la figure (7.16) (directions principales) ont même courbure $\frac{1}{R}$. Sur la première, $d\varphi = 0$,

$$ds = \sqrt{g_{rr}} dr = R d\theta = \frac{dr}{\cos \theta},$$

soit

$$K_1 = \frac{1}{R} = \frac{\partial \theta}{\partial s} \quad \text{et} \quad \cos \theta = \frac{\partial r}{\partial s} = \frac{1}{\sqrt{g_{rr}}}.$$

En dérivant par rapport à s $\sin^2 \theta = 1 - \frac{1}{g_{rr}}$ on a de plus

$$2 \sin \theta \cos \theta \frac{\partial \theta}{\partial s} = \frac{1}{g_{rr}^2} \frac{\partial g_{rr}}{\partial s} = \frac{1}{g_{rr}^2} \frac{\partial g_{rr}}{\partial r} \frac{\partial r}{\partial s}$$

Sur la seconde géodésique, $d\theta = 0$ d'où

$$ds = R \sin \theta d\varphi = r d\varphi,$$

et la courbure s'écrit

$$K_2 = \frac{1}{R} = \frac{\sin \theta}{r}.$$

En combinant ces deux dernières expressions, on obtient la courbure de Gauss :

$$K = K_1 K_2 = \frac{\sin \theta}{r} \frac{\partial \theta}{\partial s} = \frac{1}{2r g_{rr}^2} \frac{\partial g_{rr}}{\partial r}. \quad (7.8)$$

Cette dernière égalité constitue le théorème cherché. On peut le vérifier à l'aide de la métrique (7.7) :

$$\frac{\partial g_{rr}}{\partial r} = - \left(1 - \frac{r^2}{R^2} \right)^{-2} \left(-\frac{2r}{R^2} \right)$$

soit $K = \frac{1}{R^2}$.

Dans un espace homogène et isotrope (c'est le cas de la sphère), toutes les géodésiques ont même courbure en un point, et cette courbure est la même partout. L'équation métrique, pour un espace isotrope de dimension n , est alors de la forme

$$ds^2 = g_{rr}(r) dr^2 + r^2 d\Omega^2,$$

où $d\Omega$ est l'angle solide (éventuellement généralisé si la dimension spatiale dépasse 3). On constate ainsi que, grâce au théorème de Gauss, la connaissance de la courbure définit entièrement la métrique d'un espace courbe homogène, isotrope (et statique).

Chapitre 8

Le principe d'équivalence

Le principe d'équivalence, ses conséquences et ses vérifications, sont discutés dans ce chapitre, largement inspiré de Hakim.

1. Les expériences d'Eötvös-Dicke

Le problème de l'égalité de la masse inerte (m_i , quantité qui intervient dans toutes les relations dynamiques : $\mathbf{p} = m_i \mathbf{v}$, $\mathbf{F} = m_i \mathbf{a}$, ...) et de la masse gravitationnelle (m_g , présente dans les relations qui font intervenir le champ de pesanteur : $\mathbf{P} = m_g \mathbf{g}$, $\mathbf{F} = -G \frac{m_g m'_g}{r^2} \mathbf{u}_r$, ...) était considéré par Einstein comme fondamental. La théorie de la Relativité Générale est fondée sur cette identité, valable pour tous les corps, et Einstein, dans une lettre à Bergmann, estimait qu'il était préférable de réaliser de nouvelles expériences en vue de vérifier cette égalité une fois de plus et avec une précision accrue, plutôt que de mesurer à nouveau les effets classiques de sa théorie⁽⁶³⁾.

La matière est constituée d'atomes, eux-mêmes constitués d'électrons et de nucléons. Les nucléons subissent l'interaction forte, mais pas les électrons, la question de savoir si nucléons et électrons ressentent la même interaction gravitationnelle est donc importante, d'autant plus que le rapport du nombre d'électrons au nombre de nucléons dans la matière varie de 1 à 2,5 environ. Il est évident qu'une différence serait difficile à mesurer du fait de la faible masse des électrons devant celle des nucléons. De plus la masse d'un noyau diffère de la somme des masses des nucléons

⁽⁶³⁾ V.N. Rudenko, *Relativistic experiments in gravitational fields*, *Sov. Phys. Uspekhi* **21** (1978) 893.

qui le composent par l'énergie de liaison qui varie de 0 pour l'hydrogène à 0.7% de la masse pour le fer. Si l'interaction gravitationnelle, comme l'interaction forte, dépendait du nombre de nucléons plutôt que de la masse, on pourrait mesurer une différence de 0.7% entre l'interaction gravitationnelle subie par l'hydrogène et celle subie par le fer.

Le principe d'équivalence remonte dans une certaine mesure à Galileo Galilei (1564-1642) qui a découvert, au cours de ses expériences à l'aide de plans inclinés, que tous les corps tombent à un *taux* indépendant de leur masse. Il s'intéresse alors au référentiel en chute libre :

Notez d'abord qu'il y a lieu de distinguer les corps graves en mouvement de ceux qui sont à l'état de repos. Une grosse pierre posée sur une balance, non seulement pèsera davantage si on lui superpose une autre pierre, mais la seule adjonction d'une quenouillée d'étope augmentera son poids des quelques huit ou dix onces que pèsera cette quenouillée ; au contraire, si vous laissez tomber librement, d'une certaine hauteur, votre pierre, après y avoir attaché l'étope, croyez-vous que le poids ajouté à celui de la pierre accélérera son mouvement, ou bien qu'elle le retardera en la soutenant en partie ? Nous portons sur nos épaules le poids d'un objet quand nous voulons nous opposer au mouvement qu'il ferait s'il tombait, mais si nous descendons nous-mêmes à la vitesse qui serait normalement celle du corps grave dans sa chute, comment voulez-vous qu'il nous charge et qu'il pèse sur nous ?⁽⁶⁴⁾

Ces observations ont été améliorées par Christiaan Huygens (1629-1695), mais c'est Isaac Newton (1642-1727), dans ses *Principia Mathematica*, qui a attiré l'attention sur le fait que ses lois de la gravitation et du mouvement pouvaient n'être qu'approchées car les masses gravitationnelle et inerte pouvaient ne pas être précisément égales. Si c'était le cas, pour un corps en chute libre le principe fondamental de la dynamique et l'expression du poids donneraient :

$$\mathbf{F} = m_i \mathbf{a}, \quad \text{et} \quad \mathbf{F} = m_g \mathbf{g}, \quad \text{soit} \quad \mathbf{a} = \frac{m_g}{m_i} \mathbf{g}.$$

Ainsi les accélérations de corps en chute libre seraient différentes pour des matériaux dont les rapports m_g/m_i seraient eux-mêmes différents, tout comme des charges différentes subissent dans un même champ électrique \mathbf{E} des accélérations différentes, puisque $\mathbf{a} = \frac{q}{m_i} \mathbf{E}$.

De même, des pendules d'égales longueurs auraient des périodes proportionnelles à $\sqrt{m_i/m_g}$. Newton a réalisé lui-même l'expérience avec des pendules et a montré que diverses substances avaient le même rapport m_g/m_i à 10^{-3} près. Il utilise les termes *quantité de matière* pour masse inerte et *poids* ou *grave* pour masse gravitationnelle⁽⁶⁵⁾.

Que la chute de tous les graves sur la Terre s'effectue en des temps égaux (en faisant abstraction, du moins, du retard inégal suscité par une très faible résistance de l'air), d'autres l'ont observé depuis longtemps ; mais c'est avec la plus grande exactitude que l'on peut connaître l'égalité des temps avec les pendules. J'en ai fait l'expérience sur de l'or, de l'argent, du plomb, du verre, du

⁽⁶⁴⁾ G. Galilée, *Dialogues et Lettres choisies, Discours sur deux Sciences Nouvelles*, trad. P.H. Michel (Hermann, Paris 1966), p.284.

⁽⁶⁵⁾ I. Newton, *De Philosophiae Naturalis Principia Mathematica*, (Christian Bourgeois Éditeur, Paris 1985).

sable, du sel commun, du bois, de l'eau et du froment. Je comparais deux petites boîtes de bois rondes et égales. J'en remplissais l'une de bois et aussi exactement que possible, je suspendais au centre d'oscillation de l'autre le même poids d'or. Ces boîtes attachées à des fils égaux de onze pieds formaient des pendules complètement semblables, quant au poids, à la figure et à la résistance de l'air : placées l'une à côté de l'autre, elles effectuaient des oscillations semblables, allant en même temps et revenant, et ce, pendant fort longtemps. Par conséquent, la quantité de matière de l'or était à celle du bois, comme l'action de la force motrice sur l'or tout entier à l'action de la même force sur le bois tout entier ; c'est-à-dire, comme le poids de l'or au poids du bois. Et il en était de même avec les autres corps. Dans les corps de même poids, une différence de matière, inférieure au millième de la matière totale, a pu être nettement relevée de ces expériences. Désormais, on ne peut vraiment plus douter que la nature de la gravité soit la même sur les planètes que sur la Terre. Supposons, en effet, que des corps terrestres soient élevés jusqu'à l'orbe de la Lune et abandonnés avec la Lune sans mouvement initial de manière à tomber en même temps sur la Terre ; il est certain, par ce qui vient d'être montré, qu'ils décriraient comme la Lune des espaces égaux en des temps égaux et qu'ainsi ils sont à la quantité de matière de la Lune, comme leurs poids au poids de celle-ci. (...)

Corollaire 1 : Par suite, les poids des corps ne dépendent pas de leurs formes ni de leurs textures. Car, s'ils pouvaient varier avec les formes, ils devraient être plus ou moins grands selon la variété de celles-ci à matière égale : ce qui va absolument à l'encontre de l'expérience.

Corollaire 2 : Tous les corps qui sont autour de la Terre sont pesants sur la Terre et tous les poids qui sont également distants du centre de la Terre sont comme les quantités de matière contenues en eux. Telle est la qualité de tous les corps, pour lesquels on peut réaliser des expériences, et par conséquent, de par la règle 3, elle doit être affirmée de tous les corps en général.

En 1830, Friedrich Wilhelm Bessel (1717-1846) reprendra avec une plus grande précision les expériences de Newton, mais c'est en 1889 que des expériences nouvelles, extrêmement précises, seront réalisées à Budapest par Roland von Eötvös (1849-1919). Ces expériences ont fortement impressionné Einstein et ont été reprises dans leur principe de nombreuses fois au cours du XX^{ème} siècle. Le principe en est particulièrement simple : il s'agit d'une balance dans laquelle la masse inerte est liée à la force centrifuge due à la rotation diurne de la Terre et la masse gravitationnelle est liée au champ de pesanteur terrestre.

La figure 8.1 représente schématiquement le dispositif expérimental : deux masses A et B sont suspendues aux bras (de longueurs L_A et L_B) d'une balance. Sur ces masses s'exercent les poids $m_{gA}\mathbf{g}$ (direction verticale) et les forces d'inertie d'entraînement $-m_{iA}\mathbf{a}_e$ (direction perpendiculaire à l'axe de rotation de la Terre) où \mathbf{g} est uniforme en bonne approximation et $a_e = R\omega^2 \sin\theta$ si θ est la colatitude. La condition d'équilibre des forces verticales s'écrit

$$L_A(M_{gA}g - m_{iA}a_r) = L_B(M_{gB}g - m_{iB}a_r),$$

où $a_r = a_e \sin\theta$. A ces forces s'ajoute un couple de torsion (rotation autour du fil de torsion déterminée par le bilan des forces dans le plan horizontal) :

$$T = m_{iA}a_t L_A - m_{iB}a_t L_B,$$

que l'on peut encore écrire

$$T = m_{iA}a_t L_A \left[1 - \frac{m_{iB} m_{gA}g - m_{iA}a_r}{m_{iA} m_{gB}g - m_{iB}a_r} \right].$$

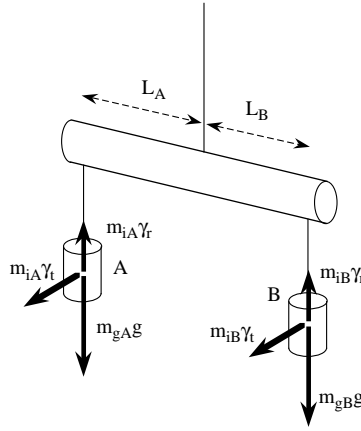


Figure 8.1 L'expérience d'Eötvös : deux masses de natures différentes sont suspendues à une balance. Le bilan des forces fait intervenir la masse gravitationnelle par l'intermédiaire du poids et la masse inerte par la force centrifuge due au mouvement de rotation diurne de la Terre.

Comme $a_r \ll g$ et $m_i \simeq m_g$, on peut simplifier cette dernière expression :

$$T \simeq m_{gA} a_t L_A \left(\frac{m_{iA}}{m_{gA}} - \frac{m_{iB}}{m_{gB}} \right).$$

Toute inégalité des rapports $\frac{m_i}{m_g}$ pour des corps de compositions différentes devrait produire une rotation de la balance autour de la verticale du lieu. Aucune rotation ne fut détectée expérimentalement et le résultat définitif d'Eötvös⁽⁶⁶⁾, établi pour des substances aussi variées que le bois, le platine, le cuivre, l'eau, le sulfate de cuivre, le suif, ... donne

$$\frac{\Delta m}{m} = \left| \frac{m_i - m_g}{m_g} \right| \leq 3.10^{-9}. \quad (8.1)$$

Après une critique minutieuse des expériences d'Eötvös, (sensibilité au gradient de \mathbf{g} , influence de l'expérimentateur qui crée un champ de gravitation du même ordre que l'incertitude absolue annoncée, ...), le groupe de Dicke (Princeton) a amélioré la précision des résultats d'Eötvös en utilisant la gravitation due au Soleil et la force centrifuge de rotation annuelle de la Terre autour du Soleil⁽⁶⁷⁾. Un résultat positif éventuel se traduirait ainsi par un mouvement périodique de période égale à 24 heures de la balance. De plus, les masses étaient disposées aux sommets d'un triangle équilatéral pour diminuer le moment quadrupolaire du pendule (figure 3.2), et les positions étaient mesurées à distance, sans intervention d'un observateur, source de perturbation gravitationnelle.

⁽⁶⁶⁾ R.v. Eötvös, D. Pekár and E. Fekete, *Beiträge zum Gezetzte der Proportionalität von Trägheit und Gravität*, *Ann. Physik* **68** (1922) 11.

⁽⁶⁷⁾ R.H. Dicke, *The Eötvös experiment*, *Sci. Am.* **205** (1961, december) 84.

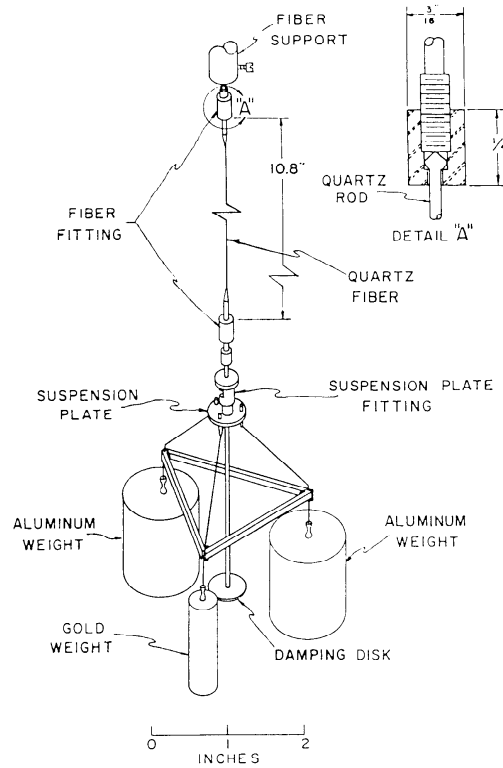


Figure 8.2 L'expérience d'Eötvös réalisée à Princeton par l'équipe de Dicke.

Le résultat final obtenu dans le cas de l'aluminium et de l'or est de⁽⁶⁸⁾

$$\frac{\Delta m}{m} = \left| \frac{m_i - m_g}{m_g} \right| \leq 1.5 \cdot 10^{-11}. \quad (8.2)$$

La précision ici est essentiellement limitée par l'existence de vibrations sismiques, aussi l'expérience a-t-elle été reprise dans un meilleur environnement sismique et avec un dispositif de moment quadrupolaire encore plus faible par Bragansky et Panov⁽⁶⁹⁾ en 1971. Ces derniers ont obtenu une borne supérieure de :

$$\frac{\Delta m}{m} = \left| \frac{m_i - m_g}{m_g} \right| \leq 0.45 \cdot 10^{-12}. \quad (8.3)$$

Plus récemment, mais avec une précision moindre, des expériences ont été réalisées au niveau microscopique⁽⁷⁰⁾. Il a ainsi été montré que les neutrons tombent

⁽⁶⁸⁾ P.G. Roll, R. Krotov and R.H. Dicke, *The Equivalence of Inertial and Passive Gravitational Mass*, *Ann. Phys. (New-York)* **26** (1964) 442.

⁽⁶⁹⁾ V.B. Bragansky and V.I. Panov, *Verification of the equivalence of inertial and gravitational mass*, *Sov. Phys. JETP* **34** (1972) 463.

⁽⁷⁰⁾ V.B. Braganski and V.N. Rudenko, *Relativistic gravitational experiments*, *Sov. Phys. Uspekhi* **13** (1970) 165.

avec la même accélération que la matière ordinaire⁽⁷¹⁾ et que la force de gravitation sur les électrons dans le cuivre est la même que pour les électrons dans le vide⁽⁷²⁾.

Le tableau 8.1 résume l'évolution des résultats expérimentaux des tests du principe d'équivalence faible⁽⁷³⁾ :

Tableau 8.1 Tests du principe d'équivalence faible : limites supérieures pour la différence relative entre masse gravitationnelle et masse inerte.

Expérience	Nom	Méthode	Substances	$\frac{\Delta m}{m}$
Newton	Newton	Pendule	variées	10^{-3}
Bessel	Bessel	Pendule	variées	5.10^{-5}
Eötvös	Eötvös, Pekár, Fekete	Balance de torsion	variées	5.10^{-9}
Potter	Potter	Pendule	variées	2.10^{-5}
Renner	Renner	Balance de torsion	variées	2.10^{-9}
Princeton	Roll, Krotkov, Dicke	Balance de torsion	Al et Au	10^{-11}
Moscou	Bragansky, Panov	Balance de torsion	Al et Pt	10^{-12}
Munich	Koester	Chute libre	neutrons	3.10^{-4}
Stanford	Worden	Suspension magnétique	niobium Terre	10^{-4}
Boulder	Keiser, Faller	Flottation	Cu et W	4.10^{-11}

2. Principe d'équivalence faible et principe d'équivalence fort

L'analyse des expériences du type de celle de Galilée, puis des expériences plus récentes d'Eötvös-Dicke-Bragansky montre que le rapport m_i/m_g est indépendant de la nature du matériau considéré. On peut alors, grâce à un choix convenable des unités, postuler l'égalité de la masse inerte et de la masse gravitationnelle. Cet énoncé constitue le principe d'équivalence faible d'Einstein, que l'on peut encore formuler de la façon suivante :

Le mouvement d'un corps d'épreuve est indépendant de sa composition.

On peut alors s'interroger sur l'existence d'expériences de mécanique susceptibles de prouver qu'un référentiel est animé d'un mouvement de chute libre. Imaginons pour cela une capsule en chute libre dans le champ de gravitation terrestre (figure 8.3).

Un objet placé dans la capsule est soumis à la même accélération que la capsule elle-même et, à condition de ne considérer qu'une région suffisamment limitée de l'espace, un observateur dans le référentiel en chute libre pourra conclure qu'il est au repos dans un référentiel libre de toute force de gravitation. Il faut bien entendu

⁽⁷¹⁾ J.W.T. Bobbs, J.A. Harvey, D. Paya and H. Hortsman, *Gravitational acceleration of free neutrons*, *Phys. Rev.* **139** (1965) B756.

⁽⁷²⁾ F.C. Witteborn and W.M. Fairbank, *Experimental comparison of the gravitational forces on freely falling electrons and metallic electrons*, *Phys. Rev. Lett.* **19** (1967) 1049.

⁽⁷³⁾ C.M. Will, *Experimental gravitation from Newton's Principia to Einstein's general relativity*, in *300 years of gravitation*, ed by S. Hawking and W. Israel (Cambridge University Press, 1987).

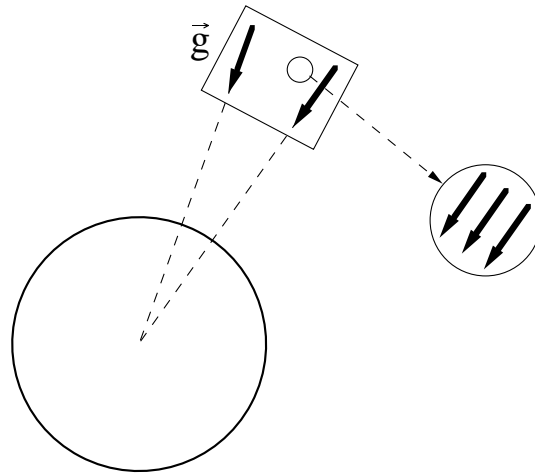


Figure 8.3 Le champ de gravitation est localement uniforme. Il y a alors équivalence du point de vue dynamique entre un référentiel immobile dans le champ de gravitation \mathbf{g} et un référentiel accéléré avec une accélération $\mathbf{a} = -\mathbf{g}$ en l'absence de force de gravitation.

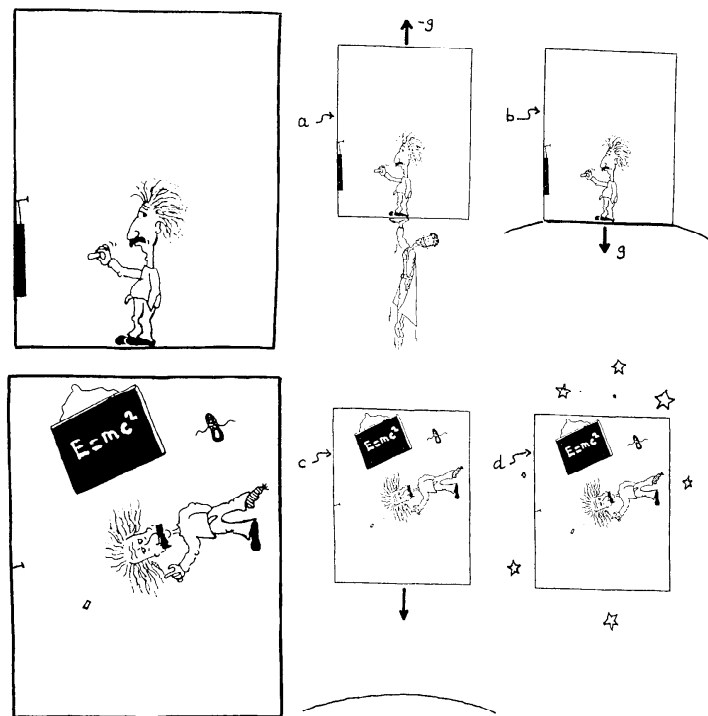


Figure 8.4 Une présentation du principe d'équivalence (l'ascenseur d'Einstein) qui compare l'effet d'un champ de gravitation à celui d'une accélération.

que le champ gravitationnel soit en bonne approximation uniforme dans toute la capsule afin d'éliminer les "effets de marée" (tidal effects)⁽⁷⁴⁾.

Une formulation plus commune consiste à dire qu'un observateur au repos dans un champ de gravitation \mathbf{g} ne pourra, par aucune expérience de mécanique locale, distinguer l'effet de la pesanteur de celui des forces d'inertie qui interviendraient dans un référentiel soumis à l'accélération $\mathbf{a} = -\mathbf{g}$ en l'absence de toute force de gravitation⁽⁷⁵⁾.

Cet énoncé constitue le principe d'équivalence faible, connu de Newton. Einstein l'a généralisé pour y inclure en particulier les lois de l'électromagnétisme, ce qui conduit au principe d'équivalence fort :

- 1 *Le résultat de toute expérience locale effectuée dans un référentiel en chute libre est indépendant de l'état de mouvement du référentiel.*
- 2 *Les résultats de telles expériences sont les mêmes dans tous les référentiels en chute libre, en tout lieu et à tout instant.*
- 3 *Les résultats des expériences locales sont compatibles avec la relativité restreinte.*

Alors que la relativité restreinte postulait l'égalité des résultats d'expériences de physique réalisées dans tout référentiel galiléen, le principe d'équivalence fort étend cette notion au cas des référentiels en chute libre, et, par extension, au cas des référentiels accélérés, à condition néanmoins de se limiter à des expériences locales. Les deux formulations du principe d'équivalence sont illustrées sur la figure 8.4⁽⁷⁶⁾.

3. Interactions fondamentales et principe d'équivalence

Conformément au principe d'équivalence, les interactions fondamentales doivent contribuer de façon identique à la masse gravitationnelle et à la masse inerte. Les résultats des expériences d'Eötvös-Dicke-Braganski permettent de préciser dans quelle mesure l'équivalence est réalisée pour les différentes interactions fondamentales.

Supposons au contraire qu'une interaction k viole, éventuellement faiblement, le principe d'équivalence. On peut alors écrire que la contribution E_k de l'interaction k à l'énergie de liaison d'un corps, ne satisfait pas à l'identité entre masse inerte et masse gravitationnelle. En introduisant un coefficient η_k de "violation du principe d'équivalence", on peut poser

$$m_g = m_i + \eta_k \frac{E_k}{c^2}.$$

Le principe fondamental de la dynamique $m_i \mathbf{a} = m_g \mathbf{g}$ conduit alors à

$$\mathbf{a} = \left(1 + \eta_k \frac{E_k}{c^2} \right) \mathbf{g}.$$

⁽⁷⁴⁾ P.G. Bergmann, *The General theory of Relativity*, in Handbuch der Physik vol. IV : "Principles of Electrodynamics and Relativity", ed. by S. Flügge (Springer Verlag, Berlin 1962).

⁽⁷⁵⁾ A. Einstein, *Principe de relativité et gravitation* (1907). Op. cit., p 115.

⁽⁷⁶⁾ figure extraite de M. Boratav et R. Kerner, *Relativité* (Ellipses), p. 202.

Deux corps distincts, si leurs énergies de liaison diffèrent, ne subiront donc pas la même accélération dans un champ de gravitation. En prenant la valeur obtenue par Braganski et Panov pour le rapport

$$\left| \frac{m_i - m_g}{m_g} \right| = |\eta_k| \frac{|E_k|}{m_g c^2} \leq 10^{-12},$$

on peut estimer le coefficient η_k :

$$\eta_k \leq 10^{-12} \frac{m_g c^2}{|E_k|}. \tag{8.4}$$

Dans le cas de l'atome d'hydrogène, $m_g \simeq 1.67 \cdot 10^{-27}$ kg, l'énergie de liaison électron-proton, d'origine électrostatique, est estimée par l'énergie de l'état fondamental $E_{1s} = -13.6$ eV, de sorte que l'on obtient

$$\eta_{es} \leq 10^{-4}.$$

Il est clair que la limite supérieure imposée à η_k par les résultats expérimentaux est d'autant plus faible que l'interaction considérée est intense. Les valeurs actuellement admises sont données dans le tableau 8.2⁽⁷⁷⁾.

Tableau 8.2 Bornes supérieures d'éventuelles violations du principe d'équivalence faible pour différents types d'interactions.

Nature de l'interaction	Coefficient de violation
interaction électrostatique (noyau)	$\eta < 4.10^{-10}$
interaction magnétostatique (noyau)	$\eta < 6.10^{-6}$
couplage magnétique nucléons	$\eta < 2.10^{-7}$
interaction électrostatique (atome)	$\eta < 5.10^{-7}$
interaction nucléaire forte	$\eta < 5.10^{-10}$
interaction faible	$\eta < 10^{-2}$
interaction gravitationnelle (nucléons)	$\eta < 10^{+27}(!)$

Dans le cas de l'interaction gravitationnelle, si l'on souhaite obtenir un résultat plus encourageant que $\eta < 10^{+27}$, du fait de la faible intensité de cette interaction, il faut envisager l'interaction entre des objets très massifs. La solution est donnée par l'effet Nordvedt⁽⁷⁸⁾. Ce dernier a montré que, dans l'hypothèse d'un coefficient de violation η_{grav} non nul, la distance $r(t)$ entre la Terre et la Lune en interaction gravitationnelle (en présence du Soleil), oscillerait autour de la solution de Kepler :

$$\delta r(t) \simeq 920 \eta_{grav} \cos(\omega_0 - \omega_S)t,$$

où ω_0 est la vitesse angulaire de rotation de la Lune autour de la Terre et ω_S celle de la Terre autour du Soleil. Des expériences de réflexion laser ont été conduites afin d'obtenir une mesure précise de la distance Terre-Lune, et le résultat semble indiquer l'absence d'effet Nordvedt⁽⁷⁹⁾

$$\eta_{grav} \leq 0.001 \pm 0.015.$$

⁽⁷⁷⁾ R. Hakim, *Gravitation Relativiste*, (InterÉditions/CNRS Editions, Paris 1994), p 166.

