



Université de Liège

Faculté des sciences

Syllabus pour le cours de

Relativité générale

par

Yves De Rop

décembre 2021

« L'amour, c'est l'espace et le temps rendus sensibles au cœur. »

Marcel Proust (1871–1922),
A la recherche du temps perdu, tome 5 : *La Prisonnière*.

A la mémoire de mon professeur et ami, Jacques Demaret.

Table des matières

Avant-propos	vi
1 Relativité restreinte	1
1.1 Quelques considérations historiques	1
1.2 La dilatation du temps	5
1.3 La contraction de l'espace	6
1.4 La relativité de la simultanéité	6
1.5 Etude d'un premier paradoxe cinématique	7
1.6 La transformation de Lorentz	8
1.7 Etude d'un second paradoxe cinématique	13
1.8 Quadrivecteur énergie-impulsion d'une particule	16
1.9 Le photon	18
1.10 Equivalence de la masse et de l'énergie	21
1.11 Une démonstration simple	23
1.12 Problèmes	24
2 Approche du champ gravitationnel uniforme et constant	27
2.1 Le voyageur de Langevin	27
2.2 Référentiel d'une particule uniformément accélérée	29
2.3 Le principe d'équivalence et le principe de covariance générale	32
2.4 Conséquences du principe d'équivalence	35
2.4.1 L'accélération en relativité générale	36
2.4.2 Nature géométrique du champ de gravitation	37
2.4.3 L'effet Einstein	38
2.5 Champ de gravitation uniforme et constant	40
2.6 Le paradoxe de Langevin	41
2.7 Les limites de la relativité générale	42
2.8 Courbure des trajectoires en théorie newtonienne de la gravitation	43
3 Le carrousel d'Einstein	46
3.1 Position du problème	46
3.2 Distances et durées physiques	48
3.3 Application : la transformation de Lorentz	51
3.4 Mesures intrinsèques sur les sections spatiales	52
3.5 Le point de vue lagrangien	52
3.6 L'effet Sagnac	55

4	La courbure de Gauss	56
4.1	Géodésiques d'une surface dans l'espace euclidien	56
4.2	Première et deuxième formes fondamentales	57
4.3	Un exemple : la sphère S^2 de rayon \mathcal{R}	59
4.4	La courbure de Gauss	59
4.5	Le <i>Theorema egregium</i>	62
4.6	Equation de déviation géodésique	62
4.7	Mesure de la circonférence d'un cercle sur une surface	64
4.8	Courbure gaussienne du tore	64
4.9	Courbure gaussienne des surfaces de révolution	65
4.10	Les géodésiques du tore	67
4.11	Champ de gravitation et courbure de l'espace	69
4.12	Les surfaces à courbure négative constante	71
	4.12.1 La pseudosphère de Beltrami	71
	4.12.2 Autres représentations	73
4.13	Le tenseur de courbure	74
5	Géométrie affine et riemannienne	75
5.1	Notion de variété différentielle	75
5.2	Notions de topologie différentielle	76
	5.2.1 Fonctions et courbes	76
	5.2.2 Vecteurs	76
	5.2.3 Formes différentielles	77
	5.2.4 Le produit intérieur	79
	5.2.5 Tenseurs	80
5.3	Notion de transport parallèle	82
5.4	Les coordonnées cartésiennes dans \mathbb{R}^n	82
5.5	Les coordonnées générales dans \mathbb{R}^n	83
	5.5.1 Champs de vecteurs	83
	5.5.2 Champs de formes	84
	5.5.3 Champs tensoriels quelconques	85
5.6	Propriétés axiomatiques de la dérivée covariante	86
5.7	Géodésiques dans un espace affine	88
5.8	Les variétés riemanniennes	89
5.9	Coordonnées localement cartésiennes	91
5.10	Les connexions riemanniennes	92
5.11	Métrique et connexion sur une surface immergée dans \mathbb{R}^3	94
5.12	Géodésiques dans un espace riemannien	96
5.13	L'ascenseur d'Einstein	97
5.14	Commutateur de deux champs de vecteurs	99
	5.14.1 Définition algébrique	99
	5.14.2 Interprétation géométrique	100
5.15	L'opérateur de courbure	101
5.16	Approche intuitive du tenseur de Riemann-Christoffel	102
5.17	Le tenseur de Riemann-Christoffel	105
	5.17.1 Définition affine	105
	5.17.2 Symétries	105
	5.17.3 Interprétation géométrique	106

5.18	Les espaces euclidiens	108
5.19	Equation de déviation géodésique	110
5.20	Les identités de Bianchi	112
5.20.1	Aspects algébriques	112
5.20.2	Interprétation géométrique	113
5.21	Courbure des variétés riemanniennes	115
5.21.1	Nouvelles propriétés de symétrie	115
5.21.2	Nombre de composantes indépendantes du tenseur de courbure . . .	116
5.21.3	La matrice de Petrov	116
5.21.4	Tenseurs associés	117
5.21.5	Les identités de Bianchi contractées	117
5.21.6	Lien avec la courbure de Gauss	118
5.22	Exercices	119
6	Le tenseur d'énergie-impulsion	123
6.1	Le statut de l'énergie volumique en relativité restreinte	123
6.2	Définitions	123
6.3	Le tenseur d'énergie-impulsion des fluides	125
6.3.1	Symétrie du tenseur d'énergie-impulsion	125
6.3.2	Conservation de l'énergie et de la quantité de mouvement	127
6.3.3	Conservation du nombre de particules	128
6.4	Les fluides parfaits	128
6.4.1	Expression du tenseur d'énergie-impulsion	128
6.4.2	Equation d'état	130
6.4.3	Equations de conservation pour les fluides parfaits non relativistes .	130
6.4.4	Considérations thermodynamiques	131
6.4.5	Fluides parfaits dans un champ de gravitation	132
6.5	Les fluides visqueux	133
6.6	Le tenseur d'énergie-impulsion du champ électromagnétique	134
6.6.1	Lois de conservation	134
6.6.2	Fluide de particules chargées de type « poussière »	136
7	Les équations d'Einstein	138
7.1	Retour sur le principe d'équivalence	138
7.2	« Die Lochbetrachtung »	140
7.3	La démarche heuristique	143
7.4	Equations du champ et équations du mouvement	144
7.5	La méthode de Hilbert	146
7.5.1	L'action du champ gravitationnel	146
7.5.2	Obtention des équations d'Einstein à l'aide d'un principe variationnel	148
7.6	L'approximation des champs faibles	148
7.6.1	Généralités	148
7.6.2	Les équations d'Einstein linéarisées	151
7.6.3	La limite newtonienne	153
7.6.4	Calcul de la constante de couplage	154
7.6.5	Commentaire	156
7.6.6	Fluide parfait dans un champ de gravitation newtonien	157
7.6.7	L'intégrale d'action	157

7.6.8	Déviations des rayons lumineux par un corps sphérique	159
7.7	Champ fort loin des sources	162
7.8	L'approximation post-newtonienne	163
7.9	La tentation de l'espace absolu	167
7.9.1	L'équation du mouvement en mécanique newtonienne	167
7.9.2	Sur la covariance de la mécanique newtonienne	169
7.9.3	Sur la covariance de la relativité restreinte	170
7.9.4	Sur la covariance de la relativité générale	171
8	Les univers de Friedmann-Lemaître	173
8.1	Les difficultés de la cosmologie newtonienne	173
8.2	Quelques citations	174
8.3	Les modèles cosmologiques spatialement homogènes et isotropes	175
8.3.1	Motivation	175
8.3.2	Notion de groupe d'isométries	176
8.3.3	Notion d'espace à symétrie maximale	177
8.3.4	Espaces avec sous-espaces à symétrie maximale	178
8.3.5	Les variétés spatiales homogènes et isotropes à trois dimensions	178
8.3.6	Les modèles cosmologiques spatialement homogènes et isotropes	180
8.4	Cosmologie et inertie	182
8.5	Evolution de l'univers et décalage des spectres	184
8.6	La loi de Hubble-Lemaître	185
8.7	Ecriture des équations d'Einstein	187
8.8	Evolution qualitative du facteur d'échelle	189
8.9	Résolution des équations d'Einstein ($\Lambda = 0$)	192
8.9.1	Cas $P = 0$ (matière incohérente)	192
8.9.2	Cas $P = \rho/3$ (rayonnement)	194
8.9.3	Exercices	196
8.10	Cosmologie physique	199
8.10.1	Histoire thermique de l'univers	199
8.10.2	L'entropie de l'univers	201
8.10.3	Les neutrinos cosmologiques	202
8.10.4	Matière et antimatière	203
9	La solution de Schwarzschild	205
9.1	L'espace-temps à symétrie sphérique spatiale	205
9.2	Solution dans le vide	207
9.3	Solution intérieure	208
9.4	Visualisation des sections spatiales	211
9.5	Equations du mouvement et intégrales premières dans le vide	213
9.6	Equation des trajectoires	215
9.7	Le potentiel effectif	216
9.8	La capture gravitationnelle	219
9.8.1	Section efficace de capture	219
9.8.2	Angle de capture	222
9.9	Les tests de la relativité générale	224
9.9.1	L'effet Einstein	224
9.9.2	Précession du périastre des planètes	225

9.9.3	Déviations des rayons lumineux par le soleil	229
9.9.4	Retard des échos RADAR	230
9.10	Particules au voisinage du rayon gravitationnel	233
9.11	Le trou noir : solution extérieure	237
9.11.1	Pathologies des coordonnées de Schwarzschild	237
9.11.2	Les coordonnées de Kruskal-Szekeres	241
9.12	Le trou noir : solution intérieure	244
9.12.1	Effondrement radial d'une sphère de poussière selon Newton	245
9.12.2	Effondrement radial d'une sphère de poussière selon Einstein	245
9.12.3	La formation de l'horizon	247
A	L'électromagnétisme en Système International	250
A.1	Notations et équations fondamentales	250
A.2	L'invariance de jauge	251
A.3	Solutions quasi-statiques des équations de Maxwell	255
B	Les coordonnées normales	259
C	Les équations d'Einstein par la méthode d'Euler-Lagrange	262
D	Quelques résultats de mécanique statistique	267
D.1	Potentiels thermodynamiques	267
D.2	Densité et énergie d'un gaz parfait	267
D.3	Quatre intégrales utiles	269
D.4	Application au gaz de photons (V et T fixés)	269
E	Physique et philosophie	271
E.1	Origine et signification de la <i>Critique de la raison pure</i>	271
E.2	Lexique	278

Avant-propos

Si les applications pratiques de la relativité générale demeurent encore assez limitées, son importance conceptuelle dans la pensée contemporaine, par contre, semble à présent indiscutable. Le lecteur intéressé par cette théorie en trouvera sur le marché d'excellents exposés, parfois écrits par des auteurs illustres. Pour fixer le niveau à atteindre au terme de celui-ci, je me suis inspiré du livre de Bernard F. Schutz, *A first course in general relativity*, Cambridge University Press, 2009, à l'exclusion du chapitre sur les ondes gravitationnelles qui fera l'objet d'un autre texte. Je me suis écarté du discours de Schutz chaque fois que j'ai cru pouvoir proposer une formulation plus générale, précise ou pédagogique. Il en a résulté ces notes, qui ne prétendent donc pas constituer un traité autonome. Je serais cependant heureux de recevoir les critiques permettant d'en améliorer le contenu ou la présentation. En voici le fil conducteur.

Le premier chapitre rappelle les principaux acquis de la relativité restreinte en insistant notamment sur les deux principes à la base de cette théorie et en tâchant de rendre aussi intuitifs que possible la dilatation du temps, la contraction de l'espace, la relativité de la simultanéité, l'espace-temps de Minkowski, le concept de photon, l'équivalence entre la masse et l'énergie. Certains paradoxes apparents sont analysés pour montrer la cohérence de la théorie.

Mais l'objectif principal du cours est de dégager les propriétés essentielles de la théorie relativiste de la gravitation, telle qu'elle fut conçue par Einstein à partir de 1907. La pierre angulaire de l'argumentation est le *principe d'équivalence* entre gravitation et inertie, reposant sur l'égalité entre masse inerte et masse gravifique et selon lequel gravitation et inertie constituent un seul et même phénomène envisagé selon deux points de vue différents. Or, en mécanique newtonienne, les forces fictives d'inertie sont la marque des référentiels non inertiels : incorporer la gravitation dans l'appareil théorique revient donc à saper le concept de référentiel inertiel et suggère l'équivalence formelle de tous les référentiels. C'est le *principe de relativité générale*. Il implique, comme conséquence étonnante, l'incompatibilité d'une description naturelle du champ de gravitation avec le cadre géométrique de la relativité restreinte.

Deux exemples sont traités en détail afin de développer progressivement une intuition du champ de gravitation en relativité générale.

Le premier exemple (chapitre 2) commence par une remarque fondamentale : en relativité restreinte, la définition d'un mouvement uniformément accéléré est délicate car l'accélération spatiale n'est pas absolue ; ce qui, incidemment, signe l'arrêt de mort de la dynamique newtonienne, où l'accélération spatiale était caractérisée par un vecteur indépendant du référentiel inertiel choisi. Ainsi, la cinématique d'une particule « uniformément accélérée » ne respecte pas les lois galiléennes. Les résultats obtenus sont utilisés pour modéliser un champ de gravitation constant et (presque) homogène. Bien qu'un tel champ ne se rencontre pas rigoureusement dans la nature, il constitue néanmoins une excellente pre-

mière approche et révèle plusieurs caractéristiques d'un champ de gravitation générique :

- inadéquation des coordonnées cartésiennes de Minkowski ; recours à des transformations non linéaires des coordonnées, ces dernières n'ayant plus de sens physique immédiat ; apparition du potentiel gravifique dans l'expression du tenseur métrique ;
- déviation des rayons lumineux ;
- décalage en fréquence de la lumière en fonction de l'altitude ;
- impossibilité de maintenir synchrones des horloges au repos situées à des altitudes différentes.

Le deuxième exemple, le référentiel en rotation autour d'un axe (« le carrousel d'Einstein »), fait l'objet du chapitre 3. En appliquant une procédure géométrique pour la mesure opérationnelle des distances spatiales et des durées, on constate l'impossibilité de définir une simultanéité globale. Les sections spatiales, c'est-à-dire le carrousel à un moment donné, ne peuvent donc plus être définies que localement. Un calcul simple révèle leur propriété la plus significative : elles possèdent une courbure (l'espace-temps, cependant, reste celui de Minkowski).

Le chapitre 4 étudie la géométrie des surfaces et, en particulier, la courbure de Gauss. Dans le cas des surfaces de révolution il est assez facile de dériver une expression analytique pour cette dernière. Une analogie, due à Poincaré, est ensuite exposée pour illustrer combien peut s'avérer pertinente l'interprétation d'une situation physique en termes de courbure de l'espace-temps. Il s'agit d'un plan soumis à un champ de température induisant des modifications de longueur des règles et de l'indice de réfraction de la lumière (le champ de température joue ici le rôle du champ de gravitation en relativité générale). Bien que l'espace soit euclidien pour un observateur isolé thermiquement, il est commode de le décrire comme une surface courbe de géométrie sphérique ou hyperbolique selon le choix du champ de température. Le principe de Fermat appliqué à la lumière impose à celle-ci de parcourir les géodésiques de cette surface et les *géodésiques nulles de l'espace-temps*. Pour finir, le chapitre évoque les surfaces à courbure négative constante.

Le chapitre 5 est consacré aux notions géométriques indispensables pour comprendre la suite. En premier lieu on y présente les objets fondamentaux de la topologie différentielle : fonctions, vecteurs, formes différentielles, tenseurs d'ordre quelconque. On aborde ensuite les variétés affines, puis les variétés riemanniennes (à ce stade du cours, l'étudiant aura déjà perçu l'importance du tenseur métrique en relativité générale). Nous introduisons ainsi le concept de géodésique, les connexions riemanniennes, l'opérateur de courbure, le tenseur de Riemann-Christoffel avec son interprétation géométrique et ses symétries, quelques considérations relatives aux espaces euclidiens, l'équation de déviation géodésique, et enfin les identités de Bianchi. Chaque fois que c'est possible, nous nous efforçons de distinguer les propriétés affines des propriétés métriques.

En guise d'introduction au chapitre 6, nous recommandons l'excellente présentation du concept de tenseur d'énergie-impulsion dans le chapitre 4 du livre de Schutz. Nous nous contentons ici d'en donner une définition générale avant d'étudier l'impact, sur ce tenseur, des lois de conservation du moment cinétique, de l'énergie et de la quantité de mouvement. Puis nous établissons son expression pour un fluide parfait et pour le champ électromagnétique.

« La matière courbe l'espace-temps qui, en retour, prescrit à la matière son mouvement ». Voilà, dans ses grandes lignes, la philosophie de la relativité générale. Le chapitre 7 l'explique en revenant d'abord sur le principe d'équivalence, réinterprété comme principe de couplage minimum entre la gravitation et les autres lois de la physique. Il conduit aux

lois de conservation de l'énergie et de la quantité de mouvement en présence d'un champ de gravitation. Du point de vue mathématique, il est naturel de formaliser le principe de relativité générale en imposant l'équivalence de tous les systèmes de coordonnées : les lois de la physique doivent être tensorielles pour des transformations *quelconques* de coordonnées. Historiquement, ce *principe de covariance générale* mit un certain temps à s'imposer, comme en témoigne l'argument du « Lochbetrachtung ». Mais si on l'admet, on arrive assez directement aux équations relativistes du champ gravitationnel, également appelées *équations d'Einstein*. Nous amenons celles-ci en suivant, d'abord une voie heuristique proche du cheminement originel, ensuite la méthode variationnelle de Hilbert. L'approximation des champs faibles assure le raccord avec la théorie de Newton et permet de calculer la constante de couplage entre la géométrie et la matière. La théorie est alors logiquement achevée. Cette arrivée au sommet fournit l'occasion de calculer les premières corrections post-newtoniennes dans un espace-temps courbé par un fluide parfait. Nous en profitons aussi pour prendre un peu de recul et jeter un regard rétrospectif sur le chemin parcouru en comparant les fondements de la mécanique newtonienne avec ceux de la relativité restreinte et de la relativité générale.

Depuis l'avènement de la relativité générale, la cosmologie peut prétendre au statut de science. Après un tour d'horizon historique et philosophique, le chapitre 8 décrit les modèles mathématiquement les plus simples, qui sont spatialement homogènes et isotropes. Certaines propriétés importantes sont discutées qui ne dépendent que des symétries du tenseur métrique, comme la loi de Hubble-Lemaître. Ensuite on examine les solutions des équations d'Einstein, essentiellement sous l'angle cinématique : notion d'âge de l'univers, lien entre décalage vers le rouge et distance physique, notion d'horizon-particule. Le chapitre se conclut sur une modeste introduction à l'histoire thermique de l'univers.

Le chapitre 9 explore la géométrie engendrée par un astre à symétrie sphérique spatiale, appelée métrique de Schwarzschild. Cette solution prédit des divergences avec la théorie de Newton, qui peuvent être testées par l'observation ou l'expérience. Son intérêt astrophysique est donc exceptionnel. D'un point de vue pédagogique, elle contribue à développer l'intuition de la gravitation relativiste, tant au niveau des champs faibles (système solaire) que forts (trous noirs).

Chapitre 1

Relativité restreinte

1.1 Quelques considérations historiques

Au début du XIX^e siècle, il ressortit des travaux de Young et Fresnel que la lumière possède les caractéristiques d'un phénomène ondulatoire. On disposait depuis longtemps déjà d'une théorie de la propagation des ondes, décrite par l'équation de d'Alembert, et notamment des ondes acoustiques. Celles-ci sont des perturbations *matérielles longitudinales*, consistant en compressions et dilatations successives du milieu qu'elles traversent, qu'il soit gazeux, liquide ou solide. Ainsi, le son se déplace facilement dans un tuyau en fonte, dans l'eau ou dans l'air. Mais il apparut que la lumière est une onde *transversale*, très rapide de surcroît (sa vitesse c vaut environ 3×10^8 m/s), et aucune substance connue à l'époque ne semblait pouvoir constituer un support possible pour expliquer sa propagation. On imagina donc un substrat, l'*éther*, doté de qualités paradoxales : assez rigide pour expliquer la vitesse de la lumière (de même qu'une vibration se meut plus rapidement dans une corde plus tendue), il devait d'autre part rester suffisamment subtil pour pénétrer tous les milieux où se propage la lumière et ne pas entraver les planètes dans leurs révolutions autour du soleil.

Malgré ces difficultés, il se trouvait dans la seconde moitié du XIX^e siècle d'assez nombreux arguments pour légitimer l'hypothèse de l'existence de l'éther, en théorie tout au moins. C'est que dans l'esprit des physiciens, l'éther ne jouait pas seulement le rôle de milieu matériel pour la propagation de la lumière : il définissait également un système d'inertie privilégié. Voici pourquoi. En 1864, Maxwell avait publié les équations fondamentales de l'électromagnétisme, où il apparaissait que *les ondes lumineuses sont de nature électromagnétique*. Ces équations n'étaient pas covariantes pour la transformation de Galilée ; mais cela n'avait rien de choquant dans le cadre d'une théorie de l'éther : on pensait en effet que les équations de Maxwell, régissant les ondes électromagnétiques, revêtaient une forme simplifiée dans le référentiel de l'éther, exactement comme l'équation de d'Alembert, régissant les ondes acoustiques, dans un repère immobile par rapport à l'air. Bien mieux, cette absence de neutralité des équations de Maxwell vis-à-vis des différents repères inertiels fit germer l'espoir selon lequel on allait enfin pouvoir *matérialiser l'espace absolu*, cette entité métaphysique alléguée par Newton pour justifier les lois de la mécanique : l'éther serait au repos par rapport à l'espace absolu.

Une conséquence optique généralement admise de l'hypothèse de l'éther était la suivante : *dans le référentiel de l'éther, la vitesse de la lumière est égale à c quel que soit le mouvement de la source*. Il s'agit là d'une idée très naturelle, que l'on peut comprendre par une analogie acoustique : par rapport au talus d'une autoroute un jour sans vent, la vitesse

du son émis par une voiture est toujours la même, que la voiture s'approche ou s'éloigne de l'observateur. Une remarque due à de Sitter, datant de 1913¹, donne un argument en faveur de cette conjecture pour la lumière. Considérons une étoile double gravitant dans un plan passant par la Terre. Imaginons que l'éther n'existe pas, et supposons valable une « *théorie d'émission* » selon laquelle la vitesse de la lumière par rapport à un repère inertiel donné est égale à sa vitesse par rapport à l'étoile émettrice, superposée à celle de l'étoile par rapport au repère. Dans ce cas, un astronome terrestre observera des rayons lumineux se propageant plus ou moins rapidement, selon qu'ils ont été émis au moment où l'étoile s'approche ou s'éloigne de la Terre. Compte tenu de l'éloignement des étoiles, la durée du trajet est donc sensiblement plus courte ou plus longue, respectivement. Par conséquent, le mouvement de l'étoile, observé depuis la Terre, semblera irrégulier et plusieurs images de l'étoile pourraient même être observées simultanément en divers points de sa trajectoire. Or, ces phénomènes n'ont jamais été signalés.

Deux effets importants étaient ainsi expliqués par une combinaison de la vitesse de la lumière avec celle de la Terre par rapport à l'éther : l'aberration en astronomie, découverte par Bradley en 1725, et l'effet Doppler-Fizeau (1842, 1848).

Au cours du XIX^e siècle, diverses expériences tentèrent de mettre en évidence les propriétés physiques de l'éther. Dès 1849, Fizeau s'était rendu célèbre en parvenant à mesurer la vitesse de la lumière par des méthodes physiques, sans avoir recours aux observations astronomiques comme c'était l'usage jusque-là. En 1851, il se demanda si l'éther était entraîné par les corps en mouvement. Pour répondre à cette question, il fit passer un faisceau lumineux dans des canalisations remplies d'un courant d'eau, de vitesse égale à V par rapport au laboratoire. Utilisant un dispositif interférométrique, il constata que la vitesse de la lumière par rapport au laboratoire était modifiée de $0.44V$. Il en déduisit que l'éther était *partiellement entraîné* par l'eau en mouvement, partiellement seulement puisque le coefficient d'entraînement vaut 0.44 et pas 1. Ce constat s'accordait avec une prédiction théorique de Fresnel qui annonçait un coefficient d'entraînement égal à $1 - 1/n^2$, n étant l'indice de réfraction du milieu en mouvement.²

Il est, d'autre part, un référentiel naturel dont la vitesse dépasse de loin les ordres de grandeur que pouvait atteindre Fizeau avec son courant d'eau (en l'occurrence, quelques mètres par seconde) : c'est la Terre elle-même, gravitant autour du soleil à la vitesse de 30 km/s. De même que sur une autoroute par temps calme les passagers d'une voiture décapotable ressentent les effets d'un vent contraire, il doit souffler par rapport à notre planète un *vent d'éther*. Michelson (1881), puis Michelson et Morley (1887) se fixèrent l'objectif d'en déterminer la vitesse, également à l'aide d'un dispositif interférométrique. Le verdict de leurs expériences fut révolutionnaire : *la vitesse mesurée était nulle*.

Diverses tentatives d'explication de ce résultat ont été proposées. On pouvait, par exemple, postuler un *entraînement total de l'éther* par la Terre. Lorentz et FitzGerald suggérèrent une autre possibilité, non moins artificielle : c'était d'admettre une *contraction ad hoc du bras de l'interféromètre* dans la direction du mouvement terrestre par rapport à l'éther immuable.

A vrai dire, le résultat de Michelson et Morley n'excluait pas l'existence de l'éther mais il relançait crûment le défi d'en détecter expérimentalement le mouvement, car *dans les faits*

1. Donc postérieure à l'article original d'Einstein. Historiquement, l'intérêt de cette remarque fut d'argumenter non pas en faveur de l'éther, mais bien de l'une de ses conséquences, à savoir cette propriété d'invariance de c qu'avant Einstein on ne parvenait à justifier autrement.

2. Voir par exemple James Lequeux, *La lumière et l'éther au XIX^e siècle*,

<https://www.refletsdelaphysique.fr/articles/refdp/pdf/2016/01/refdp201647-48p28.pdf>

tout se passait comme si les référentiels inertiels se trouvaient tous au repos par rapport à lui. En 1902 Poincaré en vint à soupçonner qu'*aucune expérience de physique* (donc non seulement de *mécanique*, mais également d'*optique*, c'est-à-dire d'*électromagnétisme*) *ne pourrait jamais mettre en évidence autre chose que les déplacements relatifs des corps matériels*. Un tel bilan ruinait l'espoir de conférer un sens positif à la notion d'espace absolu, dont l'éther était censé représenter le dernier bastion.

Ainsi, à l'orée du XX^e siècle, hypothéqué de son statut de système d'inertie privilégié, greffé de propriétés contradictoires qui soulevaient autant de difficultés qu'elles n'en levaient, retranché dans le camp des *variables cachées*, l'éther n'offrait plus à la science qu'une figure désincarnée, ravalée par les railleurs au niveau d'une simple fonction grammaticale : « *sujet du verbe onduler* » !

Les choses en étaient à peu près là lorsqu'Einstein entra en scène. On peut difficilement évaluer l'impact précis de l'expérience de Michelson et Morley sur l'évolution de sa pensée, car dans la suite lui-même s'est contredit à ce propos³. Ce sont surtout des considérations générales d'élégance et de symétrie qu'il invoque prioritairement dans un mémoire pionnier soumis le 30 juin 1905 à la revue *Annalen der Physik*, et intitulé « *Zur Elektrodynamik bewegter Körper* »⁴. Quoi qu'il en soit, il fut amené à énoncer un principe universel : *tous les référentiels inertiels sont équivalents en ce qui concerne leurs propriétés physiques*. C'est le *principe de relativité*, également appelé *premier postulat de la théorie de la relativité restreinte*. Voici comment Einstein l'exprime dans son article :

« Dans tous les systèmes de coordonnées où les équations de la mécanique sont valables⁵, ce sont également les mêmes lois de l'optique et de l'électrodynamique qui sont valables. »

A ce stade cependant, il subsistait encore un choix quant à l'interprétation de ce principe : fallait-il penser que *dans n'importe quel référentiel inertiel* la vitesse de la lumière est égale à c quel que soit le mouvement de la source, ou bien cela n'est-il vrai que *dans le référentiel de l'émetteur seulement* ? La seconde possibilité, qui trouverait un cadre naturel dans une théorie balistique d'émission corpusculaire par exemple, stigmatise l'hypothèse de l'éther, laquelle impose que la vitesse de la lumière par rapport à la source varie selon le mouvement de celle-ci. La première possibilité est encore plus déconcertante, car alors la vitesse mesurée de la lumière se trouverait indépendante, non seulement du mouvement de la source par rapport à l'éther, mais également de celui de l'observateur. Elle obligerait à *sacrifier la transformation de Galilée au profit d'autres lois de transformation*. En effet, la loi classique d'addition des vitesses se révélerait incorrecte pour la lumière : un rayon lumineux possédant une vitesse égale à c par rapport à un train, se déplacerait toujours à la vitesse c par rapport au quai.

« Je savais », écrira Einstein à Ehrenfest le 20 juin 1912, « que le principe de la constance de la vitesse de la lumière était quelque chose d'indépendant du postulat de relativité et j'ai balancé pour savoir lequel était le plus vraisemblable, le principe de la constance de c requis par les équations de Maxwell ou la constance de c uniquement pour un observateur situé sur la source de lumière. J'ai opté en faveur du premier. »⁶

3. *Albert Einstein, œuvres choisies 2*, textes choisis et présentés par F. Balibar, O. Darrigol, B. Jech et J. Stachel, Seuil, 1993, p. 22, note 39.

4. « Sur l'électrodynamique des corps en mouvement ». Une traduction française annotée se trouve dans *Albert Einstein, œuvres choisies 2, op. cit.*, p. 31. Einstein y critique l'asymétrie de l'électrodynamique de Lorentz, qui interprète l'interaction entre un aimant et un conducteur différemment selon que c'est l'aimant ou le conducteur qui se déplace *absolument* (c'est-à-dire par rapport à l'éther), alors que la valeur numérique du courant induit dans le conducteur ne dépend que du mouvement *relatif* des deux corps.

5. C'est-à-dire les systèmes d'inertie.

6. *Albert Einstein, œuvres choisies 2, op. cit.*, note 54 p. 24.

C'est le *second postulat de la relativité restreinte*.⁷ Il est nécessaire pour assurer la compatibilité des équations de Maxwell avec la symétrie exigée par le principe de relativité, car la vitesse de la lumière c apparaît dans ces équations comme coefficient numérique.

Ayant posé ces deux postulats, Einstein conclut :

« L'introduction d'un « éther lumineux » se révélera superflue, dans la mesure où, dans la conception qui va être développée, il ne sera pas introduit d'« espace au repos absolu » doté de propriétés particulières. »

Comme le fait remarquer Jacques Merleau-Ponty, il faut insister sur le fait que

« ce n'est pas tant la réalité des mouvements de l'éther qui fait problème que le sens ou le non-sens de l'expression. »⁸

Cette idée a été très tôt présente à l'esprit d'Einstein, comme en témoigne cet extrait d'une lettre à Mileva Marić datée probablement du 10 août 1899 :

« L'introduction du mot « éther » dans les théories de l'électricité a conduit à l'idée d'un milieu du mouvement duquel il est possible de parler, sans que l'on soit capable, à mon avis, d'associer une signification physique précise à ce que l'on dit. »⁹

Dès lors qu'il s'apprête à porter un regard nouveau sur des notions aussi fondamentales que l'espace et le temps, l'appareil théorique doit intégrer un processus opérationnel méticuleux de *graduation des axes* dans un référentiel inertiel, compatible avec (voire suggéré par) l'invariance de la vitesse de la lumière. En particulier, la notion de *simultanéité* de deux événements doit faire l'objet d'une analyse attentive. Historiquement d'ailleurs, il semble bien que c'est la compréhension de ce point qui fit soudainement jaillir dans l'esprit d'Einstein l'étincelle indispensable pour édifier sa théorie. Les procédures habituellement utilisées sont les suivantes.

Supposons qu'un rayon lumineux effectue un aller-retour entre deux particules fixes dans un certain référentiel inertiel. La *distance* entre ces particules est alors définie par $c/2$ fois le temps nécessaire au trajet, mesuré par une horloge attachée à la particule émettrice. Et dans un référentiel inertiel donné, deux événements sont dits *simultanés* si deux rayons lumineux émis en ces deux événements se croisent à mi-chemin (les distances ayant été au préalable déterminées comme ci-dessus). On trouvera dans l'exercice 1.1 une interprétation géométrique de ces définitions.

La relativité restreinte résout de façon élégante les paradoxes où s'enlisait la physique classique, en déduisant notamment du principe d'invariance de la vitesse de la lumière une nouvelle loi de composition des vitesses. Elle reste cependant d'un abord délicat pour notre intuition newtonienne basée notamment sur le caractère absolu de la simultanéité. Plus loin, nous montrerons qu'une analyse attentive de certains paradoxes apparents, loin

7. Formellement, il est assez facile de rendre compte de l'indépendance logique du principe de relativité et du principe de la constance de la vitesse de la lumière. Si l'on n'utilise que le premier on peut démontrer que les coordonnées cartésiennes de deux référentiels inertiels sont reliées par

$$\underline{x} = \frac{x - Vt}{\sqrt{1 - V^2/\nu^2}}, \quad \underline{t} = \frac{t - Vx/\nu^2}{\sqrt{1 - V^2/\nu^2}}$$

où ν est un paramètre réel indépendant de la vitesse relative V . Ce résultat est valable aussi bien en mécanique classique que relativiste. Cependant, la mécanique classique suppose le caractère absolu du temps, ce qui revient à considérer que $\nu \rightarrow \infty$, tandis que le principe relativiste de constance de la vitesse de la lumière impose $\nu = c$. Voir par exemple W. Rindler, *Essential Relativity*, Springer-Verlag, 1977, pp. 31-33 et 51-53.

8. J. Merleau-Ponty, *Einstein*, Flammarion, 1993, p. 163.

9. *Albert Einstein, œuvres choisies 2, op. cit.*, p. 29 et note 7.

de déforçer cette théorie, révèle au contraire sa remarquable cohérence. Mais tout d'abord, nous allons examiner les premières retombées de l'invariance relativiste avec un minimum de considérations techniques.

1.2 La dilatation du temps

Dans un train \underline{S} en translation rectiligne et uniforme de vitesse V par rapport à un quai inertiel S , une source émet un rayonnement lumineux, lequel parcourt une distance verticale d , est réfléchi sur un miroir et absorbé à son point de départ (figure 1.1). Entre le

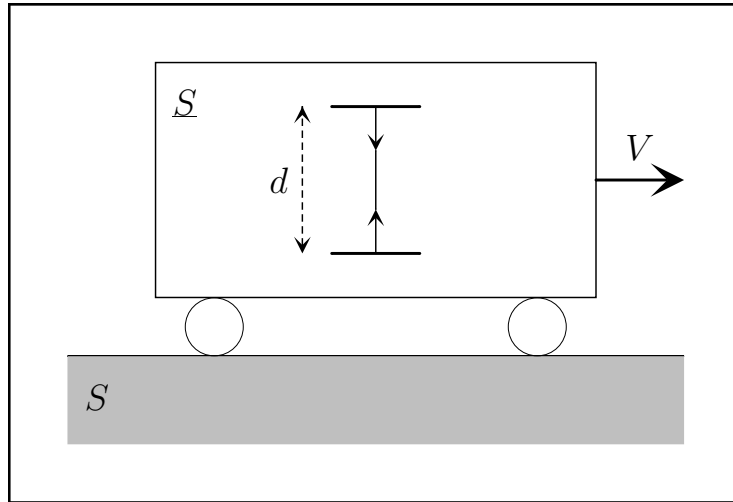


FIGURE 1.1: Dans \underline{S} , un rayon lumineux effectue un aller-retour en un temps $\Delta\underline{t} = 2d/c$.

début et la fin des opérations, une horloge \underline{H} solidaire de la source lumineuse mesure un intervalle de temps

$$\Delta\underline{t} = 2d/c. \quad (1.1)$$

Dans S , considérons deux horloges synchrones H_1 et H_2 , la première située au niveau de la source lorsqu'elle émet le signal, la seconde au niveau de la source lorsqu'elle reçoit le signal. Notons Δt le temps écoulé entre ces deux événements. La vitesse de la lumière étant égale à c aussi bien dans S que dans \underline{S} , on a

$$\Delta t = \frac{2}{c} \sqrt{d^2 + V^2 (\Delta t/2)^2} \quad \text{donc} \quad \Delta t = \frac{2d/c}{\sqrt{1 - V^2/c^2}}. \quad (1.2)$$

Comparant (1.1) et (1.2), on trouve

$$\Delta t = \frac{\Delta\underline{t}}{\sqrt{1 - V^2/c^2}}. \quad (1.3)$$

Cette relation exprime la *dilatation du temps*. Bien entendu, il s'agit là d'un effet inconnu de la physique prérelativiste. En effet, dans le cadre de la théorie de l'éther (supposé fixe par rapport à S), la vitesse de la lumière vaut c dans S et $\sqrt{c^2 - V^2}$ dans \underline{S} ; et dans le cadre d'une théorie d'émission, la vitesse de la lumière vaut c dans \underline{S} et $\sqrt{c^2 + V^2}$ dans S (loi galiléenne d'addition des vitesses). Dans les deux cas, on en déduit $\Delta\underline{t} = \Delta t$.

1.3 La contraction de l'espace

Soit une règle fixe dans un référentiel inertiel \underline{S} . Pour fixer les idées, imaginons que \underline{S} se déplace de la gauche vers la droite avec une vitesse V par rapport à un autre référentiel inertiel S . Afin de mesurer la longueur L de la règle dans S , on procède comme suit. Supposons qu'une horloge H immobile dans S voie passer l'extrémité droite de la règle en $t = 0$. Lorsque l'extrémité gauche de la règle passe à son tour devant H , celle-ci marque un temps Δt . Tout naturellement, on définit L par

$$L = V\Delta t. \quad (1.4)$$

Considérons à présent ce processus depuis \underline{S} . L'horloge H est maintenant en mouvement rectiligne et uniforme de la droite vers la gauche et de vitesse V . Supposons que l'extrémité gauche de la règle coïncide avec l'origine \underline{Q} des abscisses \underline{x} et qu'il s'y trouve une horloge \underline{H}_G , immobile dans \underline{S} . Supposons que l'extrémité droite de la règle se trouve en l'abscisse $\underline{x} = \underline{L}$ et qu'il s'y trouve une horloge \underline{H}_D synchronisée avec \underline{H}_G . Supposons enfin que \underline{H}_D marque $\underline{t} = 0$ quand H passe à son niveau et que \underline{H}_G marque $\Delta\underline{t}$ quand H passe à son niveau. La longueur \underline{L} de la règle, mesurée dans \underline{S} , est naturellement donnée par la vitesse de H fois son temps de survol :

$$\underline{L} = V\Delta\underline{t}. \quad (1.5)$$

Or, lorsque H passe devant l'extrémité gauche de la règle, elle marque le temps Δt , comme nous en avons convenu ci-dessus. Mais d'après la relation (1.3), on a

$$\Delta t = \Delta\underline{t}\sqrt{1 - V^2/c^2} \quad (1.6)$$

(attention à l'inversion des rôles joués par S et \underline{S}). Rassemblant les résultats (1.4), (1.5) et (1.6), on conclut

$$L = \underline{L}\sqrt{1 - V^2/c^2}. \quad (1.7)$$

C'est la *contraction des longueurs*, également appelée *contraction de l'espace*.

1.4 La relativité de la simultanéité

Considérons un train \underline{S} en translation rectiligne et uniforme de vitesse V (disons vers la droite) par rapport à un quai S supposé inertiel. En $\underline{t} = 0$, deux rayons lumineux sont émis simultanément dans \underline{S} depuis deux sources A et B , respectivement vers la droite et vers la gauche. Ces deux rayons parviennent donc simultanément en un récepteur M fixe dans \underline{S} , à mi-chemin entre A et B (figure 1.2).

Cependant, pour un observateur lié à S les émissions des deux rayons ne sont plus simultanées. Sinon, le rayon de droite atteindrait M *avant* le rayon de gauche car *sa vitesse est la même et il a moins de distance à parcourir* puisque M vient vers lui. Si l'horloge de S en face de A au moment de l'émission marque $t = 0$, il faut donc bien conclure que l'horloge de S située en face de B au moment de l'émission marque un temps $t^* > 0$: *la simultanéité est relative*.

Pour calculer t^* notons, d'une part, que $x_B = \underline{L}/\sqrt{1 - V^2/c^2}$ (contraction de l'espace, vue depuis \underline{S}) ; et que, d'autre part, $x_B = \underline{L}\sqrt{1 - V^2/c^2} + Vt^*$. On en déduit

$$t^* = \frac{V\underline{L}}{c^2\sqrt{1 - V^2/c^2}}. \quad (1.8)$$

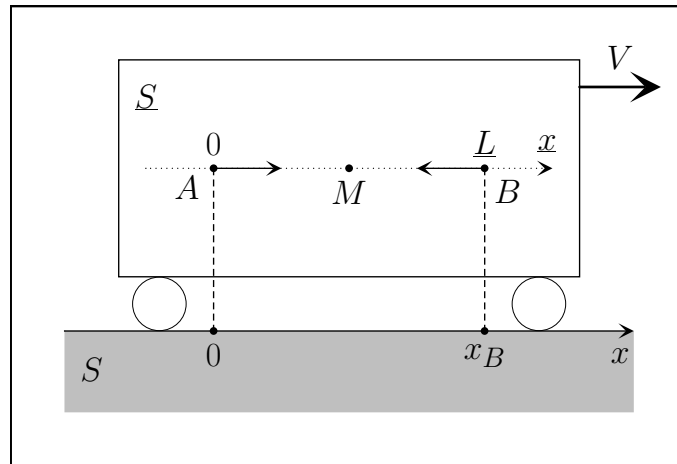


FIGURE 1.2: *La relativité de la simultanéité.* La source A émet un flash lumineux vers la droite en $ct = \underline{x} = x = 0$. La source B émet un flash lumineux vers la gauche en $ct = 0$, $\underline{x} = \underline{L}$. Cependant, du point de vue de S , ce dernier événement se produit en $ct^* \neq 0$.

1.5 Etude d'un premier paradoxe cinématique

Deux fusées F et \underline{F} , de longueur propre L chacune, évoluent dans des sens contraires sur deux droites parallèles de distance négligeable avec une vitesse relative constante V (figure 1.3).

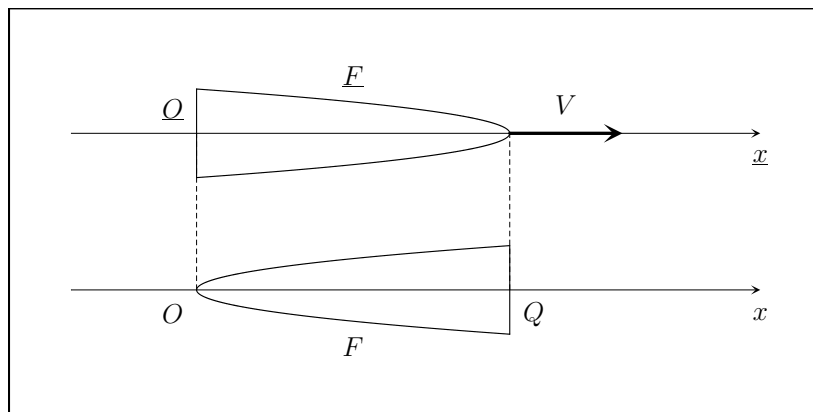


FIGURE 1.3: *La guerre spatiale.*

On peut supposer, sans perte de généralité, que l'origine des coordonnées x et \underline{x} dans les deux référentiels est située en O et en \underline{Q} respectivement et que l'horloge O de F marque $t = 0$ au moment où elle croise l'horloge \underline{Q} de \underline{F} , qui marque alors $\underline{t} = 0$.

En $t = 0$ un obus est émis depuis la queue Q de F , dans le but de détruire \underline{F} . Le point de vue de F pourrait être le suivant : « Nous allons manquer notre coup. En effet, à cause de la contraction des longueurs la fusée \underline{F} est rétrécie, sa longueur ne vaut plus que $L\sqrt{1 - V^2/c^2}$ ». Les occupants de \underline{F} pourraient, cependant, tenir des propos fort différents : « Nous allons être touchés car la fusée F , de longueur $L\sqrt{1 - V^2/c^2}$ à cause de la contraction des longueurs, est plus courte que la nôtre qui est de longueur L ».

Il en résulte un paradoxe : qui a raison ?¹⁰

Voici la réponse. En vertu de la contraction des longueurs l'abscisse de Q , mesurée dans \underline{F} en $\underline{t} = 0$, vaut effectivement $L\sqrt{1 - V^2/c^2}$. Cependant, la relativité de la simultanéité (1.8) nous enseigne que l'événement « Q émet un obus » se produit au temps

$$\underline{t}_Q = \frac{-VL}{c^2\sqrt{1 - V^2/c^2}}.$$

L'abscisse de cet événement mesurée dans \underline{F} , c'est-à-dire l'abscisse de Q en $\underline{t} = \underline{t}_Q$, vaut par conséquent

$$L\sqrt{1 - V^2/c^2} - V\underline{t}_Q = \frac{L}{\sqrt{1 - V^2/c^2}} > L.$$

C'est donc l'analyse de F qui est correcte : l'obus va manquer son but.

1.6 La transformation de Lorentz

Si l'on admet que la lumière est une onde se propageant à la vitesse c , alors l'invariance des lois de la physique, lors du passage d'un référentiel inertiel à un autre, implique l'invariance de l'opérateur de d'Alembert : si $\{ct, x, y, z\} = \{x^\alpha\}$ et $\{c\underline{t}, \underline{x}, \underline{y}, \underline{z}\} = \{\underline{x}^\alpha\}$ désignent des coordonnées cartésiennes dans deux référentiels inertiels S et \underline{S} , respectivement, on doit avoir

$$-\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = -\frac{1}{c^2}\frac{\partial^2}{\partial \underline{t}^2} + \frac{\partial^2}{\partial \underline{x}^2} + \frac{\partial^2}{\partial \underline{y}^2} + \frac{\partial^2}{\partial \underline{z}^2}. \quad (1.9)$$

Il est naturel de supposer que la loi de transformation des coordonnées est linéaire (car la lumière se propage en ligne droite dans tous les référentiels inertiels) et à coefficients constants (car l'espace-temps est homogène). Si les origines des deux systèmes d'axes coïncident, on peut ainsi postuler que

$$\vec{x} = \Lambda \vec{\underline{x}} \quad (1.10)$$

où Λ est une matrice 4×4 à éléments constants. Cette relation équivaut à

$$\vec{\nabla} = \tilde{\Lambda}^{-1} \vec{\underline{\nabla}}. \quad (1.11)$$

Soit la matrice η de composantes

$$(\eta_{\alpha\beta}) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1.12)$$

(représentation du *tenseur de Minkowski* en coordonnées cartésiennes). On peut écrire la relation (1.9) sous la forme

$$\vec{\nabla} \eta \vec{\nabla} = \vec{\underline{\nabla}} \eta \vec{\underline{\nabla}} \quad (1.13)$$

qui, compte tenu de (1.11), impose

$$\tilde{\Lambda} \eta \Lambda = \eta \quad (1.14)$$

10. E.F. Taylor et J.A. Wheeler, *Spacetime physics*, Freeman, 1992, pp. 79-80. On peut rendre la situation plus réaliste en supposant que les deux fusées évoluent sur des droites parallèles de distance d . Dans le référentiel de F l'obus, de vitesse U , est alors lancé perpendiculairement à l'axe des x depuis Q en $t = -d/U$. Un tel ajout ne contribuerait cependant qu'à voiler l'essence du problème.

donc, vu (1.10) :

$$\tilde{\vec{x}} \eta \vec{x} = \underline{\tilde{\vec{x}}} \eta \underline{\vec{x}}. \quad (1.15)$$

Ainsi, la conservation de l'opérateur de d'Alembert est équivalente à celle du « produit scalaire » associé au tenseur de Minkowski : deux quadrivecteurs \vec{A} et \vec{B} qui se transforment selon la loi (1.10) vérifient

$$\eta_{\alpha\beta} A^\alpha B^\beta = \eta_{\alpha\beta} \underline{A}^\alpha \underline{B}^\beta, \quad (1.16)$$

c'est-à-dire

$$-A^0 B^0 + A^1 B^1 + A^2 B^2 + A^3 B^3 = -\underline{A}^0 \underline{B}^0 + \underline{A}^1 \underline{B}^1 + \underline{A}^2 \underline{B}^2 + \underline{A}^3 \underline{B}^3.$$

On pose

$$\vec{A} \cdot \vec{B} = \eta_{\alpha\beta} A^\alpha B^\beta. \quad (1.17)$$

En particulier, l'élément de longueur $ds^2 = d\vec{x} \cdot d\vec{x} = \eta_{\alpha\beta} dx^\alpha dx^\beta$ est invariant :

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 = -c^2 d\underline{t}^2 + d\underline{x}^2 + d\underline{y}^2 + d\underline{z}^2 = d\underline{s}^2. \quad (1.18)$$

L'ensemble des matrices du type Λ définies par (1.14), doté de la multiplication matricielle, constitue le *groupe de Lorentz*. L'espace-temps muni du « produit scalaire » (1.17) s'appelle *espace-temps de Minkowski* (1907).

Étudions le cas particulier où les axes cartésiens spatiaux de S et \underline{S} sont parallèles entre eux et où la vitesse relative V des deux référentiels est parallèle à l'axe des x (donc des \underline{x}). C'est ce que l'on appelle, en littérature anglo-saxonne, un *boost* selon l'axe des x .

Comme les plans $y = 0$ et $\underline{y} = 0$ coïncident, on doit avoir $y = Ky$. Cette relation n'est pas affectée si l'on change le signe des axes x, z, \underline{x} et \underline{z} . Or, cette dernière opération revient à intervertir les rôles de S et \underline{S} : on peut donc également écrire $\underline{y} = Ky$, ce qui impose $K = \pm 1$. Par continuité, seul le signe positif doit être retenu. En tenant un raisonnement analogue pour l'axe des z , on obtient ainsi

$$y = \underline{y}, \quad z = \underline{z}.$$

La loi de transformation pour ct et x s'obtient dès lors facilement par la voie algébrique. Cependant, il est instructif d'interpréter la situation géométriquement, dans l'espace-temps de Minkowski à deux dimensions.

Dans S l'axe des ct a pour équation $x = Vt$, donc

$$\underline{x} = 0 \equiv ct = \frac{c}{V} x.$$

Toujours dans S , notons $ct = \xi x$ l'équation de l'axe des \underline{x} . Son vecteur directeur, de coordonnées $(\xi, 1)$, est orthogonal au vecteur directeur de l'axe des ct , de coordonnées $(c/V, 1)$. Par conséquent, $\xi = V/c$ et

$$ct = 0 \equiv ct = \frac{V}{c} x.$$

L'invariance de la vitesse de la lumière impose enfin

$$D \equiv ct = x \equiv ct = \underline{x},$$

ce qui permet de déterminer l'unité sur l'axe spatial à partir de l'unité sur l'axe temporel.

L'événement de coordonnées $(1, 0)$ dans \underline{S} admet, dans S , les coordonnées $(\alpha, V\alpha/c)$. L'invariance de l'élément de longueur impose

$$-1^2 + 0^2 = -\alpha^2 + (V\alpha/c)^2$$

donc

$$\alpha = \frac{1}{\sqrt{1 - V^2/c^2}} \quad (\text{dilatation du temps}).$$

L'événement de coordonnées $(0, 1)$ dans \underline{S} a pour coordonnées $(V\delta/c, \delta)$ dans S , donc

$$\delta = \frac{1}{\sqrt{1 - V^2/c^2}}.$$

Comme la droite $D_1 \equiv ct = c(x - \gamma)/V$ contient le point $(ct = V\delta/c, x = \delta)$, on a

$$\gamma = \delta(1 - V^2/c^2) = \sqrt{1 - V^2/c^2} \quad (\text{contraction des longueurs}).$$

De façon analogue, la droite $D_2 \equiv ct - \beta = Vx/c$ contient le point $(ct = \alpha, x = V\alpha/c)$, d'où

$$\beta = \alpha(1 - V^2/c^2) = \sqrt{1 - V^2/c^2}.$$

Ces résultats sont illustrés dans la figure 1.4 où, conformément à la tradition, on a porté les axes du temps en ordonnée.

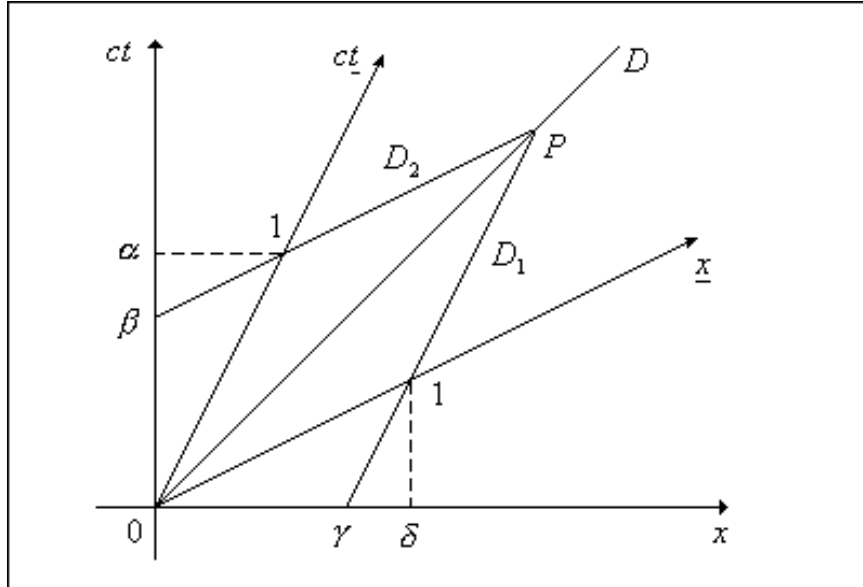


FIGURE 1.4: P est le point de coordonnées $(1, 1)$ dans \underline{S} . La droite D_1 est la parallèle à l'axe des ct' passant par P . Donc $D_1 \equiv \underline{x} = 1 \equiv ct = c(x - \gamma)/V$. De façon similaire, la droite D_2 , parallèle à l'axe des x' et passant par P , vérifie $D_2 \equiv ct' = 1 \equiv ct - \beta = Vx/c$.

Une simple généralisation permet alors de déduire la figure 1.5. On a

$$D_3 \equiv ct = \frac{c}{V} \left(x - \underline{X} \sqrt{1 - V^2/c^2} \right)$$

$$D_4 \equiv ct - c\underline{T} \sqrt{1 - V^2/c^2} = \frac{V}{c} x.$$

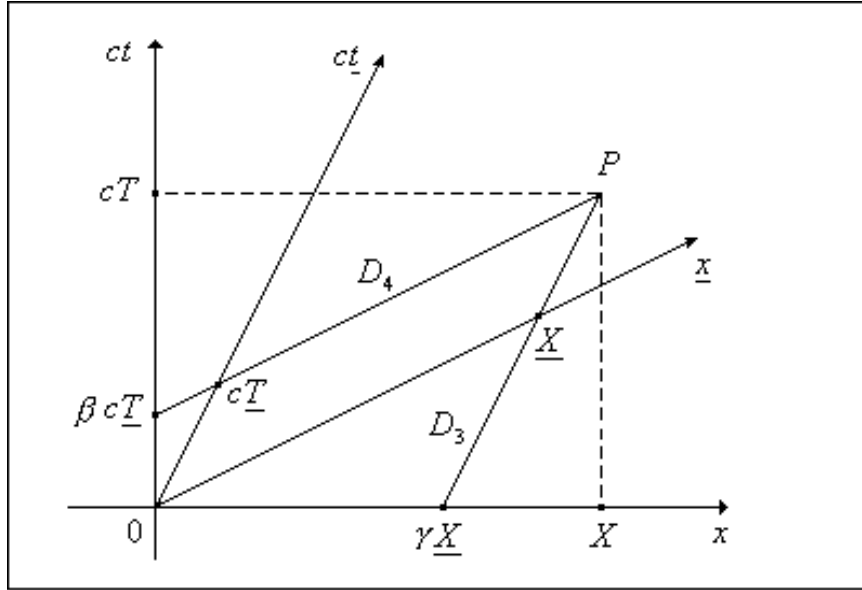


FIGURE 1.5: *Interprétation géométrique de la transformation de H.A. Lorentz. P est un point de coordonnées $(c\underline{T}, \underline{X})$ dans \underline{S} et de coordonnées (cT, X) dans S . Les droites D_3 et D_4 , contenant P , sont respectivement parallèles à l'axe des $c\underline{t}$ et des \underline{x} .*

Ces deux droites se coupent en P de coordonnées (cT, X) dans S : on en déduit la *transformation de Lorentz* dans le cas particulier d'un « boost » :

$$\underline{T} = \frac{T - VX/c^2}{\sqrt{1 - V^2/c^2}}, \quad \underline{X} = \frac{X - VT}{\sqrt{1 - V^2/c^2}}. \quad (1.19)$$

Posant

$$V = c \tanh \theta, \quad \theta \in] -\infty, +\infty[,$$

on constate que les relations (1.19) constituent une *rotation hyperbolique* dans l'espace-temps de Minkowski :

$$\begin{pmatrix} c\underline{T} \\ \underline{X} \end{pmatrix} = \begin{pmatrix} \cosh \theta & -\sinh \theta \\ -\sinh \theta & \cosh \theta \end{pmatrix} \begin{pmatrix} cT \\ X \end{pmatrix}.$$

Elles traduisent ainsi une *symétrie de l'espace-temps* (isotropie spatio-temporelle), au même titre que les rotations traditionnelles expriment l'isotropie des sections spatiales.

La composition de deux « boosts » de paramètres respectifs $V_1 = c \tanh \theta_1$ et $V_2 = c \tanh \theta_2$ donne un « boost » de paramètre $V = c \tanh \theta$ avec la *loi d'addition des vitesses*

$$\theta = \theta_1 + \theta_2,$$

c'est-à-dire

$$V = \frac{V_1 + V_2}{1 + V_1 V_2 / c^2}.$$

Exercice 1.1 *Montrer que la construction des axes et de leurs graduations est bien compatible avec la méthode décrite à la fin de la section 1.1 pour définir les distances et synchroniser les horloges.*

Solution. Cela apparaît immédiatement dans le système S , donc cela doit être vrai dans n'importe quel référentiel inertiel. Les figures 1.6 et 1.7 interprètent la situation dans \underline{S} .

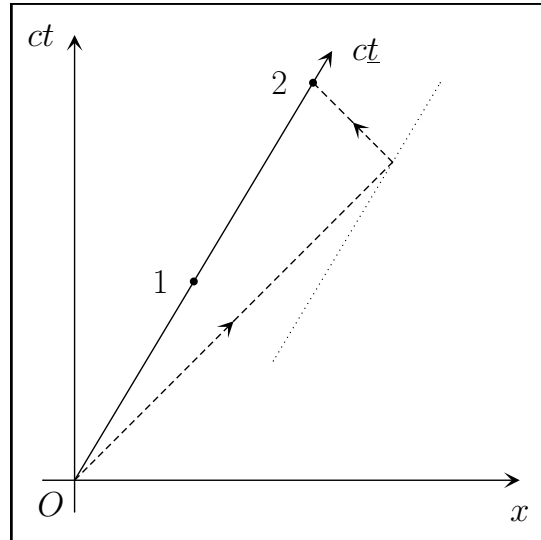


FIGURE 1.6: En pointillés on a représenté la ligne d'univers, parallèle à l'axe des ct , d'une particule dont on cherche à déterminer l'abscisse dans un référentiel inertiel \underline{S} . Pour cela, on lui fait réfléchir un rayon lumineux émis en $ct = 0$ depuis l'origine de \underline{S} (trajectoire en tirets). Si le rayon regagne l'origine en $ct = 2$, on décide que la particule est située à un mètre de l'origine de \underline{S} . On vérifiera que la ligne d'univers de la particule intersecte bien l'axe des x , tel que déterminé dans la figure 1.4, au point d'abscisse $x = 1$.

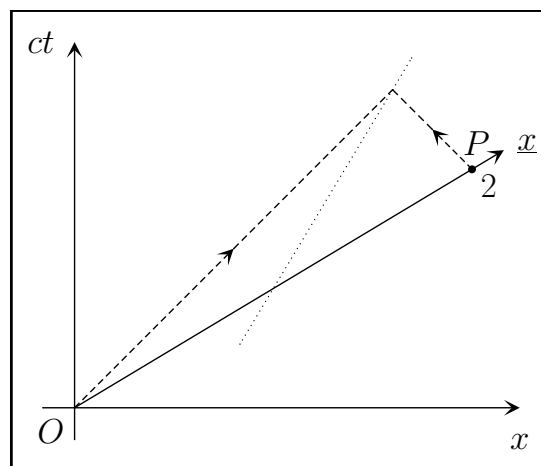


FIGURE 1.7: L'événement P est situé à deux mètres de l'origine dans \underline{S} (la mesure ayant été effectuée au préalable selon les modalités de la figure 1.6). Il est simultané à l'événement O dans \underline{S} car deux rayons lumineux émis respectivement par O et par P se rencontrent à mi-chemin dans \underline{S} , en l'occurrence sur la droite en pointillés, laquelle est identique à celle de la figure 1.6. La coordonnée temporelle de P est donc $ct = 0$. On vérifiera que P appartient, comme il se doit, à l'axe des x tel que déterminé dans la figure 1.4.

1.7 Etude d'un second paradoxe cinématique

Une poutre de longueur propre L se déplace avec une vitesse constante V parallèlement à une grange dont la porte a également une longueur propre L . Le point de vue de la grange pourrait être le suivant : « La poutre a une longueur $L\sqrt{1-V^2/c^2}$, elle pourrait donc entrer dans la grange ». Mais la poutre serait en droit d'affirmer : « La porte de la grange a une longueur $L\sqrt{1-V^2/c^2}$, la poutre ne pourrait donc pas entrer dans la grange ».

D'où provient la contradiction apparente ?¹¹

En physique relativiste il faut bien se mettre d'accord sur ce dont on parle avant de tirer des conclusions paradoxales. La principale difficulté provient, nous l'avons déjà suggéré, du fait que *la simultanéité est relative*. Lorsqu'un observateur lié à la poutre affirme que la poutre ne pourrait pas entrer dans la grange, qu'entend-il exactement par « la grange » ? Il imagine implicitement que, face aux diverses particules constituant la poutre à un instant donné, se trouvent les particules constituant la grange, également toutes à un instant donné. Mais cela est une lecture *galiléenne* du monde, qui n'induit d'ailleurs guère de contradiction puisque la transformation de Galilée n'implique pas de contraction des longueurs. En fait, la théorie de la relativité nous enseigne que les diverses particules constituant la grange sont alors envisagées à des *temps propres différents* : l'expression « la grange » désigne une juxtaposition de photos montrant les particules de la grange à des instants différents.

On peut transposer ces considérations au point de vue de la grange.

En conclusion, le paradoxe résulte d'un malentendu au niveau du vocabulaire : *les deux protagonistes ne parlent pas des mêmes événements*.

La contraction des longueurs résulte des propriétés de l'espace-temps lui-même (par l'intermédiaire du groupe de transformations préservant sa métrique, ici le groupe de Lorentz) et pas de la structure physico-chimique des particules. C'est en quelque sorte un effet de perspective lié à un découpage arbitraire de l'espace-temps, de la même façon que deux règles de longueur unité mais non parallèles se voient mutuellement plus petites.

On est cependant en droit de se demander dans quelle mesure cet effet pourrait être exploité pour *effectivement* faire entrer la poutre dans la grange. Pour cela il faut généraliser quelque peu les données du problème.

Soient $\underline{S}(ct, \underline{x}, y)$ et $S'(ct', x', y')$ les référentiels inertiels attachés respectivement à la poutre et à la grange. Appelons $S(ct, x, y)$ un référentiel inertiel intermédiaire où la poutre se déplace à vitesse V constante dans la direction des x croissants tandis que la grange se meut uniformément à vitesse U dans la direction des y décroissants. Supposons enfin que en $\underline{t} = t' = 0$ le milieu \underline{Q} de la poutre et le milieu O' de la porte de la grange coïncident en l'événement de coordonnées $ct = x = y = 0$. La situation, envisagée depuis S , est représentée à la figure 1.8.

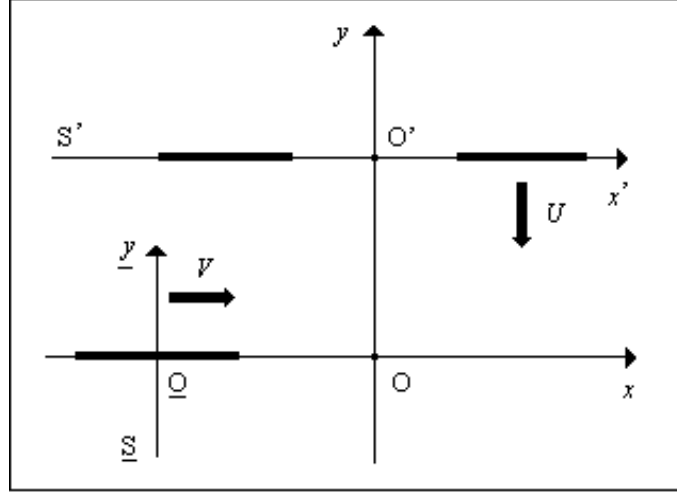
Analysons cette même situation du point de vue de \underline{S} . Posant

$$V = c \tanh \theta_1, \quad U = c \tanh \theta_2,$$

on a, d'après les transformations de Lorentz :

$$\begin{pmatrix} ct \\ \underline{x} \\ y \end{pmatrix} = \begin{pmatrix} \cosh \theta_1 & -\sinh \theta_1 & 0 \\ -\sinh \theta_1 & \cosh \theta_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \end{pmatrix}$$

11. R. Shaw, Am. J. Phys. **30**, 72 (1962).

FIGURE 1.8: La grange (référentiel S') et la poutre (référentiel \underline{S}).

$$\begin{pmatrix} ct \\ x \\ y \end{pmatrix} = \begin{pmatrix} \cosh \theta_2 & 0 & -\sinh \theta_2 \\ 0 & 1 & 0 \\ -\sinh \theta_2 & 0 & \cosh \theta_2 \end{pmatrix} \begin{pmatrix} ct' \\ x' \\ y' \end{pmatrix}$$

donc

$$\begin{pmatrix} ct \\ \underline{x} \\ \underline{y} \end{pmatrix} = \begin{pmatrix} \cosh \theta_1 \cosh \theta_2 & -\sinh \theta_1 & -\cosh \theta_1 \sinh \theta_2 \\ -\sinh \theta_1 \cosh \theta_2 & \cosh \theta_1 & \sinh \theta_1 \sinh \theta_2 \\ -\sinh \theta_2 & 0 & \cosh \theta_2 \end{pmatrix} \begin{pmatrix} ct' \\ x' \\ y' \end{pmatrix}.$$

L'équation, dans l'espace-temps, de la porte de la grange est donnée par $y' = 0$, soit

$$ct = \cosh \theta_1 \cosh \theta_2 ct' - \sinh \theta_1 x' \quad (1.20)$$

$$\underline{x} = -\sinh \theta_1 \cosh \theta_2 ct' + \cosh \theta_1 x' \quad (1.21)$$

$$\underline{y} = -\sinh \theta_2 ct'. \quad (1.22)$$

L'équation (1.20) permet d'éliminer ct' au profit de ct et x' :

$$\underline{x} = -\tanh \theta_1 ct + \frac{x'}{\cosh \theta_1} \quad (1.23)$$

$$\underline{y} = -\frac{\tanh \theta_2}{\cosh \theta_1} ct - \tanh \theta_1 \tanh \theta_2 x' \quad (1.24)$$

Au temps $t = 0$, les extrémités ($x' = \mp L/2, y' = 0$) de la porte se trouvent aux points A et B (voir figure 1.9) de coordonnées

$$(\underline{x}_A, \underline{y}_A) = -(\underline{x}_B, \underline{y}_B) = \left(\frac{-L}{2 \cosh \theta_1}, \frac{L \tanh \theta_1 \tanh \theta_2}{2} \right). \quad (1.25)$$

La largeur \underline{L} de l'ouverture, mesurée à un instant t donné dans le système \underline{S} de la poutre, est donc, comme annoncé, inférieure à L :

$$\underline{L} = L \sqrt{\frac{1}{\cosh^2 \theta_1} + \tanh^2 \theta_1 \tanh^2 \theta_2} = L \sqrt{1 - \frac{V^2}{c^2} \left(1 - \frac{U^2}{c^2}\right)} < L. \quad (1.26)$$

Cependant, en éliminant x' dans les équations (1.23) et (1.24), nous remarquons que la droite $y' = 0$ a pour équation dans le système \underline{S} :

$$\underline{y} = -\cosh \theta_1 \tanh \theta_2 \underline{ct} - \sinh \theta_1 \tanh \theta_2 \underline{x}. \quad (1.27)$$

Dans \underline{S} à un instant \underline{t} donné, la grange est ainsi *incliné* par rapport à l'axe horizontal \underline{x} d'un angle

$$\varphi_1 = \arctan(-\sinh \theta_1 \tanh \theta_2). \quad (1.28)$$

Enfin, en éliminant \underline{t} dans les équations (1.23) et (1.24), on trouve

$$\underline{y} = \frac{\tanh \theta_2}{\sinh \theta_1} \underline{x} - \frac{\tanh \theta_2}{\tanh \theta_1} x'. \quad (1.29)$$

Donc, les extrémités ($x' = \mp L/2, y' = 0$) de la porte parcourent les droites d'équation $\underline{y} = a \underline{x} \pm b$, où

$$a = \frac{\tanh \theta_2}{\sinh \theta_1}, \quad b = \frac{L}{2} \frac{\tanh \theta_2}{\tanh \theta_1}. \quad (1.30)$$

Elles sont inclinées d'un angle

$$\varphi_2 = \arctan\left(\frac{\tanh \theta_2}{\sinh \theta_1}\right) \quad (1.31)$$

par rapport à l'axe des \underline{x} qu'elles coupent au point d'abscisse

$$\mp \frac{b}{a} = \mp \frac{L}{2} \cosh \theta_1. \quad (1.32)$$

Le paradoxe est ainsi élucidé, comme l'illustre la figure 1.9.