⁽⁷⁸⁾ K. Nordvedt, *Equivalence principle for massive bodies I Phenomenology*, *Phys. Rev.* **169** (1968) 1014.

⁽⁷⁹⁾ I.I. Shapiro, C.C. Counselman, *Verification of the Equivalence Principle for Massive Bodies*, *Phys. Rev. Lett.* **36** (1976) 555.

4. Les constantes fondamentales

L'une des conséquences importantes du principe d'équivalence est que les combinaisons sans dimension que l'on peut former à partir des constantes fondamentales (le rapport m_e/m_p , la constante de structure fine α , ...) doivent avoir même valeur en tout point et à tout instant.

Pour examiner les implications des expériences d'Eötvös-Dicke-Bragansky sur une éventuelle variation dans l'espace de telles combinaisons sans dimension, on considère une expérience de pensée due à Einstein⁽⁸⁰⁾. Etant donné un noyau formé de n constituants élémentaires (supposés identiques), de masse μ , et dont la cohésion est assurée par une énergie de liaison E_b . Cette énergie de liaison dépend des constantes fondamentales (par exemple la constante de structure fine α pour l'énergie coulombienne), elles-mêmes susceptibles de varier dans l'espace. On fait subir aux n constituants un cycle de transformations au cours desquelles on estime les échanges d'énergie avec le milieu extérieur (figure 8.5).

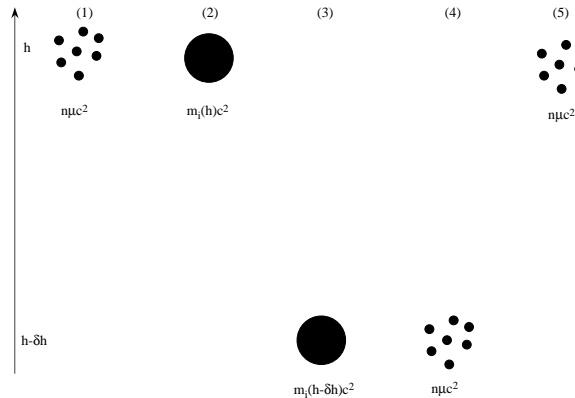


Figure 8.5 L'expérience de pensée d'Einstein permet de déduire des résultats d'expériences du type Eötvös une limite à la variation éventuelle des constantes fondamentales avec la position dans l'espace.

Lorsque les n constituants sont séparés à l'altitude h dans un champ de gravitation, l'énergie du système vaut $n\mu c^2$ (situation (1) sur la figure 8.5). A la même altitude, les n particules sont associées pour former le noyau (situation (2)) et l'énergie du système devient $m_i(h)c^2 = n\mu c^2 - |E_b(h)|$. Il reçoit donc de l'extérieur une quantité d'énergie $-|E_b(h)|$. Le noyau est ensuite abaissé de δh dans le champ de pesanteur (situation (3)). La masse inerte $m_i(h) = n\mu - |E_b(h)|/c^2$ subit une accélération a , elle reçoit donc de l'extérieur un travail $\delta W_{2,3} \simeq m_i(h)a\delta h$. Entre les phases (3) et (4), le noyau reçoit une énergie $+|E_b(h - \delta h)|$ permettant de séparer de nouveau les constituants et enfin, entre (4) et (5), ces derniers sont remontés à l'altitude initiale. La variation d'énergie potentielle $\delta E_p = n\mu g\delta h$ correspond à un travail $\delta W_{4,5} = -n\mu g\delta h$ reçu de l'extérieur. Le système se retrouvant finalement dans son état initial, le bilan énergétique des échanges avec l'extérieur est nul et il vient

$$-|E_b(h)| + m_i(h)a\delta h + |E_b(h - \delta h)| - n\mu g\delta h = 0,$$

⁽⁸⁰⁾ C.W. Misner, K.S. Thorne and J.A. Wheeler, op. cit. pp. 1060-1063.

soit, au premier ordre en δh :

$$(m_i a - n\mu g) = \frac{\partial |E_b(h)|}{\partial h}.$$

Comme m_i et $n\mu$ sont proches, on obtient

$$a - g \simeq \frac{1}{m_i} \frac{\partial |E_b(h)|}{\partial \alpha} \frac{\partial \alpha}{\partial h}. \quad (8.5)$$

Considérons les cas de noyaux d'aluminium et d'or. La contribution d'origine coulombienne à l'énergie du noyau est de 0.4% de l'énergie de masse dans le cas de l'or et de 0.1% dans le cas de l'aluminium.

$$E_{\text{coul}}(\text{Au}) = 0.004mc^2 = \eta\alpha^2,$$

où la dernière égalité provient du fait que l'énergie coulombienne est proportionnelle au carré de la constante de structure fine⁽⁸¹⁾. On a donc :

$$\frac{\partial E_{\text{coul}}(\text{Au})}{\partial \alpha} = 2\eta\alpha = \frac{2}{\alpha} \times 0.004mc^2.$$

On en déduit pour l'or et l'aluminium

$$a_{\text{Au}} - g = 0.008c^2 \frac{1}{\alpha} \frac{\partial \alpha}{\partial h},$$

$$a_{\text{Al}} - g = 0.002c^2 \frac{1}{\alpha} \frac{\partial \alpha}{\partial h},$$

soit

$$\frac{a_{\text{Au}} - a_{\text{Al}}}{g} = 0.006 \frac{c^2}{g} \frac{1}{\alpha} \frac{\partial \alpha}{\partial h}.$$

En prenant le résultat de Bragansky et Panov et le champ de pesanteur dû au Soleil $g \simeq 0.6 \cdot 10^{-2} \text{ ms}^{-2}$, on obtient :

$$\frac{1}{\alpha} \frac{\partial \alpha}{\partial h} \leq \frac{10^{-12}}{0.006} \frac{g}{c^2} = 10^{-9} \text{ m}^{-1},$$

ce qui permet de supposer en bonne approximation que la constante de structure fine ne varie pas dans l'espace.

L'éventualité d'une variation dans le temps a été testée en comparant les fréquences spectrales émises par des objets éloignés dans l'Univers (éloignement

⁽⁸¹⁾ Il suffit de se souvenir des niveaux d'énergie de l'atome d'hydrogène : $E_n = -\frac{1}{2n^2}\alpha^2 mc^2$, ou d'observer que $\langle e^2/r \rangle \sim e^2/a_0 \sim \alpha^2 mc^2$.

dans l'espace et dans le temps) et les mêmes lignes spectrales observées sur Terre, ce qui conduit à :

$$\frac{1}{\alpha} \frac{\partial \alpha}{\partial t} \leq 10^{-12} \text{année}^{-1}.$$

Par ailleurs, la comparaison des taux de désintégration β de Re^{187} présent dans les roches anciennes avec le taux de désintégration du même élément produit actuellement amène une limite plus précise :

$$\frac{1}{\alpha} \frac{\partial \alpha}{\partial t} \leq 5 \cdot 10^{-15} \text{année}^{-1}.$$

Pour tester la variation dans le temps de la constante de gravitation universelle, on considère une masse m en orbite circulaire autour d'une masse M :

$$\frac{GMm}{R^2} = \frac{mv^2}{R},$$

d'où l'on déduit la période T de la planète m :

$$T = 2\pi \sqrt{\frac{R^3}{GM}}.$$

Ainsi, si G varie dans le temps, la période T et le rayon R de l'orbite varient suivant :

$$\frac{1}{T} \frac{dT}{dt} = \frac{3}{2} \frac{1}{R} \frac{dR}{dt} - \frac{1}{2} \frac{1}{G} \frac{dG}{dt}.$$

Les orbites et périodes des planètes devraient alors varier dans le temps. En analysant les dates des éclipses dans le passé, on arrive à une variation relative de G :

$$\left| \frac{\dot{G}}{G} \right| \leq 10^{-11} \text{année}^{-1}.$$

Le tableau 8.3 résume les résultats actuels pour les diverses constantes fondamentales.

Tableau 8.3 Ordres de grandeur de la borne supérieure des variations annuelles des constantes de couplage des interactions fondamentales.

Grandeur	Borne supérieure	Auteurs
G	10^{-11}	Damour et al (1988)
e^2	10^{-14}	Davies (1972)
e^2	10^{-17}	Shlyakhter (1976)
G_{nucl}	10^{-18}	Shlyakhter (1976)
G_{nucl}	10^{-14}	Davies (1972)
G_{faible}	10^{-12}	Shlyakhter (1976)
G_{faible}	10^{-10}	Davies (1972)

5. Le principe de Mach

L'espace absolu de Newton pose des difficultés conceptuelles liées à l'impossibilité de l'existence du mouvement absolu. Leibniz et Berkeley ont argumenté que seul le mouvement relatif avait un sens :

Si nous supposons que tout l'Univers est annihilé à l'exception d'un objet, il serait impossible d'imaginer un mouvement quelconque dudit objet. (...) Imaginons deux objets comme seules choses existant dans l'Univers matériel : il serait alors impossible d'imaginer un mouvement circulaire autour d'un centre commun. Mais en supposant que le ciel d'étoiles fixes soit créé subitement, nous pourrions alors voir le mouvement de ces deux objets en observant leur position relative par rapport aux parties différentes du ciel.⁽⁸²⁾

Ainsi se pose le problème du mouvement de deux corps seuls dans l'Univers. Sont-ils en rotation autour de leur centre de masse ? Newton aurait répondu qu'il suffit de joindre les deux corps par un ressort. Si le système possède une rotation absolue, le ressort devra s'étendre pour créer une force centripète qui les maintienne sur leur trajectoire circulaire, et cette extension sera observable.

Les effets qui différencient le mouvement absolu du mouvement relatif sont les forces qui s'exercent à partir de l'axe du mouvement circulaire. Ces forces n'existent pas dans un mouvement circulaire purement relatif, mais seulement dans un mouvement circulaire véritable et absolu, et elles sont plus grandes ou plus petites conformément à la quantité de mouvement.⁽⁸³⁾

Berkeley aurait pour sa part répliqué qu'aucune tension du ressort n'aurait lieu. L'intuition de Newton est en effet fondée sur son observation de la réalité où les étoiles et les galaxies forment une toile de fond par rapport à laquelle les référentiels non inertiels sont accélérés.

En 1872, Ernst Mach (1838-1916) énonce que ce n'est pas l'accélération absolue, mais relativement à la matière éloignée, qui détermine les propriétés inertielles de la matière. L'origine de l'égalité des masses inertielle et gravitationnelle peut être comprise par cet argument:

Il ne peut exister, à mon avis, que des mouvements relatifs. Quand un corps tourne par rapport aux étoiles, les forces centrifuges sont engendrées ; quand un corps tourne par rapport à un autre corps, mais pas par rapport aux étoiles, aucune force centrifuge n'est engendrée. (...) Apparemment, il est sans importance que nous admettions que la Terre tourne autour de son axe par rapport aux étoiles fixes, ou que ce soit les étoiles dites fixes qui tournent autour de la Terre. Géométriquement, cela produit le même cas de rotation relative entre la Terre et les étoiles. Mais si nous croyons que ce sont véritablement les étoiles qui tournent autour de la Terre, on ne devrait pas observer la déviation du pendule de Foucault, ni l'aplatissement de la Terre, etc. - du moins, d'après notre conception usuelle des lois de l'inertie. Cette difficulté peut être résolue de deux façons. Soit tout mouvement est absolu, soit notre loi de l'inertie est mal exprimée. Je préfère la deuxième solution. La loi de l'inertie devrait être conçue de telle manière que les mêmes phénomènes puissent être observés comme conséquence d'une rotation relative, avec la Terre tournante et les étoiles immobiles, et vice versa. Cela suppose que la loi d'inertie, exprimée de cette façon, doit tenir compte de toutes les masses dans l'Univers.⁽⁸⁴⁾

⁽⁸²⁾ Berkeley, cité dans M. Boratav et R. Kerner, op. cit., p. 199.

⁽⁸³⁾ I. Newton, op. cit.

⁽⁸⁴⁾ Mach, cité dans M. Boratav et R. Kerner, op. cit., p. 200.

On peut résumer l'expérience de Mach de la façon suivante. Se tenant debout, par une nuit étoilée, à observer le ciel, les bras ballants le long du corps, un observateur se met à tourner autour de lui-même. Deux phénomènes se produisent alors : les étoiles se mettent à tourner et les bras de l'observateur s'écartent de l'axe de rotation. Pour Mach il est impensable que ces deux événements soient déconnectés et il doit être possible de développer une physique des référentiels accélérés où les étoiles en rotation exercent une force inertielle sur les bras. L'argument conduit à interpréter la masse inertielle d'un corps comme le résultat de son interaction gravitationnelle avec le reste de l'Univers. Cet énoncé constitue le Principe de Mach. La mécanique de laboratoire s'avère liée à la cosmologie.

Le référentiel héliocentrique, dont l'origine est située au centre de gravité du système solaire et dont les axes pointent dans les directions d'étoiles "fixes", est considéré galiléen en très bonne approximation. La relation fondamentale de la dynamique y prend donc la forme habituelle

$$\sum \mathbf{F} = m\mathbf{a},$$

où \mathbf{a} représente l'accélération (absolue) de la masse relativement au référentiel galiléen. Si l'on se place dans le référentiel terrestre, en rotation par rapport au référentiel héliocentrique, aux forces objectives (*i.e.* qui sont indépendantes du référentiel) on ajoute des forces d'inertie (force d'inertie d'entraînement et force de Coriolis) :

$$\sum \mathbf{F} - m\mathbf{a}_e - m\mathbf{a}_c = m\mathbf{a},$$

responsables par exemple de la rotation du plan d'oscillation du pendule de Foucault. Selon Mach, cet effet d'inertie peut aussi bien être interprété comme résultant de l'interaction gravitationnelle du pendule avec les étoiles fixes. On peut en donner une argumentation qualitative simple⁽⁸⁵⁾. Dans la description newtonienne, un corps de masse m soumis à une force \mathbf{F} possède une accélération $\mathbf{a} = \frac{\mathbf{F}}{m}$ par rapport à un référentiel inertiel \mathcal{R} . Considérons maintenant le mouvement de ce même corps par rapport au référentiel accéléré \mathcal{R}' de sorte que son accélération soit nulle $\mathbf{a}' = \mathbf{0}$ dans \mathcal{R}' . D'après Mach, on doit pouvoir expliquer cette accélération par l'action de forces $\mathbf{F}' = -\mathbf{F} = -m\mathbf{a}$, tout comme dans \mathcal{R} , avec $m\mathbf{a}' = \mathbf{F} + \mathbf{F}'$, où \mathbf{F}' doit provenir de l'accélération du reste de l'Univers par rapport à \mathcal{R}' , puisque cette accélération existe dans \mathcal{R}' et pas dans \mathcal{R} . Considérons la contribution à \mathbf{F}' d'une galaxie de masse M . La force gravitationnelle statique habituelle de Newton, GMm/r^2 , sommée sur tout l'Univers, a une contribution essentiellement nulle du fait de l'isotropie observable de l'Univers. De plus, cette interaction ne dépend pas de l'accélération et ne peut donc pas rendre compte de \mathbf{F}' . Si l'on demande en revanche que la gravitation soit au moins compatible avec la Relativité Restreinte, elle doit, dans sa structure, être semblable à l'électromagnétisme. On sait en effet que la force gravitationnelle statique est analogue à la force de Coulomb, mais dans le cas des charges, si q_1 et q_2 ont une accélération relative, il apparaît une force supplémentaire (rayonnement d'une charge accélérée) $F_{acc.} = q_1 q_2 a / 4\pi\epsilon_0 c^2 r$. Pour la gravitation, c'est ce type de force (appelée loi de l'induction inertielle par Sciama et devant donner lieu à des ondes gravitationnelles) qui doit être à l'origine de \mathbf{F}' :

$$F' = GMma/c^2 r.$$

⁽⁸⁵⁾ M.V. Berry, *Principles of Cosmology and Gravitation* (IOP Publishing, Cambridge 1989), pp. 37-41.

Tableau 8.4 Rayonnement gravitationnel (analogie).

Statique (courte portée)		Rayonnement (longue portée)
$\frac{qq'}{4\pi\epsilon_0 r^2}$	\rightarrow	$\frac{qq'}{4\pi\epsilon_0} \frac{a}{rc^2}$
$\frac{Gmm'}{r^2}$	\rightarrow	$Gmm' \frac{a}{rc^2}$

La décroissance lente en $\frac{1}{r}$ est satisfaisante car elle montre que les corps éloignés ont une contribution plus forte qu'à l'interaction statique. En sommant sur l'Univers observable, on obtiendra $F'^{(86)}$.

$$\begin{aligned}
 F' &= \sum_{M \in \mathcal{U}} \frac{GMma}{rc^2} = \frac{Gma\rho_{\text{gal}}}{c^2} \int_{\mathcal{U}_{\text{obs.}}} \frac{d^3r}{r} \\
 &= \frac{4\pi Gma\rho_{\text{gal}}}{c^2} \int_0^{c/H} \frac{r^2 dr}{r} \\
 &= ma \times \left\{ \frac{2\pi G\rho_{\text{gal}}}{H^2} \right\}.
 \end{aligned}$$

Au-delà de l'horizon c/H , l'Univers serait inobservable à cause d'une vitesse de récession supérieure à celle de la lumière. Avec les valeurs admises pour la densité de masse dans les galaxies $\rho_{\text{gal}} = 3 \cdot 10^{-28} \text{kg.m}^{-3}$ et la constante de Hubble $H = 55 \pm 7 \text{km/s.Mpc}$, on obtient un facteur numérique $\{\cdot\}$ de l'ordre de $1/25$, au lieu de 1. Ce désaccord n'est cependant pas dramatique car la répartition de matière en dehors des galaxies est très mal connue et dépasse vraisemblablement la quantité de matière présente dans les galaxies. On peut de plus améliorer d'un facteur 2 le coefficient numérique ci-dessous en tenant compte de la "correction relativiste" à la masse :

$$\rho_{\text{gal}} \rightarrow \rho_{\text{gal}} \frac{1}{\sqrt{1 - \frac{H^2 r^2}{c^2}}}, \quad v = Hr.$$

On trouve cette fois

$$\begin{aligned}
 F' &= \frac{4\pi Gma\rho_{\text{gal}}}{c^2} \int_0^{c/H} r dr \left(1 - \frac{H^2}{c^2} r^2\right)^{-1/2} \\
 &= ma \times \left\{ \frac{4\pi G\rho_{\text{gal}}}{H^2} \right\}.
 \end{aligned} \tag{8.6}$$

Notons toutefois que cette interprétation serait très élégante, mais qu'elle n'emporte pas à l'heure actuelle l'adhésion de la communauté.

⁽⁸⁶⁾ Dans les modèles d'Univers en expansion, on peut estimer le "rayon" de l'Univers par le rapport c/H où H est la constante de Hubble (1929). Les données spectroscopiques (effet Doppler) permettent de déterminer la vitesse d'éloignement (vitesse de récession radiale) des objets lointains (par exemple pour la galaxie NGC7619 qui se situe à $d = 65 \text{ Mpc} = 65 \times 3.0856 \cdot 10^{13} \text{km}$, $z = \frac{\lambda' - \lambda_0}{\lambda_0} = 0.0126$ conduit à $v \simeq cz = 3780 \text{ km/s}$). En supposant que cet éloignement est uniforme (et l'a été par le passé), on obtient une estimation de l'âge de l'Univers $d/v = 16.9 \cdot 10^9$ années et de la constante de Hubble $H = v/d = 58 \text{ km/s.Mpc}$.

Chapitre 9

Gravitation relativiste

Dans ce chapitre on aborde le problème de la gravitation relativiste. Contrairement au développement historique qui a vu d'abord la formulation la plus générale avec les équations d'Einstein, puis la solution particulière dans le cas statique à symétrie sphérique (Schwarzschild), nous adoptons ici une démarche simplifiée avec l'établissement de la métrique de Schwarzschild et des ses conséquences et nous mentionnons simplement quelques notions de la théorie plus générale. Ce chapitre reprend en partie les raisonnements de Kenyon ou de Berry. Les lecteurs qui veulent en savoir plus peuvent consulter l'abondante littérature sur la relativité générale et pour donner un avant-goût de ce qui les attend, je cite Landau et Lifshitz, *Théorie des champs*, §82, p.301 :

La théorie des champs de gravitation élaborée sur la base de la théorie de la relativité est appelée Relativité Générale. Elle a été créée par Einstein (qui l'a formulée définitivement en 1916) et est vraisemblablement la plus belle des théories physiques existantes. Il est remarquable qu'Einstein l'ait construite par voie purement déductive et que c'est seulement par la suite qu'elle ait été confirmée par des observations astronomiques.

1. Le décalage spectral gravitationnel

En 1907, Einstein a montré, à partir du principe d'équivalence, que les radiations électromagnétiques sont affectées par la présence d'un champ gravitationnel⁽⁸⁷⁾.

Le raisonnement utilise simplement l'effet Doppler et le principe d'équivalence qui stipule en particulier que la lumière se propage à la vitesse c dans un référentiel tangent au référentiel en chute libre (i.e. dans le référentiel animé d'une vitesse

⁽⁸⁷⁾ A. Einstein, Principe de Relativité et Gravitation, §V, op. cit., pp. 115-119.

uniforme coïncidant à l'instant donné avec la vitesse du référentiel en chute libre, et ceci à condition que la durée de l'expérience soit suffisamment brève et que l'étendue spatiale du référentiel soit petite). Considérons une capsule de longueur dr en chute libre dans le champ de gravitation $\mathbf{g} = -\frac{GM}{r^2}\mathbf{u}_r$ créé par une masse M ⁽⁸⁸⁾. Une source lumineuse placée au bas de la capsule (figure (9.1)) émet une onde électromagnétique captée ensuite par un détecteur situé au plafond de la capsule.

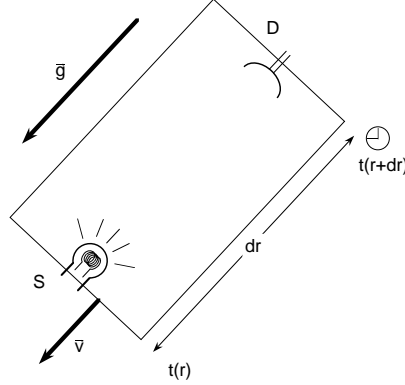


Figure 9.1 Le temps propre dépend du champ de gravitation. Pour le mettre en évidence, on considère une capsule en chute libre dans le champ de gravitation \mathbf{g} supposé uniforme. Une source située sur le sol de la capsule émet une onde électromagnétique détectée par un récepteur placé au plafond. On compare les fréquences mesurées par un observateur lié à la capsule et un observateur fixe par rapport à la source de gravitation.

Conformément au principe d'équivalence, la vitesse de propagation de l'onde c est constante dans toutes les directions dans la capsule en chute libre. Le pulse de lumière émis par la source atteint le détecteur après une durée $dt = dr/c$. Pour un observateur extérieur, immobile par rapport à la source du champ de gravitation, l'onde est détectée avec un décalage Doppler dû à l'éloignement de la source à la vitesse instantanée $v = g dt$. Au premier ordre en $\frac{v}{c}$, la variation de fréquence vaut⁽⁸⁹⁾

$$\frac{d\nu}{\nu} = -\frac{v}{c} = -\frac{g dr}{c^2}, \quad \text{ou} \quad \frac{d\nu}{\nu} = -\frac{g}{c^2} \nu.$$

À l'altitude $r + dr$, un observateur externe mesure une fréquence modifiée $\nu(r + dr) = \nu(r) + dr \frac{d\nu}{dr}$:

$$\nu(r + dr) = \nu(r) \left(1 - \frac{g dr}{c^2}\right) = \nu(r) \left(1 - \frac{GM}{r^2 c^2} dr\right).$$

Si chaque "pulse" de l'onde est utilisé comme unité de temps, les horloges situées aux distances r et $r + dr$ indiquent des temps $t(r) \sim \frac{1}{\nu(r)}$ tels que :

$$t(r + dr) = t(r) \left(1 + \frac{GM}{r^2 c^2} dr\right)$$

⁽⁸⁸⁾ I.R. Kenyon, *General Relativity* (Oxford University Press, Oxford 1990), p. 15.

⁽⁸⁹⁾ Il s'agit de l'effet Doppler longitudinal : si la source S s'éloigne à la vitesse v de l'observateur O , on a $\nu_O = \nu_S \sqrt{\frac{1-v/c}{1+v/c}} \simeq \nu_S \left(1 - \frac{v}{c} + O\left(\frac{v^2}{c^2}\right)\right)$. Au premier ordre en $\frac{v}{c}$, la relativité restreinte redonne le résultat classique de l'effet Doppler.

au premier ordre. Avec $t(r + dr) = t(r) + \frac{dt}{dr} dr$, on en déduit l'équation différentielle

$$\frac{dt}{t} = \frac{GM}{r^2 c^2} dr$$

que l'on intègre de l'infini, où le champ de gravitation n'est plus sensible, jusqu'à r , soit :

$$\begin{aligned} \ln \frac{t(\infty)}{t(r)} &= \frac{GM}{c^2} \left[-\frac{1}{r} \right]_r^\infty = \frac{GM}{rc^2} \\ t(r) &= t(\infty) \times e^{-GM/rc^2} \\ &\sim t(\infty) \left(1 - \frac{GM}{rc^2} \right). \end{aligned} \quad (9.1)$$

Cette expression est parfois écrite en fonction du potentiel gravitationnel $\psi(r) = -GM/r$:

$$t(r) = t(\infty) \left(1 + \frac{\psi(r)}{c^2} \right).$$

Comme $\psi(r)$ est toujours négatif, $t(\infty) > t(r)$, les horloges sont accélérées par le champ de gravitation.

L'équation (9.1) a été obtenue en partant de l'effet Doppler au premier ordre en $\frac{v}{c}$. A cet ordre, les résultats classique et relativiste coïncident, ce qui suppose que la vitesse d'éloignement de la source reste faible par rapport à c , ou encore, que le champ de gravitation n'est pas trop important (ce qui légitime le développement au premier ordre dans (9.1)).

Il est commode d'introduire ici la notation propre à la notion d'intervalle. $t(r)$ représente le temps propre local, plus généralement noté τ . Dans la limite des faibles champs de gravitation, on a donc au premier ordre :

$$d\tau^2 = dt^2 \left(1 - \frac{2GM}{rc^2} \right), \quad (9.2)$$

qui fournit l'une des composantes du tenseur métrique.

2. Courbure de l'espace-temps en présence d'un champ de gravitation statique à symétrie sphérique

La courbure de l'espace-temps au voisinage d'une source de champ gravitationnel s'obtient facilement grâce à la déviation géodésique (voir le chapitre sur les surfaces courbes). Une géodésique de l'espace-temps plat est une ligne droite, c'est-à-dire la trajectoire d'une particule libre dans un référentiel inertiel. Les lignes géodésiques de l'espace-temps courbe doivent être, par extension, les trajectoires de particules tests en chute libre. Considérons deux objets A et B , de masses identiques m , en chute libre dans le champ de gravitation d'une même masse M située à l'origine des coordonnées (figure 9.3).

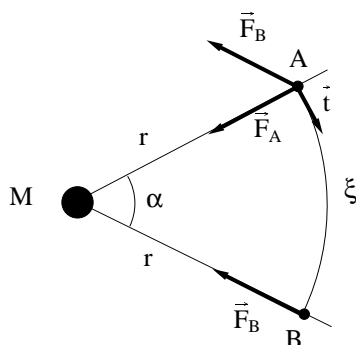


Figure 9.2 On calcule la courbure de l'espace-temps grâce à la déviation géodésique. On considère pour cela deux masses tests attirées par une même source de gravitation M et l'on calcule l'accélération relative des deux corps d'épreuve.