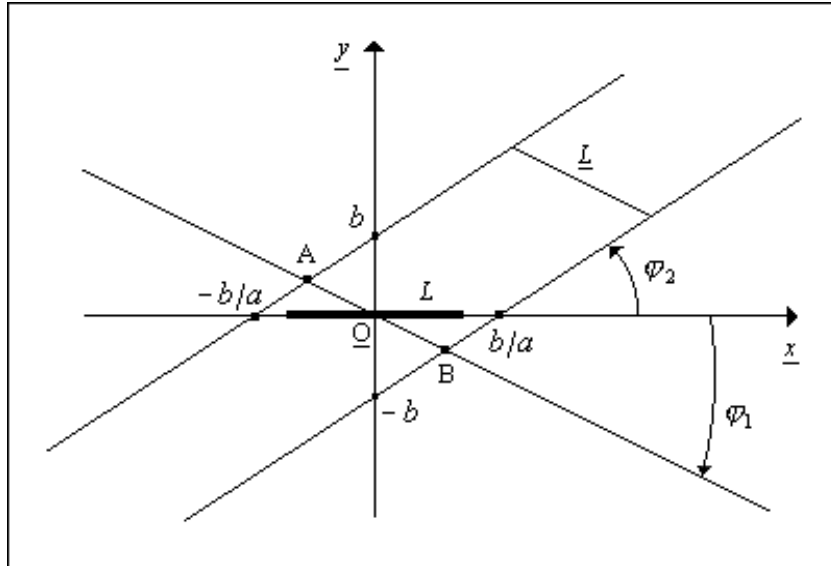


FIGURE 1.9: Projection de la situation dans un plan $\underline{t} = \text{constante}$ du référentiel \underline{S} de la poutre de longueur L . En vertu de la contraction des longueurs, la porte a une largeur $\underline{L} < L$; mais la valeur des angles φ_1 et φ_2 est telle que la poutre peut effectivement entrer dans la grange.

1.8 Quadrivecteur énergie-impulsion d'une particule

Dans l'espace-temps de Minkowski rapporté aux coordonnées cartésiennes $\{ct, x, y, z\} = \{x^\alpha\}$ et de métrique

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

considérons une particule dont la vitesse \vec{V} , de composantes $V^i = dx^i/dt$,¹² a pour norme $V = \sqrt{(V^1)^2 + (V^2)^2 + (V^3)^2}$. Observons tout d'abord que, le long de la trajectoire (parfois appelée *ligne d'univers*) de cette particule, on a

$$ds^2 = -c^2 d\tau^2, \quad (1.33)$$

où

$$d\tau = dt \sqrt{1 - V^2/c^2} \quad (1.34)$$

est le temps propre (dont vieillit la particule) correspondant à un intervalle de temps-coordonnée dt .¹³

La quantité \vec{V} ne constitue un vecteur que pour le groupe des transformations préservant les sections spatiales, euclidiennes, de l'espace-temps de Minkowski. Cependant, on peut définir un quadrivecteur apparenté, la *quadrivitesse* \vec{U} , de composantes $U^\alpha = dx^\alpha/d\tau$, c'est-à-dire, explicitement,

$$(U^\alpha) = \frac{1}{\sqrt{1 - V^2/c^2}} (c, \vec{V}).$$

Remarquons que

$$\vec{U} \cdot \vec{U} = \eta_{\alpha\beta} U^\alpha U^\beta = -c^2, \quad (1.35)$$

relation équivalente à (1.33) et qu'il faut lire comme suit : *dans l'espace-temps, toute particule matérielle se meut à la vitesse de la lumière.*

Notant m_0 la *masse propre* de la particule, on définit son *quadrivecteur énergie-impulsion* par

$$\vec{p} = m_0 \vec{U}.$$

Le carré de la norme de \vec{p} vaut

$$\vec{p} \cdot \vec{p} = (m_0)^2 \vec{U} \cdot \vec{U} = - (m_0)^2 c^2$$

et se trouve fondamentalement relié à la masse propre de la particule.

Il est naturel de fonder la dynamique relativiste sur l'équation quadrivectorielle

$$\vec{K} = \frac{d\vec{p}}{d\tau} \quad (1.36)$$

où \vec{K} est un quadrivecteur force que seule l'expérience, évidemment, peut avaliser. Cette équation se projette de la manière suivante :

$$\sum F^i V^i = \frac{d}{dt} (mc^2) \quad (\text{indice } 0) \quad (1.37)$$

$$\vec{F} = \frac{d}{dt} (m\vec{V}) \quad (\text{indices } 1, 2, 3) \quad (1.38)$$

12. Par convention, les indices latins désigneront les coordonnées spatiales 1, 2 et 3 tandis que les indices grecs courront sur les coordonnées spatio-temporelles 0, 1, 2 et 3.

13. L'hypothèse selon laquelle l'écoulement du temps propre ne dépend que de la vitesse (et pas des dérivées d'ordre supérieur, comme on serait en droit de l'imaginer *a priori*), est au cœur de la théorie de la relativité. Elle est justifiée par l'invariance lorentzienne de la quantité $dt \sqrt{1 - V^2/c^2}$ et le fait que, en vertu du principe de relativité, l'intervalle de temps propre doit apparaître comme un invariant formel.

avec ¹⁴

$$F^i = K^i \sqrt{1 - V^2/c^2} \quad \text{et} \quad m = \frac{m_0}{\sqrt{1 - V^2/c^2}}. \quad (1.39)$$

La puissance des forces est égale à la dérivée temporelle de l'énergie cinétique T . Or, par définition, celle-ci doit s'annuler lorsque la particule est au repos. D'après (1.37), on en déduit

$$\begin{aligned} T &= (m - m_0) c^2 \\ &= \frac{1}{2} m_0 V^2 + O(1/c^2). \end{aligned} \quad (1.40)$$

C'est l'un des succès de la relativité restreinte que d'avoir pu réunir en un seul quadrivecteur \vec{p} l'information relative, d'une part à l'énergie de la particule, qui n'est pas un scalaire pour le groupe des transformations de Galilée, et d'autre part à sa quantité de mouvement, qui n'est pas davantage un vecteur pour ce groupe (figure 1.10).

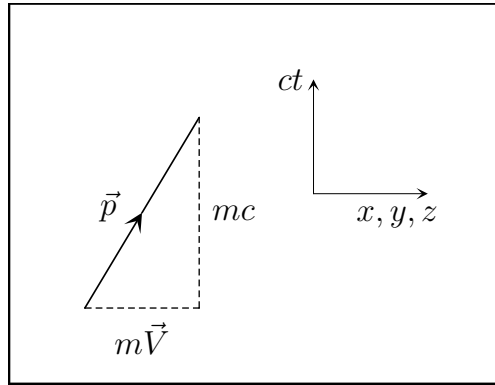


FIGURE 1.10: *Quadrivecteur énergie-impulsion d'une particule.*

On postule que, lors d'un processus de collision ponctuelle (instantanée), *le quadrivecteur \vec{p} total est conservé.*¹⁵ Cela vaut aussi bien pour les collisions élastiques (sans modification de la structure interne des particules) que pour les collisions inélastiques. Cette hypothèse, au demeurant largement vérifiée par l'expérience, est suggérée par sa limite en mécanique classique : si l'on néglige les termes du second ordre en V/c , elle conduit en effet à la conservation de la masse totale $\sum m_{0i}$ (composante temporelle) et de la quantité de mouvement totale $\sum m_{0i} \vec{V}_i$ (composantes spatiales). Remarquons d'ailleurs que, du fait de l'invariance relativiste, la conservation des composantes spatiales implique celle de la composante temporelle. En effet, désignant par $\Delta \vec{p}$ la variation du quadrivecteur au cours d'une collision, on a

$$\Delta \underline{p}^\alpha = \Lambda^\alpha_\beta \Delta p^\beta,$$

où Λ désigne génériquement une matrice régissant la transformation de Lorentz. Si donc $\Delta p^i = \Delta \underline{p}^i = 0$, on a

$$0 = \Delta \underline{p}^i = \Lambda^i_0 \Delta p^0,$$

ce qui entraîne l'annulation de Δp^0 .

14. On utilise souvent le vocable « masse relativiste » pour désigner la quantité m . Einstein n'appréciait guère cet usage, probablement parce que m n'est pas un invariant. Voir *Albert Einstein, œuvres choisies 2, op. cit.*, p. 69.

15. Voir par exemple Bernard F. Schutz, *A first course in general relativity*, Cambridge University Press, 2009, p. 43 ; Ericourgoulhon, *Relativité restreinte*, EDP Sciences/CNRS Editions, 2010, chapitre 9.

1.9 Le photon

Bien qu'auteur d'un article sur les quanta lumineux envoyé dès le mois de mars 1905 à la revue *Annalen der Physik*, Einstein ne fait guère allusion à cette hypothèse dans son article sur la relativité de juin (où il parle plutôt de *Lichtkomplex*), probablement pour ne pas donner l'illusion que ses conclusions en dépendent. Or, le concept de *photon* est entièrement compatible avec la théorie de la relativité ; on pourrait même affirmer qu'il est justifié par cette dernière. Afin de le montrer, nous allons reprendre ici les différents points de l'argumentation exposée dans l'article du mois de juin sur l'électrodynamique des corps en mouvement.

a) Etablissons la loi de transformation de la fréquence et de la direction d'une onde électromagnétique plane de pulsation ω et de normale unitaire $(\vec{n}) = (\cos \varphi, \sin \varphi, 0)$, lors d'une transformation de Lorentz de vitesse V le long de l'axe des x (c'est le cas le plus général ; voir figure 1.11).

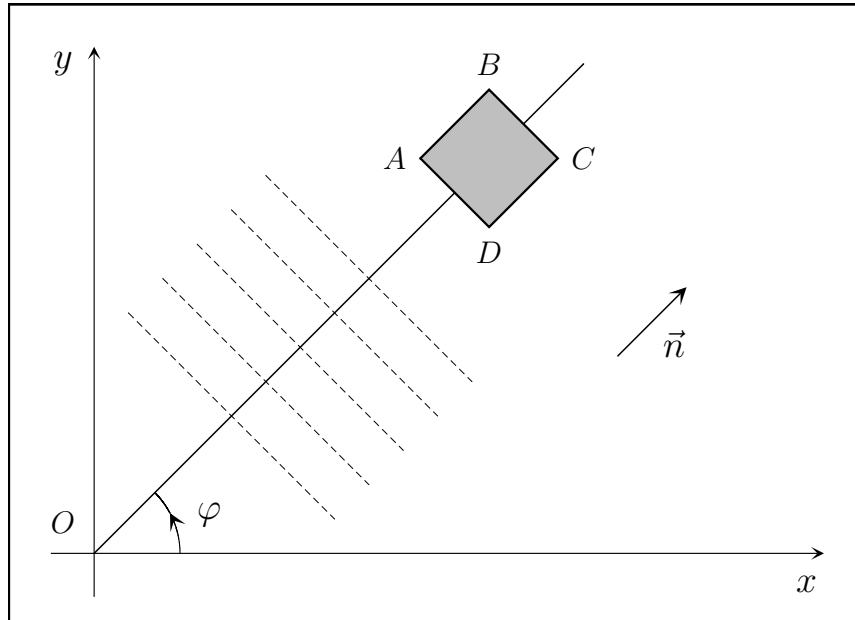


FIGURE 1.11: Onde électromagnétique plane de normale unitaire \vec{n} située dans le plan (x, y) . En tirets, les fronts d'onde.

Cette onde consiste en un champ électrique \vec{E} et un champ magnétique \vec{B} . Comme la polarisation de la lumière n'a guère d'importance dans ce contexte, nous les supposons donnés par

$$(\vec{E}) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mathcal{A} \sin [\omega (t - \vec{n} \cdot \vec{x}/c)], \quad (c\vec{B}) = \begin{pmatrix} \sin \varphi \\ -\cos \varphi \\ 0 \end{pmatrix} \mathcal{A} \sin [\omega (t - \vec{n} \cdot \vec{x}/c)]$$

où l'amplitude \mathcal{A} est une constante. La transformation de Lorentz

$$t = \gamma (\underline{t} + V\underline{x}/c^2), \quad x = \gamma (\underline{x} + V\underline{t}), \quad y = \underline{y},$$

où

$$\gamma = \frac{1}{\sqrt{1 - V^2/c^2}},$$

permet d'écrire

$$\omega(t - \vec{n} \cdot \vec{x}/c) = \underline{\omega}(\underline{t} - \underline{\vec{n}} \cdot \underline{\vec{x}}/c)$$

avec la loi de transformation de la fréquence (*effet Doppler-Fizeau*)

$$\underline{\omega} = \omega \frac{1 - V \cos \varphi/c}{\sqrt{1 - V^2/c^2}} \quad (1.41)$$

et la loi relativiste de l'*aberration* ($\underline{\vec{n}} = (\cos \underline{\varphi}, \sin \underline{\varphi}, 0)$) où

$$\cos \underline{\varphi} = \frac{\cos \varphi - V/c}{1 - V \cos \varphi/c} \quad (1.42)$$

$$\sin \underline{\varphi} = \frac{\sin \varphi \sqrt{1 - V^2/c^2}}{1 - V \cos \varphi/c}. \quad (1.43)$$

b) Pour établir la loi de transformation de l'énergie volumique du champ électromagnétique

$$W = \frac{\varepsilon_0}{2} \left(|\vec{E}|^2 + |c\vec{B}|^2 \right) = \varepsilon_0 \mathcal{A}^2 \sin^2 [\omega(t - \vec{n} \cdot \vec{x}/c)],$$

il suffit de savoir que le potentiel scalaire Φ (divisé par c) et le potentiel vecteur \vec{A} s'associent pour constituer un quadrivecteur dans l'espace-temps de Minkowski. (Fondamentalement, c'est grâce à cette circonstance que les équations de Maxwell sont covariantes pour la transformation de Lorentz : on trouvera les détails dans l'annexe A.1.) On a donc

$$\begin{pmatrix} \Phi/c \\ \underline{A}^1 \\ \underline{A}^2 \\ \underline{A}^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma V/c & 0 & 0 \\ -\gamma V/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Phi/c \\ A^1 \\ A^2 \\ A^3 \end{pmatrix}.$$

Le champ électrique et le champ magnétique dans le nouveau système se déduisent des relations

$$\vec{E} = -\vec{\nabla} \Phi - \frac{\partial \vec{A}}{\partial t}, \quad \vec{B} = \vec{\nabla} \wedge \vec{A}.$$

On obtient

$$\begin{aligned} (\vec{E}) &= \begin{pmatrix} E_1 \\ \gamma(E_2 - VB_3) \\ \gamma(E_3 + VB_2) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \gamma(1 - V \cos \varphi/c) \end{pmatrix} \mathcal{A} \sin [\underline{\omega}(\underline{t} - \underline{\vec{n}} \cdot \underline{\vec{x}}/c)] \\ (c\vec{B}) &= \begin{pmatrix} cB_1 \\ \gamma(V E_3/c + cB_2) \\ \gamma(-V E_2/c + cB_3) \end{pmatrix} = \begin{pmatrix} \sin \varphi \\ \gamma(V/c - \cos \varphi) \\ 0 \end{pmatrix} \mathcal{A} \sin [\underline{\omega}(\underline{t} - \underline{\vec{n}} \cdot \underline{\vec{x}}/c)], \end{aligned}$$

dont il résulte

$$\underline{W} = W \frac{(1 - V \cos \varphi/c)^2}{1 - V^2/c^2}.$$

c) Calculons enfin la loi de transformation de la mesure d'un volume se déplaçant avec l'onde, donc à la vitesse de la lumière. Pour fixer les idées, nous envisagerons un cube de côté 1 (voir figure 1.11). L'équation de mouvement du point A est

$$\begin{aligned} x_A &= ct \cos \varphi - 1/2 \cos \varphi - 1/2 \sin \varphi \\ y_A &= ct \sin \varphi - 1/2 \sin \varphi + 1/2 \cos \varphi. \end{aligned}$$

Pour étudier la valeur du volume dans le nouveau référentiel, il suffit d'envisager les coordonnées du cube en $\underline{t} = 0$, puisque les phénomènes sont stationnaires, ce qui revient à poser

$$t = \gamma V \underline{x}/c^2, \quad x = \gamma \underline{x}, \quad y = \underline{y}.$$

On obtient

$$\begin{aligned} \underline{x}_A &= \frac{-\cos \varphi - \sin \varphi}{2\gamma(1 - V \cos \varphi/c)} \\ \underline{y}_A &= \frac{-V/c + \cos \varphi - \sin \varphi}{2(1 - V \cos \varphi/c)}. \end{aligned}$$

Procédant de même pour les points B , C et D , on trouve, dans le nouveau référentiel en $\underline{t} = 0$:

$$\begin{aligned} \underline{x}_B &= \frac{\cos \varphi - \sin \varphi}{2\gamma(1 - V \cos \varphi/c)}, & \underline{x}_C &= -\underline{x}_A, & \underline{x}_D &= -\underline{x}_B \\ \underline{y}_B &= \frac{-V/c + \cos \varphi + \sin \varphi}{2(1 - V \cos \varphi/c)}, & \underline{y}_C &= -\underline{y}_A, & \underline{y}_D &= -\underline{y}_B. \end{aligned}$$

On constate que le côté déterminé par les sommets A et D est orthogonal au vecteur normal \underline{n} (il en va de même pour le côté déterminé par B et C); mais le parallépipède n'est plus rectangle. La mesure de son volume est donnée par

$$\left| \left(\underline{x}_B - \underline{x}_A, \underline{y}_B - \underline{y}_A, 0 \right) \wedge \left(\underline{x}_D - \underline{x}_A, \underline{y}_D - \underline{y}_A, 0 \right) \right| = \frac{\sqrt{1 - V^2/c^2}}{1 - V \cos \varphi/c}.$$

Donc, d'une façon générale, les mesures d'un volume quelconque étudié dans les deux référentiels inertiels, respectivement notées Γ et $\underline{\Gamma}$, sont reliées entre elles par

$$\underline{\Gamma} = \Gamma \frac{\sqrt{1 - V^2/c^2}}{1 - V \cos \varphi/c}.$$

Remarquons que le volume se transforme comme la longueur d'onde [vu l'équation (1.41)].

d) Aussi naïf soit-il, le meilleur modèle classique que l'on puisse se donner pour le photon est celui d'une quantité d'énergie électromagnétique E contenue dans un volume de mesure Γ donnée se déplaçant à la vitesse de la lumière. Des trois points précédents, on déduit que cette énergie $E = W \Gamma$ se transforme comme la fréquence (1.41) du champ :

$$\underline{E} = E \frac{1 - V \cos \varphi/c}{\sqrt{1 - V^2/c^2}}.$$

La loi quantique $E = \hbar \omega$, reliant l'énergie d'un grain de lumière à sa fréquence (\hbar étant la constante de Planck), est donc compatible avec la théorie de la relativité.

D'autre part, Einstein connaissait l'existence, prédite par Maxwell, d'une *pression de radiation* exercée par un complexe de rayonnement d'énergie E , ce qui est équivalent à l'existence d'une *quantité de mouvement* se propageant dans la direction normale aux fronts d'onde et ayant pour valeur E/c .¹⁶

Dans le langage de Minkowski, tout cela peut être résumé très simplement en décrivant la lumière en termes de photons possédant leur propre quadrivecteur énergie-impulsion

16. *Albert Einstein, œuvres choisies 2, op. cit., p. 71, note 17.*

d'énergie $\hbar\omega$ et d'impulsion égale, en norme, à $\hbar\omega/c$. En effet, ses composantes dans les systèmes barré et non barré sont bien reliées par la transformation de Lorentz :

$$\begin{pmatrix} \hbar\omega/c \\ \cos \varphi \hbar\omega/c \\ \sin \varphi \hbar\omega/c \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma V/c & 0 & 0 \\ -\gamma V/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hbar\omega/c \\ \cos \varphi \hbar\omega/c \\ \sin \varphi \hbar\omega/c \\ 0 \end{pmatrix}.$$

Le point d) ci-dessus assure que des particules lumineuses d'énergie $E = \hbar\omega$ sont compatibles avec la théorie de la relativité et donne la loi de transformation de la partie temporelle du quadrivecteur. Le point a) donne la loi de transformation de sa partie spatiale. Enfin, la théorie de Maxwell relie la composante temporelle et la norme de la partie spatiale.

La norme du quadrivecteur énergie-impulsion d'un photon est nulle, ce que l'on traduit en général en disant que « la masse propre d'un photon est nulle » ou encore, que toute l'énergie du photon est cinétique.

1.10 Equivalence de la masse et de l'énergie

La norme d'un vecteur n'est pas, en général, une propriété additive :

$$\vec{p} = \sum_i \vec{p}_i \not\Rightarrow |\vec{p}| = \sum_i |\vec{p}_i|.$$

Ainsi, en général, la masse propre d'un système de particules n'est pas égale à la somme des masses propres des divers constituants du système :

$$m_0 \neq \sum_i m_{0i}.$$

Par contre, l'énergie mc^2 est additive :

$$m = \sum_i m_i.$$

Or, nous avons vu que dans un processus de collision, le quadrivecteur énergie-impulsion total est conservé, donc notamment l'énergie totale :

$$\sum_i m_i = \sum_j m'_j \quad (1.44)$$

(les indices i et j caractérisent des sommations sur des nombres de particules pouvant être différents, le « prime » faisant référence à la situation après la collision). Si

$$T_i = (m_i - m_{0i}) c^2$$

représente l'énergie cinétique de la particule i avant la collision et

$$T'_j = (m'_j - m'_{0j}) c^2$$

désigne celle de la particule j après la collision, l'équation (1.44) prend la forme

$$\sum_i \left(m_{0i} + \frac{T_i}{c^2} \right) = \sum_j \left(m'_{0j} + \frac{T'_j}{c^2} \right). \quad (1.45)$$

Deux situations peuvent alors se présenter.

a) Lors d'une *collision élastique*, les particules, par définition, ne changent pas de nature. Les diverses masses propres de part et d'autre de la collision sont identiques, donc

$$\sum_i T_i = \sum_i T'_i.$$

L'énergie cinétique totale est conservée : seule sa répartition entre les particules peut varier, comme en mécanique newtonienne.

b) Pour les *collisions inélastiques*, les cas de figure sont plus nombreux. En général

$$\sum_i m_{0i} \neq \sum_j m'_{0j}.$$

Notamment, si toute l'énergie cinétique incidente $\sum T_i$ est nulle, on peut dire que toute l'énergie cinétique $\sum T'_j$ des fragments résulte de la conversion d'une quantité égale à $\sum T'_j/c^2$ de la masse propre initiale. C'est le cas de la désintégration radioactive, ou encore de l'émission, par un atome au repos, de deux photons de même énergie et de directions opposées (figure 1.12) : si

$$(\vec{p}) = (m_0c, 0), \quad (\vec{p}'_1) = (m'_0c, 0), \quad (\vec{p}'_2) = \frac{\hbar\omega}{c} (1, -1), \quad (\vec{p}'_3) = \frac{\hbar\omega}{c} (1, 1),$$

alors

$$m_0 = m'_0 + 2 \frac{\hbar\omega}{c^2}.$$

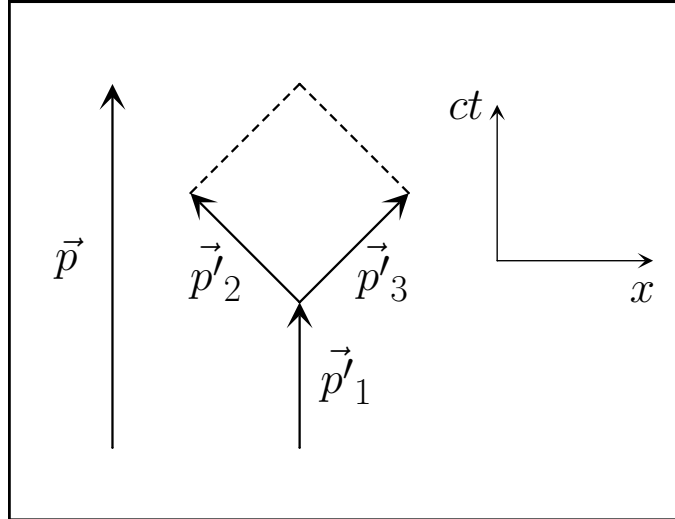


FIGURE 1.12: *Emission de deux photons par un atome.*

On peut évidemment reprendre ces considérations en envisageant cette fois l'absorption de photons. On en déduit que *le rayonnement transfère de l'inertie entre les corps qui émettent et les corps qui absorbent.*

Pour rendre compte de ces phénomènes on parle d'« équivalence de la masse et de l'énergie ». L'expression est, en fait, quelque peu équivoque puisque, justement, le concept d'énergie s'avère plus fondamental que celui de masse : l'énergie est conservée lors de n'importe quel processus de collision ; mais la somme des diverses masses propres ne l'est pas.

1.11 Une démonstration simple

La démonstration originale de l'équivalence entre la masse et l'énergie figure dans l'article « Ist die Trägheit eines Körpers von seinem Energiegehalt abhängig? »,¹⁷ soumis par Einstein le 27 septembre 1905 aux *Annalen der Physik*. Soucieux de pédagogie, il proposa en 1946 une argumentation simplifiée¹⁸ ayant l'avantage de ne pas requérir tout l'appareil formel de la relativité restreinte, mais seulement les lois suivantes :

- l'expression de la valeur E/c de la quantité de mouvement portée par un complexe de radiation (Strahlungskomplex) d'énergie E , connue depuis Maxwell ;
- l'expression de l'aberration de la lumière, connue depuis Bradley : si un référentiel se dirige avec une vitesse $V \ll c$ perpendiculairement à un rayon lumineux, alors dans ce référentiel le rayon n'est pas perpendiculaire à la direction de la vitesse V mais détermine avec cette perpendiculaire un angle α donné par $\alpha = V/c$;
- la loi de conservation de la quantité de mouvement ;
- le principe de relativité restreinte.

Reprenons ce raisonnement en le précisant quelque peu. Un corps de masse propre m_0 , initialement au repos dans un référentiel inertiel S , absorbe deux complexes lumineux de direction verticale et de sens opposés, de fréquence ω et d'énergie $E = \hbar\omega$. En vertu du principe de conservation de la quantité de mouvement, le corps demeure au repos dans S (figure 1.13 gauche). Décrivons maintenant cette situation dans un autre référentiel inertiel

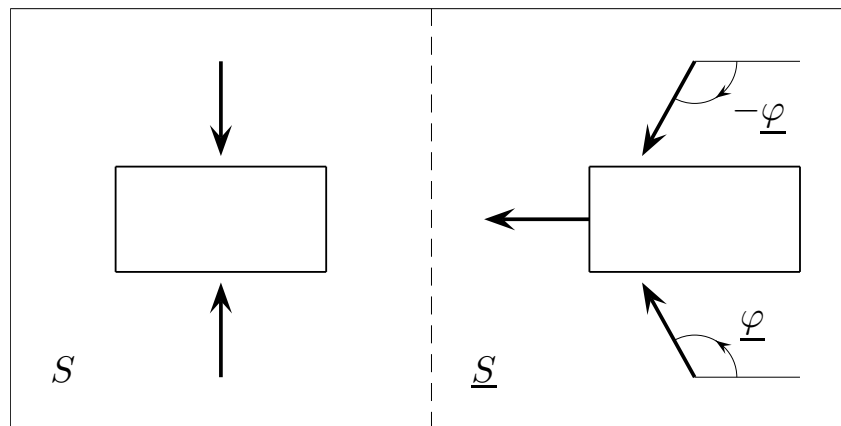


FIGURE 1.13: Absorption de deux photons par un atome, vue depuis deux référentiels inertiels S et \underline{S} .

\underline{S} de vitesse V dirigée horizontalement vers la droite par rapport à S (figure 1.13 droite). La vitesse du corps émetteur dans \underline{S} n'a pas varié au cours du processus, elle reste en permanence égale à $-V$. Cependant, sa quantité de mouvement a changé : en effet, celle de chacun des complexes lumineux possède maintenant une composante horizontale

$$\frac{\hbar\omega}{c} \cos \varphi.$$

17. « L'inertie d'un corps dépend-elle de son contenu en énergie? ». On en trouvera une traduction française dans *Albert Einstein, œuvres choisies 2, op. cit.*, pp. 60-62.

18. Traduction française dans *Albert Einstein, œuvres choisies 2, op. cit.*, pp. 69-71.

Einstein en déduit que *l'absorption de la lumière modifie la masse propre du récepteur*, qui prend désormais la valeur m'_0 . Compte tenu de l'effet Doppler (1.41) et de la loi de l'aberration (1.42), on a

$$\underline{\omega} = \frac{\omega}{\sqrt{1 - V^2/c^2}}, \quad \cos \underline{\varphi} = -\frac{V}{c}$$

et la loi de conservation de la quantité de mouvement s'écrit

$$-\frac{m_0 V}{\sqrt{1 - V^2/c^2}} - 2 \frac{\hbar \omega V/c^2}{\sqrt{1 - V^2/c^2}} = -\frac{m'_0 V}{\sqrt{1 - V^2/c^2}}.$$

On en déduit

$$\begin{aligned} m'_0 &= m_0 + 2 \frac{\hbar \omega}{c^2} \\ &= m_0 + 2 \frac{E}{c^2}. \end{aligned}$$

1.12 Problèmes

Problème 1.1 *Etudier l'absorption classique d'un photon par un atome initialement au repos.*

Solution. Il s'agit là, bien sûr, d'un processus inélastique. Soit \vec{p}_1 le quadrivecteur énergie-impulsion de l'atome avant la collision, \vec{p}_2 celui du photon et \vec{p}' celui du système « atome plus photon », c'est-à-dire du système « atome excité » (figure 1.14).

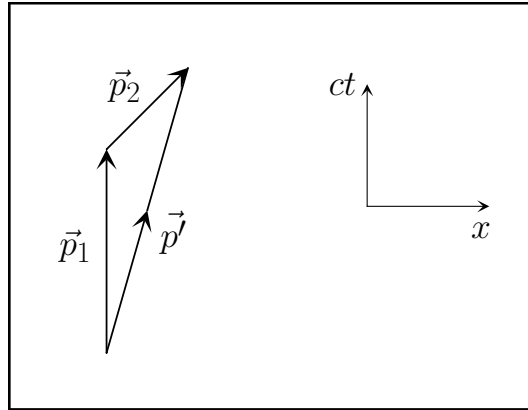


FIGURE 1.14: *Absorption d'un photon par un atome.*

Si la masse propre de l'atome avant la collision est égale à m_0 et si $m = \hbar \omega/c^2$, on a

$$(\vec{p}_1) = (m_0 c, 0), \quad (\vec{p}_2) = (m c, m c).$$

Par conséquent,

$$(\vec{p}') = [(m_0 + m) c, m c],$$

correspondant à une masse propre

$$m'_0 = m_0 \sqrt{1 + 2 \frac{m}{m_0}}.$$

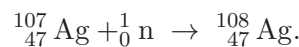
La masse propre d'un atome excité est donc supérieure à la masse propre du même atome dans son état fondamental. De plus, il possède maintenant une énergie cinétique

$$T' = (m_0 + m - m'_0) c^2.$$

Problème 1.2 (Réactions nucléaires de capture simple) *Il peut arriver qu'une particule incidente soit intégrée dans un noyau pour former un noyau nouveau, par exemple lors de la capture d'un proton :*



ou encore d'un neutron :



Dans un référentiel donné, une particule 1 vient frapper une particule 2 initialement au repos. Elles restent agglutinées après la collision. Discuter la masse propre et l'énergie cinétique du système en fonction des masses propres initiales et de l'énergie cinétique de la particule 1.

Solution. Avec des notations évidentes, on a

$$(m_1 c, m_1 \vec{V}_1) + (m_0 c, 0) = (m c, m \vec{V}).$$

Calculons le carré de la masse propre du système :

$$\begin{aligned} (m_0)^2 &= m^2 - m^2 V^2 / c^2 \\ &= (m_1 + m_0 c)^2 - (m_1)^2 (V_1)^2 / c^2 \\ &= (m_1)^2 - (m_1)^2 (V_1)^2 / c^2 + 2 m_1 m_0 c + (m_0 c)^2 \\ &= (m_0 c)^2 + 2 m_1 m_0 c + (m_0 c)^2 \\ &= (m_0 c)^2 + (m_0 c)^2 + 2 (m_0 c + T_1 / c^2) m_0 c \\ &= (m_0 c + m_0 c)^2 + 2 T_1 m_0 c / c^2. \end{aligned}$$

Par conséquent,

$$m_0 = (m_0 c + m_0 c) \sqrt{1 + \frac{2 T_1 m_0 c}{(m_0 c + m_0 c)^2 c^2}} > m_0 c + m_0 c.$$

L'inégalité est d'autant plus forte que T_1 est grand. Quant à l'énergie cinétique du système, elle vaut

$$\begin{aligned} T &= (m - m_0) c^2 \\ &= (m_1 + m_0 c - m_0) c^2 \\ &= T_1 + (m_0 c + m_0 c - m_0) c^2 < T_1. \end{aligned}$$

La collision est endothermique : le noyau résultant est, en général, instable. Ainsi, le noyau ${}^{108}_{47}\text{Ag}$ évoqué ci-dessus est radioactif.

Problème 1.3 *Discuter la valeur de la masse propre d'un système de particules en fonction des masses propres des constituants du système.*

Solution. Le quadrivecteur énergie-impulsion du système est donné par

$$(\vec{p}) = \left(\sum_i m_i c, \sum_i m_i \vec{V}_i \right).$$

Donc le carré de sa masse propre vaut

$$(m_0)^2 = \left(\sum_i m_i \right)^2 - \frac{1}{c^2} \left(\sum_i m_i \vec{V}_i \right)^2.$$

Raisonnons dans le référentiel du centre d'inertie (c'est-à-dire dans lequel la quantité de mouvement totale $\sum m_i \vec{V}_i$ est nulle). On a

$$m_0 = \sum_i m_i = \sum_i \left(m_{0i} + T_i/c^2 \right) \geq \sum_i m_{0i}.$$

L'égalité n'a lieu que si toutes les énergies cinétiques sont nulles dans le système du centre d'inertie, c'est-à-dire, en raisonnant dans un référentiel général, si toutes les particules se déplacent avec la même vitesse.

Corollaire. Soit un récipient d'eau immobile par rapport au laboratoire ($\sum m_i \vec{V}_i = 0$) : on augmente sa masse en le chauffant car cela revient à augmenter l'énergie cinétique des molécules d'eau.

Problème 1.4 (L'intégrité du photon) Déterminer sous quelles conditions un photon peut spontanément se désintégrer en deux particules.¹⁹

Solution. Supposons qu'un photon de quadrivecteur énergie-impulsion

$$(\vec{p}) = (mc, m\vec{V}), \quad \text{avec} \quad |\vec{V}| = c,$$

se désintègre en deux particules de quadrivecteurs

$$(\vec{p}_1) = (m_1 c, m_1 \vec{V}_1) \quad \text{et} \quad (\vec{p}_2) = (m_2 c, m_2 \vec{V}_2),$$

respectivement. On a

$$\vec{p} = \vec{p}_1 + \vec{p}_2,$$

donc

$$\begin{aligned} m &= m_1 + m_2 \\ m\vec{V} &= m_1 \vec{V}_1 + m_2 \vec{V}_2. \end{aligned}$$

Le carré scalaire de cette dernière équation donne successivement

$$\begin{aligned} m^2 c^2 &= (m_1)^2 (V_1)^2 + (m_2)^2 (V_2)^2 + 2 m_1 m_2 \vec{V}_1 \cdot \vec{V}_2 \\ (m_1)^2 \left[1 - (V_1)^2 / c^2 \right] &+ (m_2)^2 \left[1 - (V_2)^2 / c^2 \right] + 2 m_1 m_2 \left(1 - V_1 V_2 \cos \theta / c^2 \right) = 0, \end{aligned}$$

où θ est l'angle entre \vec{V}_1 et \vec{V}_2 . Tous les termes étant positifs, cette relation n'est possible que si $V_1 = V_2 = c$ et $\theta = 0$, impliquant $\vec{V}_1 = \vec{V}_2 = \vec{V}$. Dans ce cas extrême, le photon de départ se désintègre en deux photons de mêmes direction et sens.

Corollaire. Dans le vide, la matérialisation d'un photon en une paire électron-positron (ou, inversement, l'annihilation d'une paire électron-positron en un seul photon) est impossible. On peut d'ailleurs arriver directement à cette conclusion en remarquant que, dans le référentiel du centre d'inertie des deux leptons, le photon aurait une quantité de mouvement $m\vec{V}$ nulle, ce qui est impossible.

19. E.F. Taylor et J.A. Wheeler, *op. cit.*, p. 259.

Chapitre 2

Approche du champ gravitationnel uniforme et constant

2.1 Le voyageur de Langevin

Dans l'espace-temps de Minkowski, le *quadrivecteur accélération* \vec{W} d'une particule est tout naturellement défini comme la dérivée de son quadrivecteur vitesse \vec{U} par rapport au temps propre :

$$\vec{W} = \frac{d\vec{U}}{d\tau}.$$

Ces deux quadrivecteurs sont orthogonaux. En effet, de la constance de la norme de la quadrivitesse $\vec{U} \cdot \vec{U} = -c^2$, on tire

$$0 = \frac{d}{d\tau}(\vec{U} \cdot \vec{U}) = 2 \frac{d\vec{U}}{d\tau} \cdot \vec{U} = 2 \vec{W} \cdot \vec{U}.$$

Dès lors, \vec{W} constitue un quadrivecteur spatial ; sauf si la vitesse est constante, auquel cas \vec{W} est le quadrivecteur dont *toutes* les composantes sont nulles.

Nous allons étudier le mouvement relativiste d'une particule ponctuelle P soumise à une accélération propre constante de mesure g . Mais commençons par définir les termes.

Soit un référentiel inertiel S doté de coordonnées cartésiennes (ct, x) dans lesquelles nous effectuerons les calculs. Notons-y $V(t)$ la vitesse instantanée de P . A chaque instant t , le *référentiel comobile* $\underline{S}(t)$ (également appelé *référentiel concomitant*) est le référentiel inertiel, de vitesse $V(t)$ par rapport à S , doté de coordonnées cartésiennes $(c\underline{t}, \underline{x})$ et dont l'origine \underline{Q} coïncide avec P en $\underline{t} = 0$. Par définition, on a $\underline{U}^0 = c$, $\underline{U}^1 = 0$. Vu l'orthogonalité de \vec{U} et de \vec{W} , cela impose $\underline{W}^0 = 0$. On dit que la particule possède une « accélération propre de mesure g » si $\underline{W}^1 = g$.

Cette définition est naturelle, puisqu'elle revient à poser $d\underline{V} = g d\underline{t}$. Bien entendu, elle n'implique pas $dV = g dt$ (car $d\underline{V} \neq dV$, vu la loi d'addition des vitesses ; et en plus $d\underline{t} \neq dt$, compte tenu de la dilatation du temps). L'accélération spatiale n'est plus une notion absolue, même lorsque l'on passe d'un référentiel inertiel à un autre ! Il est capital de bien réaliser ceci : l'invariance lorentzienne a précipité l'idole newtonienne de son piédestal, et dans sa disgrâce la vieille reine déchuée a entraîné l'espace absolu.

Comme \underline{S} et S sont reliés par une transformation de Lorentz de vitesse V , on a ainsi

$$\begin{pmatrix} W^0 \\ W^1 \end{pmatrix} = \frac{1}{\sqrt{1-V^2/c^2}} \begin{pmatrix} 1 & V/c \\ V/c & 1 \end{pmatrix} \begin{pmatrix} 0 \\ g \end{pmatrix}.$$

Il en découle

$$W^1 = \frac{g}{\sqrt{1 - V^2/c^2}},$$

c'est-à-dire

$$\frac{d}{d\tau} \left(\frac{V}{\sqrt{1 - V^2/c^2}} \right) = \frac{g}{\sqrt{1 - V^2/c^2}},$$

ou encore

$$\frac{d}{dt} \left(\frac{V}{\sqrt{1 - V^2/c^2}} \right) = g. \quad (2.1)$$

Du point de vue dynamique, il s'agit là de l'équation de mouvement (1.38) d'une particule soumise à une force $F = m_0g$ constante :

$$\frac{d}{dt} \left(\frac{m_0V}{\sqrt{1 - V^2/c^2}} \right) = m_0g.$$

On en déduit, en supposant $V(0) = 0$ et $x(0) = 0$:

$$V(t) = \frac{gt}{\sqrt{1 + g^2t^2/c^2}} \quad (2.2)$$

$$x(t) = \frac{c^2}{g} \left(\sqrt{1 + g^2t^2/c^2} - 1 \right). \quad (2.3)$$

Le vieillissement de la particule est donné par

$$d\tau = \sqrt{1 - V^2(t)/c^2} dt = \frac{dt}{\sqrt{1 + g^2t^2/c^2}},$$

donc

$$\tau = \int_0^t \frac{d\lambda}{\sqrt{1 + g^2\lambda^2/c^2}} = \frac{c}{g} \operatorname{arcsinh} \left(\frac{gt}{c} \right). \quad (2.4)$$

Application numérique : le voyageur de Langevin. Considérons, dans S , un voyageur accéléré pendant un temps T puis décéléré pendant le même temps. Il repart alors en sens inverse pour se retrouver à son point de départ au bout d'un temps total égal à $4T$. Pour $g = 9.8 \text{ m/s}^2$, on obtient le tableau suivant.

destination	distance atteinte	temps terrestre $4T$	vieillessement du voyageur	vitesse maximale $/c$
	244 secondes-lumière	4 jours	4 jours - 1/2 s	0.0028
	35 heures-lumière	3 mois	3 mois - 91 min.	0.064
α Centauri	4.2 années-lumière	11.6 ans	7.1 ans	0.95
Véga	26 années-lumière	55.7 ans	13.0 ans	0.997 59
Aldébaran	68 années-lumière	139.8 ans	16.6 ans	0.999 62
Rigel	800 années-lumière	1604 ans	26.0 ans	0.999 997 08
M 31	2×10^6 années-lumière	4×10^6 ans	56.4 ans	$1 - 5 \times 10^{-13}$

On peut facilement exprimer la vitesse et la position du voyageur dans S en fonction de son temps propre τ : comme

$$ct(\tau) = \frac{c^2}{g} \sinh \left(\frac{g\tau}{c} \right), \quad (2.5)$$

on obtient

$$V(\tau) = c \tanh\left(\frac{g\tau}{c}\right) \quad (2.6)$$

$$x(\tau) = \frac{c^2}{g} \left[\cosh\left(\frac{g\tau}{c}\right) - 1 \right]. \quad (2.7)$$

2.2 Référentiel d'une particule uniformément accélérée

Nous allons étudier la *fusée uniformément accélérée*, matérialisée par un ensemble de particules dont la distance au voyageur P de la section 2.1 est *constante* dans le référentiel concomitant $\underline{S}(t)$. Chaque particule sera caractérisée par son abscisse ξ dans S en $t = 0$. Cherchons à établir sa loi de mouvement.¹

L'axe des \underline{x} est déterminé par l'ensemble des événements simultanés, dans le référentiel comobile $\underline{S}(\tau)$, à l'événement origine $P(\tau)$ de coordonnées (2.5) et (2.7) :

$$ct(\tau, 0) = \frac{c^2}{g} \sinh\left(\frac{g\tau}{c}\right), \quad x(\tau, 0) = \frac{c^2}{g} \left[\cosh\left(\frac{g\tau}{c}\right) - 1 \right]. \quad (2.8)$$

Les coordonnées $x(\tau, \xi)$ et $ct(\tau, \xi)$ de l'événement situé sur cet axe en l'abscisse $\underline{x} = \xi$ résultent de la transformation de Lorentz : posant $\gamma = 1/\sqrt{1 - V^2(\tau)/c^2}$, on trouve (voir figure 2.1)

$$x(\tau, \xi) = x(\tau, 0) + \gamma \xi, \quad ct(\tau, \xi) = ct(\tau, 0) + \gamma \xi V(\tau)/c.$$

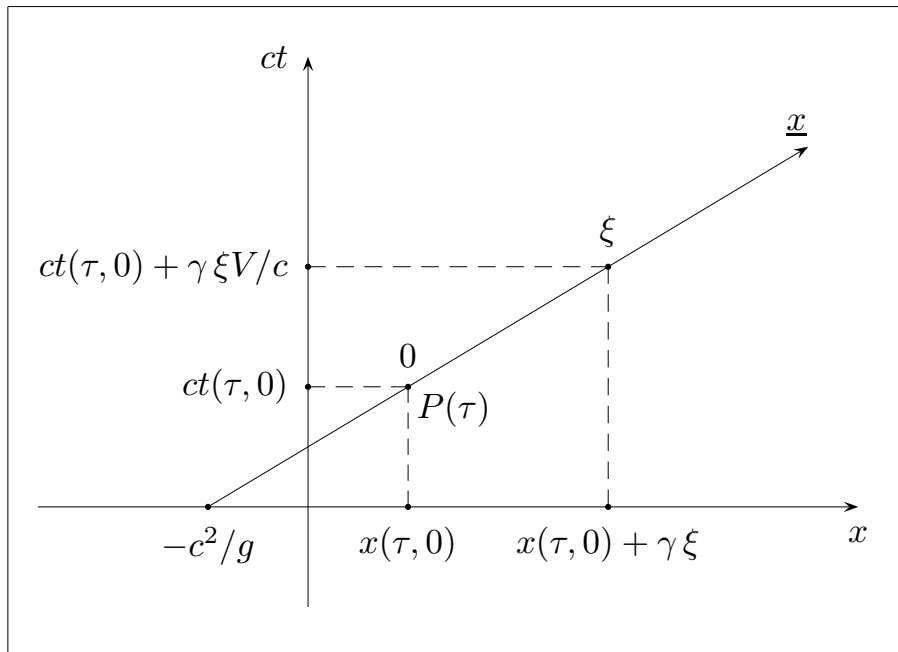


FIGURE 2.1: L'abscisse du point $\underline{x} = \xi$ est déduite des résultats de la section 1.6. Pour calculer son ordonnée, il faut se rappeler que la pente de l'axe des \underline{x} est égale à V/c .

1. C. Möller, *The theory of Relativity*, Oxford University Press, 1972, pp. 289-292 ; C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation*, Freeman, 1973, pp. 172-174.

Compte tenu de (2.8) et (2.6), on obtient

$$ct(\tau, \xi) = \left(\xi + \frac{c^2}{g} \right) \sinh \left(\frac{g\tau}{c} \right) \quad (2.9)$$

$$x(\tau, \xi) = \left(\xi + \frac{c^2}{g} \right) \cosh \left(\frac{g\tau}{c} \right) - \frac{c^2}{g}. \quad (2.10)$$

Ces relations constituent le changement de coordonnées permettant de passer du référentiel $F(c\tau, \xi)$ attaché à la particule P uniformément accélérée (la fusée) au système inertiel $S(ct, x)$ et vice-versa (voir la figure 2.2).

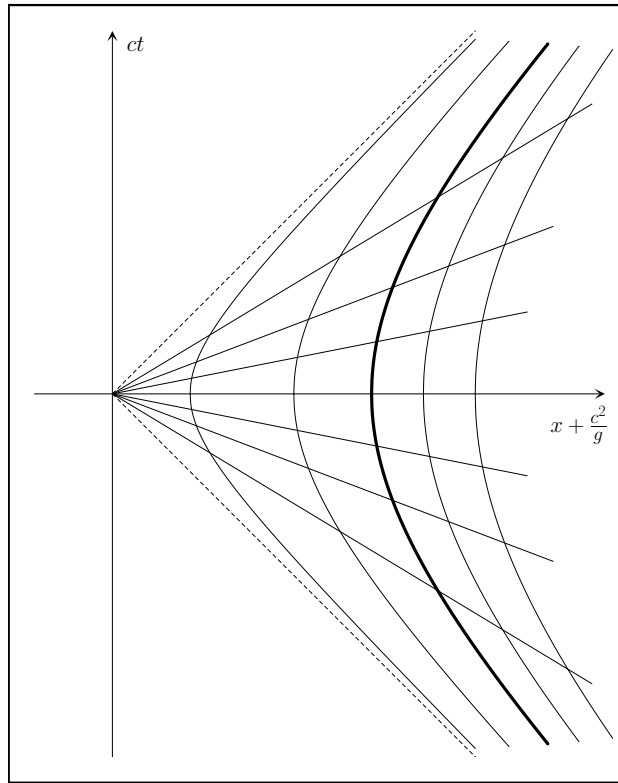


FIGURE 2.2: *Système de coordonnées attaché à la fusée uniformément accélérée, en unités $c^2/g = 1$, $c/g = 1$. De gauche à droite on a représenté les branches d'hyperbole relatives à $\xi = -0.7 c^2/g$, $\xi = -0.3 c^2/g$, $\xi = 0$ (en gras), $\xi = 0.2 c^2/g$, $\xi = 0.4 c^2/g$. De bas en haut sont figurées les droites $\tau = -0.7 c/g$, $\tau = -0.4 c/g$, $\tau = -0.2 c/g$, $\tau = 0$ (axe horizontal), $\tau = 0.2 c/g$, $\tau = 0.4 c/g$, $\tau = 0.7 c/g$.*

Remarquons que l'axe des \underline{x} a pour équation dans S

$$c[t - t(\tau, 0)] = \frac{V(\tau)}{c} [x - x(\tau, 0)],$$

c'est-à-dire

$$ct = \left(x + \frac{c^2}{g} \right) \tanh \left(\frac{g\tau}{c} \right).$$

Quel que soit τ , cet axe passe donc par le point $(ct = 0, x = -c^2/g)$.

En éliminant τ dans les relations (2.9) et (2.10), on obtient la loi $x_\xi(t)$ donnant en fonction du temps t l'abscisse x d'une particule d'étiquette constante ξ :

$$x_\xi(t) = \frac{c^2}{g} \left[\sqrt{\left(1 + \frac{g\xi}{c^2}\right)^2 + \frac{g^2 t^2}{c^2}} - 1 \right]$$

pour autant que $(1 + g\xi/c^2) > 0$. On en déduit la loi des vitesses dans S :

$$V_\xi(t) = \frac{dx_\xi(t)}{dt} = \frac{gt}{\sqrt{\left(1 + \frac{g\xi}{c^2}\right)^2 + \frac{g^2 t^2}{c^2}}}. \quad (2.11)$$

Cette quantité tend vers c si $\xi \rightarrow -c^2/g$. La loi des vitesses dépend de ξ , mais toutes les particules situées sur l'axe des \underline{x} (correspondant à un temps τ donné) possèdent la même vitesse $c \tanh(g\tau/c)$, $\forall \xi$.

L'équation (2.1) permet de calculer l'accélération propre d'une particule ξ donnée, qui se trouve être *constante* :

$$\frac{d}{dt} \frac{V_\xi(t)}{\sqrt{1 - V_\xi^2(t)/c^2}} = \frac{g}{1 + \frac{g\xi}{c^2}}. \quad (2.12)$$

Considérons à présent l'élément de longueur de l'espace-temps de Minkowski. Pour mémoire, il s'écrit, dans le système des coordonnées cartésiennes :

$$ds^2 = -c^2 dt^2 + dx^2.$$

Compte tenu des relations (2.9) et (2.10), il prend, dans les coordonnées $(c\tau, \xi)$, la forme²

$$ds^2 = - \left(1 + \frac{g\xi}{c^2}\right)^2 c^2 d\tau^2 + d\xi^2. \quad (2.13)$$

Posant ($x^0 = c\tau$, $x^1 = \xi$), on note comme suit son expression générique :

$$ds^2 = g_{\alpha\beta}(\vec{x}) dx^\alpha dx^\beta. \quad (2.14)$$

On dit que *les fonctions $g_{\alpha\beta}$ sont les composantes du tenseur de Minkowski en coordonnées générales* (ici dans le système de coordonnées de la fusée).

Le temps propre $T_\xi(\tau)$ indiqué par l'horloge attachée à la particule ξ se déduit immédiatement de (2.13) :

$$\begin{aligned} ds^2 &= -c^2 dT_\xi^2(\tau) \\ &= -c^2 \left(1 + \frac{g\xi}{c^2}\right)^2 d\tau^2, \end{aligned}$$

donc

$$T_\xi(\tau) = \tau \left(1 + \frac{g\xi}{c^2}\right), \quad (2.15)$$

ce qui le relie très simplement à la coordonnée temporelle τ (qui est, rappelons-le, le temps propre de la particule située à l'origine). Du reste on a, comme il se doit, en intégrant à ξ constant :

$$T_\xi(\tau) = \int_0^{t(\tau, \xi)} \sqrt{1 - V_\xi^2(\lambda)/c^2} d\lambda.$$

2. Les coordonnées s'avèrent singulières en $\xi = -c^2/g$ [comme la figure 2.2 et la loi des vitesses (2.11) le suggéraient déjà] mais, bien entendu, l'espace-temps est toujours celui de Minkowski. L'élément de longueur (2.13) est parfois appelé *métrique de Rindler*. Voir par exemple R. Wald, *General Relativity*, The University of Chicago Press, 1984, pp. 149-152.

2.3 Le principe d'équivalence et le principe de covariance générale

L'expérience suggère l'égalité de la masse inerte et de la masse pesante, dont il résulte que l'accélération d'un corps dans un champ de gravitation est indépendante de la nature physique ou chimique (bois, laiton...) de ce corps.³ Cette propriété était bien connue de Newton, qui l'a constatée sans pouvoir l'interpréter. Pourtant son statut n'est guère anodin, car elle jette le trouble sur la définition opérationnelle des référentiels inertiels, dont nous avons déjà noté l'importance épistémologique dans l'élaboration de la mécanique classique.

Comment, en effet, s'assurer qu'un référentiel donné est inertiel ou non ? On connaît la réponse théorique : un référentiel est inertiel si et seulement si une particule non soumise à l'action d'une force s'y meut en mouvement rectiligne et uniforme. Mais comment peut-on garantir qu'aucune force n'agit sur ce corps ? Car s'il est facile de détecter la présence d'un champ électromagnétique (essentiellement parce que le rapport de la charge à la masse n'est pas le même pour tous les corps : ainsi, des électrons et des protons trahiront l'existence d'un champ électromagnétique en y réagissant différemment), en revanche aucune expérience locale ne peut trancher en faveur de la présence ou non d'un champ gravitationnel. Cela signifie que, sauf à connaître la position et la masse de tous les astres peuplant l'univers (et à condition bien sûr d'admettre la loi en $1/r^2$), la mécanique newtonienne est, par exemple, incapable de distinguer un référentiel soi-disant inertiel, d'un référentiel non inertiel soumis à un champ de gravitation. Ainsi, l'ensemble du système solaire pourrait se trouver, soit inertiel, soit accéléré en chute libre dans le champ, constant et homogène à l'échelle de quelques heures-lumière, généré par une masse gigantesque située à une distance énorme et qui, elle, serait en translation rectiligne et uniforme par rapport à l'espace absolu : l'édifice newtonien souffre d'une ambiguïté fâcheuse quant à la définition pratique des référentiels inertiels.

Comme il s'agissait d'être concret, Newton a dû compléter sa construction à l'aide d'un ingrédient supplémentaire. Il a donc postulé un référentiel inertiel plausible (impliquant tous les autres) : les axes de Copernic, liés au soleil et pointant constamment vers des étoiles données. Il faut s'incliner devant l'éblouissante cohérence et les succès expérimentaux de la théorie ainsi parachevée.

Einstein va reprendre toutes ces considérations depuis leur base et, s'engageant dans une direction conceptuelle différente, les ériger en la pierre angulaire de sa théorie de la relativité générale, le *principe d'équivalence*, qu'il formule comme suit en 1911.⁴ Soit K un système de coordonnées au repos par rapport à un repère inertiel K_0 et dans lequel règne un champ de pesanteur de mesure g selon une direction donnée, appelons-la « vers le bas ». Soit K' un système dans une région éloignée de toute source de champ gravitationnel et soumis, par rapport à K_0 , à une accélération de mesure g vers le haut. Alors *les lois de la physique rapportées à K concordent parfaitement avec ces mêmes lois rapportées à K'* .

On peut exprimer ce principe sous une autre forme. Dans un référentiel en chute libre (sans rotation), la force « réelle » due à la gravitation est toujours exactement compensée par la force « fictive » d'entraînement due à l'inertie. Le mouvement d'une particule-test soumise uniquement à son inertie et à la gravitation y est donc rectiligne et uniforme.

3. A strictement parler, c'est le *rapport* de la masse inerte à la masse pesante qui est identique pour tous les corps, mais on peut alors choisir la constante de Cavendish pour que ce rapport soit égal à 1.

4. C'est en 1907 qu'Einstein conçut cette hypothèse qu'il devait qualifier en 1919 de « glücklichste Gedanke meines Lebens », « l'idée la plus heureuse de toute ma vie ». Voir A. Pais, « *Subtle is the Lord...* », *The Science and the Life of Albert Einstein*, Oxford University Press, 1982, pp. 177-178.

Einstein a vu là davantage qu'une simple coïncidence : conjecturant l'identité physique de la gravitation et de l'inertie, il décréta qu'*un référentiel en chute libre dans un champ de gravitation est inertiel par essence*. Cela signifie que la relativité restreinte y est applicable.

Ainsi, le principe d'équivalence postule qu'il est toujours possible de générer ou d'annuler un champ de gravitation arbitraire par un choix approprié de référentiel. Dès lors il serait tentant d'imaginer que l'on peut refouler les lois de la gravitation dans le camp de la pure cinématique, mais cette conclusion s'avèrerait erronée. En effet, il se trouve que dans la nature tout « vrai » champ de pesanteur s'annule à l'infini et ne peut donc pas être homogène : le principe d'équivalence n'est en fait valable que *localement* dans l'espace-temps. Remarquons d'ores et déjà que deux systèmes localement inertiels ne sont pas nécessairement en translation rectiligne et uniforme l'un par rapport à l'autre (pensons à deux systèmes tombant vers la Terre, aux antipodes l'un de l'autre).

Il s'ensuit que :

- un référentiel en translation rectiligne et uniforme par rapport aux axes de Copernic (donc réputé inertiel en mécanique newtonienne) mais dans lequel règne un champ de pesanteur peut désormais être appréhendé comme non inertiel ;
- et inversement, on a le droit de considérer comme inertiel un référentiel accéléré par rapport aux axes de Copernic, en considérant que le comportement des corps dans ce système résulte de l'action d'un champ de gravitation « réel ».

On comprend dès lors que le principe d'équivalence vient brouiller les cartes en vidant totalement de sa substance le concept de système inertiel global : il appelle à n'en plus conserver qu'une version *locale* au travers d'une nouvelle définition opérationnelle, audacieuse, de la classe des systèmes (localement) inertiels, désormais réduits aux référentiels en chute libre. Il ouvre la voie à une théorie où les masses et leurs interactions perdent leurs prérogatives de concepts fondamentaux au profit du concept de *champ local* (sans restriction imposée *a priori* par des conditions aux limites spatiales, comme par exemple l'annulation du champ à l'infini).

Dans le domaine de validité de la mécanique de Newton, l'expérience peut cautionner l'égalité de la masse inerte et de la masse pesante. Cependant, tel qu'il est conçu ici, le principe d'équivalence est censé s'étendre à la relativité restreinte. Or, celle-ci enseigne l'équivalence de la masse inerte m_i et de l'énergie ($E = m_i c^2$) mais ne fournit en elle-même aucune information concernant la masse gravitationnelle m_g . Le principe d'équivalence impose qu'à une variation de la masse inerte corresponde la même variation de la masse gravitationnelle. Autrement dit, *en plus de son inertie l'énergie doit également posséder un poids*.⁵ On subodore ainsi, en particulier, une piste originale pour traiter la question de l'interaction entre la lumière et un champ de gravitation : *la lumière doit tomber*, ne fût-ce d'ailleurs que pour des raisons cinématiques puisqu'une trajectoire localement rectiligne (c'est-à-dire, non accélérée au point considéré) dans un référentiel en chute libre au voisinage de la Terre est forcément courbée dans un référentiel lié à notre planète.

En frappant d'anathème le concept de système inertiel global, le principe d'équivalence enlève toute pertinence à l'idée newtonienne de « champ gravitationnel réel », *fondant ainsi gravitation et inertie en un seul et même phénomène* envisagé selon deux points de vue différents. Il s'agit là d'une démarche tout à fait analogue à l'assimilation du champ électrique et du champ magnétique à deux projections d'une entité plus large, le champ électromagnétique. Un conducteur immobile par rapport à un système de référence ressent

5. Si ce n'était pas le cas, l'accélération prise par un corps lâché dans un champ de pesanteur varierait en fonction de son énergie. Notamment, *l'accélération verticale d'un corps dépendrait de la composante horizontale de sa vitesse initiale*.

les effets du champ électrique induit par un aimant en translation uniforme et il apparaît un courant induit. D'autre part, un observateur solidaire de l'aimant ne détecte aucun champ électrique mais interprète les déplacements de charges dans le conducteur à l'aide de la force de Lorentz. La relativité restreinte permet de relier ces deux points de vue qu'elle considère comme étant totalement équivalents. De même, un observateur immobile dans un champ de pesanteur déduit l'existence d'un champ de gravitation à partir de l'observation d'objets accélérés dans son voisinage, mais un observateur en chute libre avec tous ces objets conclut à l'absence de champ de gravitation : *ces observateurs ont raison tous les deux, leurs points de vue sont totalement équivalents.*

La relativité restreinte déniait tout sens au concept de mouvement de l'éther et postulait l'équivalence physique de tous les référentiels inertiels. En destituant ceux-ci, le principe d'équivalence poursuit ce travail de nivellement et suggère d'envisager l'*équivalence formelle de tous les référentiels*, désormais conçus comme *systèmes de coordonnées*. Cette hypothèse porte le nom de *principe de relativité générale*. Ainsi, à la question : « Pourquoi y a-t-il des référentiels inertiels étendus dans tout l'espace, dotés de propriétés particulières ? », la réponse inattendue fournie par Einstein ne fut pas « pour telle ou telle raison » mais consista tout bonnement à renier l'existence de ces référentiels privilégiés !

Il faut cependant se garder d'interpréter le principe de relativité générale comme une extension littérale du principe de relativité restreinte. En effet, ce dernier ne fait, somme toute, que traduire une propriété géométrique d'isotropie spatio-temporelle en tout point de l'espace-temps absolu de Minkowski, en mettant en relief l'invariance du tenseur métrique sous les transformations de Lorentz reliant les coordonnées cartésiennes des référentiels inertiels, et donc la symétrie entre ces derniers. Mais pour un référentiel quelconque cette symétrie est purement *formelle*. Par exemple, deux billes initialement au repos dans un carrousel en rotation verront leur distance varier au cours du temps, propriété qui distingue clairement le carrousel d'un référentiel inertiel. La relativité générale se donne cependant pour mission de trouver un formalisme covariant (indifférent au choix des coordonnées), tant pour décrire le mouvement des particules que le champ d'inertie-gravitation lui-même.

L'exigence de *covariance générale* constitue une sorte de couronnement de l'entreprise qui, depuis le géocentrisme antique jusqu'à la cosmologie contemporaine en passant par Copernic et Newton, n'a cessé de miner les prérogatives attribuées à un point de vue particulier. Toutefois, elle a un prix. On se rappelle en effet que les coordonnées cartésiennes utilisées dans les référentiels inertiels de l'espace-temps de Minkowski sont déterminées par des mesures d'intervalles spatiaux et temporels effectuées selon des procédures opérationnelles soigneusement définies. L'enjeu de la transformation de Lorentz est ainsi d'assurer la liaison entre deux systèmes de coordonnées cartésiennes directement issues de l'expérience. La signification physique des entités rapportées à ces coordonnées (comme le quadricteur énergie-impulsion dont elles justifient le nom, ou encore le quadripotential vecteur en électromagnétisme) leur est étroitement associée. On doit dès lors s'attendre à ce que des coordonnées générales (quoique appropriées pour décrire les phénomènes puisqu'en définitive la physique ne s'intéresse qu'aux *faits*, c'est-à-dire aux coïncidences entre des événements, lesquelles peuvent sans ambiguïté s'exprimer dans des coordonnées quelconques) posent des problèmes d'interprétation.

Pour fixer les idées, considérons l'ensemble des quatre nombres p^α constitué, d'une part de l'énergie d'une particule divisée par c , et d'autre part des trois composantes de sa quantité de mouvement spatiale, ces quantités étant mesurées dans un certain référentiel inertiel S . Si l'on définit pareillement des \underline{p}^α relatifs à un autre référentiel inertiel \underline{S} , alors le lien entre les p^α et les \underline{p}^α est assuré par le groupe de Lorentz. Mais d'un point

de vue strictement sémantique, il faut noter que l'ensemble « énergie plus impulsion » ne se comporte comme un quadrivecteur *que* pour le groupe de Lorentz. Bien entendu, rien n'empêche de calculer les composantes dans des coordonnées arbitraires d'un certain objet \vec{p} , quadrivectoriel pour le groupe de transformations *quelconques* des coordonnées et se réduisant aux p^α dans les coordonnées cartésiennes de S . Mais alors, par exemple, la composante temporelle ainsi obtenue ne décrira pas l'énergie de la particule, mesurée par qui que ce soit.

Afin de rendre possible sa vérification par l'expérience, sans quoi elle ne pourrait évidemment pas prétendre au statut de science, la relativité générale devra donc prévoir le calcul d'*invariants*, tout résultat expérimental pouvant être conçu comme un invariant puisqu'il se ramène toujours, d'une façon ou d'une autre, à une indication chiffrée fournie par un appareil de mesure. Au rang des invariants importants figure le *temps propre*, qui assumera souvent le rôle du temps absolu de Newton, quoique différemment pour chaque particule. Pour revenir à l'énergie, remarquons que, dans le référentiel inertiel concomitant à un observateur de quadrivitesse \vec{U}_{obs} [c'est-à-dire le référentiel inertiel dans lequel les composantes de cette quadrivitesse sont $(c, 0, 0, 0)$], elle peut s'écrire sous la forme $E = -\vec{p} \cdot \vec{U}_{obs}$. Or, le produit scalaire de deux vecteurs est indépendant du système de coordonnées dans lequel on le calcule. On en déduit que, dans des coordonnées quelconques $\{x^\alpha\}$ où les composantes du tenseur de Minkowski ont pour expression $g_{\alpha\beta}$, un observateur de quadrivitesse \vec{U}_{obs} , mesurant l'énergie d'une particule de quadrivecteur « énergie-impulsion » \vec{p} , obtient le résultat

$$\begin{aligned} E &= -\vec{p} \cdot \vec{U}_{obs} \\ &= -g_{\alpha\beta} p^\alpha U_{obs}^\beta. \end{aligned} \tag{2.16}$$

2.4 Conséquences du principe d'équivalence

La théorie de la relativité s'est élaborée en deux étapes conceptuellement très différentes, dont chacune révolutionne une des deux acceptions du concept newtonien d'« espace absolu ».

Chez Newton, une section spatiale de l'espace-temps à temps constant détermine une hypersurface indépendante de l'observateur, puisque le temps est universel. C'est le premier sens (il est vrai, un peu anecdotique) de la notion d'« espace absolu ». Or, en associant l'espace et le temps par le biais d'une métrique, la relativité restreinte bouleverse la notion de simultanéité. Aussi l'espace est-il devenu *relatif à l'observateur*.

Mais, plus fondamentalement, rappelons que le repos absolu n'est pas détectable dans les équations de la mécanique newtonienne, centrées sur l'accélération. Ainsi, l'espace absolu ne se manifeste qu'au travers de ses lieutenants, les référentiels inertiels, dont aucun, bizarrement, ne peut sortir du rang en revendiquant l'immobilité absolue. En outre, l'espace dit « absolu » n'est pas qu'un simple réceptacle géométrique. Régissant l'inertie, il se dote en effet de propriétés physiques. C'est par rapport à lui qu'un corps non soumis à l'action d'une force se meut en translation rectiligne et uniforme, sans que cet espace soit lui-même soumis, en retour, à l'influence des masses et de leurs mouvements. Il s'agit là de points noirs dont hérite la relativité restreinte, dans la mesure où cette théorie ne remet pas en cause la notion de référentiel inertiel. Dès lors, elle n'apporte sur les forces fictives d'inertie aucun éclairage fondamentalement nouveau. A ce titre, il n'y a guère de différence entre l'espace absolu de Newton et l'espace-temps absolu de Minkowski. Le principe d'équivalence vient troubler cette perspective : *en associant l'inertie et la gravitation,*

il fait perdre toute signification à la notion d'« accélération par rapport à l'espace ». Il en résultera l'avènement d'un espace-temps relatif à son contenu matériel, caractéristique de la relativité générale.

2.4.1 L'accélération en relativité générale

Il est *a priori* légitime d'essayer de fonder, dans l'espace-temps de Minkowski, une théorie de la gravitation reliant le quadrivecteur accélération d'une particule à un quadrivecteur « force » dont l'expression resterait à préciser. Or, le principe d'équivalence, qu'il faut bien entendu comprendre dans son acception *relativiste* amalgamant l'espace et le temps, impose que le quadrivecteur accélération d'une particule en chute libre ait *toutes* ses composantes (et pas seulement ses composantes spatiales) nulles dans le référentiel inertiel concomitant. Mais alors, cette propriété doit être vérifiée dans tous les référentiels !

Une telle exigence semble absurde et irréalisable. A titre d'exemple, considérons en effet, dans les conditions de la balistique newtonienne (confirmée par l'expérience dont toute théorie de la gravitation doit évidemment pouvoir rendre compte en première approximation), une particule lancée verticalement depuis le sol. Appelant z l'axe vertical et t celui du temps dans le référentiel de la Terre (supposée inertielle), on sait que la trajectoire spatio-temporelle $[t, z(t)]$ est une parabole. Par conséquent la quadrivitesse de la particule, tangente à cette trajectoire, ne peut être constante !

Observons cependant que, dans les coordonnées (ct, z) , compte tenu de la valeur numérique énorme de la vitesse de la lumière, cette parabole se voit considérablement écrasée et présente l'allure d'une ligne presque droite s'écartant imperceptiblement de l'axe temporel. A vrai dire, une particule en chute libre newtonienne se déplace surtout dans le temps et peu dans l'espace, en s'éloignant à peine de la ligne d'univers réalisée par une particule immobile dans ce référentiel, c'est-à-dire d'un mouvement rectiligne non accéléré.