Du fait de leur mouvement, les deux corps A et B se rapprochent l'un de l'autre et on se propose de calculer leur accélération relative, dans une direction perpendiculaire à la géodésique (direction tangentielle),

$$F_t = m \frac{d^2 \xi}{dt^2},$$

où $\xi = \alpha r$ est l'abscisse curviligne dans la direction tangentielle. Dans le référentiel en chute libre A , la force apparente⁽⁹⁰⁾ s'exerçant sur A , projetée dans la direction tangentielle \mathbf{t} s'écrit $F_t = -F_B \sin \alpha \simeq -F_B \alpha$. En écrivant

$$F_B = \frac{GMm}{r^2}, \quad \alpha = \frac{\xi}{r},$$

on obtient l'accélération différentielle

$$\frac{d^2 \xi}{dt^2} = -\frac{GM}{r^3} \xi$$

qui ne peut disparaître que si le champ de gravitation est uniforme. Pour un mouvement supposé non relativiste, $ds^2 = c^2 d\tau^2 \simeq c^2 dt^2$, on obtient

$$\frac{d^2 \xi}{ds^2} = \frac{1}{c^2} \frac{d^2 \xi}{dt^2} = -\frac{GM}{r^3 c^2} \xi = K \xi \quad (9.3)$$

par définition de la courbure K dans un continuum hyperbolique (on modifie simplement le signe pour tenir compte de la signature des coordonnées spatiales dans l'équation métrique⁽⁹¹⁾) :

$$K = -\frac{GM}{r^3 c^2}. \quad (9.4)$$

⁽⁹⁰⁾ On néglige l'interaction entre les corps A et B , cette force apparente résulte donc de l'existence d'une accélération relative entre A et B .

⁽⁹¹⁾ Il s'agit en fait de la composante $K_{\phi t}$ du tenseur de courbure car l'accélération est transverse par rapport au mouvement de A qui se fait au cours du temps.

Cette quantité représente la courbure de l'espace-temps en présence d'un champ de gravitation statique et à symétrie sphérique. On constate que c'est la distribution de matière, à l'origine des champs gravitationnels, qui est responsable de la courbure de l'espace-temps (figure 9.4)⁽⁹²⁾.

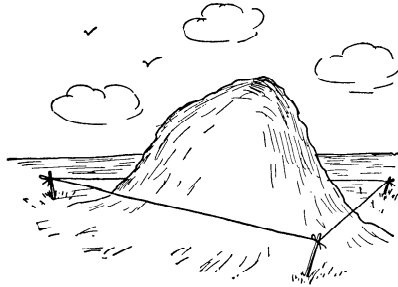


Figure 9.3 La distribution de matière est à l'origine de la courbure de l'espace-temps (tiré de Gamow, op. cit., p. 106).

3. La métrique de Schwarzschild

3.1. Tenseur métrique pour un champ statique à symétrie sphérique

Le “redshift gravitationnel” nous donne la partie temporelle de l'équation métrique au voisinage d'un objet massif. Si l'on considère un champ de gravitation à symétrie sphérique, la partie angulaire est également connue et l'on a

$$\begin{aligned} ds^2 &= c^2 d\tau^2 = c^2 \left(1 - \frac{2GM}{rc^2} \right) dt^2 + g_{rr} dr^2 - r^2 d\Omega^2 \\ &= c^2 \left(1 - \frac{2GM}{rc^2} \right) dt^2 + g_{rr} dr^2 - r^2 d\theta^2 - r^2 \sin^2 \theta d\varphi^2, \end{aligned}$$

il manque néanmoins une information sur la partie radiale.

Dans le cas d'une répartition de matière à symétrie sphérique, l'espace-temps doit être isotrope, c'est-à-dire que la courbure obtenue au paragraphe précédent doit être la même dans toutes les directions. De plus, celle-ci détermine la partie radiale du tenseur métrique par l'intermédiaire de l'“excellent théorème de Gauss”. On a donc ici, en ne considérant que l'espace tridimensionnel, donc en ignorant les signes négatifs dans les composantes du tenseur métrique (cela signifie que l'on pose ici $ds^2 = g'_{rr} dr^2 + g'_\theta d\theta^2 + g'_{\varphi\varphi} d\varphi^2$, avec $g'_{ij} = -g_{ij}$ où les g_{ij} sont finalement les composantes cherchées)

$$K = \frac{1}{2rg'^2_{rr}} \frac{dg'_{rr}}{dr} = -\frac{GM}{r^3 c^2},$$

⁽⁹²⁾ Figure extraite de G. Gamow, op. cit., p. 106

soit, par intégration entre le point courant et l'infini où $g'_{rr} = 1$ (espace euclidien),

$$\begin{aligned}\frac{dg'_{rr}}{g'_{rr}{}^2} &= -\frac{2GM}{c^2} \frac{dr}{r^2}, \\ \left[-\frac{1}{g'_{rr}} \right]_{\infty}^r &= \left[\frac{2GM}{rc^2} \right]_{\infty}^r, \\ 1 - \frac{1}{g'_{rr}} &= \frac{2GM}{rc^2},\end{aligned}$$

et finalement

$$g'_{rr} = \frac{1}{1 - \frac{2GM}{rc^2}}. \quad (9.5)$$

La métrique purement spatiale est donc de la forme

$$ds_{\text{espace}}^2 = \frac{dr^2}{1 - \frac{2GM}{rc^2}} + r^2 d\Omega^2.$$

On en déduit, en rétablissant les signes, l'équation métrique et le tenseur métrique de Schwarzschild⁽⁹³⁾ :

$$ds^2 = c^2 d\tau^2 = \left(1 - \frac{2GM}{rc^2} \right) c^2 dt^2 - \frac{dr^2}{1 - \frac{2GM}{rc^2}} - r^2 d\Omega^2, \quad (9.6a)$$

$$g_{\mu\nu} = \begin{pmatrix} 1 - \frac{2GM}{rc^2} & 0 & 0 & 0 \\ 0 & -\left(1 - \frac{2GM}{rc^2} \right)^{-1} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2 \theta \end{pmatrix}. \quad (9.6b)$$

qui redonnent bien les expressions de Minkowski dans les limites $r \rightarrow \infty$ ou $M = 0$.

Si l'on introduit le potentiel gravitationnel $\psi(r) = -GM/r$, les composantes du tenseur métrique s'écrivent

$$g_{00} = -\frac{1}{g_{11}} = 1 + \frac{2\psi(r)}{c^2}.$$

Les effets de courbure (donc de relativité générale) sont importants si $\frac{2GM}{rc^2}$ n'est pas négligeable devant 1. La quantité

$$r_g = \frac{2GM}{c^2}$$

donne l'ordre de grandeur des distances auxquelles les effets relativistes sont sensibles. A titre indicatif, à la surface de la Terre ou du Soleil, on a respectivement

$$\text{Terre : } r_{g\oplus}/R_{\oplus} = 1.4 \cdot 10^{-9},$$

$$\text{Soleil : } r_{g\odot}/R_{\odot} = 4.2 \cdot 10^{-6}.$$

⁽⁹³⁾ La métrique en présence d'un champ de gravitation statique à symétrie sphérique a été obtenue par la résolution des équations d'Einstein, moins de deux mois après la parution de l'article d'Einstein, par K. Schwarzschild, On the Gravitational Field of a Point Mass According to Einstein's theory, *Sitzber. Preuss. Akad. Wiss., Physik-Math. Kl.*, **189** (1916) 189, traduit dans *Black Holes, Selected Reprints*, Am. Ass. Physics Teachers, december 1982, p.19 ou encore dans [arXiv:physics/9905030](https://arxiv.org/abs/physics/9905030).

3.2. Equation géodésique dans la métrique de Schwarzschild

3.2.1. Retour à l'action classique

On sait qu'en mécanique classique, le lagrangien d'une particule de masse m , soumise à l'énergie potentielle $U(\mathbf{r})$, s'écrit $L = \frac{1}{2}m\mathbf{v}^2 - U(\mathbf{r})$, ce qui permet de définir l'action $S = \int_a^b L dt$, les équations du mouvement se déduisant alors du principe d'Hamilton

$$\delta S = 0, \quad \text{soit} \quad \frac{d}{dt} \frac{\partial L}{\partial \mathbf{v}} - \frac{\partial L}{\partial \mathbf{r}} = 0.$$

Pour le problème de Kepler, on remarque généralement que la conservation du moment cinétique $\boldsymbol{\sigma}$ impose une trajectoire plane. Si l'on omet pour le moment cette propriété, l'expression générale de la vitesse en coordonnées sphériques $\mathbf{v} = \dot{r}\mathbf{u}_r + r\dot{\theta}\mathbf{u}_\theta + r\sin\theta\dot{\phi}\mathbf{u}_\phi$ permet d'exprimer le lagrangien sous la forme :

$$L = \frac{1}{2}m \left[\left(\frac{dr}{dt} \right)^2 + r^2 \left(\frac{d\theta}{dt} \right)^2 + r^2 \sin^2 \theta \left(\frac{d\phi}{dt} \right)^2 \right] + \frac{GM_\odot m}{r}. \quad (9.7)$$

Les choix respectifs $q = \phi, \theta, r$ dans les équations d'Euler-Lagrange conduisent alors à

$$\begin{aligned} \frac{d}{dt} \left(mr^2 \sin^2 \theta \frac{d\phi}{dt} \right) &= 0, \\ \frac{d}{dt} \left(mr^2 \frac{d\theta}{dt} \right) &= mr^2 \sin \theta \cos \theta \left(\frac{d\phi}{dt} \right)^2, \\ \frac{d}{dt} \left(m \frac{dr}{dt} \right) &= mr \left(\frac{d\theta}{dt} \right)^2 + mr \sin^2 \theta \left(\frac{d\phi}{dt} \right)^2 - \frac{GM_\odot m}{r^2}. \end{aligned} \quad (9.8)$$

Le moment cinétique s'écrivant $\boldsymbol{\sigma} = mr^2 \left(\frac{d\theta}{dt} \right) \mathbf{u}_\phi - mr^2 \sin \theta \left(\frac{d\phi}{dt} \right) \mathbf{u}_\theta$, sa conservation conduit à

$$\begin{aligned} \frac{d\boldsymbol{\sigma}}{dt} &= -\frac{1}{\sin \theta} \frac{d}{dt} \left(mr^2 \sin^2 \theta \frac{d\phi}{dt} \right) \mathbf{u}_\theta \\ &+ \left[\frac{d}{dt} \left(mr^2 \frac{d\theta}{dt} \right) - mr^2 \sin \theta \cos \theta \left(\frac{d\phi}{dt} \right)^2 \right] \mathbf{u}_\phi = 0, \end{aligned}$$

c'est-à-dire aux deux premières équations. La troisième est équivalente à la conservation de l'énergie.

Dans le cadre de la relativité restreinte, on a vu que l'action de la particule libre est proportionnelle à la longueur de la géodésique⁽⁹⁴⁾, soit, avec la métrique de Minkowski :

$$S = -mc \int ds = \int L d\lambda,$$

$$ds = (\eta_{\mu\nu} dx^\mu dx^\nu)^{1/2} = \left(\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \right)^{1/2} d\lambda, \quad (9.9)$$

$$L = -mc \left(\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \right)^{1/2},$$

où λ est un paramètre arbitraire permettant de paramétrer la trajectoire de la particule libre. On retrouve bien entendu $L = -mc^2 \left(1 - \frac{|\mathbf{v}|^2}{c^2} \right)^{1/2}$ si λ est le temps t mesuré par un observateur.

En présence d'une énergie potentielle $U(\mathbf{r})$, l'observateur fixe pourra définir un lagrangien par l'expression⁽⁹⁵⁾

$$L = -mc^2 \left(1 - \frac{|\mathbf{v}|^2}{c^2} \right)^{1/2} - U(\mathbf{r}),$$

et les équations du mouvement découlent des équations d'Euler-Lagrange appliquées à ce lagrangien.

3.2.2. Le principe de moindre action en relativité générale

En gravitation relativiste, l'énergie potentielle sera automatiquement incluse dans les composantes $g_{\mu\nu}$ du tenseur métrique, et l'étude du mouvement d'un corps d'épreuve se réduit à un problème purement géométrique.

Tandis que, chez Newton, les équations des trajectoires doivent être postulées à part, tout à fait indépendamment de la structure de l'espace, le principe géodésique édicte ici, d'une manière extraordinairement naturelle, que les trajectoires des particules d'épreuve ne sont rien d'autre que les géodésiques de l'espace^(). L'espace une fois esquissé, après que les équations de champ ont été résolues, les trajectoires sont toutes et totalement définies, contenues, tracées, creusées dans la structure même de l'espace qui va s'en trouver précisé. (...) Qui plus est, le principe géodésique s'appuie sur le principe d'inertie qu'il généralise. A tel point que l'expression de l'un s'applique à l'autre : un corps abandonné à lui-même persiste dans son état de mouvement. Mais tandis que, pour Galilée, cet état se réduit au mouvement rectiligne uniforme, il couvre désormais l'ensemble des mouvements purement gravitationnels possibles. Ainsi les corps sont-ils abandonnés à eux-mêmes, libérés de toute contrainte sinon de celle que leur trace l'espace lui-même. L'espace qui n'est guère plus que l'ensemble des mouvements possibles⁽⁹⁶⁾.*

⁽⁹⁴⁾ F. Rohrlich, Classical Charged Particles (Addison-Wesley, Redwood City 1990), p. 277, P.J.E. Peebles, Principles of Physical Cosmology (Princeton University Press, Princeton 1993), p. 244.

⁽⁹⁵⁾ L. Brillouin, L'atome de Bohr (Presses Universitaires de France, Paris 1931), p. 105.

⁽⁹⁶⁾ J. Eisenstaedt, Trajectoires et Impasses de la Solution de Schwarzschild, Arch. Hist. Exact Sci. **37** (1988) 275, p.283.

(*) Il ne s'agit ici que des particules d'épreuve et non pas des trajectoires des corps participant au champ gravitationnel dans lequel ils naviguent.

▷ La généralisation de l'action en relativité générale conduit naturellement à la définir par l'élément curviligne le long de la courbe géodésique. Celle-ci est bien invariante par changement arbitraire de système de coordonnées⁽⁹⁷⁾.

$$S = -mc \int ds, \quad ds = (g_{\mu\nu} dx^\mu dx^\nu)^{1/2}$$

$$\delta \int ds = 0. \tag{9.10}$$

On peut alors effectuer le calcul variationnel en conservant le formalisme covariant. On note pour cela

$$\begin{aligned} \delta(ds^2) &= \delta(g_{\mu\nu} dx^\mu dx^\nu) \\ &= 2 ds \delta(ds) \\ &= dx^\mu dx^\nu \delta g_{\mu\nu} + 2g_{\mu\nu} dx^\mu \delta(dx^\nu). \end{aligned}$$

On obtient

$$\delta(ds) = \frac{1}{2} \frac{dx^\mu}{ds} dx^\nu \frac{\partial g_{\mu\nu}}{\partial x^\rho} \delta x^\rho + g_{\mu\nu} \frac{dx^\mu}{ds} d(\delta x^\nu).$$

On pose $\dot{x}^\mu = \frac{dx^\mu}{ds}$ et $g_{\mu\nu,\rho} = \frac{\partial g_{\mu\nu}}{\partial x^\rho}$, ce qui conduit à

$$\delta(ds) = \left[\frac{1}{2} \dot{x}^\mu \dot{x}^\nu g_{\mu\nu,\rho} \delta x^\rho + g_{\mu\nu} \dot{x}^\mu \frac{d}{ds}(\delta x^\nu) \right] ds$$

dont l'intégrale sur la ligne d'univers doit s'annuler,

$$\begin{aligned} \delta \int_{\text{Ligne}} ds &= \int_{\text{Ligne}} \left[\frac{1}{2} \dot{x}^\mu \dot{x}^\nu g_{\mu\nu,\rho} \right] \delta x^\rho ds + \underbrace{\int_{\text{Ligne}} \left[g_{\mu\nu} \dot{x}^\mu \frac{d}{ds}(\delta x^\nu) \right] ds}_{[g_{\mu\nu} \dot{x}^\mu \delta x^\nu]_{\partial L} - \int_L \frac{d}{ds}(g_{\mu\nu} \dot{x}^\mu) \delta x^\nu ds} \\ &= \int_{\text{Ligne}} \left[\frac{1}{2} \dot{x}^\mu \dot{x}^\nu g_{\mu\nu,\rho} - \frac{d}{ds}(g_{\mu\rho} \dot{x}^\mu) \right] \delta x^\rho ds. \end{aligned}$$

Comme d'habitude, on a annulé le terme de bords aux extrémités de la ligne d'univers. La variation est arbitraire, de sorte que l'intégrand doit être nul, soit

$$\frac{1}{2} \dot{x}^\mu \dot{x}^\nu g_{\mu\nu,\rho} - g_{\mu\rho} \ddot{x}^\mu - g_{\mu\rho,\nu} \dot{x}^\mu \dot{x}^\nu = 0.$$

⁽⁹⁷⁾ P.J.E. Peebles, op. cit., p. 245.

On arrange les termes en multipliant par $g^{\sigma\rho}$ et en notant que dans le dernier terme ci-dessus on peut permuter les indices contractés μ et ν , soit $g_{\mu\rho,\nu}\dot{x}^\mu\dot{x}^\nu = \frac{1}{2}(g_{\mu\rho,\nu}\dot{x}^\mu\dot{x}^\nu + g_{\nu\rho,\mu}\dot{x}^\nu\dot{x}^\mu)$, d'où

$$\underbrace{g^{\sigma\rho}g_{\mu\rho}\ddot{x}^\mu}_{\ddot{x}^\sigma} = g^{\sigma\rho} \left[\frac{1}{2}g_{\mu\nu,\rho} - \frac{1}{2}g_{\mu\rho,\nu} - \frac{1}{2}g_{\nu\rho,\mu} \right] \dot{x}^\mu\dot{x}^\nu.$$

En définissant le symbole de Christoffel ou connexion affine ⁽⁹⁸⁾ (contraction sur ρ),

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2}g^{\sigma\rho}[g_{\mu\rho,\nu} + g_{\nu\rho,\mu} - g_{\mu\nu,\rho}],$$

on obtient l'équation de mouvement ⁽⁹⁹⁾

$$\frac{d^2x^\sigma}{ds^2} = -\Gamma^\sigma_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds}. \quad (9.11)$$

▷ Il est souvent plus facile d'écrire explicitement les équations de Lagrange à partir du lagrangien exprimé directement en termes de coordonnées usuelles. Une planète considérée comme corps d'épreuve⁽¹⁰⁰⁾ dans le champ de gravitation du Soleil suit une géodésique de l'espace-temps de Schwarzschild. Le principe de moindre action conduit comme on l'a vu aux équations de Lagrange

$$\frac{d}{d\lambda} \left(\frac{\partial L}{\partial(dx^\mu/d\lambda)} \right) - \frac{\partial L}{\partial x^\mu} = 0 \quad (9.12)$$

⁽⁹⁸⁾ Ce n'est pas un tenseur car il est identiquement nul dans un référentiel localement inertiel et la loi linéaire de transformation des tenseurs exigerait sinon qu'il soit identiquement nul dans tout référentiel.

⁽⁹⁹⁾ On peut noter à ce sujet que la notion de dérivée covariante permet d'écrire directement l'équation géodésique. La dérivée covariante est donnée par

$$\frac{Dq^\sigma}{Ds} = \frac{dq^\sigma}{ds} + \Gamma^\sigma_{\mu\nu}q^\mu \frac{dx^\nu}{ds}.$$

En présence d'un champ de force F^μ , l'équation de mouvement devient

$$\frac{Dp^\sigma}{D\tau} = F^\sigma$$

soit dans le cas du mouvement de chute libre

$$0 = \frac{Dp^\sigma}{D\tau} = m \frac{d^2x^\sigma}{d\tau^2} + m\Gamma^\sigma_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}.$$

Précisons également des notations anciennes pour les connexions affines (symboles de Christoffel de 1ère et 2ème espèces) : $\Gamma_{\mu\nu,\rho} = [\mu\nu, \rho]$, $\Gamma^\rho_{\mu\nu} = \left\{ \begin{matrix} \rho \\ \mu\nu \end{matrix} \right\} = g^{\rho\sigma}[\mu\nu, \sigma]$.

⁽¹⁰⁰⁾ On néglige donc la courbure supplémentaire due à la planète elle-même.

que l'on simplifie en introduisant la notation traditionnelle $\dot{x}^\mu = dx^\mu/d\lambda$. En coordonnées sphériques, $x^\mu = ct, r, \theta, \phi$, on obtient explicitement

$$L = -mc \left[\left(1 - \frac{r_g}{r}\right) (c\dot{t})^2 - \left(1 - \frac{r_g}{r}\right)^{-1} \dot{r}^2 - r^2 \dot{\theta}^2 - r^2 \sin^2 \theta \dot{\phi}^2 \right]^{1/2}, \quad (9.13)$$

où $r_g = 2GM_\odot/c^2$ est le rayon gravitationnel. On peut noter que L est de la forme $L = -mcF^{1/2}$, avec

$$F = \left(\frac{ds}{d\lambda} \right)^2 = g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu.$$

Les équations de Lagrange se réduisent alors à

$$\frac{1}{2F^{1/2}} \frac{d}{d\lambda} \left(\frac{\partial F}{\partial (dx^\mu/d\lambda)} \right) - \frac{1}{2F^{1/2}} \frac{\partial F}{\partial x^\mu} = 0, \quad (9.14)$$

ce qui permet de simplifier $\frac{1}{2F^{1/2}}$ ou encore d'oublier la puissance 1/2 dans la définition de L , ce qui donne au choix

$$\begin{aligned} \frac{d}{d\lambda} \left(\frac{\partial F}{\partial (dx^\mu/d\lambda)} \right) - \frac{\partial F}{\partial x^\mu} &= 0, \\ \frac{d}{d\lambda} \left(\frac{\partial}{\partial \dot{x}^\mu} g_{\sigma\tau} \dot{x}^\sigma \dot{x}^\tau \right) - \frac{\partial}{\partial x^\mu} g_{\sigma\tau} \dot{x}^\sigma \dot{x}^\tau &= 0. \end{aligned} \quad (9.15)$$

▷ Le choix $\lambda = \tau$ correspond à une géodésique du genre temps (la particule test se déplace à une vitesse inférieure à celle de la lumière). C'est ce que l'on supposera dans la suite sauf mention contraire. En fait, que l'on choisisse τ ou un autre paramètre n'est pas très important car on cherche à se ramener à une équation de la trajectoire par élimination de cette variable λ .

▷ Lorsque l'une des coordonnées x^μ n'apparaît pas explicitement dans le lagrangien, on obtient automatiquement une intégrale première puisque la quantité $\frac{\partial L}{\partial (dx^\mu/d\tau)}$ est conservée.

- Choix $x^\mu = \phi$: l'équation de Lagrange devient dans ce cas

$$\frac{d}{d\tau} \left(\frac{\partial L}{\partial (d\phi/d\tau)} \right) = \frac{\partial L}{\partial \phi},$$

ce qui conduit à une intégrale première du mouvement (puisque L est indépendant de ϕ) correspondant à la conservation du moment cinétique :

$$\frac{d}{d\tau} \left(r^2 \sin^2 \theta \frac{d\phi}{d\tau} \right) = 0.$$

- Choix $x^\mu = \theta$:

$$\begin{aligned} \frac{d}{d\tau} \left(\frac{\partial L}{\partial (d\theta/d\tau)} \right) &= \frac{\partial L}{\partial \theta} \\ \sin \theta \cos \theta \left(\frac{d\phi}{d\tau} \right)^2 &= \frac{d^2 \theta}{d\tau^2} + \frac{2}{r} \frac{dr}{d\tau} \frac{d\theta}{d\tau}. \end{aligned}$$

• Le choix $x^\mu = ct$ donne une nouvelle intégrale première, liée cette fois à l'énergie :

$$\frac{1}{c} \frac{d}{d\tau} \left(\frac{\partial L}{\partial (dt/d\tau)} \right) = \frac{1}{c} \frac{\partial L}{\partial t}$$

$$\frac{d}{d\tau} \left[\left(1 - \frac{r_g}{r} \right) \frac{dt}{d\tau} \right] = 0.$$

• Choix $x^\mu = r$: il est cette fois plus simple de reprendre l'équation métrique elle-même, équivalente à l'intégrale première de conservation de l'énergie :

$$c^2 = \left(1 - \frac{r_g}{r} \right) c^2 \left(\frac{dt}{d\tau} \right)^2$$

$$- \left(1 - \frac{r_g}{r} \right)^{-1} \left(\frac{dr}{d\tau} \right)^2 - r^2 \left(\frac{d\theta}{d\tau} \right)^2 - r^2 \sin^2 \theta \left(\frac{d\phi}{d\tau} \right)^2.$$

La comparaison avec la mécanique classique et le traitement en relativité restreinte est immédiate si l'on se réfère aux équations du paragraphe précédent.

4. Les trois tests classiques de la théorie de la relativité générale

4.1. Vérification du décalage spectral gravitationnel

L'expression

$$d\tau^2 = dt^2 \left(1 - \frac{2GM}{rc^2} \right)$$

pour le décalage spectral gravitationnel a pu être vérifiée expérimentalement au moyen d'horloges embarquées qui réalisent l'expérience des jumeaux. Deux horloges sont synchronisées, l'une sert de référence, elle reste au sol, l'autre est embarquée dans un avion qui fait le tour de la Terre. Les indications des deux horloges sont comparées au moyen de signaux électromagnétiques. Le problème complet est extrêmement ardu, car au problème de relativité restreinte des jumeaux, s'ajoutent deux autres phénomènes, celui dû au mouvement non inertiel de l'horloge embarquée (qui peut entrer dans le traitement de relativité restreinte) et celui dû à la gravitation que l'on cherche à tester. Au total il reste une désynchronisation résiduelle mesurée représentée sur la figure ci-dessous et qui est en parfait accord avec l'expression $d\tau = dt \left(1 - \frac{GM}{rc^2} \right)$.

Des mesures beaucoup plus précises, utilisant l'effet Mössbauer, ont été réalisées par Pound et Rebka ⁽¹⁰¹⁾. Elles confirment également l'expression du décalage spectral gravitationnel.

⁽¹⁰¹⁾ R.V. Pound and G.A. Rebka, *Apparent weight of photons*, Phys. Rev. Lett. **4**, 337 (1960).

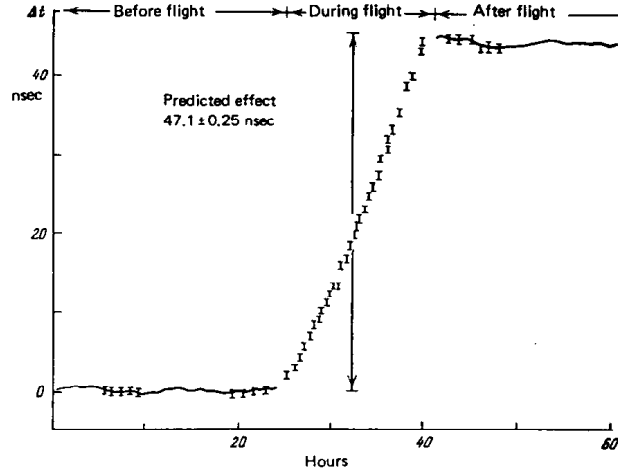


Figure 9.4 Décalage gravitationnel des fréquences des raies spectrales : désynchronisation d'une horloge embarquée par rapport à une horloge au sol.

	Voyage d'Ouest en Est	Voyage d'Est en Ouest
Différence observée	$(-59 \pm 10) \times 10^{-9}$	$(273 \pm 7) \times 10^{-9}$
Correction cinématique	$(-184 \pm 18) \times 10^{-9}$	$(96 \pm 10) \times 10^{-9}$
Reste	$(125 \pm 21) \times 10^{-9}$	$(177 \pm 12) \times 10^{-9}$
Effet gravitationnel prédit	$(144 \pm 14) \times 10^{-9}$	$(179 \pm 18) \times 10^{-9}$

Tableau 9.1 Dilatation du temps (différence mesurée par rapport à une horloge de référence) mesurée à l'aide d'horloges atomiques embarquées (*The story of general relativity*, A.P. French, in *Einstein, a centenary volume* Ed. by A.P. French, Heinemann, London, 1979).

4.2. Avance du périhélie de Mercure

On exprime la métrique de Schwarzschild dans le plan équatorial $\theta = \pi/2$ afin d'éliminer déjà une variable,

$$c^2 d\tau^2 = g_{00}(r)c^2 dt^2 - dr^2/g_{00}(r) - r^2 d\phi^2$$

avec $g_{00}(r) = 1 - 2GM_{\odot}/rc^2$. Si l'on multiplie cette équation par $m^2/d\tau^2$, on obtient

$$m^2 c^2 = g_{00}(r)m^2 c^2 \left(\frac{dt}{d\tau}\right)^2 - \frac{m^2}{g_{00}(r)} \left(\frac{dr}{d\tau}\right)^2 - m^2 r^2 \left(\frac{d\phi}{d\tau}\right)^2,$$

et l'on remarque aisément que cette équation correspond à la norme du quadrivecteur énergie-impulsion. En effet, en l'absence de gravitation, cette équation se réduit à

$$m^2 c^2 = m^2 c^2 \left(\frac{dt}{d\tau}\right)^2 - m^2 \left(\frac{dr}{d\tau}\right)^2 - m^2 r^2 \left(\frac{d\phi}{d\tau}\right)^2,$$

soit, avec $\frac{dt}{d\tau} = \gamma$,

$$\begin{aligned} m^2 c^2 &= \gamma^2 m^2 c^2 - \gamma^2 m^2 v_r^2 - \gamma^2 m^2 v_\phi^2 \\ &= \gamma^2 m^2 c^2 - \gamma^2 m^2 |\mathbf{v}|^2 \\ &= E^2/c^2 - |\mathbf{p}|^2. \end{aligned}$$

Considérons maintenant les deux intégrales premières du mouvement,

$$\begin{aligned} g_{00}(r) \frac{dt}{d\tau} &= \left(1 - \frac{2GM_\odot}{rc^2}\right) \frac{dt}{d\tau} = \text{const.} = E \\ r^2 \frac{d\phi}{d\tau} &= \text{const.}' = J. \end{aligned} \tag{9.16}$$

Nous avons noté E et J les constantes pour rappeler le contenu physique des quantités conservées. La première par exemple est liée à la composante 0 de p_μ , car $cp_0 = g_{00}p^0 = g_{00}mc \frac{dt}{d\tau}$.