Ce commentaire pourrait passer pour un lamentable sophisme (« A quoi riment donc ces *presque* et ces *à peine* ? », pourrait s'offusquer le lecteur) mais décèle pourtant un point de vue radicalement novateur, une idée de génie aux fondements de la relativité générale. Elle suggère que l'on pourrait sauver les phénomènes tout en les maintenant compatibles avec le principe d'équivalence, en basculant dans un paradigme complètement différent : *il suffirait que l'espace-temps ne soit pas tout à fait l'espace-temps de Minkowski pour que la trajectoire d'une particule newtonienne en chute libre soit rigoureusement « droite », entendons par là que sa quadrivitesse soit strictement transportée parallèlement le long de sa ligne d'univers.*

Concrètement, tout l'art fut de trouver les mathématiques adéquates en adoptant une définition plus large du parallélisme (en dotant l'espace-temps d'une connexion affine plus riche qu'en relativité restreinte, c'est-à-dire en considérant l'espace-temps comme une variété différentielle *courbe*). L'élégante solution proposée par la relativité générale se résume alors comme suit : la quadrivitesse $\vec{U}(\tau)$ d'une particule en chute libre *est* effectivement *constante* le long de la trajectoire paramétrée en fonction du temps propre τ . En d'autres termes, les quadrivecteurs $\vec{U}(\tau + d\tau)$ et $\vec{U}(\tau)$ sont *parallèles* et toutes les composantes du quadrivecteur accélération sont égales à zéro par rapport à n'importe quel référentiel. Dans l'espace-temps courbe, une particule-test en chute libre se déplace selon un mouvement rectiligne et uniforme ! Elle tombe en vertu du principe d'inertie, indépendamment de la valeur de sa masse, de même qu'en mécanique galiléenne on ne s'étonne pas que deux particules de masses différentes, glissant sans frottement dans un plan horizontal, soient dotées de mouvements identiques pourvu qu'elles aient été lancées selon les mêmes

conditions initiales.

Ainsi se révèle une antipathie foncière entre l'espace-temps de Minkowski et le principe d'équivalence qui, à ce titre, fait office de trait d'union entre la relativité restreinte et une théorie relativiste de la gravitation. Il en découle un profond changement de perspective que l'on peut récapituler comme suit. Si l'on suppose la Terre inertielle (comme on l'admet souvent en première approximation), alors une pomme attachée à sa branche n'est pas accélérée pour Newton, mais bien pour Einstein. Une fois détachée, elle se trouve en chute libre, donc accélérée selon Newton mais pas selon Einstein.

2.4.2 Nature géométrique du champ de gravitation

Il résulte de la section 2.4.1 que dorénavant l'espace-temps doit être envisagé comme une entité *dynamique* recelant une part essentielle de l'information relative au champ gravitationnel. On ne peut que s'émerveiller du recul avec lequel Einstein a traité cette question : vu le brillant succès remporté par la relativité restreinte et son interprétation géométrique, il aurait paru naturel de chercher plutôt, dans le cadre de l'espace-temps de Minkowski, un acteur apte à rendre compte du champ de gravitation et qui fût compatible avec l'invariance relativiste (rôle que ne pouvait jouer le potentiel gravitationnel de Newton). Or, Einstein a pressenti qu'il fallait modifier la scène plutôt qu'inventer un nouvel acteur. Cette intuition est, du reste, accessible par trois voies directes.

1) La gravitation ne peut pas être décrite par un champ scalaire, vectoriel ou tensoriel défini en tout point de l'espace-temps de Minkowski ; car ses effets peuvent toujours être annulés localement par un changement de coordonnées, et ce champ devrait donc être nul. En relativité générale, le champ de gravitation n'est pas « plaqué » sur l'espace-temps : *il constitue la trame de l'espace-temps*.

2) Un deuxième argument découle du confinement des référentiels inertiels à un district local (pour rappel, sont désormais réputés *localement inertiels* les référentiels en chute libre). En effet, la question fondamentale de la dynamique n'est plus : « quelles sont les forces qui font accélérer telle particule par rapport à l'espace-temps absolu, lui-même figuré par les référentiels inertiels globaux ? » Non, la question devient : « quelles sont les forces qui font accélérer cette particule par rapport au référentiel inertiel local en chute libre ? » Autrement dit, « quelles sont les forces qui empêchent la particule de suivre sa tendance universelle par défaut, qui est de tomber ? » Comme il se doit, la réponse à cette question est identique selon que l'on envisage une particule au repos par rapport à (mettons) la Terre ou bien immobile dans une fusée accélérée loin de toute masse grave : dans les deux cas, il s'agit des forces de contact électromagnétiques. Ainsi, les forces de gravitation ont évacué le champ du discours ! Où ont-elles disparu ? La relativité générale assure sa cohérence en intégrant l'interaction gravitationnelle dans une scène d'un genre nouveau. L'espace-temps n'est plus absolu, il devient relatif à la matière.

3) Une nouvelle interprétation de l'espace-temps est d'ailleurs inévitable si on se rappelle les liens intimes que, sans innovation à cet égard par rapport à l'espace absolu de Newton, l'espace-temps absolu de Minkowski maintient avec les référentiels inertiels globaux. Ces derniers incarnent la référence permettant de distinguer la « gravitation » des « forces fictives d'inertie ». Le principe d'équivalence conteste cette distinction et réfute ainsi l'existence des référentiels inertiels. Dès lors, il invite à remplacer l'espace-temps euclidien par une structure géométrique plus large (ce sera l'enjeu du théorème 5.9 : un espace est euclidien si et seulement si on peut le rapporter à des référentiels inertiels globaux).

2.4.3 L'effet Einstein

D'après le principe d'équivalence et (2.12), la métrique (2.13) décrit un champ de gravitation *constant mais non homogène*. Son échelle d'inhomogénéité est de l'ordre de c^2/g . Si l'on prend pour g la valeur de l'accélération gravifique 9.8 m/s^2 à la surface de la Terre, on calcule $c^2/g \cong 10^{16} \text{ m}$, soit à peu près une année-lumière. D'autre part, l'échelle d'inhomogénéité du champ gravitationnel « réel » à la surface de notre planète est de l'ordre de son rayon $R \cong 6 \times 10^6 \text{ m}$.

Le champ (2.13) est donc bien plus homogène que le champ de pesanteur terrestre. Son inhomogénéité ne se marque qu'à des distances très grandes par rapport à R , et ne nous gêne pas pour effectuer une prédiction concernant le comportement de la lumière et des horloges dans une région suffisamment petite (par exemple, quelques dizaines de mètres) accessible à l'expérience. Etudions ces phénomènes.

L'horloge attachée à une particule ξ donnée est immobile dans le champ de gravitation décrit par la métrique (2.13). Dès lors, l'influence de l'altitude sur son comportement est régie par l'équation (2.15). De façon un peu lapidaire, on dira que « plus l'altitude est grande, plus vite le temps s'écoule ».

Pour donner un sens plus précis à cette affirmation, considérons la situation suivante, illustrée à la figure 2.3. En $\xi = \xi_1$, deux signaux lumineux sont émis dans le sens des ξ croissants (« vers le haut »), respectivement aux instants τ_1 et $\tau_1 + \Delta\tau$. Ils sont recueillis en $\xi = \xi_2$ aux instants respectifs τ_2 et $\tau_2 + \Delta\tau$ [la métrique étant stationnaire dans les coordonnées $(c\tau, \xi)$, le même intervalle de temps coordonnée $\Delta\tau$ sépare l'émission des deux signaux en $\xi = \xi_1$ et leur réception en $\xi = \xi_2$].⁶

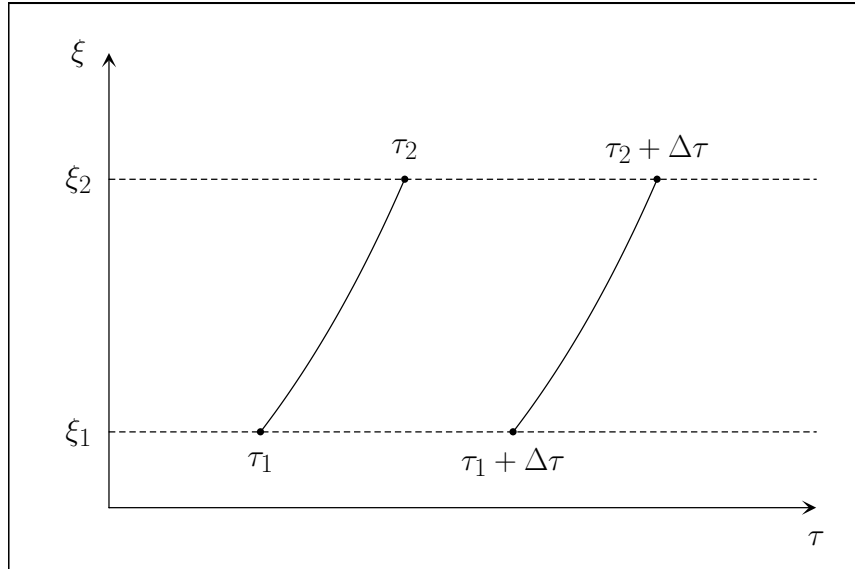


FIGURE 2.3: Propagation de signaux lumineux dans un champ de gravitation.

6. On peut d'ailleurs le vérifier par un calcul explicite. Comme l'équation de propagation d'un photon vérifie $ds^2 = 0$, on a

$$d\xi = \pm (1 + g\xi/c^2) c d\tau.$$

Dès lors, la courbe de gauche a pour équation $\xi(\tau) = (\xi_1 + c^2/g) \exp[g(\tau - \tau_1)/c] - c^2/g$ et celle de droite $\xi(\tau) = (\xi_1 + c^2/g) \exp[g(\tau - \Delta\tau - \tau_1)/c] - c^2/g$. Cela montre bien qu'elles sont translatées l'une de l'autre.

D'après (2.15), l'intervalle de temps propre séparant les deux émissions en $\xi = \xi_1$ vaut

$$(\Delta T)_1 = \left(1 + \frac{g\xi_1}{c^2}\right) \Delta\tau$$

tandis que l'intervalle de temps propre séparant les deux réceptions en $\xi = \xi_2$ vaut

$$(\Delta T)_2 = \left(1 + \frac{g\xi_2}{c^2}\right) \Delta\tau.$$

Une onde électromagnétique se déplaçant dans le champ gravitationnel (2.13) subit donc un décalage en fréquence (*effet Einstein*) donné par

$$\frac{\nu_1}{\nu_2} = \frac{1 + \frac{g\xi_2}{c^2}}{1 + \frac{g\xi_1}{c^2}}. \quad (2.17)$$

Si le rayon lumineux se propage vers le haut, il s'agit d'un décalage vers le rouge ; dans le cas contraire, il y a décalage vers le bleu.

De façon générale, on a

$$\nu(h) \left(1 + \frac{gh}{c^2}\right) = \nu_0 \quad (2.18)$$

où ν_0 est une constante. Ce résultat fut vérifié dès 1959 par R. Pound et G. Rebka suite à la découverte récente de l'effet Mössbauer. Un cristal de ^{57}Fe émet ou absorbe des raies γ très fines à 14.4 keV, autorisant des mesures extrêmement précises. L'expérience fut réalisée au laboratoire Jefferson de l'Université de Harvard, haut de 22.6 mètres. Pour cette différence d'altitude on doit avoir $|\Delta\nu/\nu| = gh/c^2 = 2.5 \times 10^{-15}$. Les radiations émises en haut du bâtiment étaient absorbées par un cristal de même nature placé au pied de l'immeuble et en mouvement judicieux vers le bas : le décalage gravitationnel vers le bleu était compensé par l'effet Doppler, ce qui a permis de l'évaluer.

Comme l'énergie d'un complexe lumineux se transforme de façon identique à sa fréquence (section 1.9), on en déduit que dans notre champ de gravitation l'énergie $E(h)$ d'un rayonnement mesurée à l'altitude h est soumise à la relation

$$E(h) \left(1 + \frac{gh}{c^2}\right) = E_0 \quad (2.19)$$

où E_0 est une constante. Pour que le principe de conservation de l'énergie soit respecté, il faut donc ajouter à l'énergie E du rayonnement une énergie potentielle de pesanteur égale à

$$\frac{E(h)}{c^2} gh,$$

que l'on peut interpréter comme l'énergie potentielle d'une masse pesante $m = E/c^2$ placée dans le potentiel gravifique $\Phi(h) = gh$. On peut ainsi écrire la relation

$$m(h)c^2 + m(h)\Phi(h) = E_0,$$

généralisant l'équation newtonienne

$$\frac{1}{2} mV^2 + m\Phi(h) = E_0.$$

Enfin, on peut combiner l'effet Einstein, dû à l'altitude, avec le phénomène de dilatation du temps résultant de la vitesse. Ajoutant deux axes y et z orthogonaux au plan (ct, x) défini précédemment (ce qui donne les degrés de liberté nécessaires pour se mouvoir à une altitude constante dans le champ de gravitation), l'élément de longueur (2.13) devient

$$ds^2 = - \left(1 + \frac{g\xi}{c^2} \right)^2 c^2 d\tau^2 + d\xi^2 + dy^2 + dz^2. \quad (2.20)$$

Si T désigne le temps propre d'une particule, on obtient

$$dT^2 = d\tau^2 \left[\left(1 + \frac{g\xi}{c^2} \right)^2 - \frac{d\xi^2 + dy^2 + dz^2}{c^2 d\tau^2} \right],$$

c'est-à-dire

$$dT = d\tau \left[1 + \frac{g\xi}{c^2} - \frac{V^2}{2c^2} + O(1/c^4) \right] \quad (2.21)$$

où V désigne la norme de la vitesse spatiale de la particule. Le résultat (2.21) sera entériné (avec d'autres notations) par la relation (7.52).

2.5 Champ de gravitation uniforme et constant

Lorsque $g\xi_1 \ll c^2$ et $g\xi_2 \ll c^2$, le rapport des fréquences dans (2.17) ne dépend que de la différence d'altitude entre le point émetteur et le point récepteur :

$$\frac{\nu_1}{\nu_2} \cong 1 + \frac{g}{c^2} (\xi_2 - \xi_1).$$

Cette remarque nous suggère la forme du tenseur métrique apte à décrire un champ de gravitation non seulement constant mais également *homogène* (ou, si l'on préfère, un « référentiel uniformément accéléré ») : pour que chaque point d'abscisse ξ soit équivalent à l'origine $\xi = 0$, l'élément de longueur doit être du type

$$ds^2 = -f^2(\xi) c^2 d\tau^2 + d\xi^2,$$

où $f(\xi)$ vérifie les deux conditions suivantes :

$$\begin{aligned} \frac{f(\xi_1)}{f(\xi_2)} &= f(\xi_1 - \xi_2) \\ f(\xi) &= 1 + \frac{g\xi}{c^2} + O\left(\frac{g\xi}{c^2}\right)^2. \end{aligned}$$

On en déduit l'équation différentielle

$$\frac{f(\xi + d\xi) - f(\xi)}{d\xi} = \frac{g}{c^2} f(\xi)$$

de solution

$$f(\xi) = \exp\left(\frac{g\xi}{c^2}\right).$$

En résumé, ces réflexions heuristiques nous amènent à retenir la métrique

$$ds^2 = - \exp\left(\frac{2g\xi}{c^2}\right) c^2 d\tau^2 + d\xi^2, \quad \mathcal{R} = \frac{c^2}{g} \quad (2.22)$$

comme modèle didactique pour la description relativiste du champ gravitationnel le plus simple, c'est-à-dire homogène et constant.

2.6 Le paradoxe de Langevin

Posons d'abord le problème. Imaginons deux frères jumeaux, l'un immobile dans un référentiel inertiel donné et l'autre en mouvement. Approximons la trajectoire de ce dernier par une ligne coudée : nous le supposons en translation rectiligne et uniforme de vitesse V pendant la première partie de son voyage, et de vitesse $-V$ ensuite. Le paradoxe est le suivant : chacun des jumeaux a le droit de se considérer immobile et d'affirmer que c'est son frère qui se déplace. Chacun peut donc décrier que l'autre vieillit moins vite (dilatation des durées). Mais, lorsqu'ils se retrouveront à l'issue du voyage, il faudra bien que l'un des deux ait raison.

Bien sûr, un dessin dans l'espace-temps de Minkowski permet de comparer les longueurs des deux trajectoires, donc leurs temps propres, et de conclure immédiatement que c'est le frère « au repos » qui a raison. L'argument de symétrie exposé ci-dessus est trop superficiel et ne prend pas en considération le fait qu'un des deux jumeaux (celui qui est au repos) se trouve en permanence immobile dans un, et un seul, référentiel inertiel ; mais pas l'autre (celui qui voyage). *A priori*, seules les conclusions du frère immobile sont donc légitimes.

Cependant, afin d'amorcer un pas vers la philosophie de la relativité générale, il est intéressant de se glisser dans la peau du voyageur et de voir comment celui-ci raisonne. La résolution du paradoxe implique donc d'analyser en détail les conséquences de l'existence du coude.

Appelons τ le temps propre au bout duquel le voyageur fait demi-tour. A ce moment-là, l'intersection de sa ligne de simultanéité avec l'axe ct passe subitement du point A au point B . Il omet ainsi de prendre en compte les événements de la vie de son frère compris entre ces deux points (figure 2.4).

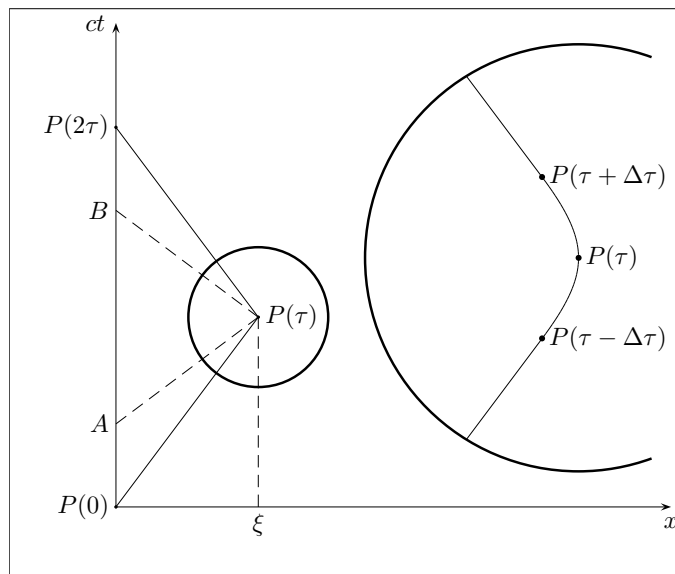


FIGURE 2.4: Trajectoire d'un voyageur P en mouvement rectiligne uniforme, sauf au niveau du coude où il passe instantanément de la vitesse V à la vitesse $-V$. Son frère, quant à lui, parcourt l'axe ct . À droite, on a représenté une situation plus réaliste où le voyageur passe progressivement de la vitesse V à la vitesse $-V$ avec une accélération constante (le cercle de droite constitue un agrandissement de celui de gauche).

On calcule facilement :

$$A = \frac{c\tau}{\gamma}, \quad B = 2c\tau\gamma - \frac{c\tau}{\gamma}, \quad \xi = V\tau\gamma, \quad \text{avec } \gamma = \frac{1}{\sqrt{1 - V^2/c^2}}.$$

Le voyageur néglige ainsi la durée de vie

$$\frac{|AB|}{c} = \frac{2V^2\tau\gamma}{c^2}$$

du jumeau resté au sol. C'est la raison pour laquelle celui-ci ne semble avoir vieilli, au moment des retrouvailles, que de la quantité

$$2 \frac{A}{c} = 2 \frac{\tau}{\gamma} < 2\tau.$$

On peut étudier continûment la situation en « émoussant » le coude et en le remplaçant par une trajectoire uniformément accélérée de mesure g pendant le temps propre $2\Delta\tau$. Du point de vue du voyageur, son frère se trouve alors en chute libre dans un champ de gravitation homogène pointant vers les x croissants, avec une trajectoire culminant à l'altitude $\xi = V\tau\gamma$.

Supposons, pour simplifier les calculs, que $V \ll c$. Dans ce cas, en notant $O(n)$ les termes en $(V/c)^n$:

$$\xi = V\tau [1 + O(2)], \quad \Delta\tau = \frac{V}{g} [1 + O(2)] \quad (2.23)$$

[relation (2.6)]. De plus, pendant toute la phase accélérée, le jumeau resté au sol se trouve approximativement à l'altitude constante ξ . Il vieillira donc d'un temps propre

$$2\Delta T = \left[1 + \frac{g\xi}{c^2} + O(2)\right] 2\Delta\tau \quad (2.24)$$

[relation (2.15)]. Supposons, en outre, que τ soit suffisamment grand pour que

$$\frac{c^2}{gV\tau} = O(2). \quad (2.25)$$

Compte tenu de (2.23) et (2.25), (2.24) devient :

$$2\Delta T = \frac{2V^2\tau}{c^2} [1 + O(2)].$$

A l'issue du voyage, le jumeau sédentaire aura donc vieilli de

$$2\tau \left[1 - \frac{V^2}{2c^2} + \frac{V^2}{c^2} + O(4)\right] = 2\tau \left[1 + \frac{V^2}{2c^2} + O(4)\right] > 2\tau$$

et se trouvera plus âgé que son frère voyageur. Le paradoxe est résolu.

2.7 Les limites de la relativité générale

Nous en savons assez pour pressentir quand la relativité générale doit remplacer la théorie de Newton. L'équation (2.13) suggère que les effets relativistes sont importants si le potentiel gravitationnel est de l'ordre de c^2 . (Cette intuition se trouvera corroborée au

chapitre 9, lorsque nous aurons établi la solution de Schwarzschild.) En d'autres termes, pour calculer les effets gravitationnels engendrés par une masse m contenue dans une région de dimension caractéristique l , il faut utiliser la relativité générale si

$$l \leq l_S = \frac{Gm}{c^2}, \quad (2.26)$$

en notant G la constante de Cavendish. Remarquons que la ligne de démarcation entre les deux théories est déterminée par une vitesse, et non une accélération. L'expérience ne semble d'ailleurs pas appuyer l'existence d'une accélération limite. Du reste, nous savions déjà que le concept d'accélération a perdu le statut que lui accordait la physique newtonienne (l'accélération d'un corps varie selon le repère inertiel par rapport auquel on la considère).

D'autre part, pour observer un objet de masse m et de dimension typique l , on peut l'éclairer avec de la lumière de longueur d'onde λ . La diffraction est négligeable si $\lambda \ll l$. Apparemment donc, on pourrait forcer un objet quelconque à réfléchir la lumière qu'on lui envoie, pourvu que l'on utilise une longueur d'onde suffisamment courte. Sauf que, au-delà d'une certaine fréquence, les photons utilisés se trouveront assez énergiques pour créer des paires particule-antiparticule de masse m ! Cela se produira pour une longueur d'onde λ_C telle que $\hbar c/\lambda_C \simeq mc^2$, où \hbar est la constante de Planck. C'est la *longueur d'onde de Compton*. Une particule de masse m et de dimension

$$l \leq l_C = \frac{\hbar}{mc} \quad (2.27)$$

doit être décrite par la théorie quantique des champs.

A l'échelle astronomique on a $l_S \gg l_C$. Pour les électrons, protons etc., on a $l_C \gg l_S$. En combinant (2.26) et (2.27), on détermine les échelles de masse et de dimension à partir desquelles la relativité générale et la théorie quantique des champs doivent simultanément être prises en compte :

$$l \simeq \frac{Gm}{c^2} \simeq \frac{\hbar}{mc}.$$

On en déduit

$$\begin{aligned} m \simeq m_P &= \sqrt{\frac{\hbar c}{G}} = 2.18 \times 10^{-8} \text{ kg}, \\ l \simeq l_P &= \sqrt{\frac{\hbar G}{c^3}} = 1.62 \times 10^{-35} \text{ m}. \end{aligned} \quad (2.28)$$

Les quantités m_P et l_P , appelées respectivement *la masse et la longueur de Planck*, caractérisent une physique encore totalement inconnue. La longueur de Planck vaut essentiellement le rayon de Schwarzschild d'un trou noir de masse égale à la masse de Planck : c'est la masse du plus petit trou noir qui puisse être décrit sans recourir à la théorie quantique des champs.

2.8 Courbure des trajectoires en théorie newtonienne de la gravitation

Considérons une particule astreinte à se mouvoir dans un champ de gravitation newtonien, homogène et constant (figure 2.5). Notons V_x et V_z les composantes x et z de la

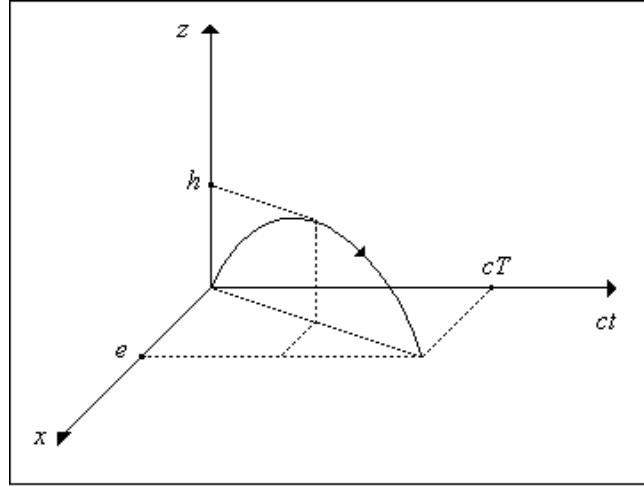


FIGURE 2.5: Trajectoire parabolique d'une particule dans un champ de gravitation newtonien, homogène et constant.

vitesse initiale. La trajectoire est caractérisée par un temps de chute

$$cT = 2cV_z/g,$$

une hauteur maximale atteinte

$$h = (V_z)^2/(2g)$$

et une distance horizontale parcourue

$$e = 2V_x V_z/g.$$

Si $V_x \ll c$, alors $e \ll cT$ et la longueur $\sqrt{e^2 + c^2 T^2}$ de la projection de la trajectoire spatio-temporelle dans le plan (x, ct) est à peu près égale à cT . D'autre part, si $V_z \ll c$ on voit que $h \ll cT$. Sous ces hypothèses de mouvement non relativiste, il est possible d'approximer la trajectoire par un arc de cercle⁷, comme le montre la figure 2.6.

En effet, on a

$$h = R[1 - \cos(\theta/2)] \cong R\theta^2/8$$

$$cT = 2R \sin(\theta/2) \cong R\theta$$

et, par conséquent,

$$R \cong \frac{c^2 T^2}{8h}, \quad \theta \cong \frac{8h}{cT}.$$

Combinant ce résultat avec les valeurs données plus haut pour h et cT , on trouve

$$R \cong \frac{c^2}{g}, \quad \theta \cong \frac{2V_z}{c}.$$

Le rayon du cercle ne dépend que du champ de gravitation, et pas des conditions initiales!⁸ Dans le cas de la Terre, il vaut à peu près une année-lumière. Attention, l'espace-temps

7. Il s'agit du cercle osculateur au sommet de la trajectoire, notion qui sera précisée au chapitre 4.

8. C.W. Misner, K.S. Thorne, J.A. Wheeler, *op. cit.*, Freeman, 1973, pp. 32-33.

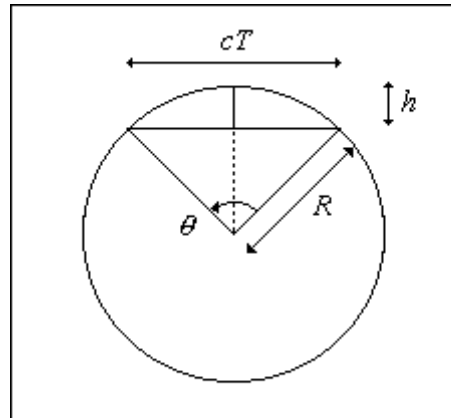


FIGURE 2.6: Si $V_z \ll c$ alors $h \ll cT$ et donc $\theta \ll 1$.

euclidien à trois dimensions dans lequel cette quantité est évaluée n'est pas l'espace-temps de Minkowski : en effet, le temps y est traité sur un pied d'égalité avec les dimensions spatiales (utilisation du théorème de Pythagore). Néanmoins, on retrouve la longueur caractéristique associée au mouvement d'une particule uniformément accélérée dans l'espace-temps de Minkowski, étudié ci-dessus. Rappelons que tous les points de la trajectoire sont alors situés, sans approximation, sur une branche d'*hyperbole* (c'est-à-dire un « cercle » dans l'espace-temps de Minkowski), à la distance c^2/g de son centre.

Chapitre 3

Le carrousel d'Einstein

3.1 Position du problème

Soit l'espace-temps de Minkowski à deux dimensions spatiales, rapporté à des coordonnées cartésiennes (ct', x', y') . Dans le système des coordonnées polaires (r', φ') , l'élément de longueur s'écrit

$$ds^2 = -c^2 dt'^2 + dr'^2 + r'^2 d\varphi'^2. \quad (3.1)$$

Supposons qu'une plate-forme tournante (figure 3.1) se meuve avec la distribution de vitesses orthoradiale $\vec{V}(r')$ de module

$$V(r') = \omega r', \quad r' < c/\omega. \quad (3.2)$$

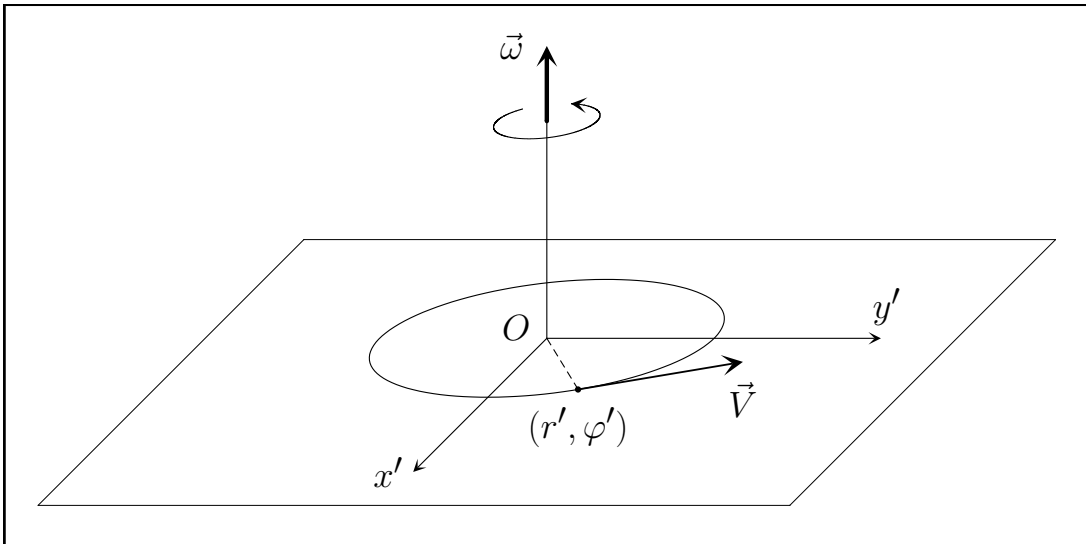


FIGURE 3.1: *Le carrousel d'Einstein. On y associe souvent le nom d'Ehrenfest.*

Nous nous proposons de doter ce carrousel d'un système de coordonnées et d'y étudier le ds^2 , en insistant sur le fait que la question ne sera envisagée ici que du point de vue *cinématique*. Nous ne nous soucierons pas de savoir comment la loi de distribution de vitesses (3.2)

peut être obtenue pratiquement. Ce point est, en effet, beaucoup plus délicat en physique relativiste qu'en physique newtonienne, où l'on peut admettre une interaction instantanée entre l'axe de rotation et les différentes particules solidaires constituant le disque solide. Nous ne nous traiterons donc pas de la *dynamique* du problème [par exemple, la répartition des tensions à l'intérieur du carrousel pour la loi (3.2)] : il nous suffit d'admettre que le disque est constitué de petits éléments rigides dont chacun a reçu la consigne de se déplacer selon la loi (3.2). En d'autres termes, l'objet de notre étude n'est pas tant le *corps matériel rigide en rotation* que le *mouvement de rotation rigide*.

Ce problème a joué un rôle important dans l'édification de la relativité générale. Einstein en parle dans plusieurs de ses écrits fondamentaux.¹ Il faut en effet remarquer qu'on y voit apparaître certains concepts de base de la théorie :

- utilisation de coordonnées générales, c'est-à-dire reliées à des coordonnées cartésiennes (inertielles) par une transformation non linéaire ;
- expression du tenseur métrique dans ces coordonnées ;
- en liaison avec cela, présence d'un champ de gravitation à bord de la plate-forme ;
- courbure (des sections spatiales).

Soit le changement de variables

$$t = t', \quad r = r', \quad \varphi = \varphi' - \omega t'. \quad (3.3)$$

Il permet d'associer des coordonnées (r, φ) constantes à toute particule fixe par rapport à la plate-forme tournante. Ces coordonnées, dites *comobiles*, sont commodes pour décrire les événements à bord de celle-ci. Cependant, t n'a manifestement pas de sens physique immédiat car nous savons que le temps propre marqué par une horloge dépend de sa vitesse, donc de sa distance à l'origine. D'autre part, on doit s'attendre à ce que la contraction des longueurs intervienne dans l'expression de la distance spatiale entre les particules. Il sera donc nécessaire d'interpréter les mesures physiques de temps et d'espace dans les coordonnées (3.3). Ce sera l'objet de la section 3.2.

Mais on peut dès à présent anticiper des subtilités résultant de la relativité de la simultanéité : par exemple, la synchronisation d'horloges s'avère impossible sur l'entièreté d'un cercle à distance R de l'origine ; sinon, on aboutirait à une contradiction. Admettons en effet que la transformation de Lorentz soit applicable localement dans la direction tangente au cercle. Alors, à un instant donné dans le référentiel « fixe », les diverses horloges du carrousel seraient censées accumuler, de proche en proche, des décalages temporels par rapport à une horloge de départ à laquelle on revient pourtant après avoir effectué un tour complet ! La figure 3.2 résume et éclaire la situation.

1. Voir notamment

A. Einstein, *La relativité*, Petite Bibliothèque Payot, Gauthier-Villars, 1963, chapitre 23 ;
 A. Einstein, *Quatre conférences sur la théorie de la relativité*, Gauthier-Villars, 1971, pp. 52-55 ;
 R. Adler, M. Bazin, M. Schiffer, *Introduction to general relativity*, International Student Edition, McGraw-Hill Kogakusha, 1975, pp. 120-124 ;
 L. Landau, E. Lifchitz, *Physique théorique, Tome 2, Théorie des champs*, Mir, 1989, § 89 ;
 G. Pascoli, *La gravitation, Que Sais-je ? 2489*, Presses Universitaires de France, 1995, pp. 49-52 ;
 J. Stachel, *The Rigidly Rotating Disk as the "Missing Link" in the History of General Relativity*, dans *Einstein from 'B' to 'Z'*, Einstein Studies, Volume 9, Birkhäuser, 2002, pp. 245-260.

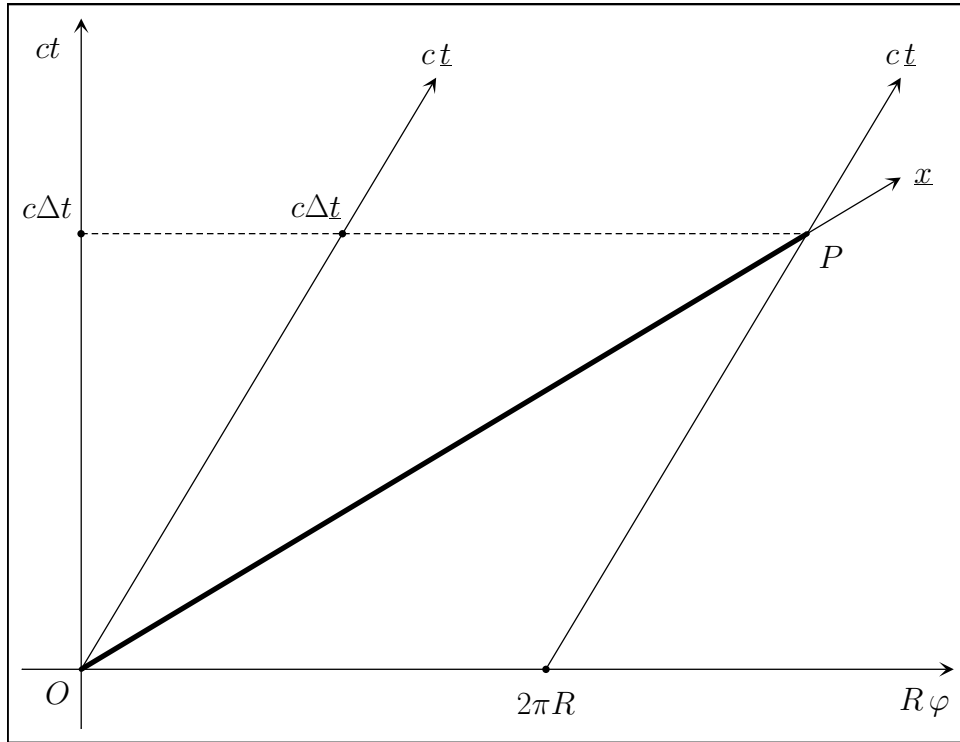


FIGURE 3.2: *Diagramme d'espace-temps pour les particules situées sur la circonférence de rayon R . Leur vitesse vaut $V = \omega R$ dans le référentiel inertiel (de coordonnée temporelle ct) par rapport auquel tourne le carrousel. Pour fixer les idées on a pris $V/c = 3/5$. Il faut imaginer que l'axe des abscisses est enroulé autour de la verticale de façon à faire coïncider les lignes d'univers (notées $c\bar{t}$) des particules d'abscisse $\varphi = 0$ et $\varphi = 2\pi$. L'axe des \underline{x} matérialise les événements simultanés, dans le référentiel du carrousel, à l'événement O situé à l'origine : on constate immédiatement que la synchronisation des horloges le long d'une ligne spatiale fermée est impossible, puisque $P \neq O$. La mesure de la circonférence est égale à $2\pi R$ dans le référentiel inertiel mais, mesurée dans le référentiel du carrousel (en gras sur l'axe des \underline{x}), elle prend la valeur $2\pi R/\sqrt{1 - V^2/c^2}$ en vertu de la contraction de l'espace. On en déduit le « défaut de synchronisation » $c\Delta\bar{t} = 2\pi RV/[c\sqrt{1 - V^2/c^2}]$. Il en résulte également $c\Delta t = 2\pi RV/[c(1 - V^2/c^2)]$, conformément à la relation (3.14) qui sera démontrée plus loin.*

3.2 Distances et durées physiques

Partant de (3.1), nous obtenons le ds^2 de l'espace-temps (toujours euclidien) dans les coordonnées (t, r, φ) :

$$ds^2 = - \left(1 - \frac{\omega^2 r^2}{c^2} \right) c^2 dt^2 + 2\omega r^2 d\varphi dt + dr^2 + r^2 d\varphi^2. \quad (3.4)$$

Remarquons d'ores et déjà que, à cause du terme en $d\varphi dt$, les lignes de temps ne sont pas orthogonales aux sections spatiales. Cependant, on peut récrire la métrique sous la forme

$$ds^2 = - \left(c\sqrt{1 - \frac{\omega^2 r^2}{c^2}} dt - \frac{\omega r^2}{c\sqrt{1 - \omega^2 r^2/c^2}} d\varphi \right)^2 + dr^2 + \frac{r^2}{1 - \omega^2 r^2/c^2} d\varphi^2. \quad (3.5)$$

La façon (3.5) de regrouper les termes constitue un cas particulier d'un processus général. En effet, dans la relation

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta \quad (3.6)$$

exprimant une métrique en coordonnées quelconques, on peut toujours regrouper les termes de la façon suivante ($i, j = 1, 2, 3$) :

$$\begin{aligned} ds^2 &= g_{00} (dx^0)^2 + 2g_{0i} dx^0 dx^i + g_{ij} dx^i dx^j \\ &= - \left(\sqrt{-g_{00}} dx^0 - \frac{g_{0i}}{\sqrt{-g_{00}}} dx^i \right)^2 + \left(g_{ij} - \frac{g_{0i}g_{0j}}{g_{00}} \right) dx^i dx^j. \end{aligned} \quad (3.7)$$

Nous allons montrer que les deux termes apparaissant dans cette formulation sont dotés d'une signification physique [découpage spatio-temporel de l'espace-temps dans un référentiel réalisé par des corps matériels, c'est-à-dire tel que chaque triplet de coordonnées (x^1, x^2, x^3) joue le rôle de coordonnées spatiales comobiles attribuées à une particule donnée].

Il existe un opérateur de projection sur l'espace orthogonal à un vecteur temporel \vec{U} (en l'occurrence, ici, la quadrivitesse d'un observateur donné). Autrement dit, on peut décomposer un vecteur quelconque \vec{W} en un vecteur parallèle et un vecteur perpendiculaire à \vec{U} :

$$\vec{W} = \vec{W}^\parallel + \vec{W}^\perp,$$

où (le point représente le produit scalaire de Minkowski)

$$\begin{aligned} \vec{W}^\parallel &= -\frac{1}{c^2} (\vec{W} \cdot \vec{U}) \vec{U}, \\ \vec{W}^\perp &= \vec{W} - \vec{W}^\parallel = \vec{W} + \frac{1}{c^2} (\vec{W} \cdot \vec{U}) \vec{U}. \end{aligned} \quad (3.8)$$

On a bien

$$\vec{W}^\perp \cdot \vec{U} = \vec{W} \cdot \vec{U} + \frac{1}{c^2} \vec{W} \cdot \vec{U} (-c^2) = 0.$$

L'équation (3.6) peut ainsi s'exprimer sous la forme

$$\begin{aligned} ds^2 &= d\vec{x} \cdot d\vec{x} \\ &= d\vec{x}^\parallel \cdot d\vec{x}^\parallel + d\vec{x}^\perp \cdot d\vec{x}^\perp \\ &= -\frac{1}{c^2} (d\vec{x} \cdot \vec{U})^2 + \left[d\vec{x} \cdot d\vec{x} + \frac{1}{c^2} (d\vec{x} \cdot \vec{U})^2 \right]. \end{aligned} \quad (3.9)$$

Par construction, le deuxième terme représente les distances spatiales mesurées dans le référentiel d'un observateur de quadrivitesse \vec{U} ; et son complémentaire, le premier terme, correspond aux intervalles de temps mesurés par cet observateur.

Si l'observateur est « immobile », c'est-à-dire si ses coordonnées spatiales x^i sont constantes, alors seule la composante temporelle $U^0 = dx^0/d\tau$ de sa quadrivitesse est non nulle. Pour la calculer, il suffit de se rappeler que

$$-c^2 = g_{\alpha\beta} U^\alpha U^\beta = g_{00} (U^0)^2.$$

Il s'ensuit (en supposant que x^0 croît avec τ)

$$(U^\alpha) = \frac{c}{\sqrt{-g_{00}}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Par conséquent,

$$-\frac{1}{c^2} \left(d\vec{x} \cdot \vec{U} \right)^2 = - \left(\sqrt{-g_{00}} dx^0 - \frac{g_{0i}}{\sqrt{-g_{00}}} dx^i \right)^2,$$

ce qui justifie l'équivalence entre (3.9) et (3.7). On vérifiera d'ailleurs aisément que

$$\begin{aligned} d\vec{x} \cdot d\vec{x} + \frac{1}{c^2} \left(d\vec{x} \cdot \vec{U} \right)^2 &= \left(g_{\alpha\beta} - \frac{g_{0\alpha}g_{0\beta}}{g_{00}} \right) dx^\alpha dx^\beta \\ &= \left(g_{ij} - \frac{g_{0i}g_{0j}}{g_{00}} \right) dx^i dx^j. \end{aligned}$$

Les propriétés géométriques des sections spatiales physiques d'une métrique quelconque sont ainsi conditionnées par les quantités

$$\gamma_{ij} = g_{ij} - \frac{g_{0i}g_{0j}}{g_{00}}.$$

Il est facile de démontrer qu'elles constituent un *tenseur tridimensionnel* pour les transformations spatiales de coordonnées. Par ailleurs, la matrice des γ_{ij} est hermitienne définie positive, puisque la forme quadratique $dl^2 = \gamma_{ij} dx^i dx^j$ doit être strictement positive $\forall dx^1, dx^2, dx^3$.

D'autre part, on qualifera de *simultanés* deux événements infiniment voisins $\mathcal{X}(x^0, x^i)$ et $\mathcal{Y}(x^0 + dx^0, x^i + dx^i)$ tels que

$$\sqrt{-g_{00}} dx^0 - \frac{g_{0i}}{\sqrt{-g_{00}}} dx^i = 0. \quad (3.10)$$

Cette définition locale de la simultanéité permet de synchroniser des horloges de proche en proche. Cependant, si ce processus est toujours possible le long d'une ligne ouverte arbitraire, il est par contre génériquement impossible sur une ligne fermée. En d'autres termes,

$$\begin{aligned} \oint dx^0 &= \oint -\frac{g_{0i}}{g_{00}} dx^i \\ &= \Delta x^0. \end{aligned} \quad (3.11)$$

A priori cette quantité est non nulle.²

Appliquée à la métrique (3.4), la condition de simultanéité (3.10) s'écrit

$$\sqrt{1 - \omega^2 r^2 / c^2} dt - \frac{\omega r^2}{c^2 \sqrt{1 - \omega^2 r^2 / c^2}} d\varphi = 0. \quad (3.13)$$

2. Signalons une exception pouvant apparaître dans le contexte des métriques stationnaires, c'est-à-dire indépendantes de la coordonnée temporelle x^0 . En effet, s'il existe une fonction $f(x^1, x^2, x^3)$ telle que

$$\frac{g_{0i}}{g_{00}} dx^i = df, \quad (3.12)$$

alors $\Delta x^0 = 0$. En général une telle fonction n'existe pas : une relation du type (3.12) exige l'égalité des dérivées croisées, comme le prescrit la théorie de la primitivation dans \mathbb{R}^n . Mais si c'est le cas, le changement de variable temporelle $X^0 = x^0 + f(x^i)$ (consistant à adapter l'origine des temps en tout point spatial) ramène la métrique à la forme statique

$$ds^2 = g_{00} (dX^0)^2 + \gamma_{ij} dx^i dx^j.$$

Ce résultat est très naturel. Considérons en effet, dans le référentiel « fixe », deux événements de coordonnées respectives (t', R, φ') et $(t' + dt', R, \varphi' + d\varphi')$; et admettons que la transformation de Lorentz soit applicable localement dans la direction tangente au cercle de rayon R . Alors, l'intervalle de temps physique séparant les deux événements, mesuré dans le référentiel en rotation, est donné par

$$\frac{dt' - \frac{\omega R}{c^2} R d\varphi'}{\sqrt{1 - \omega^2 R^2/c^2}} = \sqrt{1 - \omega^2 R^2/c^2} dt - \frac{\omega R^2}{c^2 \sqrt{1 - \omega^2 R^2/c^2}} d\varphi,$$

en accord avec (3.13). Utilisant (3.13), on peut calculer le « défaut de synchronisation » $\Delta x^0 = c\Delta t$ par la relation (3.11) lorsqu'on intègre sur une circonférence de rayon R :

$$c\Delta t = \frac{2\pi\omega R^2}{c(1 - \omega^2 R^2/c^2)}. \quad (3.14)$$

3.3 Application : la transformation de Lorentz

Les conclusions du raisonnement de la section 3.2 donnant les distances spatiales physiques dans un système de coordonnées générales, ainsi que le critère de simultanéité entre deux événements, permettent de retrouver simplement la transformation de Lorentz.

Soit un référentiel inertiel $S(ct, x)$. Considérons un autre référentiel \underline{S} en translation uniforme de vitesse V par rapport à S , matérialisé par des particules d'abscisse constante $x' = x - Vt$; et choisissons-y pour coordonnées provisoires, sans signification physique,

$$t' = t, \quad x' = x - Vt. \quad (3.15)$$

On a

$$\begin{aligned} ds^2 &= -c^2 dt^2 + dx^2 \\ &= -\left(c dt' \sqrt{1 - V^2/c^2} - dx' \frac{V/c}{\sqrt{1 - V^2/c^2}} \right)^2 + \left(\frac{dx'}{\sqrt{1 - V^2/c^2}} \right)^2. \end{aligned}$$

Deux événements (t', x') et $(t' + dt', x' + dx')$ seront *simultanés* dans \underline{S} si

$$c dt' \sqrt{1 - V^2/c^2} - dx' \frac{V/c}{\sqrt{1 - V^2/c^2}} = 0.$$

Cette relation prend la forme $d\underline{t} = 0$ si l'on pose

$$d\underline{t} = dt' \sqrt{1 - V^2/c^2} - dx' \frac{V/c^2}{\sqrt{1 - V^2/c^2}},$$

qui est intégrable et permet de définir une coordonnée temporelle physique \underline{t} reliée linéairement à t' et x' :

$$\underline{t} = t' \sqrt{1 - V^2/c^2} - x' \frac{V/c^2}{\sqrt{1 - V^2/c^2}}. \quad (3.16)$$

D'autre part, et toujours dans \underline{S} , la *distance physique* entre deux événements simultanés de coordonnées spatiales respectives x' et $x' + dx'$ est donnée par

$$d\underline{x} = \frac{dx'}{\sqrt{1 - V^2/c^2}},$$

également intégrable :

$$\underline{x} = \frac{x'}{\sqrt{1 - V^2/c^2}}. \quad (3.17)$$

Les relations (3.15) peuvent être injectées dans (3.16) et (3.17). On est ainsi naturellement amené à introduire dans \underline{S} des *coordonnées physiques* \underline{t} et \underline{x} reliées aux coordonnées t et x par

$$\underline{t} = \frac{t - Vx/c^2}{\sqrt{1 - V^2/c^2}}, \quad \underline{x} = \frac{x - Vt}{\sqrt{1 - V^2/c^2}}, \quad (3.18)$$

c'est-à-dire la transformation de Lorentz.

3.4 Mesures intrinsèques sur les sections spatiales

A bord de la plate-forme tournante de métrique (3.5), l'élément de longueur sur les sections spatiales vaut

$$dl^2 = dr^2 + \frac{r^2}{1 - \omega^2 r^2/c^2} d\varphi^2 \quad (0 \leq r < c/\omega). \quad (3.19)$$

La longueur totale d'un cercle de rayon R pour cette métrique vaut

$$\begin{aligned} L &= \int_0^{2\pi} \frac{R}{\sqrt{1 - \omega^2 R^2/c^2}} d\varphi \\ &\cong 2\pi R \left(1 + \frac{\omega^2 R^2}{2c^2} \right) \quad \text{si } \omega R \ll c. \end{aligned}$$

Ainsi, à bord de la plate-forme, le rapport de la circonférence d'un cercle centré à l'origine à son diamètre est supérieur au nombre π .

Calculons de même l'aire d'un disque de rayon R pour la métrique (3.19) :

$$\begin{aligned} A &= \int_0^R dr \int_0^{2\pi} \frac{r}{\sqrt{1 - \omega^2 r^2/c^2}} d\varphi \\ &= \frac{2\pi c^2}{\omega^2} \left(1 - \sqrt{1 - \frac{\omega^2 R^2}{c^2}} \right) \\ &\cong \pi R^2 \left(1 + \frac{\omega^2 R^2}{4c^2} \right) \quad \text{si } \omega R \ll c. \end{aligned}$$

De ces calculs on peut déjà conclure que la métrique (3.19) caractérise un espace *courbe* (non euclidien).

3.5 Le point de vue lagrangien

Une particule libre se meut en ligne droite. Or, le temps propre mesuré entre les deux extrémités A et B d'un segment temporel est supérieur au temps propre mesuré le long de n'importe quelle autre courbe joignant A et B (on pourra s'en persuader en considérant des lignes brisées de part et d'autre du segment). Il est donc naturel de considérer les équations du mouvement d'une particule libre comme résultant du principe variationnel

$$\delta \int_A^B d\tau = 0,$$

c'est-à-dire, en coordonnées cartésiennes,

$$\delta \int_{t_A}^{t_B} \sqrt{1 - V^2/c^2} dt = 0.$$

En première approximation, l'intégrand vaut $1 - V^2/2c^2$. Le terme indépendant ne contribue pas puisque les extrémités sont fixes. Pour retrouver asymptotiquement le lagrangien $m_0 V^2/2$ de la mécanique classique, on posera donc

$$L = -m_0 c^2 \sqrt{1 - V^2/c^2}$$

(Planck, 1906). Dans des coordonnées générales, le principe d'action extrémale s'écrira ainsi

$$0 = -m_0 c \delta \int_A^B \sqrt{-g_{\mu\nu}(x^\alpha) dx^\mu dx^\nu} = -m_0 c \delta \int_{\lambda_A}^{\lambda_B} \sqrt{-g_{\mu\nu}(x^\alpha) \dot{x}^\mu \dot{x}^\nu} d\lambda,$$

où λ est un paramètre quelconque (on peut garder la coordonnée temporelle x^0 mais ce choix ne s'impose plus naturellement) et où le point désigne la dérivée par rapport à λ . On a donc

$$L(x^\alpha, \dot{x}^\alpha) = -m_0 c \sqrt{-g_{\mu\nu}[x^\alpha(\lambda)] \dot{x}^\mu \dot{x}^\nu}. \quad (3.20)$$

La condition d'extrémalité fournit alors les équations d'Euler-Lagrange

$$\frac{d}{d\lambda} \frac{\partial L}{\partial \dot{x}^\mu} - \frac{\partial L}{\partial x^\mu} = 0. \quad (3.21)$$

Un choix naturel pour λ est le temps propre τ .

Notons la *différence fondamentale de philosophie entre l'approche de Newton et celle d'Einstein*. En mécanique classique, tous les référentiels ne sont pas équivalents. Lorsque l'on passe d'un système de coordonnées à un autre, le lagrangien change en général de forme : il apparaît des termes associés aux forces fictives d'inertie ; sauf si les deux référentiels sont inertiels, auquel cas le lagrangien (effectif) conserve sa forme. En mécanique relativiste, on peut utiliser n'importe quel système de coordonnées sans avoir à modifier le formalisme, qui est covariant.³ Cette propriété est rendue possible par l'existence d'un élément de longueur invariant, et donc d'un tenseur métrique. Associée au principe d'équivalence, elle constitue un ingrédient capital pour la compréhension de la relativité générale : comme le mouvement d'une particule en coordonnées quelconques est piloté par les composantes du tenseur métrique, ces dernières se révèlent un candidat idéal pour décrire le champ de gravitation.

A titre d'exemple, étudions le mouvement d'une particule ponctuelle libre dans le système de coordonnées du carrousel de métrique (3.4). On a

$$L = -m_0 c \sqrt{c^2 \left(1 - \frac{\omega^2 r^2}{c^2}\right) \left(\frac{dt}{d\tau}\right)^2 - 2\omega r^2 \frac{dt}{d\tau} \frac{d\varphi}{d\tau} - \left(\frac{dr}{d\tau}\right)^2 - r^2 \left(\frac{d\varphi}{d\tau}\right)^2}.$$

Le lagrangien ne dépendant explicitement ni de t ni de φ , on trouve immédiatement deux intégrales premières :

$$-c^2 \left(1 - \frac{\omega^2 r^2}{c^2}\right) \frac{dt}{d\tau} + \omega r^2 \frac{d\varphi}{d\tau} = K_1$$

3. Le rôle des forces fictives d'inertie est alors joué par les symboles de Christoffel, de nature non tensorielle.

$$\omega r^2 \frac{dt}{d\tau} + r^2 \frac{d\varphi}{d\tau} = K_2$$

avec K_1 et K_2 constants, dont on tire

$$\frac{dt}{d\tau} = C_1$$

$$r^2 \left(\frac{d\varphi}{d\tau} + \omega C_1 \right) = C_1 C_2$$

où C_1, C_2 sont des constantes reliées à K_1, K_2 . Quant à l'équation d'Euler-Lagrange pour r , elle s'écrit

$$\frac{d^2 r}{d\tau^2} = \omega^2 r \left(\frac{dt}{d\tau} \right)^2 + 2\omega r \frac{dt}{d\tau} \frac{d\varphi}{d\tau} + r \left(\frac{d\varphi}{d\tau} \right)^2.$$

En supposant que les origines de τ et de t coïncident, on a $t = C_1 \tau$ et

$$\frac{d\varphi}{dt} = \frac{C_2}{r^2} - \omega \quad (3.22)$$

$$\frac{d^2 r}{dt^2} = \frac{(C_2)^2}{r^3}. \quad (3.23)$$

Or, le mouvement d'une particule libre est rectiligne et uniforme : dans les coordonnées (x, y) , on a

$$x(t) = X_0 + V_x t \quad \text{et} \quad y(t) = Y_0 + V_y t,$$

donc, dans les coordonnées (r, φ) :

$$r(t) = \sqrt{[(V_x)^2 + (V_y)^2] t^2 + 2t (X_0 V_x + Y_0 V_y) + (X_0)^2 + (Y_0)^2}$$

$$\varphi(t) = \arctan \frac{Y_0 + V_y t}{X_0 + V_x t} - \omega t.$$

Cette solution satisfait bien aux équations (3.22)-(3.23) avec $C_2 = X_0 V_y - Y_0 V_x$.

Il est instructif de calculer le lagrangien newtonien comme limite du lagrangien relativiste pour des vitesses petites devant c :

$$\begin{aligned} L d\tau &= -m_0 c^2 \sqrt{1 - \frac{\omega^2 r^2}{c^2}} \sqrt{1 - \frac{1}{c^2 (1 - \omega^2 r^2 / c^2)} \left[2\omega r^2 \frac{d\varphi}{dt} + \left(\frac{dr}{dt} \right)^2 + r^2 \left(\frac{d\varphi}{dt} \right)^2 \right]} dt \\ &\cong -m_0 c^2 \left(1 - \frac{\omega^2 r^2}{2c^2} \right) \left\{ 1 - \frac{1 + \omega^2 r^2 / c^2}{2c^2} \left[2\omega r^2 \frac{d\varphi}{dt} + \left(\frac{dr}{dt} \right)^2 + r^2 \left(\frac{d\varphi}{dt} \right)^2 \right] \right\} dt \\ &\cong \left\{ -m_0 c^2 + \frac{m_0}{2} \left[\omega^2 r^2 + 2\omega r^2 \frac{d\varphi}{dt} + \left(\frac{dr}{dt} \right)^2 + r^2 \left(\frac{d\varphi}{dt} \right)^2 \right] \right\} dt \\ &= \left[-m_0 c^2 + \frac{1}{2} m_0 V^2 + m_0 \vec{V} \cdot (\vec{\omega} \wedge \vec{r}) + \frac{1}{2} m_0 (\vec{\omega} \wedge \vec{r})^2 \right] dt. \end{aligned}$$

On retrouve bien le lagrangien d'une particule libre dans le référentiel non inertiel du carrousel.⁴ Les équations du mouvement en découlent immédiatement : ce sont, de façon assez prévisible, les équations (3.22)-(3.23).

4. L. Landau, E. Lifchitz, *Physique théorique, Tome 1, Mécanique*, Mir, 1982, § 39.

3.6 L'effet Sagnac

Il serait dommage de clôturer ce chapitre sans évoquer un effet découvert par G. Sagnac en 1913.

Imaginons qu'une source lumineuse, attachée au carrousel en rotation, envoie deux signaux en $t = 0$, l'un dans le sens de la rotation et l'autre à contresens. Pour simplifier, on supposera que les photons se propagent dans le vide et sont contraints, par un jeu de miroirs, de parcourir une trajectoire circulaire de rayon R (donc non géodésique).

Il est évident que les deux signaux ne rejoindront pas l'émetteur simultanément. En effet, la source fuit le signal qui se meut dans le sens de la rotation, et va à la rencontre du signal qui se propage à contresens. Les temps de parcours, que nous noterons respectivement t_1 et t_2 , sont donc différents. On les calcule facilement dans un référentiel inertiel lié au centre du carrousel (il s'agit d'un simple exercice de cinématique où la relativité restreinte n'intervient que dans la mesure où la vitesse des photons est égale à c quel que soit leur sens de propagation⁵) : étant donné que

$$t_1 = \frac{2\pi R + \omega R t_1}{c} \quad \text{et} \quad t_2 = \frac{2\pi R - \omega R t_2}{c},$$

on obtient

$$t_1 = \frac{2\pi R}{c - \omega R}, \quad t_2 = \frac{2\pi R}{c + \omega R}.$$

Il s'ensuit la valeur du délai temporel

$$\begin{aligned} \Delta t &= t_1 - t_2 \\ &= \frac{4\pi R^2 \omega}{c^2 - \omega^2 R^2}. \end{aligned}$$

L'intervalle de temps propre $\Delta\tau$ effectivement mesuré par la source vaut ainsi

$$\Delta\tau = \frac{4\pi R^2 \omega}{c \sqrt{c^2 - \omega^2 R^2}}. \quad (3.24)$$

Lorsque les rayons se combinent, on obtient donc des franges d'interférence permettant de déterminer la vitesse angulaire du carrousel par rapport à l'espace-temps absolu de Minkowski. On réalise ainsi, en quelque sorte, un gyroscope électromagnétique.

A titre d'exercice, on retrouvera ces résultats en raisonnant sur la métrique (3.4) : comme $ds^2 = 0$, la durée de chacun des trajets est obtenue à partir de la solution positive de l'équation

$$0 = - \left(1 - \frac{\omega^2 R^2}{c^2} \right) c^2 dt^2 + 2\omega R^2 d\varphi dt + R^2 d\varphi^2,$$

c'est-à-dire

$$t = \int_0^\Phi \frac{\omega R^2 d\varphi + Rc |d\varphi|}{c^2 - \omega^2 R^2},$$

dans laquelle il faut poser $\Phi = 2\pi$ si le mouvement se fait dans le sens du carrousel et $\Phi = -2\pi$ sinon.

5. Si la lumière se déplace dans une fibre optique, il faut également prendre en compte la loi d'addition des vitesses (effet Fizeau).

Chapitre 4

La courbure de Gauss

4.1 Géodésiques d'une surface dans l'espace euclidien

Dans l'espace euclidien à trois dimensions, considérons une courbe $\mathcal{C}(s)$. Le paramètre s désigne la longueur de l'arc mesuré le long de cette courbe, dont nous supposons qu'elle définit en tout point une et une seule droite tangente $T(s)$ et donc, un et un seul plan normal.

Soit T la tangente à la courbe en un point $P(s)$ et T' la parallèle en $P(s)$ à la tangente au point $P(s + \Delta s)$. Le *plan osculateur* à la courbe en P est la limite, si elle existe, du plan défini par T et T' lorsque $\Delta s \rightarrow 0$ (figure 4.1).

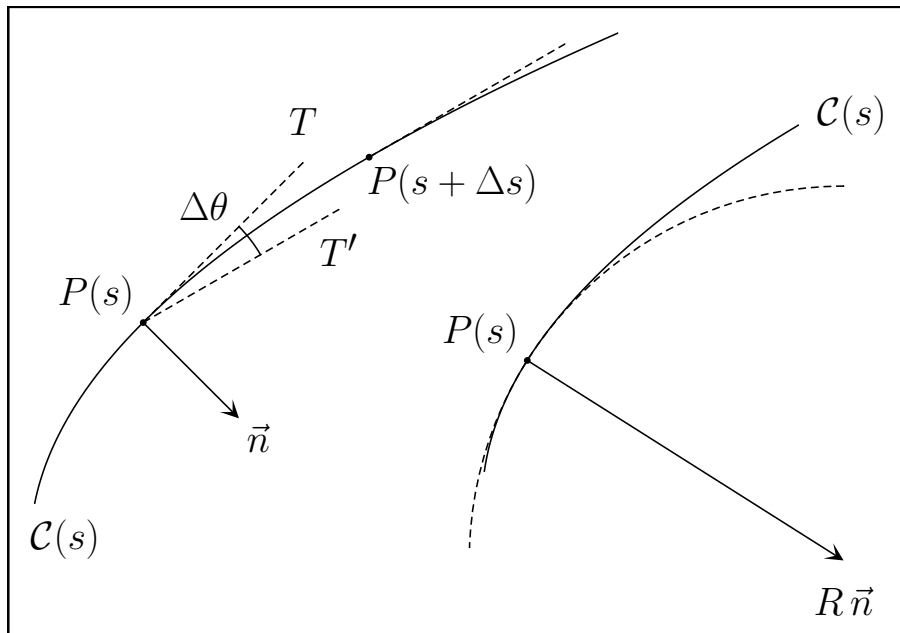


FIGURE 4.1: A gauche, plan osculateur TT' et normale principale \vec{n} en un point $P(s)$ d'une courbe $\mathcal{C}(s)$. A droite, en pointillés, cercle osculateur de rayon R .

Notant $\Delta\theta$ l'angle entre T et T' , on définit la *courbure* en P par

$$\frac{1}{R} = \lim_{\Delta s \rightarrow 0} \left| \frac{\Delta\theta}{\Delta s} \right|$$

si la limite existe. Son inverse R est le *rayon de courbure*.

Considérons l'intersection du plan osculateur et du plan normal à la courbe en P . On appelle *normale principale* à la courbe en P son vecteur unitaire \vec{n} dirigé vers la concavité de la courbe. Le *cercle osculateur* en P est le cercle de rayon R situé dans le plan osculateur et dont le centre se trouve en $R\vec{n}$.

Dans le cas particulier d'une fonction $y = f(x)$ dans le plan, ce dernier est évidemment osculateur. La direction de la normale principale est alors celle de la normale à la courbe. On calcule facilement

$$\frac{1}{R} = \frac{|d^2y/dx^2|}{\sqrt{[1 + (dy/dx)^2]^3}}.$$

Le vecteur unitaire tangent pointant dans le sens des s croissants s'écrit

$$\vec{t}(s) = \frac{d}{ds} \overrightarrow{OP}(s)$$

où O est une origine arbitraire. Il est clair que $d\vec{t}/ds$, d'une part appartient au plan osculateur, d'autre part est orthogonal à \vec{t} : il est donc proportionnel à \vec{n} . On démontre la *première formule de Frenet*

$$\frac{d\vec{t}}{ds} = \frac{\vec{n}}{R}.$$

Soit une surface immergée dans l'espace euclidien à trois dimensions et de normale unitaire \vec{N} en un point P . Considérons les courbes appartenant à cette surface. Localement en P , on appelle *géodésique* une courbe $\mathcal{C}(s)$ pour laquelle

$$\vec{n} = \pm \vec{N},$$

et donc telle que la variation du vecteur tangent $d\vec{t}/ds$ est orthogonale à la surface. Intuitivement, on dira que « la composante de courbure dans la surface » d'une géodésique est nulle. La notion de géodésique réalise ainsi le concept de « ligne la plus droite possible sur la surface ». ¹

4.2 Première et deuxième formes fondamentales

Appelons $\overrightarrow{OP}(u, v)$ une surface immergée dans l'espace euclidien à trois dimensions rapporté à des coordonnées cartésiennes (x, y, z) . La *première forme fondamentale* est l'élément de longueur mesuré le long de la surface, naturellement déduit de la métrique euclidienne $ds^2 = dx^2 + dy^2 + dz^2$ avec la contrainte $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$. Elle se présente donc sous la forme

$$ds^2 = g_{11}(u, v) du^2 + 2g_{12}(u, v) du dv + g_{22}(u, v) dv^2. \quad (4.1)$$

Soit $\overrightarrow{OP}[u(s), v(s)]$ une courbe définie sur la surface, le paramètre s désignant toujours la longueur de l'arc mesuré le long de la courbe. On a

$$\begin{aligned} \vec{t}(s) &= \frac{d}{ds} \overrightarrow{OP}[u(s), v(s)] \\ &= \frac{\partial \overrightarrow{OP}}{\partial u} \frac{du}{ds} + \frac{\partial \overrightarrow{OP}}{\partial v} \frac{dv}{ds} \end{aligned}$$

1. Le plus court chemin entre deux points donnés sur une surface s'avère d'ailleurs toujours réalisé par une géodésique. Par contre, n'importe quelle géodésique reliant deux points donnés sur la surface ne constitue pas nécessairement la ligne de plus court chemin.

$$\frac{d\vec{t}(s)}{ds} = \frac{\partial^2 \overrightarrow{OP}}{\partial u^2} \left(\frac{du}{ds}\right)^2 + 2 \frac{\partial^2 \overrightarrow{OP}}{\partial u \partial v} \frac{du}{ds} \frac{dv}{ds} + \frac{\partial^2 \overrightarrow{OP}}{\partial v^2} \left(\frac{dv}{ds}\right)^2 + \frac{\partial \overrightarrow{OP}}{\partial u} \frac{d^2 u}{ds^2} + \frac{\partial \overrightarrow{OP}}{\partial v} \frac{d^2 v}{ds^2}.$$

Compte tenu des propriétés

$$\vec{N} \cdot \frac{\partial \overrightarrow{OP}}{\partial u} = \vec{N} \cdot \frac{\partial \overrightarrow{OP}}{\partial v} = 0,$$

on peut écrire

$$\vec{N} \cdot \frac{d\vec{t}}{ds} ds^2 = b_{11}(u, v) du^2 + 2 b_{12}(u, v) du dv + b_{22}(u, v) dv^2 \quad (4.2)$$

où, dans le paramétrage (u, v) choisi, les coefficients

$$b_{11} = \vec{N} \cdot \frac{\partial^2 \overrightarrow{OP}}{\partial u^2}, \quad b_{12} = \vec{N} \cdot \frac{\partial^2 \overrightarrow{OP}}{\partial u \partial v}, \quad b_{22} = \vec{N} \cdot \frac{\partial^2 \overrightarrow{OP}}{\partial v^2}$$

ne dépendent que des caractéristiques de la surface au point P considéré, y compris, bien entendu, la façon dont elle est immergée dans l'espace euclidien. L'expression (4.2) a reçu le nom de *deuxième forme fondamentale de la surface* $\overrightarrow{OP}(u, v)$.

Théorème 4.1 (Meusnier) *En un point P d'une surface donnée de normale unitaire \vec{N} , toutes les courbes de direction \vec{t} admettent des quantités $\vec{N} \cdot \vec{n}/R$ identiques.*

Démonstration. D'après la première formule de Frenet,

$$\frac{\vec{N} \cdot \vec{n}}{R} = \vec{N} \cdot \frac{d\vec{t}}{ds}. \quad (4.3)$$

En comparant (4.1), (4.2) et (4.3), on obtient

$$\frac{\vec{N} \cdot \vec{n}}{R} = \frac{b_{11} du^2 + 2 b_{12} du dv + b_{22} dv^2}{g_{11} du^2 + 2 g_{12} du dv + g_{22} dv^2} \quad (4.4)$$

où les coefficients $b_{ij}(u, v)$ et $g_{ij}(u, v)$ ne dépendent que des caractéristiques, intrinsèques aussi bien qu'extrinsèques, de la surface dans le paramétrage adopté au point considéré, et où le quotient du/dv est déterminé par la connaissance de \vec{t} .