Reprenons l'équation initiale issue de l'expression de la métrique, en remplaçant $dt/d\tau$ par $E/g_{00}(r)$, $d\phi/d\tau$ par J/r^2 et $dr/d\tau$ par $(dr/d\phi)J/r^2$. Il vient

$$m^2 c^2 = g_{00}(r) m^2 c^2 \left(\frac{E}{g_{00}(r)}\right)^2 - \frac{m^2}{g_{00}(r)} \frac{J^2}{r^4} \left(\frac{dr}{d\phi}\right)^2 - m^2 r^2 \left(\frac{J}{r^2}\right)^2.$$

On utilise le changement de variable habituel dans le traitement du problème de Kepler, $u = 1/r$, soit $dr/d\phi = -(1/u^2) du/d\phi$, qui conduit à

$$\begin{aligned} \left(1 - \frac{2GM_\odot u}{c^2}\right) m^2 c^2 &= m^2 c^2 E^2 - m^2 J^2 \left(\frac{du}{d\phi}\right)^2 \\ &\quad - \left(1 - \frac{2GM_\odot u}{c^2}\right) m^2 J^2 u^2. \end{aligned}$$

Pour obtenir l'équation de la trajectoire, on dérive par rapport à ϕ , et après simplification il subsiste la même équation que dans le cas newtonien, modifiée par un terme correctif non linéaire,

$$\underbrace{\frac{d^2 u}{d\phi^2} + u}_{\text{cas newtonien}} = \underbrace{\frac{GM_\odot}{J^2} + \frac{3GM_\odot u^2}{c^2}}_{\text{correction}}.$$

Cette équation n'admet pas de solution exacte, mais on peut utiliser un traitement perturbatif. Le terme non linéaire écarte la solution d'une solution harmonique périodique en ϕ . On pose alors $\xi = \Omega\phi$ avec $\Omega = 1$ pour la solution newtonienne, soit $\Omega = 1 + \omega_1 + \dots$. La solution newtonienne (de $u''(\phi) + u(\phi) = GM_\odot/J^2 = 1/\ell$) est notée $u_0(\phi) = \ell^{-1}(1 + e \cos \phi)$. On développe alors en série perturbative

$$\begin{aligned} u(\xi) &= u_0(\xi) + u_1(\xi) + \dots \\ \xi &= \Omega\phi \\ \Omega &= 1 + \omega_1 + \dots \end{aligned} \tag{9.17}$$

où toute la dépendance en $\cos \phi$ est contenue dans le terme d'ordre zéro, ω_1 représente une petite correction responsable du déplacement du périhélie, qui doit apparaître dans le terme d'ordre un.

On a donc $\frac{d}{d\phi} = \Omega \frac{d}{d\xi}$, soit

$$\Omega^2 u''(\xi) + u(\xi) = \frac{1}{\ell} + \frac{3GM_{\odot}}{c^2} u^2(\xi).$$

Cette équation à l'ordre zéro redonne bien

$$\begin{aligned} u_0''(\xi) + u_0(\xi) &= \ell^{-1}, \\ \xi &= \phi, \\ u_0(\xi) &= \ell^{-1}(1 + e \cos \xi). \end{aligned} \tag{9.18}$$

A l'ordre un on obtient

$$\begin{aligned} 2\omega_1 \underbrace{u_0''(\xi)}_{-e\ell^{-1} \cos \xi} + u_1''(\xi) + u_1(\xi) &= \frac{3GM_{\odot}}{c^2} \underbrace{u_0^2(\xi)}_{\left(\frac{1+e \cos \xi}{\ell}\right)^2}, \\ u_1''(\xi) + u_1(\xi) &= \frac{3GM_{\odot}}{\ell^2 c^2} (1 + 2e \cos \xi + e^2 \underbrace{\cos^2 \xi}_{\frac{1+\cos 2\xi}{2}}) \\ &\quad + \frac{2\omega_1 e}{\ell} \cos \xi. \end{aligned}$$

On annule le terme proportionnel à $\cos \xi$, déjà présent dans la solution à l'ordre zéro, soit

$$\frac{3GM_{\odot}}{\ell^2 c^2} \left(2e + \frac{2\omega_1 e}{\ell} \frac{\ell^2 c^2}{3GM_{\odot}} \right) = 0,$$

soit

$$\omega_1 = -\frac{3GM_{\odot}}{\ell c^2}, \tag{9.19}$$

d'où l'avance du périhélie

$$\Delta\phi = 2\pi\omega_1 = \frac{6\pi GM_{\odot}}{\ell c^2} = \frac{6\pi GM_{\odot}}{a(1-e^2)c^2}. \tag{9.20}$$

En utilisant les valeurs numériques typiques, à savoir $G = 6.67 \times 10^{-11}$ U.S.I., $M_{\odot} = 2.10^{30}$ kg et pour Mercure, $a = 57.9$ millions de km, $e = 0.205$, $T = 0.240$ années on obtient un résultat de l'ordre de

$$\begin{aligned} &5.05 \times 10^{-7} \text{ rad/révolution,} \\ &2.89 \times 10^{-5} \text{ } ^{\circ}\text{/révolution,} \\ &0.012^{\circ}\text{/siècle,} \\ &43.2''\text{/siècle.} \end{aligned}$$

On a donné au chapitre 4 un bilan des diverses contributions à l'avance du périhélie de Mercure. Ces résultats sont rappelés dans le tableau qui suit ⁽¹⁰²⁾. La somme des contributions des différentes planètes est de $531.63''$ d'arc par siècle, valeur extrêmement proche du résultat de Clemence⁽¹⁰³⁾ de $531.54''$.

La comparaison entre observation et prédiction relativiste pour d'autres planètes est également indiquée dans le tableau suivant ⁽¹⁰⁴⁾.

Cause	Avance du périhélie de Mercure
Mercure (???)	0.025 ± 0.00
Vénus	277.856 ± 0.68
Terre	90.038 ± 0.08
Mars	2.536 ± 0.00
Jupiter	153.584 ± 0.00
Saturne	7.302 ± 0.01
Uranus	0.141 ± 0.00
Neptune	0.042 ± 0.00
Applatissage du Soleil	0.010 ± 0.02
Précession générale	5025.645 ± 0.50
Somme	5557.18 ± 0.85
Avance observée	5599.74 ± 0.41
Différence	42.56 ± 0.94

Tableau 9.2 Contribution des planètes externes et de l'applatissage du Soleil à l'avance du périhélie de Mercure. Les résultats sont tirés de l'article de Clemence de 1947, qui présente aussi les résultats concernant l'avance du périhélie de la Terre.

Corps	Avance observée	Prédiction relativiste
Mercure	43.11 ± 0.45	43.03
Vénus	8.4 ± 4.8	8.6
Terre	5.0 ± 1.2	3.8
Icare	9.8 ± 0.8	10.3

Tableau 9.3 Avance du périhélie de plusieurs corps du système solaire.

4.3. Déflexion des rayons lumineux

L'un des grands succès de la théorie de la relativité générale a été de prédire correctement la déviation des rayons lumineux par les champs de gravitation. L'équivalence masse-énergie de la relativité restreinte conduit à associer au photon une masse $h\nu/c^2$, ce qui permet alors de reprendre le calcul classique de la déflexion d'une particule massive en l'appliquant au photon. Un photon passant ainsi au

⁽¹⁰²⁾ Des données très précises concernant Mercure peuvent être trouvées dans *Mercury: the planet and its orbit*, A. Balogh and G. Giampieri, Rep. Prog. Phys. **65** (2002) 529-560.

⁽¹⁰³⁾ G.M. Clemence, The Relativity Effect in Planetary Motions, *Rev. Mod. Phys.* **19** (1947) 361, G.M. Clemence, Relativity Effect in Planetary Motions, *Proc. Am. Phil. Soc.* **93** (1949) 532.

⁽¹⁰⁴⁾ *The story of general relativity*, A.P. French, in *Einstein, a centenary volume* Ed. by A.P. French, Heinemann, London, 1979.

voisinage immédiat du Soleil devrait être dévié d'une quantité

$$\Delta\phi = \frac{2GM_{\odot}}{R_{\odot}c^2}. \tag{9.21}$$

Cette valeur de la déflexion de la lumière a été initialement proposée par Söldner en 1801⁽¹⁰⁵⁾, à une époque où la conception corpusculaire prévalait encore. Le premier calcul effectué par Einstein à l'aide du principe d'équivalence en 1911 conduira à la même expression⁽¹⁰⁶⁾, mais sa théorie complète de la relativité générale (1916) fournira en fait une valeur double⁽¹⁰⁷⁾.

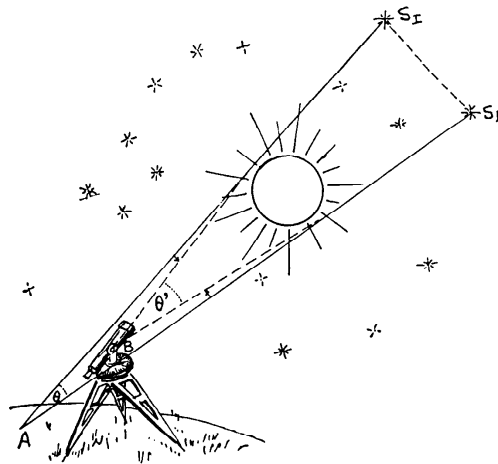


Figure 9.5 Comment tracer un triangle dans l'Univers ? (tiré de Gamow, op. cit.)

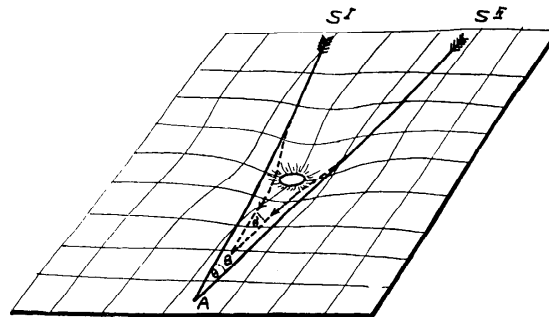


Figure 9.6 Comment tracer un triangle dans l'Univers ? (tiré de Gamow, op. cit.)

⁽¹⁰⁵⁾ S.L. Jaki, Johann Georg von Söldner and the Gravitational Bending of Light, *Found. Phys.* **8** (1978) 927.

⁽¹⁰⁶⁾ A. Einstein, De l'influence de la pesanteur sur la propagation de la lumière. In : Oeuvres choisies 2, *Relativités I (Relativités restreinte et générale)*, textes choisis par F. Balibar, O. Darrigol, B. Jech et J. Stachel (Seuil/CNRS, Paris 1993), pp. 134-142.

⁽¹⁰⁷⁾ A. Einstein, Les fondements de la théorie de la Relativité Générale, op. cit., pp. 179-227.

Reprenons ce calcul à l'aide de la métrique de Schwarzschild. Pour le photon, $ds^2 = 0$, et l'équation de la trajectoire devient simplement

$$\frac{d^2 u}{d\phi^2} + u = \frac{3GM_{\odot} u^2}{c^2}.$$

Une solution approchée en l'absence de second membre est donnée par

$$u(\phi) = b^{-1} \cos \phi$$

où b est le paramètre d'impact (homogène à une distance) et l'origine des angles est choisie pour ne pas avoir de déphasage. Cette solution est simplement la ligne droite passant à distance b du Soleil. Reportée dans l'équation de départ, elle donne

$$\frac{d^2 u}{d\phi^2} + u = \frac{3GM_{\odot} \cos^2 \phi}{b^2 c^2}$$

dont une intégrale particulière, cherchée de la forme $u_0 = A + B \cos^2 \phi$, est

$$u_0(\phi) = \frac{GM_{\odot}(2 - \cos^2 \phi)}{b^2 c^2}$$

de sorte que

$$u(\phi) = b^{-1} \cos \phi + \frac{GM_{\odot}(2 - \cos^2 \phi)}{b^2 c^2}.$$

Lorsque $u \rightarrow 0$ et $\cos \phi \rightarrow 0$ (dans ce cas $\cos^2 \phi \ll \cos \phi$), on a

$$\cos \phi = -2GM_{\odot}/bc^2.$$

On en déduit deux valeurs possibles, $\phi \rightarrow \pi/2 + 2GM_{\odot}/bc^2$ ou $\phi \rightarrow 3\pi/2 - 2GM_{\odot}/bc^2$, soit par différence,

$$\Delta\phi = 4GM_{\odot}/bc^2.$$

Pour un passage rasant au voisinage du Soleil ($b = R_{\odot}$), on obtient le résultat $1.75''$ finalement confirmé expérimentalement après bien des aieas par les expéditions dirigées par Eddington en 1919, lors d'une éclipse totale du Soleil, observable depuis les côtes du Brésil (Sobral, $1.98 \pm 0.12''$) et d'Afrique (Principe, $1.61 \pm 0.30''$). Les deux figures 9.5 et 9.6 représentent schématiquement l'effet de déflexion des rayons lumineux par le Soleil. La figure 9.7 donne des résultats expérimentaux bien plus récents, qui confirment sans ambiguïté le résultat d'Einstein.

Les mirages gravitationnels constituent une manifestation spectaculaire de la déflexion gravitationnelle des rayons lumineux. Si un objet massif non visible ⁽¹⁰⁸⁾ est interposé entre un observateur et une source, les rayons lumineux peuvent parvenir à l'observateur en contournant l'objet invisible de plusieurs directions apparentes,

⁽¹⁰⁸⁾ Le terme visible est ici employé au sens large pour la longueur d'onde du rayonnement utilisé.

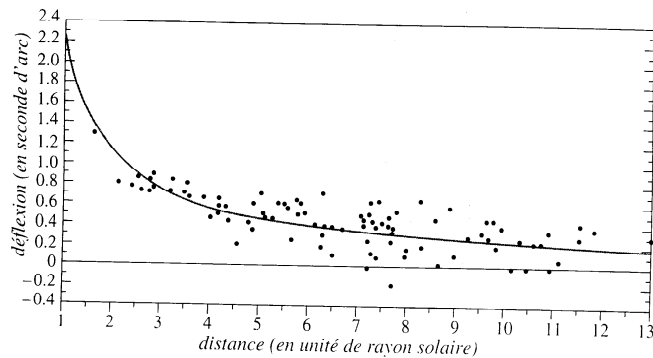


Figure 9.7 Vérifications expérimentales de la déviation des rayons lumineux (tiré de Hakim, op. cit. p.195).

produisant le phénomène de mirage gravitationnel. Les deux figures suivantes en fournissent une illustration, on voit tout d'abord 4 objets distincts très proches les uns des autres, puis après analyse des spectres émis qui s'avèrent parfaitement identiques, on constate qu'il s'agit en fait d'un seul objet.

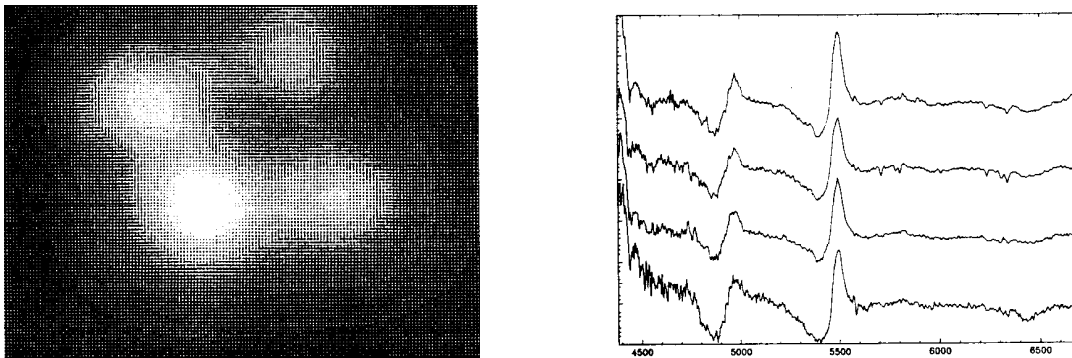


Figure 9.8 Mirage gravitationnel : quatre sources extrêmement voisines, mais les 4 spectres observés sont caractéristiques d'une seule et même source.

Dans certains cas, on peut observer des sources déformées, donnant l'illusion d'une croix ou d'un anneau, qu'on appelle croix d'Einstein ou anneau d'Einstein.

4.4. Un quatrième test, retard des échos radar

Un quatrième test a été proposé bien après les trois premiers. Il s'agit du retard des échos radar. On compare le temps mis par la lumière pour effectuer un trajet aller-retour de la Terre à une planète donnée en condition ordinaire et lorsque le Soleil est intercalé sur le trajet. Dans ce cas, il apparaît un retard lié à un effet gravitationnel ⁽¹⁰⁹⁾ et les résultats expérimentaux sont une fois encore en accord remarquable avec les prévisions de la relativité générale.

⁽¹⁰⁹⁾ On peut par exemple consulter Kenyon.

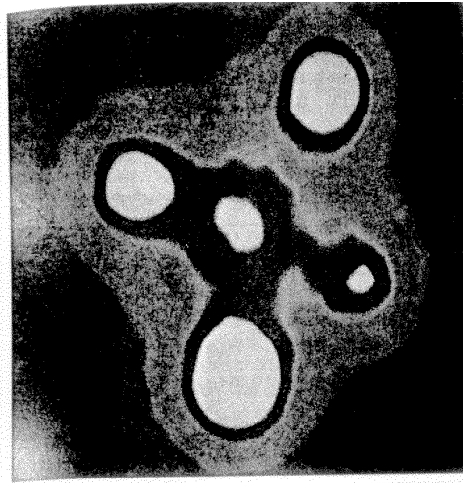


Figure 9.9 Une autre manifestation des mirages gravitationnels : la croix d'Einstein (tiré de *La Science au présent*, T.1, Encyclopædia Universalis, Paris 1992, p.81).

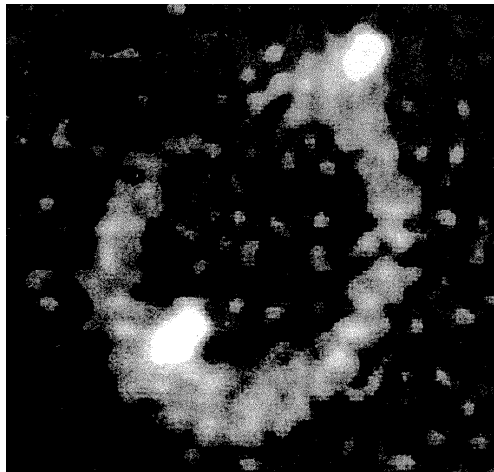


Figure 9.10 Encore une manifestation des mirages gravitationnels, un anneau d'Einstein (tiré de *La Science au présent*, T.1, Encyclopædia Universalis, Paris 1992, p.83).

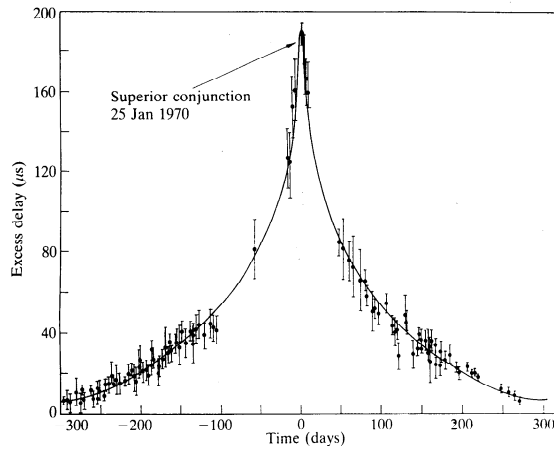


Figure 9.11 Vérifications expérimentales de la déviation des rayons lumineux : retard des echos radar (tiré de Kenyon, op. cit. p.97).

5. Les équations d'Einstein

On présente dans cette section la forme générale des équations d'Einstein. Il ne s'agit pas de démontrer les expressions mises en jeu, mais arrivé à ce stade où l'essentiel des éléments est défini, il serait dommage de ne pas pousser encore un peu le formalisme. Le lecteur intéressé pourra consulter n'importe quel ouvrage, notamment celui de Kenyon, particulièrement agréable (et si j'ose dire facile) à lire (chap. 7).

Les équations d'Einstein constituent une généralisation de l'équation de Poisson pour le potentiel gravitationnel $\psi(\mathbf{r})$,

$$\vec{\nabla}^2 \psi(\mathbf{r}) = 4\pi G \rho(\mathbf{r}),$$

défini en fonction de la densité de masse locale $\rho(\mathbf{r})$. Le rôle du potentiel $\psi(\mathbf{r})$ est joué par le tenseur métrique (rappelons par exemple dans le cas de la métrique de Schwarzschild que $g_{00} = -1/g_{11} = 1 + 2\psi/c^2$) et celui de la densité de masse par le tenseur énergie impulsion ⁽¹¹⁰⁾. L'équation d'Einstein est donc une équation différentielle pour les composantes du tenseur métrique, dont les sources sont données par le tenseur énergie impulsion.

5.1. Tenseur de Riemann, tenseur énergie impulsion, tenseur d'Einstein

▷ On cherche tout d'abord à caractériser la courbure d'un espace arbitraire (l'espace-temps) au moyen d'un tenseur. Rappelons qu'il est toujours possible de trouver en un point quelconque un référentiel localement tangent qui soit inertiel (le référentiel en chute libre), c'est-à-dire tel que

$$g_{\mu\nu} = \eta_{\mu\nu}$$

$$\frac{\partial g_{\mu\nu}}{\partial x^\rho} = 0.$$

⁽¹¹⁰⁾ Dans les ouvrages en anglais on utilise en général le terme de *stress energy tensor*.

L'information sur la courbure provient donc des dérivées d'ordre supérieur du tenseur métrique. Ce qui distingue notamment un espace courbe d'un espace plat, c'est que dans le premier, les référentiels localement tangents en x et en $x + \Delta x$ diffèrent, de sorte que les variations du tenseur métrique sont du second ordre,

$$g_{\mu\nu}(x + \Delta x) = \eta_{\mu\nu} + \frac{1}{2}g_{\mu\nu,\rho\sigma}(x)\Delta x^\rho\Delta x^\sigma$$

$$g_{\mu\nu,\rho}(x + \Delta x) = g_{\mu\nu,\rho\sigma}(x)\Delta x^\sigma.$$

L'information cherchée réside dans un tenseur du 4ème ordre, le *tenseur de Riemann* $R^\mu{}_{\nu\rho\sigma}$, défini par les connexions affines ⁽¹¹¹⁾,

$$R^\mu{}_{\nu\rho\sigma} = \Gamma^\mu{}_{\nu\sigma,\rho} - \Gamma^\mu{}_{\nu\rho,\sigma} + \Gamma^\mu{}_{\lambda\rho}\Gamma^\lambda{}_{\nu\sigma} - \Gamma^\mu{}_{\lambda\sigma}\Gamma^\lambda{}_{\nu\rho}.$$

Dans un référentiel inertiel, les connexions affines sont nulles mais pas leurs dérivées et il vient $R^\mu{}_{\nu\rho\sigma} = \Gamma^\mu{}_{\nu\sigma,\rho} - \Gamma^\mu{}_{\nu\rho,\sigma}$. On définit aussi la forme totalement covariante $R_{\mu\nu\rho\sigma} = g_{\mu\lambda}R^\lambda{}_{\nu\rho\sigma}$ telle que dans un référentiel inertiel $2R_{\mu\nu\rho\sigma} = g_{\mu\sigma,\nu\rho} - g_{\nu\sigma,\mu\rho} + g_{\nu\rho,\mu\sigma} - g_{\mu\rho,\nu\sigma}$. Notons que le tenseur de Riemann possède 256 composantes dont en fait seulement 20 sont indépendantes en raison des symétries.

▷ Énergie et impulsion sont deux aspects d'un même objet, la 4-impulsion p_μ telle que $E^2 - c^2|\mathbf{p}|^2 = m^2c^4$. Voyons maintenant comment construire le tenseur énergie impulsion en généralisant la densité de masse ρ . Considérons un nuage de poussière. Dans le référentiel \mathcal{R} dans lequel il est au repos, il possède une densité d'énergie $\rho c^2 = mnc^2 = E/V$ s'il est constitué de grains de masses m et de densité n (dans \mathcal{R}). Ici V est le volume occupé par le gaz et E son énergie. Dans un autre référentiel \mathcal{R}' , on aura une augmentation de l'énergie $E \rightarrow E' = \gamma E$ et une contraction du volume, $V \rightarrow V' = \gamma^{-1}V$ à cause de la transformation des longueurs longitudinales, $r_{\parallel} \rightarrow r'_{\parallel} = \gamma^{-1}r_{\parallel}$ de sorte que $\rho \rightarrow \rho' = \gamma^2\rho$. Il est donc évident que ρc^2 ne peut ni être un scalaire ni une composante de quadrivecteur, mais doit être la composante d'un tenseur de rang 2 ⁽¹¹²⁾. On définit ainsi le tenseur énergie impulsion $T^{\mu\nu}$ pour un gaz dilué,

$$T^{\mu\nu} = \rho v^\mu v^\nu.$$

Dans le référentiel \mathcal{R} , on a la quantité cherchée $T^{00} = \rho c^2$. Si l'on prend la pression en compte, l'expression devient

$$T^{\mu\nu} = (\rho + p/c^2)v^\mu v^\nu - g^{\mu\nu}p.$$

Une définition qui s'applique à des systèmes plus complexes que le nuage de poussière considéré ici est que $T^{\mu\nu}$ représente le flux de la composante μ du 4-vecteur énergie impulsion p^μ le long de la direction ν . Il obéit à une loi de conservation ⁽¹¹³⁾,

$$T^{\mu\nu}{}_{;\nu} = 0.$$

⁽¹¹¹⁾ On rappelle que $\Gamma^\sigma{}_{\mu\nu} = \frac{1}{2}g^{\sigma\rho}[g_{\mu\rho,\nu} + g_{\nu\rho,\mu} - g_{\mu\nu,\rho}]$.

⁽¹¹²⁾ C'est pour cette raison que l'on ne peut pas construire une théorie de la gravitation calquée sur l'électromagnétisme, car la densité de charge se transformant avec un seul facteur γ , elle est la composante d'un 4-vecteur, la 4-densité de courant j^μ et le tenseur de Faraday obéit à $\partial_\mu F^{\mu\nu} = \mu_0 j^\nu$. Dans une théorie analogue de la gravitation, le second membre est nécessairement un tenseur de rang 2, ce qui complique l'équation cherchée.

⁽¹¹³⁾ En relativité générale, cette loi de conservation est étendue à un référentiel arbitraire en passant à la dérivée covariante, $T^{\mu\nu}{}_{;\nu} = 0$.

Finalement, on retiendra que $T^{\mu\nu}$

- est un tenseur de rang 2,
- s'annule en l'absence de matière,
- a toutes ses divergences nulles,
- est symétrique.

▷ Einstein a identifié le tenseur énergie impulsion comme la source de la courbure de l'espace-temps et a ainsi suggéré la forme la plus simple possible de relation entre l'énergie $T_{\mu\nu}$ et la courbure $G_{\mu\nu}$, soit $KT_{\mu\nu} = G_{\mu\nu}$ où K est une constante à déterminer pour restaurer la limite newtonienne et $G_{\mu\nu}$, le *tenseur d'Einstein*, décrit la courbure. $G_{\mu\nu}$ est de rang 2 et de divergence nulle (à cause de ces mêmes propriétés satisfaites par $T_{\mu\nu}$). Il doit être construit à partir du tenseur de Riemann. Celui-ci est de rang 4, il faut donc une double contraction ⁽¹¹⁴⁾, c'est le *tenseur de Ricci*

$$R_{\mu\nu} = R^{\rho}{}_{\mu\rho\nu} = g^{\rho\sigma} R_{\sigma\mu\rho\nu}.$$

Le tenseur de Ricci n'est pas de divergence nulle, ce que l'on rétablit par soustraction

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R,$$

où $R = g^{\mu\nu}R_{\mu\nu}$ est le *scalaire de Ricci*. L'équation d'Einstein peut maintenant s'écrire

$$G_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}.$$

Une généralisation apportée par Einstein lors d'applications en cosmologie consiste en l'introduction d'une nouvelle constante fondamentale, la constante cosmologique Λ , de sorte que l'équation d'Einstein devient alors

$$G_{\mu\nu} - \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}.$$

Cette constante a pour effet de produire une courbure de l'espace-temps même en l'absence de toute forme de matière ou de radiation ⁽¹¹⁵⁾. Elle permet d'obtenir un univers statique en cosmologie.