□

Corollaires.

1. *Si, en un point P d'une surface, deux courbes admettent le même vecteur tangent \vec{t} et la même normale principale \vec{n} non orthogonale à \vec{N} , alors elles coïncident en P jusqu'au deuxième ordre inclus.*

En effet, elles ont alors également des courbures identiques. Or, en vertu de la première formule de Frenet, les quantités \vec{t} , \vec{n} et R régissent l'allure d'une courbe jusqu'au deuxième ordre inclus en un point donné.

2. *Les courbes géodésiques en un point P d'une surface coïncident jusqu'au deuxième ordre inclus avec les courbes résultant de l'intersection de la surface avec le faisceau de plans déterminé par la normale unitaire \vec{N} en ce point.*

En effet, par définition ces dernières courbes sont planes, donc \vec{N} appartient à leur plan osculateur. De plus, \vec{N} est orthogonal au plan tangent à la surface en P , donc au vecteur tangent \vec{t} de chacune des courbes passant par ce point et, par conséquent, \vec{N} appartient également à leur plan normal. Ainsi, ces courbes ont une normale unitaire proportionnelle à \vec{N} : elles sont donc localement géodésiques, et le corollaire 1 permet de conclure.

4.3 Un exemple : la sphère S^2 de rayon \mathcal{R}

Choisissant pour paramètres les coordonnées sphériques ($u = \theta, v = \varphi$), on calcule aisément la première forme fondamentale :

$$ds^2 = \mathcal{R}^2 (d\theta^2 + \sin^2 \theta d\varphi^2)$$

et les coefficients de la deuxième forme fondamentale :

$$b_{11} = -\mathcal{R}, \quad b_{12} = 0, \quad b_{22} = -\mathcal{R} \sin^2 \theta.$$

Etudions les courbes de vecteur unitaire tangent \vec{t} en un point P de la sphère (cette surface étant homogène et isotrope, la position P et la direction \vec{t} n'ont, en fait, guère d'importance). Soit \vec{n} un vecteur unitaire orthogonal à \vec{t} : le plan (\vec{t}, \vec{n}) intersecte la sphère selon un cercle de vecteur tangent \vec{t} et de normale unitaire \vec{n} en P . On peut choisir le pôle nord pour que ce cercle constitue un *parallèle* de colatitude θ constante et de rayon de courbure $R = \mathcal{R} \sin \theta$ (voir figure 4.2). L'angle déterminé par \vec{N} et \vec{n} est alors égal à $\pi/2 + \theta$. Le théorème 4.1 se trouve bien vérifié car

$$\frac{\vec{N} \cdot \vec{n}}{R} = -\frac{1}{\mathcal{R}} = \text{constante.}$$

La géodésique tangente à \vec{t} en P correspond à $\vec{n} = -\vec{N}$ et n'est autre qu'un *grand cercle* (c'est-à-dire, dont le centre coïncide avec celui de la sphère).

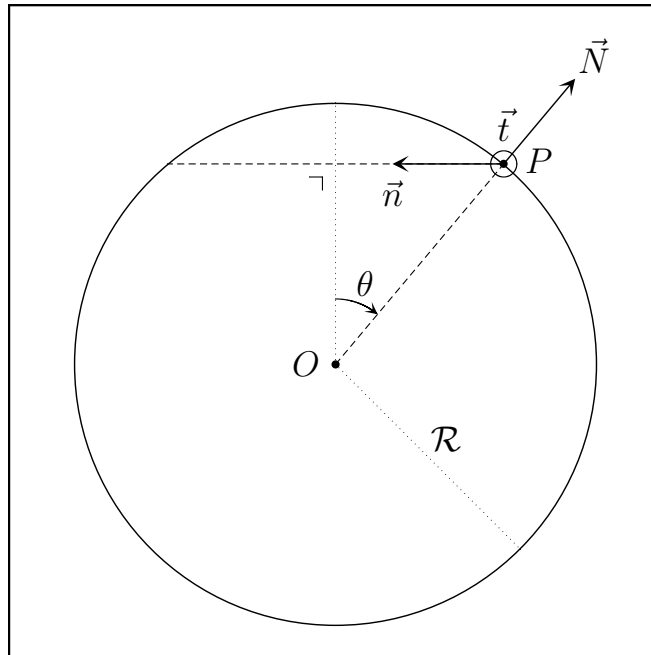


FIGURE 4.2: Le vecteur \vec{t} est perpendiculaire au plan de la page. L'intersection du plan (\vec{t}, \vec{N}) avec la sphère est un grand cercle.

4.4 La courbure de Gauss

Considérons le plan tangent en un point P d'une surface et choisissons-y un système de coordonnées cartésiennes u, v . Dans les coordonnées (u, v, z) où z est dirigé selon la

normale, l'équation de la surface peut localement s'écrire

$$z(u, v) = \left(\frac{\partial z}{\partial u}\right)_0 u + \left(\frac{\partial z}{\partial v}\right)_0 v + \frac{1}{2} \left(\frac{\partial^2 z}{\partial u^2}\right)_0 u^2 + \left(\frac{\partial^2 z}{\partial u \partial v}\right)_0 uv + \frac{1}{2} \left(\frac{\partial^2 z}{\partial v^2}\right)_0 v^2 + O(3).$$

Les deux premiers termes du membre de droite étant nuls (u et v sont dans le plan tangent), cette relation prend la forme

$$\begin{aligned} z &= \frac{1}{2} (A u^2 + 2B uv + C v^2) + O(3) \\ &= \frac{1}{2} \begin{pmatrix} u & v \end{pmatrix} \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + O(3). \end{aligned}$$

La matrice

$$M = \begin{pmatrix} A & B \\ B & C \end{pmatrix},$$

symétrique et réelle, peut être diagonalisée par une matrice orthogonale (en pratique, une rotation autour de l'axe des z) appliquant les coordonnées (u, v) sur des coordonnées (x, y) : appelant K_1 et K_2 les valeurs propres de M , on a

$$\frac{1}{2} \begin{pmatrix} u & v \end{pmatrix} \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{2} \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} K_1 & 0 \\ 0 & K_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (4.5)$$

On obtient ainsi²

$$z = \frac{1}{2} (K_1 x^2 + K_2 y^2) + O(3). \quad (4.6)$$

Vu l'invariance du déterminant et de la trace lors d'une transformation par une matrice orthogonale, K_1 et K_2 sont calculables par les relations

$$K_1 K_2 = AC - B^2$$

$$K_1 + K_2 = A + C.$$

Si $K_1 \neq K_2$, les quantités $|K_1|$ et $|K_2|$ correspondent aux valeurs extrémales parmi les courbures de toutes les géodésiques passant par P . En effet, posant $x = r \cos \varphi$, $y = r \sin \varphi$ dans (4.6), on peut écrire

$$z = \frac{r^2}{2} (K_1 \cos^2 \varphi + K_2 \sin^2 \varphi) + O(3).$$

2. Il est possible de déterminer l'angle de rotation permettant de passer des axes (u, v) aux axes (x, y) . Si l'on note

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

la relation (4.5) implique

$$\begin{aligned} K_1 &= A \cos^2 \theta + B \sin(2\theta) + C \sin^2 \theta \\ 0 &= -\frac{A}{2} \sin(2\theta) + B \cos(2\theta) + \frac{C}{2} \sin(2\theta) \\ K_2 &= A \sin^2 \theta - B \sin(2\theta) + C \cos^2 \theta. \end{aligned}$$

On en déduit

$$\theta = \frac{1}{2} \arctan \left(\frac{2B}{A-C} \right).$$

La courbe située dans le plan vertical d'angle polaire φ_0 admet donc pour courbure la quantité

$$\frac{1}{R(\varphi_0)} = |K_1 \cos^2 \varphi_0 + K_2 \sin^2 \varphi_0|$$

qui est extrémale pour $\varphi_0 = k\pi$, k entier [auquel cas $1/R(\varphi_0) = |K_1|$] ou $\varphi_0 = (2k+1)\pi/2$ [auquel cas $1/R(\varphi_0) = |K_2|$]. Les quantités K_1 et K_2 sont appelées *courbures principales* de la surface en P et les géodésiques G_1 et G_2 associées sont les *géodésiques principales* (figure 4.3).

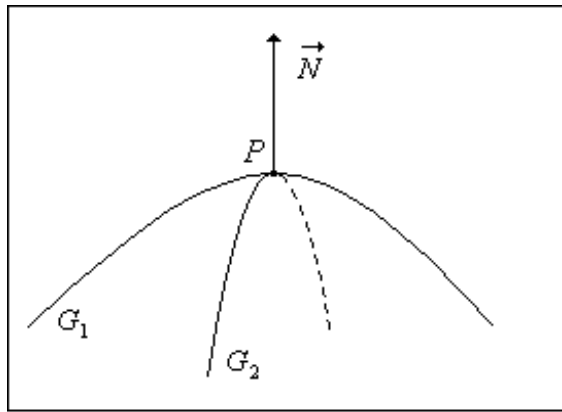


FIGURE 4.3

Le produit

$$K = K_1 K_2$$

a reçu le nom de *courbure intrinsèque*, ou *courbure de K.F. Gauss*. Si $K > 0$, la « quadrique osculatrice » à la surface est un parabolôïde elliptique et les centres de courbure des deux cercles osculateurs principaux sont situés du même côté du plan tangent. Si $K < 0$ (auquel cas la surface se comporte localement comme un parabolôïde hyperbolique), ils se trouvent de part et d'autre de ce plan. Si $K = 0$, en supposant que $K_1 + K_2 \neq 0$ (cylindre parabolique), un centre de courbure est rejeté à l'infini.

Exercice 4.1 Dans l'espace rapporté aux coordonnées (u, v, z) , soit la quadrique d'équation

$$z = \frac{1}{2}u^2 + \sqrt{3}uv - \frac{1}{2}v^2.$$

Calculer les courbures principales et la courbure de Gauss de cette surface à l'origine.

Exercice 4.2 Dans l'espace rapporté aux coordonnées (u, v, z) , soit la surface d'équation

$$z(u, v) = \sin(uv).$$

Etablir l'équation de la quadrique approximant cette surface jusqu'au deuxième ordre inclus au voisinage de l'origine. De quel type de quadrique s'agit-il? En déduire les courbures principales et la courbure de Gauss de la surface à l'origine.

4.5 Le *Theorema egregium*

Considérons dans l'espace euclidien tridimensionnel une pièce de caoutchouc *souple* mais *non élastique*. Lors d'une déformation elle conserve ses caractéristiques *métriques*, c'est-à-dire les longueurs de ses arcs et les angles entre ses courbes. Il est clair que, si la donnée de K_1 et K_2 en un point est équivalente à la connaissance locale, jusqu'au deuxième ordre inclus, de l'équation de la surface dans l'espace où elle se trouve plongée, leur seul produit $K = K_1 K_2$ est, par contre, tout à fait insuffisant à cet égard.

Gauss a démontré que cette quantité K demeure en fait *inchangée* au cours d'une déformation de la surface au sens défini ci-dessus. Plus précisément, K constitue une description de la géométrie *intrinsèque* à la surface, indépendamment de la façon dont cette dernière est déformée, et peut être calculée à l'aide du tenseur métrique uniquement.³ A ce résultat Gauss a donné le nom de « *Theorema egregium* », le théorème remarquable.⁴

Ainsi, un cône, un cylindre, un tapis bosselé sont des surfaces de courbure intrinsèque nulle au même titre qu'un mur bien droit. Un morceau de ballon d'enfant conserve la courbure constante de la sphère dont il est issu pour toute déformation isométrique qu'on lui fait subir.

La démonstration résultera de considérations plus générales exposées dans la section 5.21.6 et l'exercice 5.14.⁵ Nous allons néanmoins donner, dès à présent, deux exemples expliquant comment, pratiquement, on peut évaluer la courbure gaussienne à partir de mesures intrinsèques à la surface. Il s'agit de l'*équation de déviation géodésique* et la *mesure de la circonférence d'un cercle tracé sur la surface*.

4.6 Equation de déviation géodésique

Sur une surface à deux dimensions, considérons deux géodésiques voisines G_1 et G_2 . Choisissons la longueur d'arc s comme paramètre le long de chacune de ces deux courbes. Appelons $\vec{\eta}(s)$ le vecteur infinitésimal orthogonal à G_1 dont l'origine se trouve sur G_1 et l'extrémité sur G_2 , et soit $\eta(s)$ sa norme (figure 4.4).

On démontre l'équation suivante, dite *équation de déviation géodésique* :

$$\frac{d^2\eta(s)}{ds^2} = -K \eta(s). \quad (4.7)$$

3. C'est la *courbure moyenne* $(K_1 + K_2)/2$, aussi appelée *courbure extrinsèque*, qui renseigne sur la façon dont la surface est immergée dans l'espace euclidien, en exprimant la variation de sa normale unitaire.

4. *Si superficies curva in quacunqve aliam superficiem explicatur, mensura curvaturae in singulis punctis invariata manet* (Si une surface courbe est appliquée sur une autre surface courbe quelconque, la mesure de courbure en chaque point reste invariable). Cité par L. Boi, *Le problème mathématique de l'espace, une quête de l'intelligible*, Springer, 1995, p. 101.

5. Voici la relation permettant de calculer la courbure de Gauss à partir de l'expression générale de la métrique

$$ds^2 = g_{11}(x^1, x^2) (dx^1)^2 + 2g_{12}(x^1, x^2) dx^1 dx^2 + g_{22}(x^1, x^2) (dx^2)^2.$$

Posant $g = g_{11} g_{22} - g_{12} g_{12}$, on a :

$$\begin{aligned} K(x^1, x^2) &= \frac{1}{2g} \left[2 \frac{\partial^2 g_{12}}{\partial x^1 \partial x^2} - \frac{\partial^2 g_{11}}{(\partial x^2)^2} - \frac{\partial^2 g_{22}}{(\partial x^1)^2} \right] - \frac{g_{22}}{4g^2} \left[\left(\frac{\partial g_{11}}{\partial x^1} \right) \left(2 \frac{\partial g_{12}}{\partial x^2} - \frac{\partial g_{22}}{\partial x^1} \right) - \left(\frac{\partial g_{11}}{\partial x^2} \right)^2 \right] \\ &\quad - \frac{g_{11}}{4g^2} \left[\left(\frac{\partial g_{22}}{\partial x^2} \right) \left(2 \frac{\partial g_{12}}{\partial x^1} - \frac{\partial g_{11}}{\partial x^2} \right) - \left(\frac{\partial g_{22}}{\partial x^1} \right)^2 \right] \\ &\quad + \frac{g_{12}}{4g^2} \left[\left(\frac{\partial g_{11}}{\partial x^1} \right) \left(\frac{\partial g_{22}}{\partial x^2} \right) - 2 \left(\frac{\partial g_{11}}{\partial x^2} \right) \left(\frac{\partial g_{22}}{\partial x^1} \right) + \left(2 \frac{\partial g_{12}}{\partial x^1} - \frac{\partial g_{11}}{\partial x^2} \right) \left(2 \frac{\partial g_{12}}{\partial x^2} - \frac{\partial g_{22}}{\partial x^1} \right) \right]. \end{aligned}$$

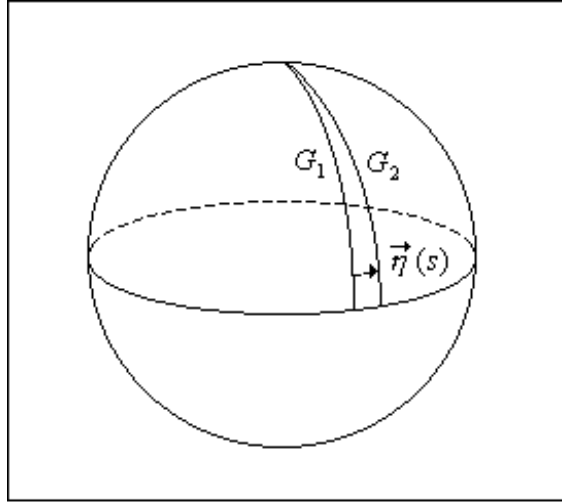


FIGURE 4.4: Illustration de la déviation géodésique sur une sphère de rayon \mathcal{R} . Soit deux grands cercles G_1 et G_2 se coupant au pôle nord et de longitudes respectives φ_0 et $\varphi_0 + d\varphi$. Leur écart le long du parallèle de colatitude θ est donné par $\eta(\theta) = \mathcal{R} \sin \theta d\varphi$. En termes de la longueur d'arc s évaluée depuis leur intersection, on a donc $\eta(s) = \mathcal{R} \sin(s/\mathcal{R}) d\varphi$. L'équation de déviation géodésique (4.7) se trouve vérifiée sur cette surface de courbure gaussienne égale à $1/\mathcal{R}^2$. Notons que $\eta(s) < s d\varphi$.

Si la courbure K est constante, on en connaît les solutions exactes :

$$\begin{aligned} \eta(s) &= a \sin \left[\sqrt{K} (s - s_0) \right] && \text{pour } K > 0 \\ \eta(s) &= a \sinh \left[\sqrt{-K} (s - s_0) \right] && \text{pour } K < 0 \\ \eta(s) &= a (s - s_0) && \text{pour } K = 0. \end{aligned}$$

Dans le dernier cas, on retrouve le théorème de Thalès si la constante a est différente de zéro. Les géodésiques sont alors des droites au sens commun du terme.

Si la courbure n'est pas constante, on peut résoudre localement l'équation (4.7) en posant que K est égal à une constante K_0 plus une petite quantité de l'ordre de s , ce que l'on note $K = K_0 + O(s)$. Cherchons $\eta(s)$ sous la forme d'un développement en série de puissances

$$\eta(s) = A s + B s^2 + C s^3 + O(s^4)$$

[nous avons supposé que $\eta(0) = 0$, c'est-à-dire que G_1 et G_2 se croisent en $s = 0$]. En injectant cette quantité dans l'équation différentielle (4.7), on obtient

$$2B + 6C s + O(s^2) + [K_0 + O(s)] [A s + B s^2 + C s^3 + O(s^4)] = 0,$$

c'est-à-dire

$$2B + (6C + K_0 A) s + O(s^2) = 0.$$

Il en résulte

$$\begin{aligned} B &= 0, & C &= -\frac{K_0 A}{6}, \\ \eta(s) &= A \left[s - \frac{K_0}{6} s^3 + O(s^4) \right]. \end{aligned} \tag{4.8}$$

La constante de primitivation A dépend des directions relatives des deux géodésiques en $s = 0$.

4.7 Mesure de la circonférence d'un cercle sur une surface

Considérons toutes les géodésiques passant par une origine commune P sur une surface à deux dimensions et soit s la longueur de l'arc mesuré le long des géodésiques à partir de ce point. Au premier ordre en s on peut raisonner dans le plan tangent à l'origine : si $d\varphi$ est l'angle déterminé par les directions initiales de deux géodésiques voisines, on a $\eta(s) \cong s d\varphi$. De (4.8) on déduit alors $A = d\varphi$, donc

$$\eta(s) = d\varphi \left[s - \frac{K_0}{6} s^3 + O(s^4) \right].$$

Par conséquent, la longueur $C(s)$ de la circonférence de rayon infinitésimal s mesuré sur la surface vaut

$$C(s) = 2\pi \left[s - \frac{K_0}{6} s^3 + O(s^4) \right]$$

et

$$K_0 = \frac{3}{\pi} \lim_{s \rightarrow 0} \frac{2\pi s - C(s)}{s^3}. \quad (4.9)$$

4.8 Courbure gaussienne du tore

Soit a le rayon du grand cercle et b celui du petit cercle (avec $b < a$; voir figure 4.5).

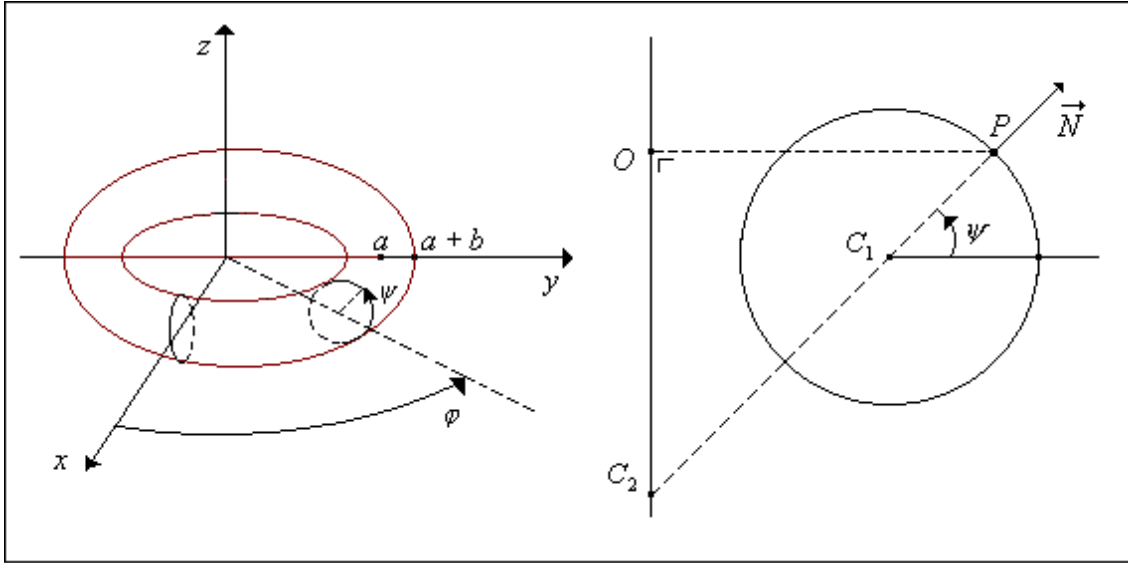


FIGURE 4.5: *A gauche, le tore vu en perspective. A droite, une coupe selon un plan vertical passant par l'axe des z .*

Compte tenu de la symétrie de révolution, la normale au tore en un point donné est toujours située dans le plan vertical passant par ce point (le plan de la page dans la figure de droite). Elle est donc dirigée selon la droite C_1P . Par conséquent, toutes les géodésiques passant par P ont leur centre de courbure sur cette droite.

Or, la section du tore par un plan vertical, c'est-à-dire le petit cercle, constitue manifestement une géodésique principale de rayon de courbure minimal

$$\overline{C_1P} = R_1 = b.$$

L'autre géodésique principale doit donc être cherchée dans la direction perpendiculaire au plan de la page, appelons-la \vec{t} . Or, d'après le théorème 4.1, pour toutes ces courbes on a

$$\frac{|\vec{N} \cdot \vec{n}|}{R} = \text{constante} = \frac{|\cos \psi|}{OP}.$$

Par conséquent, le rayon de courbure R_2 de la géodésique cherchée, pour laquelle $\vec{N} = \pm \vec{n}$, vaut

$$\overline{C_2P} = R_2 = \frac{\overline{OP}}{|\cos \psi|} = \frac{a + b \cos \psi}{|\cos \psi|} \tag{4.10}$$

et son centre de courbure C_2 se trouve sur l'axe.

En résumé, les courbures principales du tore prennent les valeurs

$$K_1 = -\frac{1}{b}, \quad K_2 = \frac{-\cos \psi}{a + b \cos \psi}$$

(on vérifie facilement les signes) et sa courbure de Gauss est donnée par

$$K = \frac{\cos \psi}{b(a + b \cos \psi)}. \tag{4.11}$$

4.9 Courbure gaussienne des surfaces de révolution

En chaque point P , une surface de révolution peut être approximée, jusqu'à l'ordre deux inclus, par le « tore osculateur » généré par la rotation, autour de l'axe, du cercle osculateur en ce point (figure 4.6).

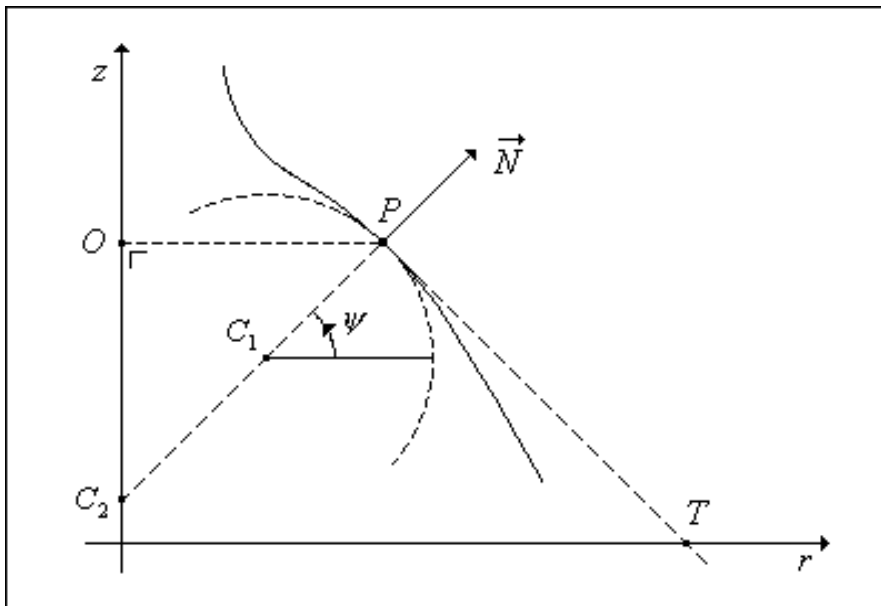


FIGURE 4.6

Si C_1 désigne le centre de ce cercle, on a

$$\overline{C_1P} = R_1$$

et, compte tenu de (4.10),

$$\overline{C_2P} = R_2 = \frac{\overline{OP}}{|\cos \psi|} = \frac{r}{|\cos \psi|}.$$

Donc

$$|K| = \frac{1}{R_1 R_2} = \frac{|\cos \psi|}{R_1 r}. \quad (4.12)$$

Calculons cette quantité en fonction de l'équation de la surface $z = z(r)$. Posant $z' = dz/dr$ et $z'' = d^2z/dr^2$, on a

$$\frac{1}{R_1} = \frac{|z''|}{\sqrt{[1 + (z')^2]^3}}.$$

D'autre part, $\cot \psi = -z'$, donc

$$|\cos \psi| = \frac{1}{\sqrt{1 + 1/(z')^2}}$$

d'où, par (4.12),

$$K(r) = \frac{z'' z'}{r [1 + (z')^2]^2}. \quad (4.13)$$

On vérifie aisément que le signe de cette expression est correct ($K > 0$ si C_1 et C_2 sont du même côté de la tangente à la courbe et $K < 0$ dans le cas contraire).

L'élément de longueur d'une surface de révolution, en coordonnées polaires (r, φ) , se déduit facilement de la métrique euclidienne en coordonnées cylindriques (r, φ, z) :

$$\begin{aligned} ds^2 &= dr^2 + r^2 d\varphi^2 + dz^2 \\ &= [1 + (z')^2(r)] dr^2 + r^2 d\varphi^2 \\ &= g_{rr}(r) dr^2 + r^2 d\varphi^2, \end{aligned} \quad (4.14)$$

où l'on a posé

$$g_{rr}(r) = 1 + (z')^2(r).$$

On peut ainsi écrire (4.13) sous la forme

$$K(r) = \frac{g'_{rr}(r)}{2r [g_{rr}(r)]^2}. \quad (4.15)$$

Pour une surface quelconque d'équation $z = z(x, y)$, on démontre que la relation généralisant (4.13) est

$$K(x, y) = \frac{(\partial_x \partial_x z)(\partial_y \partial_y z) - (\partial_x \partial_y z)^2}{[1 + (\partial_x z)^2 + (\partial_y z)^2]^2} \quad (4.16)$$

où ∂_x et ∂_y désignent les opérateurs de dérivation partielle, respectivement par rapport à x et par rapport à y .

Exercice 4.3 Soit, en coordonnées cylindriques (r, φ, z) , la surface de révolution d'équation $z(r) = \cos r$. Esquisser cette surface. En se fondant sur la signification géométrique du signe de la courbure gaussienne, en déduire pour quelles valeurs de r la courbure gaussienne est positive, nulle ou négative.

Exercice 4.4 Même question pour la surface d'équation $z(r) = \exp(-r^2)$.

Exercice 4.5 Même question pour la surface d'équation $z(r) = r^2 \exp(-r)$.

Exercice 4.6 Calculer la courbure gaussienne des sections spatiales du carrousel d'Einstein, dont la métrique (3.19) peut s'écrire sous la forme

$$ds^2 = \mathcal{R}^2 \left[\frac{d\rho^2}{(1+\rho^2)^3} + \rho^2 d\varphi^2 \right] \quad \left(\mathcal{R} = c/\omega, \quad 0 \leq \rho = \frac{r/\mathcal{R}}{\sqrt{1-(r/\mathcal{R})^2}} \right).$$

Solution. On obtient facilement, par (4.15),

$$K(\rho) = \frac{-3(1+\rho^2)^2}{\mathcal{R}^2}.$$

La courbure est partout négative, ainsi que permettaient de le prévoir les résultats de la section 3.4.

4.10 Les géodésiques du tore

A titre d'exemple, calculons explicitement l'équation des géodésiques sur le tore. Soit, sur cette surface, une courbe $\mathcal{C}(s)$ de paramètre égal à la longueur d'arc s et de vecteur tangent

$$\begin{aligned} \vec{t}(s) &= \frac{d\vec{OP}(s)}{ds} \\ &= \frac{\partial \vec{OP}}{\partial \varphi} \frac{d\varphi}{ds} + \frac{\partial \vec{OP}}{\partial \psi} \frac{d\psi}{ds}, \end{aligned}$$

résultat que nous noterons sous la forme

$$\vec{t} = t^\varphi \vec{e}_\varphi + t^\psi \vec{e}_\psi$$

(sans sommation sur φ ni ψ). Comme

$$x = (a + b \cos \psi) \cos \varphi, \quad y = (a + b \cos \psi) \sin \varphi, \quad z = b \sin \psi,$$

on calcule facilement

$$\begin{aligned} \vec{e}_\varphi &= (a + b \cos \psi)(-\sin \varphi \vec{e}_x + \cos \varphi \vec{e}_y) \\ \vec{e}_\psi &= -b \sin \psi (\cos \varphi \vec{e}_x + \sin \varphi \vec{e}_y) + b \cos \psi \vec{e}_z. \end{aligned}$$

Par définition, le long d'une géodésique la dérivée du vecteur tangent par rapport à la longueur d'arc est proportionnelle à la normale unitaire \vec{N} . On a donc

$$\frac{d\vec{t}}{ds} - \vec{N} \left(\vec{N} \cdot \frac{d\vec{t}}{ds} \right) = 0. \quad (4.17)$$

L'intérêt de ces courbes dépasse le contexte purement géométrique. En effet, en mécanique newtonienne une particule astreinte à se mouvoir sans frottement sur une surface est soumise en permanence à une force normale. De plus, comme celle-ci ne travaille pas, la

norme de la vitesse est conservée et la longueur d'arc est proportionnelle au temps absolu. L'accélération $d^2\overrightarrow{OP}/ds^2$ est ainsi proportionnelle à \vec{N} , ce qui est équivalent à la relation (4.17). Cette propriété anticipe le point de vue de la relativité générale, où les particules libres (entendons par là, soumises seulement à leur inertie et à la gravitation) empruntent les géodésiques de l'espace-temps courbe.

Définissons la *dérivée covariante* du champ vectoriel $\vec{t}(s)$ par rapport à s , par la relation

$$\nabla_s \vec{t}(s) = \frac{d\vec{t}}{ds} - \vec{N} \left(\vec{N} \cdot \frac{d\vec{t}}{ds} \right).$$

Etant donné que $\vec{N} \cdot \vec{e}_\varphi = \vec{N} \cdot \vec{e}_\psi = 0$, on a

$$\nabla_s \vec{t} = \frac{dt^\alpha}{ds} \vec{e}_\alpha + t^\alpha \nabla_s \vec{e}_\alpha,$$

avec sommation sur l'indice muet α qui représente successivement φ et ψ . En pratique, on calcule les dérivées covariantes par la relation

$$\begin{aligned} \nabla_s &= t^\beta \left[\frac{\partial}{\partial x^\beta} - \vec{N} \left(\vec{N} \cdot \frac{\partial}{\partial x^\beta} \right) \right] \\ &= t^\beta \nabla_\beta \end{aligned}$$

avec sommation sur β (les quantités x^β prenant successivement les valeurs des coordonnées intrinsèques $x^1 = \varphi$ et $x^2 = \psi$), et en adoptant la notation ∇_β pour ∇_{x^β} . En fin de compte, il vient donc

$$\begin{aligned} \nabla_s \vec{t} &= \frac{dt^\alpha}{ds} \vec{e}_\alpha + t^\alpha t^\beta \nabla_\beta \vec{e}_\alpha \\ &= \frac{d^2 x^\alpha}{ds^2} \vec{e}_\alpha + \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} \nabla_\beta \vec{e}_\alpha \end{aligned} \quad (4.18)$$

On calcule facilement :

$$\begin{aligned} \frac{\partial \vec{e}_\varphi}{\partial \psi} &= \frac{\partial \vec{e}_\psi}{\partial \varphi} = b \sin \psi (\sin \varphi \vec{e}_x - \cos \varphi \vec{e}_y) \\ \frac{\partial \vec{e}_\varphi}{\partial \varphi} &= -(a + b \cos \psi) (\cos \varphi \vec{e}_x + \sin \varphi \vec{e}_y) \\ \frac{\partial \vec{e}_\psi}{\partial \psi} &= -b \cos \psi (\cos \varphi \vec{e}_x + \sin \varphi \vec{e}_y) - b \sin \psi \vec{e}_z, \end{aligned}$$

D'autre part, la normale unitaire au tore est donnée par

$$\vec{N} = \cos \psi \cos \varphi \vec{e}_x + \cos \psi \sin \varphi \vec{e}_y + \sin \psi \vec{e}_z.$$

Donc

$$\vec{N} \cdot \frac{\partial \vec{e}_\varphi}{\partial \psi} = \vec{N} \cdot \frac{\partial \vec{e}_\psi}{\partial \varphi} = 0, \quad \vec{N} \cdot \frac{\partial \vec{e}_\varphi}{\partial \varphi} = -(a + b \cos \psi) \cos \psi, \quad \vec{N} \cdot \frac{\partial \vec{e}_\psi}{\partial \psi} = -b$$

et, par conséquent,

$$\nabla_\psi \vec{e}_\varphi = \nabla_\varphi \vec{e}_\psi = -\frac{b \sin \psi}{a + b \cos \psi} \vec{e}_\varphi \quad (4.19)$$

$$\nabla_\varphi \vec{e}_\varphi = \frac{(a + b \cos \psi) \sin \psi}{b} \vec{e}_\psi \quad (4.20)$$

$$\nabla_\psi \vec{e}_\psi = 0. \quad (4.21)$$

Les équations des géodésiques sur le tore s'obtiennent en annulant l'équation (4.18) et en y injectant les résultats (4.19), (4.20) et (4.21) :

$$0 = \ddot{\varphi} - \frac{2b \sin \psi}{a + b \cos \psi} \dot{\varphi} \dot{\psi} \quad (4.22)$$

$$0 = \ddot{\psi} + \frac{(a + b \cos \psi) \sin \psi}{b} (\dot{\varphi})^2, \quad (4.23)$$

où le point désigne la dérivée par rapport à s .