⁽¹¹⁴⁾ On peut montrer que d'autres contractions donnent des résultats équivalents.

⁽¹¹⁵⁾ C'est pour cette raison d'ailleurs qu'Eddington argumentait sur le rôle indispensable de la constante cosmologique. C'est la seule façon, d'après lui, d'introduire en physique une longueur intrinsèque, le rayon de courbure de l'Univers dans la limite d'un Univers vide, comme c est une constante fondamentale de la physique, la vitesse de propagation de particules non massives dans le vide. (A. Eddington, *The Expanding Universe*, Cambridge University Press, Cambridge 1933, chap. IV.).

5.2. Limite newtonienne de l'équation d'Einstein

Avant de considérer la limite newtonienne, donnons une forme plus pratique à l'équation d'Einstein. Pour cela on prend la trace de l'équation d'Einstein,

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu} + \Lambda g_{\mu\nu}$$

soit

$$R^\mu{}_\mu - \frac{1}{2}g^\mu{}_\mu R = \frac{8\pi G}{c^4}T^\mu{}_\mu + \Lambda g^\mu{}_\mu.$$

Par définition, $R = R^\mu{}_\mu$ et comme $g^\mu{}_\mu$ est un scalaire invariant, il prend même valeur dans tous les référentiels, notamment dans un référentiel inertiel localement plat, $g^\mu{}_\mu = \eta^\mu{}_\mu = \eta^0_0 + \eta^1_1 + \eta^2_2 + \eta^3_3 = \eta_{00} - \eta_{11} - \eta_{22} - \eta_{33} = 4$. En notant $T = T^\mu{}_\mu$ la trace du tenseur énergie impulsion, il vient finalement

$$-R = \frac{8\pi G}{c^4}T + 4\Lambda$$

et l'équation d'Einstein devient

$$R_{\mu\nu} = \frac{8\pi G}{c^4}(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}) - \Lambda g_{\mu\nu}.$$

Considérons maintenant le cas d'un champ gravitationnel faible, pour lequel le tenseur métrique dévie peu de celui de Minkowski,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$$

avec $\eta_{00} = 1$, $\eta_{ii} = -1$ et les termes non diagonaux sont nuls. On se place également dans la limite des vitesses faiblement relativistes, c'est-à-dire que les composantes spatiales des moments sont petites devant les composantes temporelles (énergie). La partie importante de l'équation d'Einstein dans la limite classique est donc

$$R_{00} = \frac{8\pi G}{c^4}(T_{00} - \frac{1}{2}Tg_{00}) - \Lambda g_{00}.$$

Calculons $R_{00} = R^\mu{}_{0\mu 0} = g^{\mu\nu}R_{\nu 0\mu 0}$,

$$R_{00} \simeq \frac{1}{2}g^{\mu\nu}(g_{\mu 0,0\nu} - g_{00,\mu\nu} + g_{0\nu,\mu 0} - g_{\mu\nu,00}).$$

Les dérivées secondes valent $g_{\mu\nu,\rho\sigma} = h_{\mu\nu,\rho\sigma}$ et on ignore les dérivées temporelles (encore une fois, c est plus grande que toutes les vitesses mises en jeu) de sorte qu'il reste uniquement un terme, $-h_{00,\mu\nu} = -h_{00,ij}$ et

$$\begin{aligned} R_{00} &\simeq \frac{1}{2}\eta^{\mu\nu}(-h_{00,\mu\nu}) \\ &= -\frac{1}{2}\eta^{ij}h_{00,ij} \\ &= +\frac{1}{2}h_{00,ii}. \end{aligned}$$

Dans le cas d'un champ de gravitation à symétrie sphérique, h_{00} est donné par la métrique de Schwarzschild,

$$h_{00} = 2 \frac{\psi}{c^2}$$

soit

$$R_{00} = \frac{1}{c^2} \nabla^2 \psi.$$

Si l'on reporte cette expression dans la composante 00 de l'équation d'Einstein avec de plus $T_{00} = \rho c^2$ et $T \simeq T_{00}$, il vient

$$\frac{1}{c^2} \nabla^2 \psi \simeq \frac{8\pi G}{c^4} (\rho c^2 - \frac{1}{2} \rho c^2) - \Lambda$$

c'est-à-dire finalement l'équation de Poisson au terme en Λ près,

$$\nabla^2 \psi(\mathbf{r}) \simeq 4\pi G \rho(\mathbf{r}) - \Lambda c^2.$$

On peut noter que la constante cosmologique (si elle est positive) joue le rôle d'une force répulsive linéaire (ou d'une pression négative). L'ordre de grandeur est de $|\Lambda| \leq 10^{-52} \text{m}^{-2}$ de sorte que sur Terre l'accélération due à la constante cosmologique est à peu près 10^{-22} fois celle qu'exerce le soleil.

5.3. Action d'Einstein - Hilbert

Une formulation alternative de l'équation d'Einstein repose sur la définition d'une action, proposée initialement par Hilbert. Il faut tout d'abord exprimer un élément d'intégration invariant dans un système de coordonnées arbitraire. Dans l'espace de Minkowski, on a vu que l'élément $d^4x = dx^0 dx^1 dx^2 dx^3$ est un invariant en raison de la compensation des facteurs γ qui interviennent dans cdt et dans dr_{\parallel} . Si l'on cherche à former un élément d'intégration invariant dans un système de coordonnées quelconque, il apparait le déterminant du tenseur métrique. Considérons les systèmes de coordonnées x et x' . On a

$$d^4x = \left| \frac{\partial x}{\partial x'} \right| d^4x'$$

où le jacobien de la transformation, $\left| \frac{\partial x}{\partial x'} \right|$, est le déterminant de la matrice des dérivées partielles $\frac{\partial x^\mu}{\partial x'^\nu}$. De la loi de transformation

$$g'_{\mu\nu} = \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma}$$

on déduit que $g' = g \left| \frac{\partial x}{\partial x'} \right|^2$ où g est le déterminant (négatif) de $g_{\mu\nu}$, soit pour l'élément d'intégration $d^4x = \sqrt{g'/g} d^4x'$ et finalement l'élément invariant est donné par

$$\sqrt{-g} d^4x = \sqrt{-g'} d^4x'.$$

L'action d'Einstein-Hilbert est formée en intégrant sur $\sqrt{-g}d^4x$ un scalaire invariant décrivant la courbure de l'espace-temps. Nous avons vu que le scalaire de Ricci est défini par contractions successives du tenseur de Riemann, il forme donc le bon lagrangien et l'action est donnée par

$$S_{\text{EH}} = -\frac{c^4}{16\pi G} \int R\sqrt{-g}d^4x.$$

Les équation d'Euler Lagrange associées conduisent à $G_{\mu\nu} = 0$ où $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ est bien le tenseur d'Einstein. Lorsque la gravité interagit avec la matière il faut ajouter un nouveau terme invariant décrivant les champs de matière et de radiation dont la variation par rapport aux composantes du tenseur métrique fait apparaître le tenseur énergie impulsion.

$$S = \int \sqrt{-g}d^4x \left(\mathcal{L}_m - \frac{c^4}{16\pi G}R \right),$$

et la forme précise de \mathcal{L}_m dépend du problème considéré, mais définit le tenseur énergie impulsion par ⁽¹¹⁶⁾

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}} \left(\frac{\partial(\sqrt{-g}\mathcal{L}_m)}{\partial g^{\mu\nu}} - \frac{\partial}{\partial x^\rho} \frac{\partial(\sqrt{-g}\mathcal{L}_m)}{\partial g^{\mu\nu},\rho} \right).$$

La variation de l'action conduit à

$$\delta S = \int \sqrt{-g}d^4x \delta g^{\mu\nu} \left(\frac{1}{2}T_{\mu\nu} - \frac{c^4}{16\pi G}(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R) \right),$$

et il vient

$$G_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}.$$

L'introduction de la constante cosmologique nécessite un terme supplémentaire dans l'action,

$$\Lambda \int \sqrt{-g}d^4x.$$

5.4. Conclusion

En définitive, bien que l'essentiel des résultats prévus par la relativité générale dans le système solaire soient en fait des corrections minimales à la gravitation newtonienne, ils sont tous confirmés expérimentalement. On trouve cependant des manifestations spectaculaires (en dehors du système solaire), comme les trous noirs, dont l'existence longtemps mise en doute est maintenant certaine.

Sur le plan conceptuel en revanche, la relativité générale constitue un changement radical de paradigme, d'où son importance en tant que théorie. Il faut également noter que la relativité générale n'est pas la seule théorie relativiste de la gravitation, mais qu'il existe de nombreuses théories concurrentes fondées sur le même principe d'équivalence initial.

⁽¹¹⁶⁾ voir par exemple S. Weinberg, *Gravitation and Cosmology*, Wiley, New-York 1972, chap. 12 ; P.J.E. Peebles, *Principles of Physical Cosmology* Princeton University Press, Princeton 1993, p.259 ; A.K. Raychaudhuri, S. Banerji and A. Banerjee, *General Relativity, Astrophysics, and Cosmology* Springer-Verlag, New-York 1992, p.47.

Mis à jour le 14 Mai 2012Mis à jour le 14 Mai 2012

Chapitre 10

Cosmologie relativiste

On se propose de donner dans ce chapitre quelques notions très élémentaires de cosmologie, tout d'abord de cosmologie newtonienne, puis de quelques raffinements introduits par la relativité. Les lecteurs intéressés par davantage de physique sont encouragés, ne serait-ce que pour butiner, à se reporter à la bibliographie ⁽¹¹⁷⁾ .

1. Introduction

La cosmologie est l'étude de l'Univers dans son ensemble. C'est un champ disciplinaire qui a été initié par Einstein ⁽¹¹⁸⁾ lorsqu'il s'est penché sur les implications

⁽¹¹⁷⁾ Notamment les ouvrages de Kenyon, de Berry ou de Longair, mais aussi le remarquable Harrison ou encore à un niveau plus élevé, Peebles ou Rich.

⁽¹¹⁸⁾ D'autres noms sont associés à la naissance de la cosmologie, de Sitter, Friedmann, Lemaître, Robertson, Walker, Eddington, Gamow, Hoyle et d'autres, mais c'est encore Einstein qui a joué un rôle précurseur, même s'il a assez rapidement abandonné ce domaine d'étude. J.P. Luminet résume l'évolution de la cosmologie de la façon suivante : ... *Friedmann et Lemaître sont de plus en plus reconnus comme des novateurs s'inscrivant dans la lignée de Ptolémée, Copernic, Kepler, Galilée, Newton et Einstein. (...) Les contributions respectives des hommes de science ayant participé à l'élaboration du nouveau paradigme cosmologique se clarifient enfin : Einstein a créé la théorie de la relativité générale et écrit les équations gouvernant les propriétés physico-géométriques de l'univers ; Friedmann a découvert les solutions non statiques de ces équations, décrivant la variation temporelle de l'espace, et entrevu son possible commencement par une singularité ; Lemaître a relié l'expansion théorique de l'espace au mouvement observé des galaxies, et jeté les bases physiques du Bing Bang ; Hubble, enfin, a démontré la nature extragalactique des nébuleuses spirales et confirmé expérimentalement la loi de proportionnalité entre leur vitesse de récession et leur distance.* in A. Friedmann, G. Lemaître, *Essais de cosmologie*, édité par J.P. Luminet, Seuil, Paris 1997.

de la relativité générale à très grande échelle.

Le fait que ce soit la relativité générale qui soit pertinente à l'échelle de l'Univers provient de ce que la gravitation est la seule interaction effective entre constituants de l'Univers dans cette limite. Le programme consiste en principe à extraire des équations d'Einstein la géométrie et la dynamique de l'Univers ⁽¹¹⁹⁾. Il faut pour cela connaître les sources des champs, soit les composantes de $T^{\mu\nu}$ de même que certaines conditions aux limites. La difficulté réside dans le fait que l'on ignore à peu près tout de la distribution de matière et d'énergie, des contraintes et des vitesses dans l'Univers. Il est vrai qu'à l'époque actuelle, les observations astronomiques nous renseignent en partie sur ces distributions, mais elles semblent si irrégulières et discontinues à première vue qu'une description mathématique s'avère horriblement compliquée. De plus les observations actuelles sont limitées au cône de lumière, elles sont donc très locales. Deux hypothèses peuvent ici être avancées : celle d'un Univers hiérarchique (discontinu et imbriqué à toute les échelles) et celle d'un Univers continu à des échelles extrêmement grandes, au-delà des amas et super-amas galactiques, de l'ordre de quelques centaines de millions d'années-lumière. A cette échelle, la description "hydrodynamique" prévaut (les galaxies sont l'analogie des molécules d'un liquide pour lequel à une échelle assez grande devant l'échelle moléculaire, une description continue en termes de densité $\rho(\mathbf{r})$, de champs et d'équations différentielles est valide). C'est cette vision hydrodynamique qui convainc actuellement la communauté.

Le problème des conditions aux limites est aussi assez particulier. En électromagnétisme en général, on résout un problème en supposant une distribution de charge finie et en imposant aux champs de s'annuler à l'infini. Il serait étrange de considérer ici que toute la distribution de matière et d'énergie de l'Univers est concentrée dans notre "voisinage" et baigne dans un grand vide, ce qui donne au problème des conditions aux limites une importance particulière.

En fondant la cosmologie en 1917, Einstein a introduit des hypothèses naturelles, mais assez arbitraires pour résoudre ces difficultés. L'absence de conditions aux limites est compensé en supposant un Univers homogène et isotrope. L'idée d'un Univers homogène (c'est le principe cosmologique) est guidée par le souci de ne pas être anthropique en attribuant une place privilégiée à la Terre. Par ailleurs on sait maintenant, notamment grâce aux observations fines du fond de rayonnement cosmique que l'Univers est essentiellement isotrope et s'il doit être isotrope partout, il est alors également homogène. Einstein a également supposé un Univers statique, ce en quoi il avait tort, mais cela l'a conduit à introduire la constante cosmologique Λ qui se révèle maintenant indispensable dans tous les modèles cosmologiques ⁽¹²⁰⁾.

⁽¹¹⁹⁾ La discussion qui suit est empruntée à A.K. Raychaudhuri, S. Banerji and A. Banerjee, *General Relativity, Astrophysics, and Cosmology* Springer-Verlag, New-York 1992.

⁽¹²⁰⁾ Les anecdotes autour de cette constante sont d'ailleurs très instructives. Einstein a tout d'abord introduit la constante cosmologique comme la simplification la plus simple qu'il puisse faire aux équations de la gravitation (respectant les symétries du problème, mais introduisant une nouvelle constante fondamentale de la physique) de sorte que l'Univers soit statique. Il a d'ailleurs combattu les solutions non statiques de Friedmann. Quelques années plus tard, les observations, notamment de Hubble, mettaient en évidence l'évolution de l'Univers (en l'occurrence son expansion), ce qui a pour un temps laissé penser à Einstein que la constante cosmologique était la "plus grosse erreur de sa vie" (voir cependant à ce sujet A. Harvey and E. Schucking, *Einstein's mistake and the cosmological constant*, Am. J. Phys. **68**, 723 (2000) ou A. Pais (*'Subtle is the Lord...': The Science and the Life of Albert Einstein*, Oxford University Press, New-York 1982) qui cite une lettre d'Einstein à Weyl dans laquelle après avoir pris connaissance des données astronomiques en faveur de l'expansion de l'Univers, il s'exprime ainsi "If there is no quasi-static world, then away with the cosmological term!"). Cette constante est de nos jours une pièce maîtresse des modèles d'évolution de l'Univers, même si sa détermination précise est un problème totalement ouvert (La constante

2. Récession des galaxies, hypothèse de Hubble

La mesure du déplacement vers le rouge par effet Doppler, de raies caractéristiques émises par des sources lointaines (galaxies, quasars...) est la preuve d'un Univers en expansion, et donne un moyen de mesurer la vitesse de récession de ces objets éloignés. En faisant appel à des modèles cosmologiques, on peut tirer des informations sur la distance de ces sources à la Terre et même une estimation de l'âge de l'Univers depuis le Big Bang.

On mesure le déplacement vers le rouge des raies d'un spectre connu dont les longueurs d'onde propres (mesures et émission effectuées sur Terre) sont λ_0 , et celles émises par la source et détectées sur Terre, λ' . Cette mesure est caractérisée par le paramètre

$$z = \frac{\lambda' - \lambda_0}{\lambda_0}. \quad (10.1)$$

Si la vitesse d'éloignement de la source par rapport à la Terre n'est pas trop grande, cz est une bonne approximation de la vitesse de récession radiale $|\mathbf{v}|$. En effet au premier ordre en $|\mathbf{v}|/c$,

$$\lambda' \simeq \lambda_0(1 + |\mathbf{v}|/c), \quad (10.2)$$

soit $\frac{\lambda' - \lambda_0}{\lambda_0} = z = |\mathbf{v}|/c$. On pourrait utiliser ici les expressions exactes de l'effet Doppler,

$$\lambda' = \lambda_0 \sqrt{\frac{1 + |\mathbf{v}|/c}{1 - |\mathbf{v}|/c}},$$

d'où l'on tire

$$|\mathbf{v}| = c \frac{(z + 1)^2 - 1}{(z + 1)^2 + 1}.$$

Des mesures spectrales faites dans le cas de la galaxie NGC7619 donnent par exemple $z = 0.0126$. La valeur approchée de la vitesse radiale $|\mathbf{v}|$ donne $cz \simeq 3779 \text{ km.s}^{-1}$ alors que la valeur exacte est 3756 km.s^{-1} . Des mesures effectuées indépendamment par d'autres méthodes basées sur la luminosité donnent une distance de $d = 65 \text{ Mpc}$ (1Mpc vaut 3.26 a.l.) entre la Terre et NGC7619. Partant du modèle du Big Bang (hypothèse qu'à l'origine, tous les objets cosmiques se trouvaient en un même point, donc que leurs distances relatives étaient nulles), et admettant que la vitesse de récession radiale est et a toujours été uniforme, on peut estimer que cette distance est $d = |\mathbf{v}|t$ où t est une estimation de l'âge de l'Univers. On obtient ainsi $t \simeq 16.9 \times 10^9 \text{ ans}$.

cosmologique a mis du temps à s'imposer. Encore assez récemment, elle ne faisait pas l'unanimité ; Feynman en 1963 par exemple écrivait dans R.P. Feynman, *Lectures on gravitation*, Penguin Books, 1995, qu'elle lui semblait inutile et Landau non plus n'était pas favorable à l'introduction de cette constante cosmologique, pas plus qu'à aucune modification de la théorie de la relativité générale qu'il considérait "la plus magnifique des théories physiques existantes" cité par V.L. Ginzburg, *Physics Today* May 1989, p.54.).

Le modèle du Big Bang a pour origine les nombreuses observations faites par Edwin Hubble ⁽¹²¹⁾ à partir de 1929 : le calcul précédent, répété pour de nombreuses sources cosmiques donne toujours, aux erreurs expérimentales près, le même résultat. Une des conséquences du modèle est que plus la source est éloignée, plus sa vitesse de récession par rapport à la Terre est grande, et que la relation entre cette distance d et la vitesse radiale $|\mathbf{v}|$ est linéaire :

$$|\mathbf{v}| = Hd. \quad (10.3)$$

Le coefficient de proportionnalité H est en fait une fonction du temps, et est appelé constante de Hubble ⁽¹²²⁾. La valeur numérique précédente donne $H = |\mathbf{v}|/d \simeq 57.8 \text{ km.s}^{-1}\text{Mpc}^{-1}$, compatible avec la fourchette admise actuellement de

$$40 < H < 100 \text{ km.s}^{-1}\text{Mpc}^{-1}. \quad (10.4)$$

Notons par ailleurs que la Terre ne devant pas occuper de position privilégiée dans l'espace, on devra s'attacher à retrouver une loi de même forme en choisissant une origine arbitraire.

3. Cosmologie newtonienne

3.1. Principe cosmologique

La cosmologie est l'étude de l'Univers considéré dans son ensemble. Elle est principalement fondée sur la relativité générale, mais on peut établir une cosmologie newtonienne qui constitue une première approximation et coïncide avec les modèles relativistes à pression nulle. La description de l'Univers ainsi obtenue est convenable pour les 10 à 15 milliards d'années passées.

Comme on vient de le voir, Hubble a remarqué l'existence d'une loi linéaire entre le décalage spectral vers le rouge, signe d'un mouvement de récession générale des galaxies et la distance à la Terre. En 1929, Hubble estimait une valeur de l'ordre de $H \simeq 500 \text{ km.s}^{-1}\text{Mpc}^{-1}$, mais on pense maintenant que ce résultat est surestimé d'un facteur 10 environ. Par ailleurs, le principe cosmologique veut que nous n'occupions pas une place privilégiée dans l'Univers, la loi de Hubble doit donc valoir pour tout observateur. Les principes de base sont les suivants,

- Il existe un temps universel, le temps cosmique, de sorte que l'état de l'Univers puisse être décrit en fonction de ce temps de la même manière pour tout observateur.
- Les lois de la nature sont les mêmes partout et tout le temps.
- L'Univers à très grande échelle est homogène et isotrope.

L'homogénéité n'est valable qu'à des échelles $> 200 \text{ Mpc}$ et l'isotropie est suggérée par la grande isotropie du fond de rayonnement cosmique à 3 K.

⁽¹²¹⁾ En fait la première apparition de la loi de Hubble est due à Lemaître en 1927, mais son article, publié en français, est longtemps resté méconnu et sa traduction en anglais par Eddington ne comporte pas ce passage où la loi de proportionnalité distance-vitesse est établie.

⁽¹²²⁾ La dénomination de constante de Hubble est justifiée dans le sens où la variation dans le temps est extrêmement faible aux échelles du système solaire et par ailleurs la valeur est la même en tout point de l'espace à un instant donné. L'image du ballon de baudruche présentée plus loin donne un sens à la dépendance temporelle.

3.2. Loi de Hubble et ballon de baudruche

Conformément au principe cosmologique, on cherche un modèle qui rende compte de la validité en tout point de la loi de Hubble. Considérons la surface d'un ballon de baudruche sphérique gonflé, dont la rayon varie au cours du temps et qui représente l'espace tridimensionnel ordinaire. La surface du ballon ($2d$ et non euclidienne) est plongée dans \mathbb{R}^3 (euclidien), mais seules les deux dimensions de la surface du ballon représentent l'espace tridimensionnel ordinaire. Si C est le centre du ballon, O_1 , O_2 et M des points de sa surface, $CM = a(t)$. La distance curviligne (sur la surface) entre O_1 et M vaut $O_1M(t) = a(t)\theta_1 = l_1(t)$, soit une vitesse de M par rapport à O_1 ,

$$|\mathbf{v}_1| = \frac{dO_1M}{dt} = \dot{a}(t)\theta_1 = \frac{\dot{a}(t)}{a(t)}l_1$$

qui n'est rien d'autre que la loi de Hubble avec

$$H \equiv H(t) = \dot{a}(t)/a(t). \quad (10.5)$$

Le même raisonnement exactement s'applique à un observateur O_2 qui conclut à la même loi de variation vitesse-distance,

$$|\mathbf{v}_2| = \frac{dO_2M}{dt} = Hl_2.$$

On constate qu'il n'y a donc pas de "centre" dans cet espace bidimensionnel, puisque tous les points y jouent le même rôle. Le rayon $a(t)$ du ballon (défini en dehors de la surface) joue le rôle de facteur d'échelle. Dans ce modèle simplifié, \mathbf{v} représente la vitesse de récession des galaxies.

On peut également éliminer la vitesse pour obtenir l'évolution des distances dans l'Univers au cours du temps :

$$\begin{aligned} |\mathbf{v}(t)| &= H(t)l(t), \\ dl(t) &= |\mathbf{v}(t)| dt \\ &= H(t)l(t) dt, \\ l(t) &= l(t_0) \exp\left(\int_{t_0}^t H(t) dt\right). \end{aligned}$$

Dans la pratique, H ne varie notablement qu'à l'échelle du temps cosmique et on peut écrire

$$l(t) = l(t_0)e^{H(t-t_0)}$$

de sorte qu'une conséquence de la loi de Hubble est l'augmentation exponentielle des distances dans l'Univers.

L'objet de ce cours étant de donner quelques éléments liés à l'approche relativiste, nous ne développerons pas beaucoup la cosmologie newtonienne ⁽¹²³⁾, bien que certains résultats puissent être déduits d'une analyse classique plus poussée ⁽¹²⁴⁾.

⁽¹²³⁾ Sur la pertinence de l'approche newtonienne en cosmologie, on peut se reporter à C. Callan, R.H. Dicke and P.J.E. Peebles, *Cosmology and newtonian mechanics*, Am. J. Phys. **33**, 105 (1965).

⁽¹²⁴⁾ On peut consulter Harrisson par exemple.

3.3. Dynamique de l'Univers plat

Considérons un Univers homogène et isotrope (non nécessairement statique). Chaque observateur voit la même densité moyenne $\rho(t)$ à un instant donné. Choisissons l'origine des coordonnées au niveau d'une galaxie quelconque autour de laquelle, en vertu du principe cosmologique, la symétrie est sphérique. Considérant une particule test de masse m , solidaire du "fluide cosmique", distante de \mathbf{r} de l'origine et supposant qu'aucune autre force que gravitationnelle n'agit sur cette masse test, elle est en mouvement inertiel et on peut écrire d'une part l'énergie cinétique de la galaxie test $\frac{1}{2}m|\dot{\mathbf{r}}|^2$ et d'autre part son énergie potentielle gravitationnelle due à la masse $M(r) = \frac{4}{3}\pi\rho(t)r^3$ comprise dans la sphère de rayon r , $-GMm/r$. En écrivant $\frac{1}{2}mC$ l'énergie totale constante de la galaxie test, la loi de conservation s'écrit

$$\frac{1}{2}|\dot{\mathbf{r}}|^2 - \frac{4\pi G\rho(t)|\mathbf{r}|^2}{3} = \frac{1}{2}C. \quad (10.6)$$

Si l'on considère le cas $C = 0$, on a $\dot{\mathbf{r}} = \sqrt{\frac{8\pi G\rho(t)}{3}}\mathbf{r}$, c'est-à-dire la loi de Hubble avec H donné par

$$H(t) = \sqrt{\frac{8\pi G\rho(t)}{3}}. \quad (10.7)$$

Inversant cette relation, on constate que la densité de l'Univers est proportionnelle à H^2 , $\rho(t) = 3H^2(t)/8\pi G$. Dans la théorie relativiste, un Univers d'énergie totale nulle correspond à un Univers plat (minkowskien), initialement étudié par Einstein et de Sitter ⁽¹²⁵⁾. En mettant un indice 0 aux quantités estimées à l'époque actuelle, on voit que la densité critique de matière nécessaire pour avoir un Univers plat est donnée par $\rho_c = 3H_0^2/8\pi G$. Si la densité actuelle ρ_0 dépasse la densité critique, l'énergie totale de l'Univers est négative ($C < 0$) et l'Univers est fini alors que si $\rho_0 < \rho_c$, l'énergie totale est positive et l'Univers est infini. Dans les deux cas, la valeur actuelle de la constante de Hubble traduit le fait que l'Univers est actuellement en phase d'expansion. On note que C représente deux fois la valeur de la densité d'énergie qui doit être la même partout en raison de l'homogénéité.

⁽¹²⁵⁾ Ce qu'Einstein a initialement conçu comme une difficulté de ce modèle est le fait qu'il soit non statique (puisque obéissant à la loi de Hubble). Il a alors introduit comme on l'a vu la constante cosmologique. La limite newtonienne des équations d'Einstein conduit ainsi à l'équation de Poisson modifiée $\nabla^2\psi(\mathbf{r}) = 4\pi G\rho(\mathbf{r}) - \Lambda c^2$. La contribution $\psi_\Lambda(\mathbf{r})$ de la constante cosmologique au potentiel gravitationnel est alors solution de $\frac{1}{r}\frac{d^2}{dr^2}(r\psi_\Lambda(\mathbf{r})) = -\Lambda c^2$, soit encore $\psi_\Lambda(\mathbf{r}) = -\frac{1}{6}\Lambda c^2 r^2$. Il correspond à ce potentiel un champ de gravitation répulsif si $\Lambda > 0$, $\mathbf{G}_\Lambda(\mathbf{r}) = +\frac{1}{3}\Lambda c^2 \mathbf{r}$ (c'est aussi la force par unité de masse). La constante Λ joue ainsi le rôle d'une force source d'expansion qui peut s'opposer à la gravitation et permet de restaurer un Univers statique. En effet, pour une énergie totale nulle (Univers plat), la conservation de l'énergie devient

$$\frac{1}{2}|\dot{\mathbf{r}}|^2 - \frac{4\pi G\rho|\mathbf{r}|^2}{3} - \frac{1}{6}\Lambda c^2|\mathbf{r}|^2 = 0,$$

de sorte que la constante de Hubble peut être nulle si l'on impose la valeur de Λ , $\dot{r}/r = H = \sqrt{8\pi G\rho/3(1 + \Lambda c^2/8\pi G\rho)^{1/2}} = 0$.

Notons finalement qu'il est possible de donner une solution paramétrique de l'équation de conservation de l'énergie en posant (dans le cas $C < 0$)

$$\begin{aligned} r(\eta) &= A(1 - \cos \eta), \\ t(\eta) &= B(\eta - \sin \eta). \end{aligned}$$

En substituant $\dot{r} = \frac{dr/d\eta}{dt/d\eta}$ dans la conservation de l'énergie, les expressions des constantes A et B en découlent. Pour $C > 0$, il faut poser $r(\eta) = A(\cosh \eta - 1)$ et $t(\eta) = B(\sinh \eta - \eta)$.

4. Cosmologie relativiste

La relativité générale est indispensable à l'établissement d'une théorie physique de l'Univers, appelée la cosmologie ou cosmologie physique. Les aspects cinématiques sont essentiellement décrits par la métrique de Robertson et Walker ⁽¹²⁶⁾ quant à la dynamique, elle nécessite de résoudre les équations d'Einstein et de se doter en plus d'une équation d'état représentant un *fluide cosmique*. On obtient alors les équations de Friedmann. L'Univers est toujours traité dans son ensemble à grande échelle, comme un milieu continu.

4.1. L'Univers comme un fluide

On considère l'Univers à très grande échelle dans une approximation de type hydrodynamique. Il est alors supposé homogène et isotrope, et la distribution de matière est assimilée à un fluide de densité égale à la densité moyenne de l'Univers. Une galaxie est vue dans cette approche comme un élément de volume du "fluide cosmique". Le temps est mesuré par des horloges au repos par rapport à ce fluide, elles sont supposées synchronisées en imposant une date donnée lorsque le fluide avait une certaine densité fixée. L'Univers ayant des propriétés identiques partout, cette procédure fournit un moyen opérationnel de mesure (et comparaison) du temps, appelé temps cosmique.

4.2. Le modèle de Robertson et Walker

4.2.1. La métrique de Robertson et Walker

Cherchons la métrique d'un Univers homogène, isotrope, et dont la courbure (qui pourra être positive, négative ou nulle), $K(t) = k/a^2(t)$ est la même en tout point en vertu du principe cosmologique et ne dépend que du temps cosmique t . $a(t)$ est un facteur d'échelle qui fixe l'amplitude de la courbure et k définit le signe, $k = +1, 0$ ou -1 . La métrique cherchée est a priori de la forme

$$ds^2 = c^2 dt^2 - dl^2$$

où dl^2 est la métrique tridimensionnelle de l'espace ordinaire, éventuellement courbe.

⁽¹²⁶⁾ On obtiendra cette métrique par un raisonnement analogue au problème du ballon de baudruche !

Lorsque k est positif, on obtient un espace à courbure positive et dans ce cas, par analogie avec la surface de la sphère, plongée dans \mathbb{R}^3 , on introduit une coordonnée supplémentaire, w , pour plonger l'espace tridimensionnel à courbure positive dans \mathbb{R}^4 ,

$$\underbrace{x^2 + y^2 + z^2}_{r^2} + w^2 = a^2(t).$$

On a donc $r^2 + w^2 = a^2(t)$, soit $(r dr)^2 = (w dw)^2$ à un instant donné, ou encore

$$dw^2 = \frac{r^2}{a^2(t) - r^2} dr^2.$$

Deux points voisins sont distants de

$$\begin{aligned} dl^2 &= dr^2 + r^2 d\Omega^2 + dw^2 \\ &= \frac{a^2(t)}{a^2(t) - r^2} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2. \end{aligned}$$

On a bien un espace homogène (la courbure est la même partout) et isotrope (dépendance angulaire de la métrique en $r^2 d\Omega^2$). On introduit en général une variable sans dimension $\sigma = r/a(t)$, de sorte que

$$dl^2 = a^2(t) \left(\frac{d\sigma^2}{1 - \sigma^2} + \sigma^2 d\Omega^2 \right).$$

On en déduit la métrique de Robertson et Walker, introduite indépendamment par ces deux auteurs en 1936, en généralisant à des courbures éventuellement nulle ou négatives en remplaçant a par a/\sqrt{k} ,

$$ds^2 = c^2 dt^2 - a^2(t) \left(\frac{d\sigma^2}{1 - k\sigma^2} + \sigma^2 d\Omega^2 \right). \quad (10.8)$$

4.2.2. Référentiel comobile

Les coordonnées intervenant dans la métrique de Robertson-Walker sont $x^\mu = (ct, \sigma, \theta, \varphi)$ où t est le temps cosmique et σ la coordonnée spatiale radiale ramenée au facteur d'échelle de l'Univers. Tout point se déplaçant dans l'Univers de sorte que ses coordonnées spatiales $x^i = \sigma, \theta, \varphi$ soient constantes suit une ligne d'Univers particulière. En effet on a dans ce cas $\dot{x}^i = \text{const.}$, soit

$$\frac{dx^i}{ds} = \frac{d^2 x^i}{ds^2} = 0,$$

et comme la métrique est diagonale,

$$\begin{aligned} \Gamma^i_{00} &= \frac{1}{2} g^{i\rho} (2\partial_0 g_{0\rho} - \partial_\rho g_{00}) \\ &= \frac{1}{2} g^{ii} (2\partial_0 \underbrace{g_{0i}}_0 - \partial_i \underbrace{g_{00}}_1) = 0. \end{aligned}$$

On en déduit automatiquement que le point en question obéit à l'équation

$$\underbrace{\frac{d^2 x^i}{ds^2}}_0 + \underbrace{\Gamma^i}_{\substack{\text{nul si} \\ \mu\nu=00}} \underbrace{\frac{dx^\mu}{ds} \frac{dx^\nu}{ds}}_{\substack{\text{nul si} \\ \mu\nu \neq 00}} = 0$$

c'est-à-dire ⁽¹²⁷⁾

$$\frac{Dp^i}{Ds} = 0. \quad (10.9)$$

Les coordonnées σ, θ, φ sont appelées coordonnées comobiles et le référentiel dans lequel elles restent constantes est le référentiel comobile. L'origine du référentiel comobile obéit à l'équation de la particule libre dans un référentiel inertiel, le référentiel comobile est donc un référentiel en chute libre. Par ailleurs, dans ce référentiel la métrique de Robertson-Walker se réduit à $ds^2 = c^2 dt^2$ c'est-à-dire que le temps cosmique est le temps propre du référentiel comobile.

4.2.3. Paramètres mesurables

Le facteur d'échelle $a(t)$ est avec la courbure k la quantité importante dans la métrique de Robertson et Walker. Il est important de caractériser ces paramètres en fonction de paramètres mesurables. Il y a tout d'abord la constante de Hubble, dont la valeur actuelle est

$$H_0 \equiv \dot{a}(t_0)/a(t_0). \quad (10.10)$$

La valeur du "redshift" z est également accessible,

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{ém}}}{\lambda_{\text{ém}}}.$$

Pour cela, considérons que nous sommes à l'origine du référentiel comobile, $\sigma = 0$, et qu'un signal lumineux issu d'une galaxie située en $\sigma_{\text{ém}}$ émis à l'instant $t_{\text{ém}}$ nous parvient à t_{obs} . Le signal émis à la période suivante est émis à l'instant $t_{\text{ém}} + \Delta t_{\text{ém}}$ et nous arrive à $t_{\text{obs}} + \Delta t_{\text{obs}}$. Ces deux signaux parcourent des intervalles de genre lumière,

$$0 = dt^2 - a^2(t) d\sigma^2 / c^2 (1 - k\sigma^2)$$

de sorte qu'il vient les deux égalités

$$\int_{t_{\text{ém}}}^{t_{\text{obs}}} \frac{dt}{a(t)} = \frac{1}{c} \int_0^{\sigma_{\text{ém}}} \frac{d\sigma}{\sqrt{1 - k\sigma^2}} \quad \text{et} \quad \int_{t_{\text{ém}} + \Delta t_{\text{ém}}}^{t_{\text{obs}} + \Delta t_{\text{obs}}} \frac{dt}{a(t)} = \frac{1}{c} \int_0^{\sigma_{\text{ém}}} \frac{d\sigma}{\sqrt{1 - k\sigma^2}}.$$

On voit que cela entraîne l'égalité des deux premiers membres, et comme les intervalles Δt sont arbitrairement petits devant le facteur d'échelle, on peut négliger la dépendance temporelle de ce dernier dans les intégrales et il vient simplement

$$\frac{\Delta t_{\text{obs}}}{a(t_{\text{obs}})} - \frac{\Delta t_{\text{ém}}}{a(t_{\text{ém}})} = 0.$$

⁽¹²⁷⁾ Voir la note sur la dérivée covariante au chapitre 9.

En multipliant les Δt par c pour avoir les longueurs d'onde, on obtient le redshift

$$z = a(t_{\text{obs}})/a(t_{\text{ém}}) - 1. \quad (10.11)$$

Il est conventionnel de donner un nom aux premiers coefficients du développement de Taylor du facteur d'échelle,

$$\begin{aligned} a(t_{\text{ém}}) &= a(t_{\text{obs}}) + (t_{\text{ém}} - t_{\text{obs}})\dot{a}(t_{\text{obs}}) + \frac{1}{2}(t_{\text{ém}} - t_{\text{obs}})^2\ddot{a}(t_{\text{obs}}) + \dots \\ &= a(t_0)(1 + H_0(t_{\text{ém}} - t_0) - \frac{1}{2}q_0H_0^2(t_{\text{ém}} - t_0)^2 + \dots) \end{aligned}$$

en notant simplement t_0 pour t_{obs} et ajoutant des indices 0 aux valeurs présentes. q_0 est le paramètre (sans dimension) de décélération,

$$q_0 \equiv -\ddot{a}(t_0)a(t_0)/\dot{a}(t_0)^2 = -H_0^2\ddot{a}(t_0)/a(t_0). \quad (10.12)$$

On peut ainsi réexprimer le redshift

$$z = H_0(t_0 - t_{\text{ém}}) + (1 + \frac{1}{2}q_0)H_0^2(t_0 - t_{\text{ém}})^2.$$

Les paramètres H_0 , q_0 , de même que la courbure k , la valeur du rapport Ω de la densité de l'Univers à la densité critique et la valeur de la constante cosmologique Λ sont les paramètres qui définissent l'Univers. Les valeurs présentes, actuellement admises ⁽¹²⁸⁾ sont

$$\begin{aligned} H_0 &= 72 \pm 8 \text{ km s}^{-1}\text{Mpc}^{-1} && \text{Hubble Key Project} \\ \tau_0 &= H_0^{-1} = 1.36 \pm 0.15 \cdot 10^{10} \text{ années} \\ -0.02 &\leq kc^2/a_0^2H_0^2 \leq 0.02 && \text{fluctuations du rayonnement cosmique à 3 K} \\ \Omega_0 &= 8\pi G\rho_0/3H_0^2 = 0.3 \pm 0.1 && \text{structures à grande échelle} \\ \lambda_0 &= \Lambda/3H_0^2 = 0.7 \pm 0.2 && \text{supernovæ Ia} \end{aligned}$$

4.3. Dynamique : les équations de Friedmann

Pour obtenir les équation de Friedmann ⁽¹²⁹⁾ qui régissent la dynamique de l'Univers, on part du tenseur énergie impulsion du fluide cosmique,

$$T_{\mu\nu} = (\rho + p/c^2)v_\mu v_\nu - pg_{\mu\nu} \quad (10.13)$$

et de la métrique de Robertson-Walker pour écrire les équations d'Einstein.

⁽¹²⁸⁾ Tirées de M. Rowan-Robinson, *Cosmology*, Oxford University Press, 4ème édition, Oxford 2004.

⁽¹²⁹⁾ Pour un commentaire sur l'orthographe de Friedmann, voir la note 11 p. 83 dans A. Friedmann, G. Lemaître, *Essais de cosmologie*, édité par J.P. Luminet, Seuil, Paris 1997.

Les composantes du tenseur métrique valent

$$\begin{aligned} g_{00} &= g^{00} = 1, \\ g_{11} &= 1/g^{11} = -a^2/(1 - k\sigma^2), \\ g_{22} &= 1/g^{22} = -a^2\sigma^2, \\ g_{33} &= 1/g^{33} = -a^2\sigma^2 \sin^2 \theta. \end{aligned}$$

Les sources sont définies par $T_{\mu\nu}$. Dans le référentiel comobile, $v^\mu = (c, 0, 0, 0)$ et le tenseur énergie impulsion se simplifie en composantes diagonales,

$$\begin{aligned} T_{00} &= \rho c^2, \\ T_{11} &= pa^2/(1 - k\sigma^2), \\ T_{22} &= pa^2\sigma^2, \\ T_{33} &= pa^2\sigma^2 \sin^2 \theta. \end{aligned}$$

La géométrie est déterminée par le tenseur d'Einstein. Les connexions affines se déduisent de $g_{\mu\nu}$,

$$\begin{aligned} \Gamma^1_{01} &= \Gamma^1_{10} = g^{11}\Gamma_{110} = \frac{1}{2}g^{11}g_{11,0} \\ &= \frac{1}{2}(-(1 - k\sigma^2)/a^2)(-2a\dot{a}/c(1 - k\sigma^2)) = \dot{a}/ac. \end{aligned}$$

La composante R^1_{010} du tenseur de Riemann est donnée par

$$\begin{aligned} R^1_{010} &= -\Gamma^1_{01,0} - \Gamma^1_{10}\Gamma^1_{01} \\ &= (\dot{a}^2/a^2c^2 - \ddot{a}/ac^2) - \dot{a}^2/a^2c^2 = -\ddot{a}/ac^2. \end{aligned}$$

De même les autres composantes valent $R^2_{020} = R^3_{030} = -\ddot{a}/ac^2$ et $R^0_{000} = 0$. On en déduit pour le tenseur de Ricci

$$R_{00} = R^0_{000} + R^1_{010} + R^2_{020} + R^3_{030} = -3\ddot{a}/ac^2$$

et de même $R_{11} = T/(1 - k\sigma^2)$, $R_{22} = T\sigma^2$ et $R_{33} = a\sigma^2 \sin^2 \theta$ avec $T = 2k + a\ddot{a}/c^2 + 2\dot{a}^2/c^2$. Le scalaire de Ricci vaut ⁽¹³⁰⁾

$$\text{Ricci} = g^{\mu\nu}R_{\mu\nu} = -6S/a^2$$

avec $S = k + a\ddot{a}/c^2 + \dot{a}^2/c^2$. Finalement, le tenseur d'Einstein, donné par

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\text{Ricci},$$

⁽¹³⁰⁾ Attention à ne pas confondre avec le facteur d'échelle a .

a pour composantes 00 et 11

$$G_{00} = 3\dot{a}^2/a^2c^2 + 3k/a^2,$$

$$G_{11} = -\frac{k + 2a\ddot{a}/c^2 + \dot{a}^2/c^2}{1 - k\sigma^2}.$$

On en déduit les composantes correspondantes des équations d'Einstein,

$$3\frac{\dot{a}^2}{a^2} + 3k\frac{c^2}{a^2} - c^2\Lambda = 8\pi G\rho,$$

$$-2\frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2} - k\frac{c^2}{a^2} + c^2\Lambda = 8\pi Gp/c^2. \quad (10.14)$$

Les autres composantes n'apportent pas d'information supplémentaire. Ces équations ont été découvertes par Friedmann en 1922 dans le cas $p = 0$ puis généralisées à pression non nulle par Lemaître en 1927. Il reste à résoudre la problème de la dépendance de la pression avec la densité et pour cela il faut encore une équation d'état $p = p(\rho)$. A l'époque actuelle, la pression est faible et elle peut être négligée dans la dernière équation,

$$-2\frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2} - k\frac{c^2}{a^2} + c^2\Lambda = 0.$$

Si la constante cosmologique s'annule aussi, l'évolution de l'Univers est régie par les deux équations

$$3\frac{\dot{a}^2}{a^2} + 3k\frac{c^2}{a^2} = 8\pi G\rho,$$

$$-2\frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2} - k\frac{c^2}{a^2} = 0,$$

qui ont une interprétation simple dans la limite newtonienne. *Considérons une sphère comobile avec le fluide cosmique, de rayon σa . Si la courbure et les vitesses locales sont faibles, la dynamique newtonienne et la gravitation newtonienne peuvent s'appliquer, de sorte que l'énergie cinétique par unité de masse d'un point de la sphère vaut $\frac{1}{2}(\sigma\dot{a})^2$. De même, la densité d'énergie gravitationnelle due au contenu de la sphère vaut $-M(\sigma a)G/\sigma a$ avec $M(\sigma a) = \frac{4}{3}\pi\rho\sigma^3a^3$ la masse de la sphère comobile ⁽¹³¹⁾. La conservation de l'énergie totale locale s'écrit*

$$\frac{1}{2}\sigma^2\dot{a}^2 - \frac{4}{3}\pi G\rho\sigma^2a^2 = \text{const.}$$

Si l'on choisit d'écrire la constante $\text{const.} = -\frac{1}{2}k\sigma^2c^2$, on obtient

$$3\frac{\dot{a}^2}{a^2} + 3k\frac{c^2}{a^2} = 8\pi G\rho.$$

⁽¹³¹⁾ Ce qui est à l'extérieur ne contribue pas en symétrie sphérique en raison du théorème de Gauss, ou plutôt du théorème de Birkhoff dans le contexte de la relativité générale.

On voit que k détermine le signe de la courbure de l'Univers, mais dans la limite newtonienne, c'est également k qui détermine le signe de l'énergie totale de la matière dans l'Univers (tableau 10.1).

Tableau 10.1 Courbure et énergie totale de l'Univers pour des modèles à pression nulle et sans constante cosmologique.

k	courbure	énergie totale
+1	positive	négative
0	plate	nulle
-1	négative	positive

Si l'on multiplie cette dernière équation par a^3 et que l'on différentie par rapport à t , il vient $6\ddot{a}aa + 3\dot{a}^3 + 3kc^2\dot{a} = \frac{d}{dt}(8\pi G\rho a^3)$, soit, en utilisant aussi la seconde équation de Friedmann qui montre que le membre de gauche s'annule,

$$\frac{d}{dt} \left(\frac{4\pi G\rho a^3}{3} \right) = 0.$$

Cette équation exprime que la masse totale dans la sphère comobile est constante. En d'autres termes, les équations de Friedmann à pression nulle et sans constante cosmologique sont équivalentes, dans la limite newtonienne, à la conservation de l'énergie et à l'équation de continuité.

Nous avons vu que les valeurs admises des paramètres cosmologiques sont en faveur d'un Univers plat $k = 0$ avec une constante cosmologique positive. Dans ce cas, en tenant compte de la conservation de la masse dans la sphère comobile, $\rho(t)a(t)^3 = \rho_0 a_0^3$, l'équation de Friedmann à pression nulle s'écrit

$$\dot{a}^2(t) = \frac{8\pi G\rho_0 a_0^3}{3a(t)} + \frac{1}{3}c^2\Lambda a^2(t). \quad (10.15)$$

On constate que $a(t)$ est une fonction croissante (Univers en expansion). De plus, pour a assez grand, seul le second terme au deuxième membre contribue, donnant lieu à une croissance exponentielle.

$$a(t) \sim \exp \left(\sqrt{\frac{c^2\Lambda}{3}} t \right). \quad (10.16)$$

On peut en fait obtenir une solution exacte en cherchant $a(t) = a_0 A \sinh^\alpha Bt$ que l'on injecte dans l'équation de Friedmann pour obtenir ⁽¹³²⁾

$$a(t) = a_0 \left(\frac{8\pi G\rho_0}{c^2\Lambda} \right)^{1/3} \sinh^{2/3} \left(\frac{1}{2} \sqrt{3c^2\Lambda} t \right).$$

Etant donné l'incertitude sur les valeurs des paramètres cosmologiques, les ouvrages de cosmologie établissent en général une classification des modèles d'Univers et de leur dynamique (modèles plats ($k = 0$), modèles sans constante cosmologique ($\Lambda = 0$), modèles vides ($\rho = 0$), etc). Nous nous limitons ici à un Univers plat avec constante cosmologique positive et densité de matière sous-critique.

⁽¹³²⁾ On identifie alors $\dot{a}^2 = (AB\alpha)^2 a_0^2 \sinh^{2(\alpha-1)} Bt \cosh^2 Bt$ au second membre de l'équation de Friedmann $\frac{8\pi G\rho_0 a_0^3}{3A} \sinh^{-\alpha} Bt + \frac{1}{3}c^2\Lambda A^2 a_0^2 \sinh^{2\alpha} Bt$, d'où les valeurs des constantes.

4.4. Le “problème” de la constante cosmologique

On s'intéresse à nouveau dans cette section à l'histoire de la constante cosmologique⁽¹³³⁾, évoquée précédemment dans les notes de bas de page. On expose tout d'abord les éléments factuels, puis on donne les estimations quantitatives d'un désaccord maintenant célèbre⁽¹³⁴⁾.

Rappelons qu'Einstein a modifié les équations de la gravitation relativiste,

$$G_{\mu\nu} = \frac{8\pi G}{c^2} T_{\mu\nu},$$

par l'introduction au premier membre d'un terme $-\Lambda g_{\mu\nu}$ qui respecte la covariance relativiste⁽¹³⁵⁾ et compense l'effet de gravitation (si $\Lambda > 0$) permettant ainsi de trouver des solutions pour un Univers statique à condition de fixer pour Λ une valeur bien précise dépendant de la densité de l'Univers. La constante cosmologique est homogène à l'inverse d'une longueur au carré et fixe une échelle universelle de longueurs : la courbure d'un Univers vide de matière (la solution de de Sitter). Rapidement, Friedmann puis Lemaître établiront que l'Univers statique d'Einstein est instable. Au contraire, les équations d'Einstein admettent (avec ou sans Λ) des solutions dynamiques pour un Univers en expansion ou en contraction. L'observation de l'expansion (à partir du décalage systématique vers le rouge des galaxies éloignées) par Slipher et Hubble fera dire à Einstein que la constante cosmologique est inutile. Après Einstein, de nombreux physiciens théoriciens adopteront la même attitude à l'égard de la constante cosmologique et s'en désintéresseront. Les astronomes et cosmologistes seront en revanche attachés à l'existence du terme cosmologique, notamment Eddington, car sans sa présence les modèles conduisent à un Univers trop jeune, plus jeune même que la Terre, et seule l'introduction du paramètre supplémentaire Λ peut “vieillir” suffisamment l'Univers.

Récemment (1997) des mesures effectuées sur des supernovae lointaines (supernova cosmology project) ont montré sans doute possible que l'expansion de l'Univers s'accélère. Pour rendre compte d'une telle expansion accélérée, il est indispensable d'avoir recours à la constante cosmologique, ce qui a suscité un regain d'intérêt pour ce Λ , sa détermination expérimentale et son interprétation théorique.

Le terme cosmologique $-\Lambda g_{\mu\nu}$ peut être rejeté au second membre sous la forme $+\Lambda g_{\mu\nu} = \frac{8\pi G}{c^2} \rho_\Lambda c^2 g_{\mu\nu}$ avec $\rho_\Lambda = \frac{\Lambda}{8\pi G}$ pour donner aux équations d'Einstein la forme

$$G_{\mu\nu} = \frac{8\pi G}{c^2} (T_{\mu\nu} + \rho_\Lambda c^2 g_{\mu\nu}).$$

Nous sommes en mesure d'estimer facilement une borne supérieure de ρ_Λ en notant que la gravitation newtonienne est une très bonne approximation à l'échelle du système solaire. Comme les effets de la constante cosmologique sont susceptibles de se faire sentir aux grandes distances, on aura une valeur maximale en écrivant que l'effet de ρ_Λ est négligeable devant celui de l'attraction gravitationnelle du Soleil à l'échelle de Pluton par exemple. Le champ de gravitation exercé par le Soleil

⁽¹³³⁾ S. Weinberg, *The cosmological constant problem*, Rev. Mod. Phys. **61**, 1 (1989), P.J.E. Peebles and B. Ratra, *The cosmological constant and dark energy*, Rev. Mod. Phys. **75**, 559 (2003), N. Straumann, *On the cosmological constant problems and the astronomical evidence for a homogeneous energy density with negative pressure*, astro-ph/0203330.

⁽¹³⁴⁾ Je conseille également le site <http://super.colorado.edu/~michaele/Lambda> pour une très agréable présentation des problèmes suscités par la constante cosmologique.

⁽¹³⁵⁾ A l'origine, Einstein a utilisé la notation λ .

au niveau de l'orbite de Pluton, $r_P = 5.9 \cdot 10^{12} \text{m}$, vaut $G_\odot(r_P) = GM_\odot/r_P^2$. Celui qu'exercerait une masse volumique uniforme ρ_Λ à cette même distance s'exprime par $G_\Lambda(r_P) = GM_\Lambda(r_P)/r_P^2 = 4\pi\rho_\Lambda Gr_P/3$. En écrivant que cette seconde contribution est négligeable devant la première, on obtient la borne supérieure suivante à la densité de masse associée à la constante cosmologique,

$$\rho_\Lambda \ll \frac{3M_\odot}{4\pi r_P^3}.$$

En introduisant les valeurs numériques on trouve⁽¹³⁶⁾

$$\rho_\Lambda \ll 2.3 \cdot 10^{-9} \text{kg.m}^{-3} = 6.56 \cdot 10^{-27} \text{f}^{-4}.$$

Des estimations beaucoup plus strictes se déduisent d'observations astronomiques au-delà du simple système solaire. et on admet une valeur de l'ordre de

$$\rho_\Lambda \ll 10^{-26} \text{kg.m}^{-3} = 2.9 \cdot 10^{-44} \text{f}^{-4}.$$

Revenons aux équations d'Einstein en présence du terme cosmologique. Le dernier terme $\rho_\Lambda c^2 g_{\mu\nu}$ est exactement la forme que prendrait le tenseur énergie-impulsion $T_{\mu\nu}$ d'un fluide au repos⁽¹³⁷⁾ de pression $p_\Lambda = -\rho_\Lambda c^2$. Il se trouve que cette équation d'état est précisément celle que l'on attend de la densité (d'énergie) associée au vide quantique⁽¹³⁸⁾. En effet, si l'on exige que le vide quantique

⁽¹³⁶⁾ Digression sur les unités "naturelles" (R.J. Adler, B. Casey and O.C. Jacob, *Vacuum catastrophe: an elementary exposition of the cosmological constant problem*, Am. J. Phys. **63**, 620 (1995)). On utilise fréquemment en physique des particules les unités naturelles dans lesquelles les constantes c et \hbar valent toutes les deux 1. Cela permet en outre de n'avoir plus qu'une seule unité fondamentale au lieu des trois dimensions usuelles M, L et T. Choisissons par exemple la longueur comme dimension fondamentale et le fermi ($1 \text{ f} = 10^{-15} \text{m}$) comme unique unité. On peut alors introduire comme unité de temps le "jiffy" (j), durée que met la lumière pour parcourir 1 f (Notons qu'historiquement, le jiffy a été introduit par le chimiste Lewis qui avait déjà forgé le terme de photon. C'était une unité de temps fantaisiste associée à la durée nécessaire à la lumière pour parcourir 1 cm.). On a alors $c = 1 \text{f.j}^{-1}$ soit encore $1 \text{ j} = 10^{-15}/3 \cdot 10^8 \text{ s}$. Si l'on pose maintenant $c = 1$ (sans dimension), alors les longueurs se mesurent en f, mais également les durées car il faut que $1 \text{ j} = 1 \text{ f}$ pour que c soit adimensionnée. Considérons maintenant le cas de \hbar . Plutôt que les J.s, on peut choisir d'exprimer les durées en j (temporairement avant de les éliminer de nouveau au profit des f) et les énergies en MeV, bien plus adaptés à l'échelle de la physique subatomique. On a donc $\hbar = 1.053 \cdot 10^{-34} \text{J.s} = 1.053 \times 3 \cdot 10^{-11} \text{J.j} = 197 \text{ MeV.j}$. On peut fixer une nouvelle unité d'énergie, le "blip" (b), telle que $1 \text{ b} = 197 \text{ MeV}$. On a $\hbar = 1 \text{ b.j}$. Si l'on impose maintenant que $\hbar = 1$ (sans dimension), cela nécessite que $1 \text{ b} = 1 \text{ j}^{-1} = 1 \text{ f}^{-1}$. On n'a plus que le fermi comme unité fondamentale. Les fréquences (E/\hbar) se mesurent comme les énergies en f^{-1} , les masses (E/c^2) également, ... Les conversions utiles sont $1 \text{ m} = 10^{15} \text{f}$, $1 \text{ s} = 3.10^{23} \text{f}$, $1 \text{ J} = 3.17 \cdot 10^{10} \text{f}^{-1}$. Pour exprimer une densité par exemple, on procède comme suit, $1 \text{ J} = 1 \text{ kg.m}^2 \text{s}^{-2} = 3.17 \cdot 10^{10} \text{f}^{-1}$, soit $1 \text{ kg} = 3.17 \cdot 10^{10} \times 10^{-30} \times 9 \cdot 10^{46} \text{f}^{-1} = 28.5 \cdot 10^{26} \text{f}^{-1}$. On a ensuite $1 \text{ kg.m}^{-3} = 28.5 \cdot 10^{26} \times 10^{-45} \text{f}^{-4} = 28.5 \cdot 10^{-19} \text{f}^{-4}$.

⁽¹³⁷⁾ Rappelons que pour un fluide $T_{\mu\nu} = (\rho + p/c^2)v_\mu v_\nu - pg_{\mu\nu}$.

⁽¹³⁸⁾ En mécanique quantique, les inégalités d'Heisenberg interdisent à tout système physique, particule ou de manière générale degré de liberté quelconque, d'être au repos. Il s'ensuit que même dans l'état de plus basse énergie subsiste une énergie résiduelle appelée énergie de point zéro. On parle des fluctuations du vide quantique et pour des raisons fondamentales (invariance de Lorentz), la densité d'énergie associée correspond à une équation d'état $p = -\rho c^2$.

satisfasse aux contraintes relativistes, le tenseur énergie-impulsion associé, $T_{\mu\nu} = (\rho + p/c^2)v_\mu v_\nu - pg_{\mu\nu}$, doit être covariant par transformation de Lorentz. Notamment en considérant le référentiel tangent au référentiel propre du vide ($v_\mu = (c, \vec{0})$) dans lequel la métrique est minkowskienne, $g_{\mu\nu} = \eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$, on a

$$T_{\mu\nu} = (\rho c^2 + p) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} - p \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \rho c^2 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}$$

soit finalement

$$T_{\mu\nu} = \rho c^2 \eta_{\mu\nu}, \quad \text{si } p = -\rho c^2.$$

Cette équation d'état assure que le vide est invariant de Lorentz. Cette particularité fait du vide quantique un concept très différent de la matière ordinaire ou du rayonnement. Dans les autres domaines de la physique où seules les différences d'énergie interviennent, l'énergie du vide n'a pas d'effet direct mais la situation est très différente lorsque la gravitation intervient, puisque toutes les formes d'énergie contribuent à la structure de l'espace-temps par le biais du tenseur énergie-impulsion dans les équations d'Einstein. Cette observation très encourageante donne l'espoir de relier la constante cosmologique (et les échelles de longueur de l'ordre de 10^{23} m) aux interactions fondamentales qui régissent la physique des particules (et des échelles de l'ordre de 10^{-15} m). La constante cosmologique trouverait ainsi une interprétation.

L'illusion est de courte durée, car la valeur estimée de la contribution de la densité d'énergie du vide est 10^{120} fois plus importante que l'estimation associée à la constante cosmologique mesurée par observations astronomiques. Pour la densité d'énergie du vide quantique, on peut fournir une borne inférieure en précisant que la théorie quantique des champs est vérifiée jusqu'à des échelles d'énergie de l'ordre de 10^{14} GeV, puisque l'on mesure de telles énergies pour des photons issus du rayonnement cosmique. Le vide possède donc une énergie $\frac{1}{2}\hbar\omega$ par degré de liberté (oscillateurs harmoniques) et l'on doit sommer sur ces oscillateurs pour obtenir

$$E_{\text{vide}} = \sum_k \frac{1}{2} \hbar \omega_k,$$

où les ω_k sont tels que $\hbar^2 \omega_k^2 = \hbar^2 k^2 c^2 + m^2 c^4$, soit en unités naturelles $\omega_k^2 = k^2 + m^2$. Dans la limite des très hautes énergies (limite ultrarelativiste) on peut négliger la masse. En considérant le vide établi dans une boîte cubique de taille L^3 , la somme sur les modes peut se calculer par une intégrale

$$\sum_k \rightarrow \frac{L^3}{(2\pi)^3} \int_0^{k_{\text{max}}} 4\pi k^2 dk$$

où le préfacteur tient compte de la quantification des valeurs permises pour k pour garantir des conditions aux limites périodiques. On a alors

$$\begin{aligned}
E_{\text{vide}} &= \sum_k \frac{1}{2} \omega_k = \frac{L^3}{(2\pi)^3} \int_0^{k_{\text{max}}} 4\pi k^2 dk \frac{1}{2} (k^2 + m^2)^{1/2} \\
&\simeq \frac{L^3}{4\pi^2} \int_0^{k_{\text{max}}} k^3 dk \\
&= L^3 \frac{k_{\text{max}}^4}{16\pi^2}.
\end{aligned}$$

La coupure à 10^{14} GeV correspond à $5 \cdot 10^{14} \text{f}^{-1}$, soit une densité d'énergie associée au vide de

$$\rho_{\text{vide}} = \lim_{L \rightarrow \infty} \frac{E_{\text{vide}}}{L^3} > 4.2 \times 10^{56} \text{f}^{-4}.$$

Cette valeur est évidemment sans aucun rapport avec celle de ρ_Λ . De nombreux physiciens pensent même que le paramètre de coupure pertinent jusqu'où appliquer les théories quantiques des champs est plutôt l'énergie de Planck de l'ordre de 10^{19} GeV, ce qui donne à la densité d'énergie de vide une borne inférieure encore plus dramatique,

$$\rho_{\text{vide}} > 10^{76} \text{f}^{-4}!$$

Ce désaccord est la plus grande contradiction connue en physique et probablement dans toutes les sciences. Pour tenter de lever la difficulté, de nombreuses hypothèses - répondant aux jolis noms de "supersymétrie", "quintessence" - ont été avancées, sans succès jusqu'à présent. Les théories de supersymétrie par exemple élimineraient la difficulté d'un vide possédant une énergie considérable, car par un mécanisme de compensation entre les fermions et les bosons en nombres identiques dans le cadre supersymétrique, l'énergie du vide est rigoureusement nulle. De nombreux théoriciens estiment très prometteuse cette voie à tel point qu'on en revient à l'idée d'une constante cosmologique nulle ! A la question : *How can the cosmological constant be so close to zero but not zero?*, Edward Witten ⁽¹³⁹⁾ répond par exemple : *I really don't know. It's very perplexing that astronomical observations seem to show that there is a cosmological constant. It's definitely the most troublesome, for my interests, definitely the most troublesome observation in physics in my lifetime. In my career that is.*

⁽¹³⁹⁾ Edward Witten est l'un des physiciens théoriciens les plus en vogue, récompensé par les mathématiciens déjà avec la médaille Fields.

Chapitre 11

Théories quantiques relativistes

Il n'est pas question dans ce cours de s'étendre sur les aspects quantiques, mais pour compléter le programme d'étude annoncé consistant à revoir les théories physiques pour les rendre compatibles avec les principes de la relativité, il n'est pas inutile de considérer le cas de la mécanique quantique. Ceci est d'autant plus logique que la théorie quantique relativiste, proposée par Dirac en 1928, a suivi de très près l'équation de Schrödinger ⁽¹⁴⁰⁾ et que Schrödinger lui-même, comme de Broglie à la même période, avait tenté une formulation relativiste. On peut consulter évidemment de très nombreux ouvrages sur cette question, par exemple Bethe ou Bjorken et Drell ⁽¹⁴¹⁾. L'une des conséquences célèbres de la nécessité d'une formulation covariante en physique quantique est la prédiction par Dirac des antiparticules. Cela conduit notamment à revoir complètement le formalisme pour inclure les possibilités de création et annihilation de particules, ce qui a exigé la construction des théories quantiques de champs. Il ne sera pas fait mention de ces aspects ici, le lecteur intéressé peut consulter l'ouvrage récent de Patricia et John Schwarz ⁽¹⁴²⁾.

1. Equation de Schrödinger

Guidés par la "dualité ondes-particules", ancienne appellation donnée à une série

⁽¹⁴⁰⁾ E. Schrödinger, *L'équation non relativiste des ondes de de Broglie*, Ann. Phys. **79**, 361-376 (1926), traduit dans *Sources et évolution de la physique quantique*, J. Leite-Lopes et B. Escoubès, Masson, Paris 1995, p.99.

⁽¹⁴¹⁾ H.A. Bethe, *Intermediate quantum mechanics*, Benjamin, New-York 1964 ; J.D. Bjorken and S.D. Drell, *Relativistic quantum mechanics*, Mc Graw Hill, New-York 1964.

⁽¹⁴²⁾ P.M. Schwarz and J.H. Schwarz, *Special relativity, from Einstein to strings*, Cambridge University Press, Cambridge 2004.

de concepts dus entre autres à de Broglie ⁽¹⁴³⁾ et résumés dans les relations

$$E = h\nu = \hbar\omega,$$

$$\mathbf{p} = \frac{hc}{\lambda} \hat{\mathbf{u}}_k = \hbar\mathbf{k},$$

nous cherchons une formulation ondulatoire de la mécanique.

Le cas le plus simple est celui de la particule libre, d'énergie $H = \frac{\mathbf{p}^2}{2m}$, que l'on peut logiquement représenter par une onde plane $u_{\mathbf{k}}(\mathbf{r}, t) = A_{\mathbf{k}} e^{i(\mathbf{k}\mathbf{r} - \omega t)}$ dans la mesure où une telle onde est totalement délocalisée et donc a priori à même de représenter une particule qui a une égale probabilité de se trouver en tout point de l'espace. A quelle équation obéit l'onde plane ? On pourrait bien entendu écrire l'équation des ondes, $\vec{\nabla}^2 u_{\mathbf{k}} - \frac{1}{v^2} \frac{\partial^2 u_{\mathbf{k}}}{\partial t^2} = 0$ (soit $\omega = v|\mathbf{k}|$), mais on cherche ici une équation du premier ordre en dérivée temporelle pour satisfaire au déterminisme. Utilisant les correspondances rappelées ci-dessus, on écrit plutôt

$$u_{\mathbf{k}}(\mathbf{r}, t) = A_{\mathbf{k}} e^{i(\frac{\mathbf{p}}{\hbar}\mathbf{r} - \frac{E}{\hbar}t)}$$

d'où on extrait E en faisant

$$Eu_{\mathbf{k}}(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} u_{\mathbf{k}}(\mathbf{r}, t)$$

et \mathbf{p} de même

$$\mathbf{p}u_{\mathbf{k}}(\mathbf{r}, t) = -i\hbar \vec{\nabla} u_{\mathbf{k}}(\mathbf{r}, t).$$

L'équation $E = \mathbf{p}^2/2m$ devient alors

$$i\hbar \frac{\partial}{\partial t} u_{\mathbf{k}}(\mathbf{r}, t) = -\frac{\hbar^2}{2m} \vec{\nabla}^2 u_{\mathbf{k}}(\mathbf{r}, t).$$

Pour une particule soumise à un potentiel $V(\mathbf{r})$, on généralise pour obtenir l'équation de Schrödinger pour la fonction d'onde $\psi(\mathbf{r}, t)$,

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = \left(-\frac{\hbar^2}{2m} \vec{\nabla}^2 + V(\mathbf{r}) \right) \psi(\mathbf{r}, t). \quad (11.1)$$

Nous ne discuterons pas davantage ici les conséquences de cette équation qui constitue l'approche non relativiste traditionnelle de la mécanique quantique. Retenons simplement que la physique est décrite par une fonction d'onde $\psi(\mathbf{r}, t)$ dont l'équation de Schrödinger régit la dynamique.

⁽¹⁴³⁾ L. de Broglie, *La longueur d'onde associée à la matière*, CRAS **177**, 507-510 (1923), reproduit dans J. Leite-Lopes et B. Escoubès, op. cit., p.92.

2. Equation de Klein-Gordon

Une généralisation immédiate de l'équation de Schrödinger a été proposée très rapidement (en 1926), indépendamment par Gordon, Fock, Klein, Kudar, de Donder et Van Dungen ⁽¹⁴⁴⁾. Il suffit d'effectuer les correspondances

$$E \rightarrow i\hbar \frac{\partial}{\partial t}$$

$$\mathbf{p} \rightarrow -i\hbar \vec{\nabla}$$

dans l'expression relativiste donnant l'énergie cinétique (au terme mc^2 près),

$$E^2 - |\mathbf{p}|^2 c^2 = m^2 c^4$$

pour obtenir l'équation de Klein-Gordon de la particule libre,

$$\vec{\nabla}^2 \psi(\mathbf{r}, t) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \psi(\mathbf{r}, t) = \frac{m^2 c^2}{\hbar^2} \psi(\mathbf{r}, t). \quad (11.2)$$

On retrouve l'équation des ondes habituelle, mais avec un terme de masse ⁽¹⁴⁵⁾. Là encore, nous ne nous étendons pas sur les conséquences cette équation (qui décrit les bosons massifs relativistes de spin zéro), mais indiquons seulement qu'il n'est pas possible de définir une densité de probabilité positive conservée au cours du temps ⁽¹⁴⁶⁾, ce qui invalide en grande partie la théorie basée sur l'équation de Klein-Gordon.

3. Equation de Pauli

L'étude de systèmes sous champ magnétique a mis en évidence de nouvelles particularités de la mécanique quantique. Outre la correspondance classique (qu'on appelle couplage minimal) qui fait passer de H à $H + q\phi(\mathbf{r}, t)$ et de \mathbf{p} à $\mathbf{p} - q\mathbf{A}(\mathbf{r}, t)$ sous un champ électromagnétique $(\mathbf{A}(\mathbf{r}, t), \phi(\mathbf{r}, t))$ Pauli a proposé d'introduire un degré de liberté interne supplémentaire, appelé le spin, ayant dans le cas de l'électron des propriétés analogues à celles d'un moment cinétique $\frac{1}{2}$. La fonction d'onde $\psi(\mathbf{r}, t)$ est remplacée par un spineur, objet à deux composantes (spineur de Pauli)

$$\chi = \begin{pmatrix} \psi_{\uparrow}(\mathbf{r}, t) \\ \psi_{\downarrow}(\mathbf{r}, t) \end{pmatrix}$$

pour décrire les degrés de liberté de spin en plus des degrés de liberté d'espace. On obtient alors l'équation de Pauli,

$$i\hbar \frac{\partial}{\partial t} \chi = \left(\frac{1}{2m} (-i\hbar \vec{\nabla} + q\mathbf{A}(\mathbf{r}, t))^2 + V(\mathbf{r}) + q\phi(\mathbf{r}, t) \right) \chi - \frac{q\hbar}{2m} \boldsymbol{\sigma} \mathbf{B}(\mathbf{r}, t) \chi,$$

⁽¹⁴⁴⁾ Voir S. Schweber, *Relativistic quantum field theory*, Harper and Row, New-York 1961.

⁽¹⁴⁵⁾ C'est-à-dire que si l'on faisait $m = 0$, on retrouverait l'équation des ondes dont on sait qu'elle satisfait à la covariance relativiste.

⁽¹⁴⁶⁾ C'est dû à l'apparition de dérivées secondes par rapport au temps dans l'équation dynamique, ce qui ne permet pas de fixer a priori à la fois le signe de ψ et celui de $\partial\psi/\partial t$ et peut ainsi conduire à une probabilité négative. On pourra consulter par exemple H.A. Bethe and R. Jackiw, *Intermediate quantum mechanics*, Westview Press 1986, Boulder, chap. 21.

ou sous forme développée,

$$i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \psi_{\uparrow}(\mathbf{r}, t) \\ \psi_{\downarrow}(\mathbf{r}, t) \end{pmatrix} = \left(\frac{1}{2m} (-i\hbar \vec{\nabla} + q\mathbf{A}(\mathbf{r}, t))^2 + V(\mathbf{r}) + q\phi(\mathbf{r}, t) \right) \mathbb{1} \begin{pmatrix} \psi_{\uparrow}(\mathbf{r}, t) \\ \psi_{\downarrow}(\mathbf{r}, t) \end{pmatrix} - \frac{q\hbar}{2m} \boldsymbol{\sigma} \mathbf{B}(\mathbf{r}, t) \begin{pmatrix} \psi_{\uparrow}(\mathbf{r}, t) \\ \psi_{\downarrow}(\mathbf{r}, t) \end{pmatrix}, \quad (11.3)$$

où $\boldsymbol{\sigma} \mathbf{B}(\mathbf{r}, t) = \sigma_x B_x + \sigma_y B_y + \sigma_z B_z$, avec les matrices de Pauli

$$[\sigma_x] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad [\sigma_y] = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad [\sigma_z] = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

La quantité $\mathbf{s} = \frac{1}{2}\hbar \boldsymbol{\sigma}$ est appelée spin. Dans le cas de l'électron, on a $q = -|q_e|$ dans l'équation de Pauli. Les matrices de Pauli forment avec l'identité une base pour la représentation des matrices 2×2 , car

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{2}(a+d)\mathbb{1} + \frac{1}{2}(a-d)\sigma_z + \frac{1}{2}(b+c)\sigma_x + \frac{1}{2}i(b-c)\sigma_y.$$

Bien entendu, cette équation n'est pas relativiste, mais elle fait apparaître la nécessité, pour inclure le spin dans la théorie, de chercher une équation relativiste agissant sur des objets à plusieurs composantes.

4. Equations de Weyl

4.1. Recherche des équations pour des particules sans masse de spin $\frac{1}{2}$

On cherche maintenant une équation relativiste pour décrire des fermions sans masse de spin $\frac{1}{2}$. On doit s'attendre à retrouver l'équation des ondes sous la forme

$$\vec{\nabla}^2 \begin{pmatrix} \cdot \\ \cdot \end{pmatrix} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \begin{pmatrix} \cdot \\ \cdot \end{pmatrix} = 0,$$

c'est-à-dire agissant sur des objets ayant a priori deux composantes, des spineurs. On notera en général $\chi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ un spineur à deux composantes. Il faut donc une équation matricielle (matrices 2×2) comportant des dérivées du premier ordre, soit pour être compatible avec $\partial_{\mu} = (\frac{1}{c} \frac{\partial}{\partial t}, \vec{\nabla})$,

$$\mathcal{D}_{\pm} \chi = \left(\frac{1}{c} \frac{\partial}{\partial t} \mathbb{1}_{2 \times 2} \pm \boldsymbol{\sigma} \cdot \vec{\nabla} \right) \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}.$$

On a utilisé le fait mentionné plus haut que les matrices de Pauli, avec l'identité, permettent d'exprimer n'importe quelle matrice 2×2 . Pour retrouver l'équation des ondes, formons

$$\begin{aligned} \mathcal{D}_+ \mathcal{D}_- \chi &= \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbb{1}_{2 \times 2} - \underbrace{(\boldsymbol{\sigma} \cdot \vec{\nabla})(\boldsymbol{\sigma} \cdot \vec{\nabla})}_{\vec{\nabla}^2 \mathbb{1}_{2 \times 2}} \right) \chi \\ &= \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 \right) \mathbb{1}_{2 \times 2} \chi = 0, \end{aligned}$$

ce qui est bien l'équation cherchée. On dispose donc de deux équations également valables,

$$\begin{aligned} \mathcal{D}_+ \chi^{(+)} &= \left(\frac{1}{c} \frac{\partial}{\partial t} \mathbb{1}_{2 \times 2} + \boldsymbol{\sigma} \cdot \vec{\nabla} \right) \chi^{(+)} = 0, \\ \mathcal{D}_- \chi^{(-)} &= \left(\frac{1}{c} \frac{\partial}{\partial t} \mathbb{1}_{2 \times 2} - \boldsymbol{\sigma} \cdot \vec{\nabla} \right) \chi^{(-)} = 0. \end{aligned} \tag{11.4}$$

Ce sont les équations de Weyl. Elles agissent sur des champs de spineurs à deux composantes $\chi^{(+)}$ et $\chi^{(-)}$ distincts que l'on interprétera un peu plus loin.

4.2. Equation covariante

4.2.1. Spineurs et matrices de Dirac

On peut introduire un objet à quatre composantes (spineur de Dirac)

$$\psi = \begin{pmatrix} \chi^{(+)} \\ \chi^{(-)} \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}$$

(un spineur à quatre composantes dont deux sont en fait associées aux particules et deux aux antiparticules comme on va le voir) et des matrices 4×4 (appelées les matrices de Dirac, ici dans la représentation chirale)

$$\begin{aligned} \gamma^\mu &= (\gamma^0, \boldsymbol{\gamma}), \\ \gamma^0 &= \begin{pmatrix} 0 & \mathbb{1}_{2 \times 2} \\ \mathbb{1}_{2 \times 2} & 0 \end{pmatrix}, \quad \boldsymbol{\gamma} = \boldsymbol{\gamma}^i = \begin{pmatrix} 0 & -\boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix} \end{aligned}$$

de sorte que les équations de Weyl se combinent sous une forme unique,

$$\gamma^\mu \partial_\mu \psi = \begin{pmatrix} 0 & \frac{1}{c} \frac{\partial}{\partial t} \mathbb{1}_{2 \times 2} \\ \frac{1}{c} \frac{\partial}{\partial t} \mathbb{1}_{2 \times 2} & 0 \end{pmatrix} \begin{pmatrix} \chi^{(+)} \\ \chi^{(-)} \end{pmatrix} + \begin{pmatrix} 0 & -\boldsymbol{\sigma} \cdot \vec{\nabla} \\ \boldsymbol{\sigma} \cdot \vec{\nabla} & 0 \end{pmatrix} \begin{pmatrix} \chi^{(+)} \\ \chi^{(-)} \end{pmatrix} = 0.$$

Les matrices de Dirac obéissent à

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}.$$

4.2.2. Energies négatives

Si l'on fait agir $\gamma^\mu \partial_\mu$ sur la solution représentant la particule libre

$$\psi = \begin{pmatrix} \chi^{(+)} \\ \chi^{(-)} \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{pmatrix} e^{\frac{i}{\hbar} p_\mu x^\mu},$$

on obtient

$$i\hbar\gamma^\mu \partial_\mu \psi = \begin{pmatrix} 0 & \frac{E}{c} \mathbb{1}_{2 \times 2} - \boldsymbol{\sigma} \cdot \mathbf{p} \\ \frac{E}{c} \mathbb{1}_{2 \times 2} + \boldsymbol{\sigma} \cdot \mathbf{p} & 0 \end{pmatrix} \psi = 0. \quad (11.5)$$

L'équation est homogène ; pour avoir une solution non nulle il faut donc annuler le déterminant,

$$\begin{aligned} \text{Det} & \begin{pmatrix} 0 & \frac{E}{c} \mathbb{1}_{2 \times 2} - \boldsymbol{\sigma} \cdot \mathbf{p} \\ \frac{E}{c} \mathbb{1}_{2 \times 2} + \boldsymbol{\sigma} \cdot \mathbf{p} & 0 \end{pmatrix} \\ &= \left| \frac{E}{c} \mathbb{1}_{2 \times 2} - \boldsymbol{\sigma} \cdot \mathbf{p} \right| \times \left| \frac{E}{c} \mathbb{1}_{2 \times 2} + \boldsymbol{\sigma} \cdot \mathbf{p} \right| \\ &= \begin{vmatrix} \frac{E}{c} - p^3 & -p^1 + ip^2 \\ -p^1 - ip^2 & \frac{E}{c} + p^3 \end{vmatrix} \times \begin{vmatrix} \frac{E}{c} - p^3 & +p^1 - ip^2 \\ p^1 + ip^2 & \frac{E}{c} + p^3 \end{vmatrix} \\ &= [(E/c)^2 - (p^3)^2 - ((p^1)^2 + (p^2)^2)]^2 \\ &= \left(\frac{E^2}{c^2} - \mathbf{p}^2 \right)^2 = 0, \end{aligned}$$

soit finalement

$$E = \pm |\mathbf{p}|c.$$

L'équation de Weyl pour la particule libre admet quatre solutions indépendantes, dont deux avec une énergie positive $E = |\mathbf{p}|c$ (ce sont les deux composantes de spin de la particule non massive que nous souhaitons modéliser) et deux (associées aux deux composantes de spin de l'antiparticule) avec une énergie négative $E = -|\mathbf{p}|c$.

4.2.3. Forme covariante de l'équation de Weyl

Il est instructif de pousser plus avant la forme covariante de l'équation de Weyl, $i\hbar\gamma^\mu \partial_\mu \psi = 0$, pour chercher à définir le carré invariant. Sur l'équation de Weyl, manifestement covariante,

$$i\hbar\gamma^\mu \frac{\partial \psi}{\partial x^\mu} = 0,$$

faisons agir l'opérateur $i\hbar\gamma^\nu\partial_\nu = i\hbar\gamma^\nu(\partial/\partial x^\nu)$:

$$\begin{aligned} 0 &= -\hbar^2\gamma^\nu\gamma^\mu\frac{\partial^2}{\partial x^\nu\partial x^\mu}\psi \\ &= -\frac{1}{2}\hbar^2(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu)\partial_\nu\partial_\mu\psi \\ &= -\hbar^2g^{\mu\nu}\partial_\nu\partial_\mu\psi \\ &= (i\hbar\partial^\mu)(i\hbar\partial_\mu)\psi \\ &= p^\mu p_\mu\psi \end{aligned}$$

où l'on a défini

$$p^\mu = i\hbar\frac{\partial}{\partial x_\mu} = i\hbar\partial^\mu. \quad (11.6)$$

On a obtenu ainsi une forme très satisfaisante,

$$p^\mu p_\mu\psi = \left(\frac{E^2}{c^2} - \mathbf{p}^2\right)\mathbb{1}_{4\times 4}\psi = 0. \quad (11.7)$$

Il est important de noter que l'équation de Weyl, sous sa forme manifestement covariante,

$$i\hbar\gamma^\mu\partial_\mu\psi = 0$$

n'est pas un simple scalaire de Lorentz car γ^μ est un objet à quatre composantes (comme un 4-vecteur), mais ces composantes sont des matrices 4×4 et $\partial_\mu\psi$ est un objet à 4 composantes scalaires, de sorte que la contraction donne au total un objet à 4 composantes scalaires. C'est le prix qu'il a fallu payer pour quantifier la théorie, puisqu'en physique quantique on travaille sur des objets qui en général ne commutent pas.

5. Equation de Dirac

5.1. Equation covariante pour des fermions massifs de spin $\frac{1}{2}$

Pour obtenir l'équation relativiste de l'électron ⁽¹⁴⁷⁾ et plus généralement pour des fermions massifs de spin $\frac{1}{2}$, il suffit de généraliser l'équation de Weyl covariante. Partons de la contraction, ce qui est plus immédiat,

$$[p^\mu p_\mu - m^2 c^2 \mathbb{1}_{4\times 4}]\psi = \left(\frac{E^2}{c^2} - \mathbf{p}^2 - m^2 c^2\right)\mathbb{1}_{4\times 4}\psi = 0.$$

⁽¹⁴⁷⁾ P.A.M. Dirac, *L'équation d'onde relativiste de l'électron*, Proc. Roy. Soc. **A 117**, 610-624 (1928), traduit dans J. Leite-Lopes et B. Escoubès, op. cit., p.194.

Cela suggère de coupler les équations pour $\chi^{(+)}$ et $\chi^{(-)}$ par un terme de masse pour obtenir l'équation covariante de Dirac ⁽¹⁴⁸⁾

$$i\hbar\gamma^\mu \frac{\partial\psi}{\partial x^\mu} = \gamma^\mu p_\mu \psi = mc\psi, \quad (11.8)$$

(on omet l'opérateur identité $\mathbb{1}_{4\times 4}$ et les $\mathbb{1}_{2\times 2}$ par la suite), soit, sous forme développée,

$$i\hbar \begin{pmatrix} 0 & \frac{1}{c} \frac{\partial}{\partial t} - \boldsymbol{\sigma} \cdot \vec{\nabla} \\ \frac{1}{c} \frac{\partial}{\partial t} + \boldsymbol{\sigma} \cdot \vec{\nabla} & 0 \end{pmatrix} \begin{pmatrix} \chi^{(+)} \\ \chi^{(-)} \end{pmatrix} = mc \begin{pmatrix} \chi^{(+)} \\ \chi^{(-)} \end{pmatrix}.$$

Dans le cas d'une particule libre, on a

$$\begin{pmatrix} 0 & E - c\boldsymbol{\sigma} \cdot \mathbf{p} \\ E + c\boldsymbol{\sigma} \cdot \mathbf{p} & 0 \end{pmatrix} \begin{pmatrix} \chi^{(+)} \\ \chi^{(-)} \end{pmatrix} = mc \begin{pmatrix} \chi^{(+)} \\ \chi^{(-)} \end{pmatrix}.$$

Il existe d'autres représentations des matrices de Dirac, notamment la représentation de Dirac obtenue par transformation unitaire à partir de la représentation chirale, $\gamma^\mu \rightarrow U\gamma^\mu U^\dagger$ où $U = 2^{-1/2} \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & -\mathbb{1} \end{pmatrix}$,

$$\begin{aligned} \gamma^\mu &= (\gamma^0, \boldsymbol{\gamma}), \\ \gamma^0 &= \begin{pmatrix} \mathbb{1}_{2\times 2} & 0 \\ 0 & -\mathbb{1}_{2\times 2} \end{pmatrix}, \quad \boldsymbol{\gamma} = \gamma^i = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}. \end{aligned}$$

Cette fois on obtient dans le cas de la particule libre l'équation

$$\begin{pmatrix} E - mc^2 & -c\boldsymbol{\sigma} \cdot \mathbf{p} \\ c\boldsymbol{\sigma} \cdot \mathbf{p} & -E - mc^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = 0$$

avec $\psi = 2^{-1/2} \begin{pmatrix} \chi^{(+)} + \chi^{(-)} \\ \chi^{(+)} - \chi^{(-)} \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}$, où u et v sont des spineurs à deux composantes. L'équation aux valeurs propres conduit à

$$\text{Det} \begin{pmatrix} E - mc^2 & -c\boldsymbol{\sigma} \cdot \mathbf{p} \\ c\boldsymbol{\sigma} \cdot \mathbf{p} & -E - mc^2 \end{pmatrix} = c^2 \mathbf{p}^2 - (E^2 - m^2 c^4) = 0,$$

ou encore pour les particules et les antiparticules,

$$E = \pm \sqrt{|\mathbf{p}|^2 c^2 + m^2 c^4}. \quad (11.9)$$

⁽¹⁴⁸⁾ Une notation "slash" a été introduite pour simplifier légèrement l'écriture. On définit

$$\not\phi \equiv \gamma^\mu a_\mu,$$

ce qui permet d'écrire l'équation de Dirac par exemple comme

$$(i\hbar \not\partial - mc)\psi = (\not{p} - mc)\psi = 0.$$

5.2. Equation de Dirac covariante sous champ

L'équation de Dirac sous champ s'obtient par couplage minimal,

$$p_\mu \rightarrow p_\mu - qA_\mu,$$

soit $i\hbar\gamma^\mu\partial_\mu\psi = mc\psi$ qu'on écrit encore $\gamma^\mu p_\mu\psi = mc\psi$, que l'on transforme en

$$\gamma^\mu(p_\mu - qA_\mu)\psi = i\hbar\gamma^\mu\partial_\mu\psi - q\gamma^\mu A_\mu\psi = mc\psi. \quad (11.10)$$

Pour comparer cette forme à l'équation de Klein-Gordon sous champ par exemple, on multiplie par $\gamma^\nu(p_\nu - qA_\nu)$:

$$\gamma^\mu\gamma^\nu(p_\mu - qA_\mu)(p_\nu - qA_\nu)\psi = m^2c^2\psi,$$

où l'on introduit les matrices ⁽¹⁴⁹⁾

$$\sigma^{\mu\nu} = \frac{1}{2}i(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu),$$

$$\gamma^\mu\gamma^\nu = g^{\mu\nu} - i\sigma^{\mu\nu}.$$

Il faut faire attention au fait que les grandeurs en général ne commutent pas en mécanique quantique. On peut alors ré-écrire le premier membre de l'équation de Dirac au carré,

$$\begin{aligned} (g^{\mu\nu} - i\sigma^{\mu\nu})(p_\mu - qA_\mu)(p_\nu - qA_\nu) &= (p_\mu - qA_\mu)(p^\mu - qA^\mu) \\ &\quad - \frac{1}{2}i(\sigma^{\mu\nu} - \sigma^{\nu\mu})(p_\mu - qA_\mu)(p_\nu - qA_\nu) \\ &= (p_\mu - qA_\mu)(p^\mu - qA^\mu) \\ &\quad - \frac{1}{2}i\sigma^{\mu\nu}[p_\mu - qA_\mu, p_\nu - qA_\nu]. \end{aligned}$$

Le commutateur se développe $[p_\mu - qA_\mu, p_\nu - qA_\nu] = q[p_\nu, A_\mu] - q[p_\mu, A_\nu] = -i\hbar q(\partial_\mu A_\nu - \partial_\nu A_\mu)$ et on obtient finalement

$$(p_\mu - qA_\mu)(p^\mu - qA^\mu)\psi + \frac{1}{2}q\hbar\sigma^{\mu\nu}F_{\mu\nu}\psi = m^2c^2\psi, \quad (11.11)$$

ce qui fait apparaître très naturellement le tenseur de Faraday, et où le spin est naturellement inclus dans le couplage avec $F_{\mu\nu}$.

⁽¹⁴⁹⁾ L'introduction de ces matrices permet d'introduire l'équation de Dirac à partir d'un raisonnement assez simple. On a vu avec l'équation de Klein-Gordon que la dérivée seconde par rapport à t posait quelques difficultés. Au lieu de partir de $p_\mu p^\mu = m^2c^2$ dans le cas classique (non quantique), qui fait apparaître cette dérivée seconde par la substitution $p_\mu \rightarrow +i\hbar\partial_\mu$, essayons de former une contraction du type $\gamma^\mu p_\mu\psi = mc\psi$. Il faut alors que γ^μ soit un quadri-vecteur contravariant sans dimension physique, c'est-à-dire une certaine généralisation de 1. Si alors on a la propriété $\gamma^\mu p_\mu \gamma^\nu p_\nu = \gamma^\mu \gamma^\nu p_\mu p_\nu = g^{\mu\nu} p_\mu p_\nu = p_\mu p^\mu$, on retrouve l'équation cherchée sous la forme $p_\mu p^\mu \psi = m^2c^2\psi$. On voit que cela impose aux composantes de γ^μ d'être des matrices 4×4 et on peut finalement écrire $p_\mu p^\mu \psi = -\hbar^2 \partial_\mu \partial^\mu \psi = m^2c^2\psi$. En toute rigueur, il faut noter que $\gamma^\mu \gamma^\nu \neq g^{\mu\nu}$, mais $\gamma^\mu \gamma^\nu = g^{\mu\nu} - i\sigma^{\mu\nu}$ avec $\sigma^{\mu\nu}$ antisymétrique, de sorte que

$$\begin{aligned} \gamma^\mu \gamma^\nu p_\mu p_\nu &= (g^{\mu\nu} - i\sigma^{\mu\nu})p_\mu p_\nu = p_\mu p^\mu - \frac{1}{2}i(\sigma^{\mu\nu} - \sigma^{\nu\mu})p_\mu p_\nu \\ &= p_\mu p^\mu - \frac{1}{2}i(\sigma^{\mu\nu} p_\mu p_\nu - \sigma^{\nu\mu} p_\nu p_\mu) = p_\mu p^\mu - \frac{1}{2}i\sigma^{\mu\nu}[p_\mu, p_\nu] = p_\mu p^\mu. \end{aligned}$$

5.3. Une formulation alternative

On trouve fréquemment dans la littérature une formulation alternative (ou plutôt des notations alternatives pour les matrices de Dirac). On définit les 4 matrices hermitiennes

$$\beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \quad \boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix},$$

c'est-à-dire que l'on a $\beta = \gamma^0$ et $\beta\boldsymbol{\alpha} = \boldsymbol{\gamma}$, $\beta^\dagger = \beta$, $\boldsymbol{\alpha}^\dagger = \boldsymbol{\alpha}$. On a ensuite directement

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \psi &= H\psi = (c\boldsymbol{\alpha} \frac{\hbar}{i} \vec{\nabla} + \beta mc^2) \psi \\ &= (c\boldsymbol{\alpha} \mathbf{p} + \beta mc^2) \psi \end{aligned} \quad (11.12)$$

ou encore, par action sur une onde plane,

$$(E - c\boldsymbol{\alpha} \mathbf{p}) \psi = \beta mc^2 \psi.$$

6. Synthèse

On peut résumer les diverses équations obtenues qui sont compatibles avec la théorie de la relativité ⁽¹⁵⁰⁾. Le prototype en est finalement fourni par l'équation de Klein-Gordon, $(\square + m^2 c^2 / \hbar^2) \phi = 0$, équation du second ordre pour un champ scalaire massif, susceptible de décrire des bosons massifs de spin nul. Les autres équations, agissant sur des objets plus complexes, comme des spineurs, doivent redonner l'équation de Klein-Gordon pour chacune des composantes.

Pour le champ de Maxwell dans le vide, décrivant des bosons non massifs de spin 1 (les photons), on a d'une part les équations de Maxwell

$$\partial_\mu F^{\mu\nu} = 0,$$

et d'autre part la définition de $F^{\mu\nu}$ en fonction de A^μ qui conduit à l'équation de propagation,

$$\square A^\mu = 0$$

compatible avec l'équation de Klein-Gordon pour des particules sans masse. On peut généraliser au cas de bosons massifs de spin 1, avec l'équation de Proca,

$$\partial_\mu F^{\mu\nu} + \frac{m^2 c^2}{\hbar^2} A^\nu = 0,$$

équivalente à l'équation de propagation,

$$\left(\square + \frac{m^2 c^2}{\hbar^2} \right) A^\mu = 0.$$

⁽¹⁵⁰⁾ L.H. Ryder, *Quantum field theory*, Cambridge University Press, Cambridge 1985.

Dans le cas des fermions de spin $\frac{1}{2}$, on a tout d'abord le cas de particules non massives, décrit par l'équation de Weyl,

$$i\hbar\gamma^\mu\partial_\mu\chi = 0$$

$$\square\chi = 0.$$

Finalement, le cas massif est donné par l'équation de Dirac,

$$(i\hbar\gamma^\mu\partial_\mu - mc)\psi = 0$$

$$\left(\square + \frac{m^2c^2}{\hbar^2}\right)\psi = 0.$$

Éléments de bibliographie

1. Incontournables...

L.D. Landau et E. Lifshitz, *Théorie des champs*, 3ème édition, Editions MIR, Moscou 1970.

Texte de référence, la relativité restreinte y est présentée dans les chapitres 1 à 4 et 8 (pour le rayonnement des charges accélérées), la relativité générale dans les chapitres 10 à 12.

Notes de cours de David Sénéchal,
en particulier *Mécanique I* et *Ondes électromagnétiques* que l'on peut trouver sur

<http://www.physique.usherb.ca/~dsenech/>

Dans *Mécanique I*, la présentation de la cinématique et de la dynamique relativiste est à un niveau élémentaire, alors que dans *Ondes électromagnétiques*, les formalismes covariant et lagrangien sont introduits.

J.D. Jackson, *Classical electrodynamics*, Wiley, New York 1999 (3ème édition), version française : *Électrodynamique classique*, Dunod, Paris 2001.

Il s'agit de l'ouvrage classique de référence en électromagnétisme. Il est très complet. Les chapitres 11 et 12 sont consacrés à la théorie de la relativité restreinte. A noter une édition française récente parue chez Dunod.

I.R. Kenyon, *General relativity*, Oxford University Press, New York 1990.

Ouvrage très accessible sur la relativité générale. Mérite vraiment qu'on s'y attarde (à conserver sur sa table de nuit !).

M.V. Berry, *Principles of cosmology and gravitation*, Institute of Physics Publishing, Bristol 1989.

Passionnant ouvrage d'introduction à la relativité générale et à la cosmologie.

R. Hakim, *Gravitation relativiste*, InterÉditions, CNRS Éditions, Paris 1994.r

Texte très bien écrit, parfait pour qui n'est pas spécialiste de la relativité générale et souhaite apprendre ! Exposé très clair avec de nombreuses références.

P.M. Schwarz and J.H. Schwarz, *Special relativity, from Einstein to strings*, Cambridge University Press, Cambridge 2004.

Ouvrage récent que j'ai envie de qualifier d'incontournable. Il développe dans sa seconde partie et à un niveau accessible les concepts modernes qui incluent la relativité et la physique quantique, comme la théorie quantique des champs ou les cordes, par l'un des spécialistes du domaine.

M. Longair, *Theoretical concepts in physics*, Cambridge University Press, Cambridge 2003.

Auteur qui revendique et assume l'aspect théorique de la physique et nous offre là un ouvrage très intéressant.

2. D'excellents compléments

- M. Lambert, *Relativité restreinte et électromagnétisme*, Ellipses, Paris 2000.
Très bonne présentation de la théorie de la relativité restreinte incluant les formalismes covariant et lagrangien et l'électromagnétisme.
- W. Pauli, *Theory of relativity*, Dover, New York 1958.
C'est un texte classique tout d'abord écrit à la demande de Sommerfeld par Wolfgang Pauli alors âgé de 21 ans, puis à peine repris et complété en 1955, à l'occasion du cinquantenaire des premiers articles d'Einstein sur la relativité. C'est devenu un texte de référence considéré comme très complet. A noter qu'il emploie d'abord la métrique de Minkowski avant d'introduire la distinction entre les tenseurs contravariants et covariants comme dans les premières éditions de Landau et Lifshitz.
- R.H. Price, *General relativity primer*, Am. J. Phys. **50**, 300 (1982).
Concis, mais intéressant et accessible.
- F. Rohrlich, *Classical charged particles*, Addison-Wesley, Redwood 1965.
Ouvrage d'un niveau avancé. Nombreuses discussions historiques.
- N.A. Doughty, *Lagrangian interaction*, Addison-Wesley, Singapore 1990.
Texte accessible sur la mécanique classique, l'électromagnétisme, la relativité, la gravitation, les symétries, l'invariance de jauge, ... Comprend de nombreux encarts historiques.
- P.G. Bergmann, *Introduction to the theory of relativity*, Dover, New York 1976.
Texte de référence sur la relativité restreinte et générale.
- R.P. Feynman, *Lectures on gravitation*, Penguin Books, 1995.
Une présentation originale de la théorie de la gravitation par Feynman, mais assez difficile à lire.
- A. Rougé, *Introduction à la physique subatomique*, Ellipses, Paris 1997.
Comprend un rapide résumé très instructif sur la relativité restreinte, notamment appliquée aux problèmes de collisions et d'électrodynamique.
- Notes de cours de Jean-Michel Raymond, *Electromagnétisme et relativité* que l'on peut trouver (avec d'autres cours excellents sur des thèmes divers) sur <http://cel.ccsd.cnrs.fr/>
- Notes de cours de J.L. Vuillemier (Neuchatel), *Relativité*, <http://www.unine.ch/phys/enseignement/physmod/re10.pdf>
- R.P. Feynman, R.B. Leighton and M. Sands, *The Feynman lectures on physics*, vol. I et II, Addison-Wesley, Reading 1977.
Sans commentaire.
- J. Vanderlinde, *Classical electromagnetic theory*, Wiley, New York 1993.
Ouvrage que l'auteur situe à un niveau intermédiaire entre le Landau et Lifshitz et le Jackson.
- W. Greiner, *Classical electrodynamics*, Springer, New York 1998.
Voir en particulier les chapitres 22 et 23.
- L. Eyges, *The classical electromagnetic field*, Dover, New York 1972.
- V.N. Rudenko, *Relativistic experiments in gravitational fields*, Sov. Phys. Usp. **21**, 893 (1978).

3. Des textes plus élémentaires

P. Lorrain et D.R. Corson, *Champs et ondes électromagnétiques*, Armand Colin, Paris 1979.

Présentation au niveau du premier cycle, centrée sur l'électromagnétisme. Illustrations soignées.

C. Gruber, *Mécanique*, Presses polytechniques romandes, Lausanne 1988.

J.H. Smith, *Introduction à la relativité*, InterÉditions, Paris 1979.

Texte d'introduction célèbre, comportant de nombreuses discussions d'expériences de pensée.

V. Ougarov, *Théorie de la relativité restreinte*, Editions de Moscou, Moscou 1979.

La cinématique et la dynamique sont présentées de façon exhaustive. L'interprétation géométrique de Bondi (calcul k) également. La notation employée pour les quadrivecteurs est celle de Minkowski.

M. Hulin, N. Hulin et L. Mousselin, *Relativité restreinte*, Dunod, Paris 1992.

De niveau DEUG. Comprend des exercices corrigés.

C. Grossetête, *Relativité restreinte*, Ellipses, Paris 1985.

De niveau DEUG. Comprend de nombreux exercices et problèmes corrigés.

M. Boratav et R. Kerner, *Relativité*, Ellipses, Paris 1991.

A. Einstein, *La théorie de la relativité restreinte et générale*, Gauthier-Villars, Paris 1976.

Texte de vulgarisation célèbre d'Einstein.

D. Bohm, *The special theory of relativity*, Addison-Wesley, Redwood 1989.

Ouvrage élémentaire bien connu. Détaille en particulier le calcul k d'Hermann Bondi.

4. Pour prolonger la réflexion...

Des grands classiques qui font référence, à feuilleter plus ou moins attentivement :

S. Weinberg, *Gravitation and Cosmology*, Wiley, New-York 1972.

Parmi les textes qui font autorité en matière de relativité générale, c'est probablement le plus célèbre, ...

C.W. Misner, K.S. Thorne and J.A. Wheeler, *Gravitation*, Freeman, San Francisco 1973.

... avec celui-ci qui est une mine d'informations.

P.A.M. Dirac, *General theory of relativity*, Princeton University Press, Princeton 1996.

Ouvrage d'une concision étonnante sur la relativité générale. Présentation très formelle.

P.J.E. Peebles, *Principles of Physical Cosmology* Princeton University Press, Princeton 1993.

A.K. Raychaudhuri, S. Banerji and A. Banerjee, *General Relativity, Astrophysics, and Cosmology* Springer-Verlag, New-York 1992.

- J. Leite Lopes, *Théorie Relativiste de la Gravitation* Masson, Paris 1993.
- W. Rindler, *Introduction to Special Relativity*, Oxford University Press, Oxford 1991.
Ouvrage très complet sur la relativité restreinte, qui met davantage l'accent sur les aspects fondamentaux et théoriques que sur les expériences.
- C. Møller, *The theory of Relativity*, Oxford University Press, Oxford 1952.
Fait partie des classiques, relativité restreinte et générale.
- R. Adler, M. Bazin and M. Shiffer, *Introduction to general relativity*, MacGraw Hill, New-York 1965.
Egalement un classique sur la relativité générale.
- R.A. Mould, *Basic relativity*, Springer, New-York 1994.
Présentation détaillée, avec de nombreuses applications de géométrie différentielle à 2 dimensions.

Des ouvrages accessibles sur la relativité, la gravitation, la cosmologie. . .

- E. Harrison, *Cosmology*, 2nd edition, Cambridge University Press, Cambridge 2000.
Un ouvrage passionnant sur la cosmologie, de l'antiquité aux théories actuelles.
Très accessible, la relativité y joue un rôle en de multiples endroits.
- J. Rich, *Fundamentals of cosmology*, Springer, Berlin 2001.
La cosmologie moderne. Certaines parties sont accessibles assez facilement.
- A. Einstein, *Œuvres choisies*, Vol. 2 et 3, Seuil, CNRS Éditions, Paris 1993.
Représente une somme ! Contient de nombreux textes d'Einstein et correspondances traduits en français et commentés.
- G.F.R. Ellis and R.M. Williams, *Flat and curved space-times*, 2nd edition, Oxford University Press, Oxford 2000.
Dans le même esprit que Harrison. Très agréable à parcourir.
- W. Rindler, *Relativity, Special, general and cosmological*, Oxford University Press, Oxford 2001.
- A. Guth et P. Steinhardt, *l'Univers inflationniste* in *La Nouvelle Physique*, ed. par P. Davies, Sciences Flammarion, Paris 1993.
Bel ouvrage de vulgarisation sur les grands domaines de la physique.
- M. Rowan-Robinson, *Cosmology*, Oxford University Press, 4ème édition, Oxford 2004.
Ouvrage parfaitement accessible et très à jour.
- A. Friedmann, G. Lemaître, *Essais de cosmologie*, édité par J.P. Luminet, Seuil, Paris 1997.
Présentation (et traduction si nécessaire) des textes fondateurs de Friedmann et de Lemaître sur la cosmologie.

5. Aspects historiques - épistémologiques

A.I. Miller, *Albert Einstein's special theory of relativity*, Springer, New York 1998.
C'est un ouvrage très documenté sur l'aspect historique de la genèse de la relativité restreinte.

Twentieth Century Physics, Ed. by L.M. Brown, A. Pais and B. Pippard, IOP and AIP Press, 1995, Vol. I, Chap. 4.

H.A. Lorentz, A. Einstein, H. Minkowski and H. Weyl, *The principle of relativity*, Dover, New York 1952.

Recueil de textes originaux des fondateurs traduits en anglais.

A. Einstein, *Réflexions sur l'électrodynamique, l'éther, la géométrie et la relativité*, Gauthier-Villars, Paris 1972.

Contributions célèbres d'Einstein traduites en français.

J. Eisenstaedt, *Histoire et singularités de la solution de Schwarzschild (1915-1923)*, Arch. Hist. Exact Sci. **27**, 157 (1982).

J. Eisenstaedt, *La relativité générale à l'étiage (1925-1955)*, Arch. Hist. Exact Sci. **35**, 115 (1986).

J. Eisenstaedt, *Trajectoires et impasses de la solution de Schwarzschild*, Arch. Hist. Exact Sci. **37**, 275 (1988).

J. Eisenstaedt, *Einstein et la relativité générale*, CNRS Editions, Paris 2002.

Des textes historiques (articles originaux, extraits de correspondance) traduits en français et commentés sous la direction de Françoise Balibar :

A. Einstein, *Œuvres choisies, Tome 1, Quanta*, Seuil/CNRS, Paris 1989.

A. Einstein, *Œuvres choisies, Tome 2, Relativités I*, Seuil/CNRS, Paris 1993.

A. Einstein, *Œuvres choisies, Tome 3, Relativités II*, Seuil/CNRS, Paris 1993.

A. Einstein, *Œuvres choisies, Tome 4, Correspondance française*, Seuil/CNRS, Paris 1989.

A. Einstein, *Œuvres choisies, Tome 5, Science, Éthique, Philosophie*, Seuil/CNRS, Paris 1991.

6. Des articles pédagogiques sur des sujets ponctuels

R.J. Adler, B. Casey and O.C. Jacob, *Vacuum catastrophe: an elementary exposition of the cosmological constant problem*, Am. J. Phys. **63**, 620 (1995).

D. Atkatz, *Quantum cosmology for pedestrians* Am. J. Phys. **62**, 619 (1994).

R.P. Bickerstaff and G. Patsakos, *Relativistic generalizations of mass*, Eur. J. Phys. **16**, 63 (1995).

Ph. Brax, *The cosmological constant problem*, Contemp. Phys. **45**, 227 (2004).

Exposé à jour sur les problèmes suscités par la valeur de la constante cosmologique et la matière noire.

- D.R. Brill and D. Goel, *Light bending and perihelion precession: A unified approach*, Am. J. Phys. **67**, 316 (1999).
- C. Callan, R.H. Dicke and P.J.E. Peebles, *Cosmology and newtonian mechanics*, Am. J. Phys. **33**, 105 (1965).
- C.F. Chyba, *Kaluza-Klein unified field theory and apparent four-dimensional space-time*, Am. J. Phys. **53**, 863 (1985).
- R. P. Comer and J.D. Lathrop, *Principle of equivalence and deflection of light by the sun*, Am. J. Phys. **46**, 801 (1978).
- V. Faraoni, *A new solution for inflation*, Am. J. Phys. **69**, 372 (2001).
Une présentation simplifiée de la théorie de l'inflation, où un champ scalaire conduit à des effets analogues à la constante cosmologique et produit une expansion exponentielle de l'Univers.
- V. Faraoni, *Solving for the dynamics of the universe*, Am. J. Phys. **67**, 732 (1999).
- N. Gauthier, *Periastron precession in general relativity*, Am. J. Phys. **55**, 85 (1987).
- V.K. Fredericks, *Einstein's principle of relativity* Physics-Uspekhi **42**, 1235 (1999), reprinted from Uspekhi Fizicheskikh Nauk bf 2, 162 (1921).
- S.K. Graves, J.E. Gaiser and J.D. French, *General relativistic trajectories using small computers*, Am. J. Phys. **50**, 86 (1982).
- O. Gron, *A new standard model of the universe*, Eur. J. Phys. **23**, 135 (2002).
Une présentation assez détaillée de la dynamique du modèle de Friedmann-Lemaître.
- A. Harvey, *Cosmological models*, Am. J. Phys. **61**, 901 (1993).
- A. Harvey and E. Schucking, *Einstein's mistake and the cosmological constant*, Am. J. Phys. **68**, 723 (2000).
Le titre est explicite. Très intéressant.
- J. Higbie, *Gravitational lens*, Am. J. Phys. **49**, 652 (1981).
- K.K. Nandi and A. Islam, *On the optical-mechanical analogy in general relativity*, Am. J. Phys. **63**, 251 (1995).
L'utilisation de l'analogie optico-mécanique avec la métrique de Schwarzschild pour déterminer la déflexion des rayons lumineux et le retard des échos radar.
- N. McCubbin, *Beauty in physics: the legacy of Paul Dirac*, Contemp. Phys. **45**, 319 (2004).
Un article en hommage à Dirac, avec notamment une présentation simplifiée très agréable de l'équation de Dirac. Insiste également sur le rôle très important joué par Heisenberg.
- W. Moreau, R. Neutze and D.K. Ross, *The equivalence principle in the Schwarzschild geometry*, Am. J. Phys. **62**, 1037 (1994).
- J.V. Narlikar, *Spectral shifts in general relativity*, Am. J. Phys. **62**, 903 (1994).
- J.W. Norbury, *From Newton's laws to the Wheeler-DeWitt equation*, Eur. J. Phys. **19**, 143 (1998).
Présentation détaillée de la cosmologie newtonienne, jusqu'à une introduction à la gravité quantique avec l'équation de Wheeler-DeWitt.
- H.C. Ohanian, *What is principle of equivalence?*, Am. J. Phys. **45**, 903 (1977).

- M.G. Olsson, *General relativistic perihelia precession and the anharmonic oscillator*, Am. J. Phys. **56**, 89 (1988).
- R. Perrin, *Twin paradox: a complete treatment from the point of view of each twin*, Am. J. Phys. **47**, 317 (1979).
- P. Pesic and S. P. Boughn, *The Weyl-Cartan theorem and the naturalness of general relativity*, Eur. J. Phys. **24**, 261 (2003).
Le théorème de Weyl-Cartan (démontré indépendamment à la même période par ces deux auteurs) stipule que le scalaire de Ricci est le seul invariant relativiste construit à partir du tenseur métrique, de ses dérivées premières et qui dépende linéairement de ses dérivées secondes. Cela entraîne que la théorie d'Einstein est la seule construite de cette manière à partir du tenseur métrique.
- W. Rindler, *General relativity before special relativity: an unconventional overview of relativity theory*, Am. J. Phys. **62**, 887 (1994).
- S. Sonego and M. Pin, *Deriving relativistic momentum and energy*, Eur. J. Phys. **26**, 33 (2005).
Utilisation d'arguments très généraux pour obtenir la forme de E et \mathbf{p} , met en lumière les propriétés de conservation à la base de la définition de ces quantités.
- B. Stefanski and D. Bedford, *Vacuum gravity*, Am. J. Phys. **62**, 638 (1994).
- A.J. Schwarz and N.A. Doughty, *Kaluza-Klein unification and the Fierz-Pauli weak field limit* Am. J. Phys. **60**, 150 (1992).
- N. Straumann, *The mystery of the cosmic vacuum energy density and the accelerated expansion of the Universe*, Eur. J. Phys. **20**, 419 (1999).
Présentation historique du rôle de la constante cosmologique (et des hésitations qu'elle a suscitées) et des problèmes que pose sa valeur actuelle extrêmement faible.
- W.G. Unruh and G.I. Opat, *The Bohr-Einstein "weighting of energy"*, Am. J. Phys. **47**, 743 (1979).
- Y. Zarmi, *The Bertrand theorem revisited*, Am. J. Phys. **70**, 446 (2002).
Le théorème de Bertrand précise que les potentiels coulombien et harmonique sont les seuls qui admettent des trajectoires fermées (en dynamique newtonienne).