Par définition de la dérivée covariante, cet opérateur déploie les quantités $\nabla_\beta \vec{V}$ sur l'espace tangent à la surface au point considéré (\vec{V} est un champ vectoriel quelconque et β se rapporte à des coordonnées définies sur la surface). En particulier, pour les vecteurs de base $\{\vec{e}_\alpha\}$ de l'espace tangent on peut donc écrire

$$\nabla_\beta \vec{e}_\alpha = \Gamma_{\alpha\beta}^\mu \vec{e}_\mu. \quad (4.24)$$

Les fonctions $\Gamma_{\alpha\beta}^\mu$, symétriques sur les deux indices du bas, sont appelées *symboles de Christoffel*. Dans le cas du tore, les symboles de Christoffel non nuls prennent les valeurs

$$\Gamma_{\varphi\psi}^\varphi = \Gamma_{\psi\varphi}^\varphi = \frac{-b \sin \psi}{a + b \cos \psi}, \quad \Gamma_{\varphi\varphi}^\psi = \frac{\sin \psi (a + b \cos \psi)}{b}. \quad (4.25)$$

4.11 Champ de gravitation et courbure de l'espace

Une analogie, inspirée d'une remarque de Poincaré,⁶ va nous permettre de mieux comprendre l'interprétation du champ de gravitation en termes de courbure.

Considérons, en coordonnées polaires (r, φ) , un plan soumis à un champ de température absolue proportionnelle à $1 + r^2/\mathcal{R}^2$, avec \mathcal{R} constant. Imaginons que tous les corps y admettent le même coefficient de dilatation, de sorte que la longueur de chacun se trouve instantanément multipliée dans toutes les directions par le facteur $1 + r^2/\mathcal{R}^2$. Attention, nous supposons ici que le champ de température n'affecte pas l'écoulement du temps.

Des êtres bidimensionnels vivant sur le plan, attribuant naturellement la longueur unité à une règle étalon quelles que soient sa position et son orientation, en déduiraient l'élément de longueur

$$dl^2 = \frac{1}{(1 + r^2/\mathcal{R}^2)^2} (dr^2 + r^2 d\varphi^2). \quad (4.26)$$

Le changement de variables⁷

$$r = \mathcal{R} \frac{1 - \cos \theta}{\sin \theta} \quad \Leftrightarrow \quad \sin \theta = \frac{2r/\mathcal{R}}{1 + r^2/\mathcal{R}^2} \quad (0 < r, 0 < \theta < \pi)$$

transforme cet élément de longueur en

$$dl^2 = (\mathcal{R}/2)^2 (d\theta^2 + \sin^2 \theta d\varphi^2).$$

6. H. Poincaré, *La science et l'hypothèse*, Flammarion, 1968, pp. 88-91. On y trouvera un remarquable exposé sur le conventionnalisme en géométrie et en mécanique.

Voir également R.P. Feynman, R.B. Leighton, M. Sands, *The Feynman Lectures on Physics*, Volume II, Addison-Wesley, 1975, § 42-1.

7. Ce changement de variables correspond à la projection stéréographique, à partir du « pôle nord », d'une sphère de rayon $\mathcal{R}/2$ sur un plan tangent à la sphère au « pôle sud ». La coordonnée θ représente l'angle zénital, compté à partir du pôle sud, tandis que φ désigne la longitude habituelle.

Dans ces coordonnées, il apparaît clairement que des mesures locales intrinsèques à la surface mèneraient ses habitants à conclure que la géométrie de leur monde est celle d'une *sphère* de rayon $\mathcal{R}/2$. (Remarquons que la courbure ne concerne que les dimensions spatiales.)

On remarquera d'ailleurs que le cercle centré à l'origine et de rayon physique l est le lieu des points de coordonnée radiale

$$r(l) = \mathcal{R} \tan(l/\mathcal{R}).$$

Sa circonférence a pour mesure

$$C(l) = \frac{2\pi r(l)}{1 + r^2(l)/\mathcal{R}^2} = \pi \mathcal{R} \sin \frac{2l}{\mathcal{R}}$$

et la courbure gaussienne à l'origine est bien donnée par

$$K_0 = \frac{3}{\pi} \lim_{l \rightarrow 0} \frac{2\pi l - C(l)}{l^3} = \frac{4}{\mathcal{R}^2}.$$

Les géodésiques du plan chauffé, déterminées à l'aide des règles étalons dilatées, sont les grands cercles de la sphère.

Du point de vue physique, il est cependant possible de choisir un autre étalon de longueur, fondé sur la distance parcourue par la lumière en l'unité de temps.⁸ Si cette dernière est uniforme pour un observateur extérieur, les habitants de la plate-forme quant à eux la trouveront variable et pourront dès lors hésiter entre deux types de géométrie : *non euclidienne* s'ils choisissent pour étalon leur « mètre » dilaté et *euclidienne* si l'étalon de longueur est dérivé de la vitesse de la lumière.

Mais supposons que, du point de vue extérieur au plan, la vitesse de la lumière soit donnée par

$$v(r) = c(1 + r^2/\mathcal{R}^2),$$

donc qu'il existe un indice de réfraction

$$n(r) = \frac{1}{1 + r^2/\mathcal{R}^2}$$

tel que $v(r) = c/n(r)$. Alors, compte tenu de la dilatation de leurs règles, les observateurs bidimensionnels mesureront en tout point une vitesse de la lumière égale à c .

Quel sera le mouvement de la lumière dans ces conditions? D'après le principe de Fermat, le trajet de la lumière entre deux points donnés A et B doit être tel qu'il minimise le temps de parcours entre ces points, donc le chemin optique car

$$t_{AB} = \int_A^B \frac{\sqrt{dr^2 + r^2 d\varphi^2}}{v(r)} = \frac{1}{c} \int_A^B n(r) \sqrt{dr^2 + r^2 d\varphi^2}.$$

Or, le chemin optique est précisément égal à la distance mesurée sur la plate-forme à l'aide des règles dilatées puisque

$$\int_A^B n(r) \sqrt{dr^2 + r^2 d\varphi^2} = \int_A^B dl.$$

8. C'est précisément le point de vue actuellement adopté en métrologie puisque, depuis 1983, le mètre est défini comme la 299 792 458^e partie de la distance parcourue par la lumière en une seconde.

La lumière parcourt donc les géodésiques de l'espace courbe de métrique donnée par (4.26).

Dans ces conditions, le caractère sphérique de l'espace s'imposerait aux habitants de la plate-forme comme l'interprétation la plus naturelle de toutes les mesures qu'ils peuvent y effectuer.

Ces résultats peuvent être interprétés dans un cadre spatio-temporel : les trajectoires de la lumière décrites ci-dessus sont, en fait, les *géodésiques nulles* de la métrique

$$ds^2 = -c^2 dt^2 + \frac{1}{(1 + r^2/\mathcal{R}^2)^2} (dr^2 + r^2 d\varphi^2).$$

Un théorème⁹ stipule en effet que la projection des géodésiques nulles de la métrique *statique*

$$ds^2 = -g_{00}(x^k) c^2 dt^2 + g_{ij}(x^k) dx^i dx^j \quad (4.27)$$

($g_{00} > 0$ et g_{ij} définie positive) sur les sections spatiales demeure géodésique pour la métrique spatiale

$$dl^2 = \frac{g_{ij}}{g_{00}} dx^i dx^j.$$

Ainsi, les géodésiques nulles minimisent la distance spatiale l entre deux points fixes donnés et donc le temps de parcours car, par (4.27),

$$dl = \sqrt{\frac{g_{ij}}{g_{00}} dx^i dx^j} = c dt.$$

Le champ de température est une métaphore pour le champ de gravitation en relativité générale. L'analogie n'est pas mauvaise dans la mesure où, comme la pesanteur, on suppose que la dilatation thermique affecte tous les corps indépendamment de leur nature. De plus, dans la théorie d'Einstein la vitesse de la lumière, mesurée localement, demeure toujours égale à une constante universelle c .

Un observateur isolé thermiquement de la plate-forme, par exemple au travers d'une plaque isolante et transparente, préférerait sans doute maintenir la géométrie euclidienne (quitte à sacrifier le caractère géodésique des trajectoires de la lumière) car elle est mathématiquement plus simple. Si l'on découvrait un jour des corps insensibles à la gravitation, la relativité générale s'en trouverait ruinée, depuis le principe d'équivalence jusqu'à l'interprétation de la gravitation en termes de courbure de l'espace-temps.

4.12 Les surfaces à courbure négative constante

4.12.1 La pseudosphère de Beltrami

La *tractrice* est la courbe décrite par un chien attaché à une laisse rigide de longueur \mathcal{R} et tiré par une personne se déplaçant le long d'une droite (la laisse est donc tangente à la trajectoire du chien en tout point). Pour fixer les idées, supposons que la personne progresse selon l'axe des z positifs et que la laisse soit perpendiculaire à cet axe en $z = 0$ (voir figure 4.7).

Établissons l'équation de cette courbe. On obtient sans peine

$$\frac{dz}{dx} = -\frac{\sqrt{\mathcal{R}^2 - x^2}}{x} \quad (4.28)$$

9. C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation*, Freeman, 1973, p. 1108.
R. Adler, M. Bazin, M. Schiffer, *Introduction to general relativity*, International Student Edition, McGraw-Hill Kogakusha, 1975, pp. 221-222.

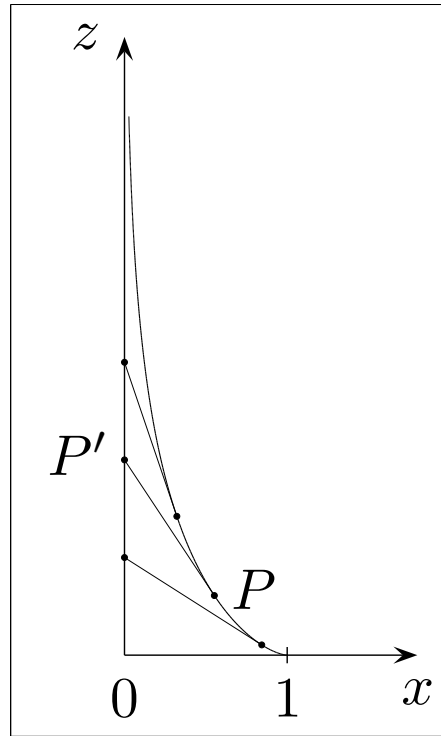


FIGURE 4.7: En chaque point P de la tractrice, la droite tangente intersecte l'axe vertical en un point P' tel que la longueur du segment PP' soit constamment égale à \mathcal{R} . La courbe est tangente à l'axe des x au point d'abscisse \mathcal{R} (ici on a pris $\mathcal{R} = 1$) et admet l'asymptote verticale $x = 0$.

$$z(x) = \mathcal{R} \ln \frac{\mathcal{R} + \sqrt{\mathcal{R}^2 - x^2}}{x} - \sqrt{\mathcal{R}^2 - x^2}.$$

Considérons maintenant la surface de révolution obtenue en faisant tourner la tractrice autour de son asymptote. Si y est un troisième axe orthogonal au plan de la page et si (r, φ) désignent les coordonnées polaires dans le plan (x, y) , alors cette surface a pour équation

$$z(r) = \mathcal{R} \ln \frac{\mathcal{R} + \sqrt{\mathcal{R}^2 - r^2}}{r} - \sqrt{\mathcal{R}^2 - r^2}. \quad (4.29)$$

Par (4.13) on en déduit sa courbure gaussienne :

$$K = -\frac{1}{\mathcal{R}^2}.$$

Cette courbure s'avère *négative et constante*. La tractrice de révolution, également appelée *pseudosphère de E. Beltrami*,¹⁰ fournit un modèle (il est vrai, singulier en $z = 0$) pour la géométrie de Lobatchevski, commode car cette surface est définie dans notre espace euclidien intuitif à trois dimensions. Son élément de longueur se déduit facilement des équations (4.14) et (4.28) :

$$ds^2 = \frac{\mathcal{R}^2}{r^2} dr^2 + r^2 d\varphi^2. \quad (4.30)$$

10. Déjà découverte par Gauss en 1827 : voir L. Boi, *op. cit.*, p. 107. On en trouvera une étude détaillée, tant du point de vue mathématique que de son importance historique, dans le chapitre 6 de cet ouvrage.

4.12.2 Autres représentations

Si, dans (4.30), on pose

$$r = \frac{\mathcal{R}}{y}, \quad \varphi = x$$

(attention, x et y n'ont plus rien à voir avec les coordonnées cartésiennes), on obtient la *métrique du demi-plan de Poincaré* :

$$ds^2 = \frac{\mathcal{R}^2}{y^2} (dx^2 + dy^2). \quad (4.31)$$

Le nouveau changement de variables

$$x = \frac{2r \sin \varphi}{1 + 2r \cos \varphi + r^2}, \quad y = \frac{1 - r^2}{1 + 2r \cos \varphi + r^2},$$

que l'on inverse facilement :

$$r \cos \varphi = \frac{1 - x^2 - y^2}{1 + 2y + x^2 + y^2}, \quad r \sin \varphi = \frac{2x}{1 + 2y + x^2 + y^2},$$

mène à la métrique du *disque de Poincaré*

$$ds^2 = \frac{4\mathcal{R}^2}{(1 - r^2)^2} (dr^2 + r^2 d\varphi^2). \quad (4.32)$$

Dans l'esprit de la section 4.11, cette expression constitue l'élément de longueur obtenu en arpentant le disque plan $0 \leq r < 2\mathcal{R}$ à l'aide de règles dilatées par un champ de température proportionnelle à $1 - r^2/4\mathcal{R}^2$. En effet, posant

$$r = \underline{r}/2\mathcal{R}$$

et laissant tomber les « barre », on trouve

$$ds^2 = \frac{1}{(1 - r^2/4\mathcal{R}^2)^2} (dr^2 + r^2 d\varphi^2)$$

[cf. équation (4.26)]. En dépit de son nom, la métrique (4.32) régit une surface de mesure infinie. L'artiste néerlandais Escher l'a régulièrement représentée dans ses gravures (figure 4.8).

Si enfin, dans (4.32), on pose

$$r = \frac{\cosh \theta - 1}{\sinh \theta},$$

c'est-à-dire

$$\sinh \theta = \frac{2r}{1 - r^2},$$

on obtient

$$ds^2 = \mathcal{R}^2 (d\theta^2 + \sinh^2 \theta d\varphi^2). \quad (4.33)$$

C'est la métrique engendrée par un hyperboloïde à deux nappes immergé dans un espace à trois dimensions pseudo-euclidien (exercice 5.11). Sa généralisation à trois dimensions spatiales joue un rôle important dans le contexte des modèles cosmologiques spatialement homogènes et isotropes (section 8.3).

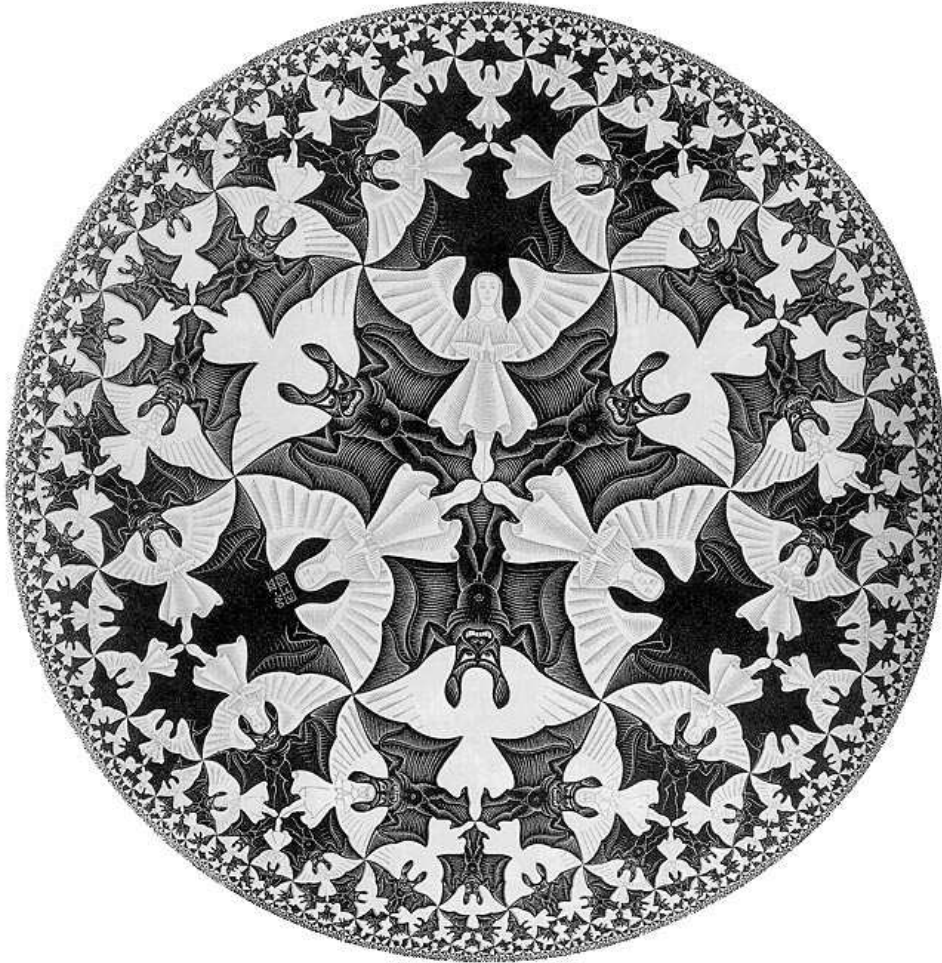


FIGURE 4.8: M.C. Escher, Limite Circulaire IV, 1960. Remarquer l'exploitation des géodésiques sur le disque de Poincaré, réalisées par les lignes radiales et les arcs de cercle « orthogonaux » à la frontière du disque. Celle-ci se situe en fait à distance infinie du centre.

4.13 Le tenseur de courbure

Le nombre de dimensions de l'espace-temps étant égal à quatre, il convient de généraliser les découvertes de Gauss relatives à la courbure des surfaces. Ce fut l'œuvre de son étudiant Riemann, qui en 1854 exposa le fruit de ses recherches dans sa leçon inaugurale à l'université de Göttingen « Über die Hypothesen, welche der Geometrie zu Grunde liegen », « Sur les hypothèses qui servent de fondement à la géométrie ». Le problème s'avère singulièrement plus complexe car, à trois dimensions ou plus, un seul nombre ne suffit plus pour décrire la courbure. Toute l'information relative à cette dernière est en fait contenue dans un tenseur à quatre indices calculable en fonction du tenseur métrique¹¹, le *tenseur de Riemann-Christoffel*, également appelé *tenseur de courbure*, et que l'on peut succinctement caractériser comme suit : *un espace est euclidien en un point si et seulement si le tenseur de courbure s'y annule*. Il sera étudié au chapitre 5.

11. Fondamentalement il s'agit d'une grandeur affine mais en relativité générale on utilise des connexions riemanniennes.

Chapitre 5

Géométrie affine et riemannienne

5.1 Notion de variété différentielle

Les *variétés différentielles* constituent le cadre mathématique naturel de la relativité générale. Nous allons les définir brièvement.

Un *espace topologique* est un couple (E, \mathcal{T}) où E est un ensemble de points et \mathcal{T} une collection de sous-ensembles de E , appelés *ouverts*, satisfaisant les propriétés suivantes :

- l'ensemble vide et E sont des ouverts ;
- toute union d'ouverts est un ouvert ;
- toute intersection finie d'ouverts est un ouvert.

\mathcal{T} est appelée une *topologie* sur E . Une *base* d'une topologie \mathcal{T} est un ensemble d'ouverts tel que tout ouvert de \mathcal{T} soit une union d'éléments de cet ensemble.

Un *voisinage* d'un point $P \in E$ pour la topologie \mathcal{T} est n'importe quelle partie de E contenant un ouvert contenant P .

La donnée d'une topologie permet de définir la notion de *continuité* : une fonction f appliquant l'espace topologique $E = \text{dom } f$ sur un espace topologique E' est continue en un point $P \in E$ ssi, pour tout ouvert $O' \ni f(P)$, $f^{-1}(O')$ est un voisinage de P [l'ensemble $f^{-1}(O')$ est défini par $\{P \in E : f(P) \in O'\}$]. Elle est dite *continue sur E* (ou tout simplement : *continue*) ssi, pour tout ouvert O' de E' , $f^{-1}(O')$ est un ouvert de E .

Un espace topologique est *séparé*, ou encore *de Hausdorff*, si deux points distincts quelconques admettent des voisinages disjoints. L'espace \mathbb{R}^n , avec la topologie définie par les boules ouvertes, en constitue un bon exemple.

Une *variété topologique* M^n à n dimensions est un espace topologique séparé à base dénombrable, localement homéomorphe à \mathbb{R}^n (un *homéomorphisme* est une bijection continue dont l'inverse est également continue) ; en d'autres termes, $\forall P \in M^n$, il existe un ouvert $O \ni P$ et un homéomorphisme $h : O \rightarrow O'$ où O' est un ouvert de \mathbb{R}^n . On dit que (h, O) est une *carte locale* (ou un *système de coordonnées locales*) de domaine $O \subset M^n$. Une famille de cartes $\{(h_\alpha, O_\alpha)\}$ constitue un *atlas* pour M^n si l'ensemble des O_α recouvre M^n .

Les points se trouvant à l'intersection de deux ouverts O_α et O_β peuvent ainsi recevoir des coordonnées dans O'_α , via l'homéomorphisme h_α , et dans O'_β , via l'homéomorphisme h_β . Le passage des premières coordonnées aux secondes, qui est donc une transformation de $\mathbb{R}^n \rightarrow \mathbb{R}^n$, est assuré par les fonctions $h_{\alpha\beta} = h_\beta \circ h_\alpha^{-1}$ (voir exemple en figure 5.1). La transformation inverse, c'est-à-dire le changement de coordonnées dans l'autre sens, est assurée par les $h_{\alpha\beta}^{-1} = h_{\beta\alpha}$. On impose en général que toutes ces transformations soient infiniment continûment dérivables : cela revient à exiger que les fonctions $h_{\alpha\beta}$ soient des *difféomorphismes* de classe C^∞ .

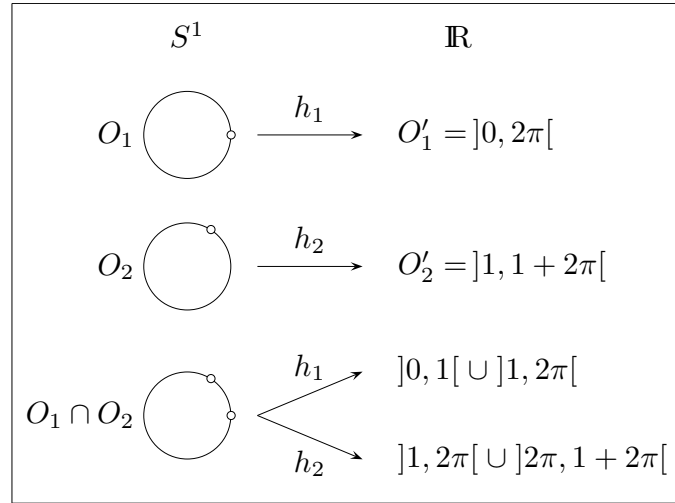


FIGURE 5.1: Le recouvrement de la sphère S^1 nécessite au moins deux ouverts O_1 et O_2 . Les points de $O_1 \cap O_2$ n'ont pas la même coordonnée réelle θ selon qu'on l'attribue par l'intermédiaire de l'homéomorphisme h_1 ou de h_2 . Cependant, la fonction $h_2 \circ h_1^{-1} :]0, 1[\cup]1, 2\pi[\rightarrow]1, 2\pi[\cup]2\pi, 1 + 2\pi[$ applique $]0, 1[\rightarrow]2\pi, 1 + 2\pi[$ et $]1, 2\pi[\rightarrow]1, 2\pi[$. On a donc, respectivement : $\theta' = \theta + 2\pi$ et $\theta' = \theta$. Ces deux transformations sont C^∞ .

On dit alors que la variété topologique M^n est dotée d'une *structure différentiable* de classe C^∞ , et on l'appelle *variété différentielle* V^n de classe C^∞ .

5.2 Notions de topologie différentielle

5.2.1 Fonctions et courbes

Une *fonction* (ou encore *champ scalaire*) est une application f

$$V^n \rightarrow \mathbb{R} : P \rightarrow f(P).$$

Concrètement, on utilise l'homéomorphisme local h pour construire sa *représentation* f_h

$$\mathbb{R}^n \rightarrow \mathbb{R} : \vec{x} \rightarrow f_h(\vec{x}) = f \circ h^{-1}(\vec{x}),$$

ce qui permet de pratiquer toutes les opérations habituelles de l'analyse. Dans la suite, afin d'alléger l'écriture nous ne distinguerons pas entre la fonction et sa représentation.

Une *courbe* est une application \mathcal{C} de la droite réelle vers la variété différentielle

$$\mathbb{R} \rightarrow V^n : \lambda \rightarrow P = \mathcal{C}(\lambda).$$

Nous l'étudierons au moyen de sa représentation \mathcal{C}_h

$$\mathbb{R} \rightarrow \mathbb{R}^n : \lambda \rightarrow \vec{x}(\lambda) = \mathcal{C}_h(\lambda) = h \circ \mathcal{C}(\lambda).$$

5.2.2 Vecteurs

a) On définit le *vecteur tangent* en un point d'une courbe $\mathcal{C}(\lambda)$ par l'opérateur de dérivation directionnelle selon la courbe en ce point :

$$\vec{V} = \frac{d}{d\lambda}. \quad (5.1)$$

Il peut se déployer sur les vecteurs $\vec{e}_\alpha = \partial/\partial x^\alpha$ tangents aux lignes de coordonnées :

$$\vec{V} = \frac{dx^\alpha(\lambda)}{d\lambda} \frac{\partial}{\partial x^\alpha} = V^\alpha(\lambda) \vec{e}_\alpha.$$

Le long de la courbe $\mathcal{C}(\lambda)$ on peut ainsi écrire

$$x^\alpha(\lambda) = x^\alpha(0) + \lambda V^\alpha(0) + \frac{\lambda^2}{2} \frac{dV^\alpha}{d\lambda}(0) + O(\lambda^3). \quad (5.2)$$

(Attention, les quantités $dV^\alpha/d\lambda = d^2x^\alpha/d\lambda^2$ ne sont pas les composantes d'un vecteur.)

Cette interprétation des vecteurs, tout en offrant l'avantage d'être intrinsèque à la variété différentielle, entraîne les lois de transformation habituelles si l'on passe à de nouvelles coordonnées \underline{x}^β :

$$\vec{e}_\beta = \frac{\partial x^\alpha}{\partial \underline{x}^\beta} \vec{e}_\alpha \quad (5.3)$$

$$\underline{V}^\beta = \frac{\partial \underline{x}^\beta}{\partial x^\alpha} V^\alpha \quad (5.4)$$

où, pour simplifier l'écriture, nous écrivons \underline{V}^β pour V^β .

L'ensemble des combinaisons linéaires des \vec{e}_α en un point donné de la variété différentielle constitue un espace vectoriel appelé *espace tangent* en ce point. Bien entendu, il n'est pas interdit de conserver l'image intuitive du vecteur conçu comme une flèche dans l'espace tangent ; ou encore, dans l'espace \mathbb{R}^n localement homéomorphe à la variété.¹

b) Inversement, soit un champ vectoriel

$$\vec{V}(\vec{x}) = V^\alpha(\vec{x}) \frac{\partial}{\partial x^\alpha}.$$

Les *courbes intégrales* $\mathcal{C}(\lambda)$ de ce champ sont les courbes admettant, en chacun de leurs points, un élément \vec{V} du champ comme vecteur tangent. Elles sont donc déterminées par l'exigence

$$\vec{V}(\vec{x}) = \frac{d}{d\lambda},$$

que l'on satisfait en résolvant le système de n équations différentielles autonomes

$$\frac{dx^\alpha(\lambda)}{d\lambda} = V^\alpha[\vec{x}(\lambda)].$$

5.2.3 Formes différentielles

La *différentielle* d'une fonction $f(\vec{x})$ est traditionnellement définie par

$$df = \frac{\partial f}{\partial x^\alpha} dx^\alpha. \quad (5.5)$$

Elle permet de calculer la variation de f sous l'effet d'un déplacement infinitésimal de coordonnées dx^α . Le *gradient* d'une fonction f en un point P donné, noté $(\tilde{d}f)_P$, est une entité engendrée à partir de l'*extrapolation linéaire* de f en P , c'est-à-dire de la fonction

$$f_P(\vec{x}) = f(P) + [x^\alpha - x^\alpha(P)] (\partial f / \partial x^\alpha)_P. \quad (5.6)$$

1. La droite tangente extrapolant linéairement la courbe $\mathcal{C}(\lambda)$ au point considéré a pour équations paramétriques $\vec{x}(\lambda) = \vec{x}(0) + \lambda \vec{V}(0)$. Sur cette droite, $\vec{V}(0)$ est la flèche reliant le point $\vec{x}(0)$ au point $\vec{x}(1)$.

Pour expliquer la nature géométrique du gradient, il est utile d'en donner une image mentale, à l'instar de la flèche pour les vecteurs : dans l'espace tangent en P , le gradient est incarné par l'ensemble des hyperplans H_k sur chacun desquels la fonction extrapolée vaut respectivement $f(P) + k$ (où k est un entier relatif), ordonnés dans le sens des k croissants (voir la figure 5.2 pour une illustration dans \mathbb{R}^2). Dans les coordonnées x^α associées à la base des \vec{e}_α , l'hyperplan H_k a pour équation $[x^\alpha - x^\alpha(P)] (\partial f / \partial x^\alpha)_P = k$.

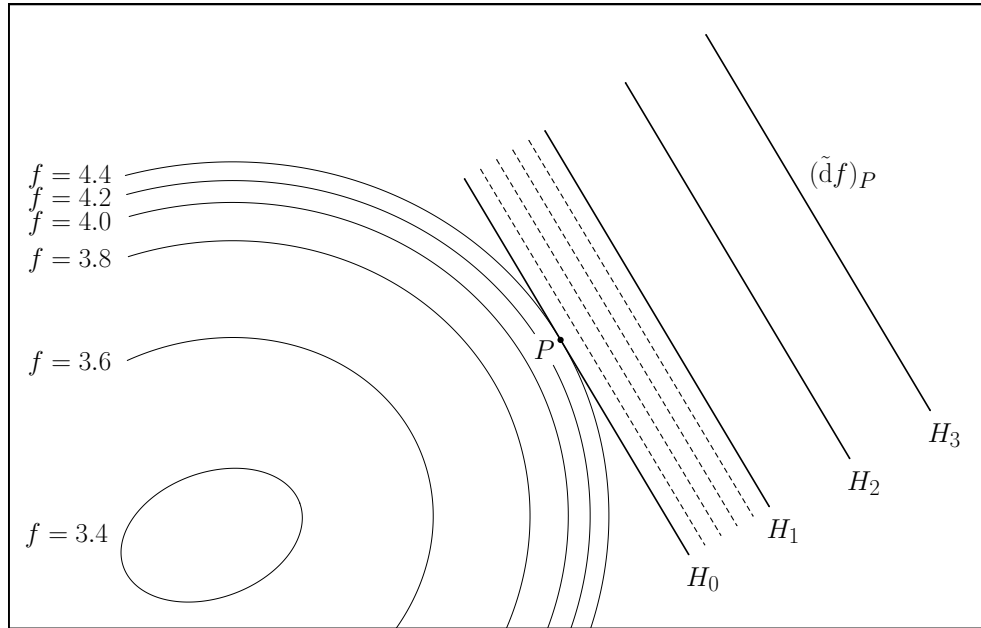


FIGURE 5.2: Dans \mathbb{R}^2 , le gradient $(\tilde{d}f)_P$ d'une fonction f en un point P est représenté par l'ensemble des droites H_k (ordonnées dans le sens des k croissants, k étant un entier relatif) extrapolant dans tout le plan la tendance à se répartir des courbes de niveau de f au point P . Comme la variation de f au premier ordre est nulle selon la direction tangente à la courbe de niveau contenant P , l'entièreté de la droite tangente n'est autre que H_0 , où la fonction extrapolée vaut $f(P)$. Chacune des droites H_k est parallèle à H_0 et réalise le lieu des points où la fonction extrapolée vaut $f(P) + k$. Remarquons que cette construction est purement géométrique : elle ne nécessite aucun système de coordonnées.

Il en découle notamment une représentation visuelle pour le gradient des coordonnées, qui sont les fonctions $f = x^\alpha$: leur gradient, noté $\tilde{d}x^\alpha$ (nous sous-entendrons : au point P), est schématisé dans l'espace tangent par l'ensemble des hyperplans $x^\alpha - x^\alpha(P) = k$, où k est un entier relatif, ordonnés dans le sens des k croissants.

On peut développer cette présentation en construisant un espace vectoriel sur la base des gradients $\tilde{d}x^\alpha$ des coordonnées. On l'appelle *espace cotangent* au point considéré de la variété différentielle, et ses éléments, obtenus par combinaisons linéaires à coefficients réels Ω_α des $\tilde{d}x^\alpha$, ont reçu le nom de *formes différentielles*. L'image mentale associée à l'élément générique $\tilde{\Omega} = \Omega_\alpha \tilde{d}x^\alpha$ de cet espace est la famille d'hyperplans d'équations $[x^\alpha - x^\alpha(P)] \Omega_\alpha = k$ (où $k = \dots, -2, -1, 0, 1, 2, \dots$).

En particulier, pour le gradient d'une fonction $f(\vec{x})$ on a $\Omega_\alpha = (\partial f / \partial x^\alpha)_P$ et on peut écrire en tout point

$$\tilde{d}f = \frac{\partial f}{\partial x^\alpha} \tilde{d}x^\alpha, \quad (5.7)$$

qui constitue le pendant géométrique des relations (5.5) ou (5.6).

Il est facile de calculer les formes de base relatives à un nouveau système de coordonnées $\tilde{x}(\vec{x})$, en application de (5.7) :

$$\tilde{d}\underline{x}^\beta = \frac{\partial \underline{x}^\beta}{\partial x^\alpha} \tilde{d}x^\alpha. \quad (5.8)$$

On en déduit la relation géométrique $\tilde{\Omega} = \Omega_\alpha \tilde{d}x^\alpha = \underline{\Omega}_\beta \tilde{d}\underline{x}^\beta$ où, en conséquence de (5.8),

$$\underline{\Omega}_\beta = \frac{\partial x^\alpha}{\partial \underline{x}^\beta} \Omega_\alpha. \quad (5.9)$$

[Comme dans la relation (5.4), nous écrivons $\underline{\Omega}_\beta$ pour $\Omega_{\underline{\beta}}$.]

Signalons enfin que les formes de base $\tilde{d}x^\alpha$ sont souvent notées $\tilde{\omega}^\alpha$.²

5.2.4 Le produit intérieur

De même que les opérateurs de dérivation ont été interprétés comme vecteurs dans l'espace tangent, l'éclairage des formes différentielles nous a permis d'appréhender les variations au premier ordre des fonctions comme éléments de l'espace cotangent. Très naturellement il en résulte une interaction des formes et des vecteurs en un point donné, appelée *produit intérieur*, notée et définie comme suit :

$$\begin{aligned} \langle \tilde{d}f, \frac{d}{d\lambda} \rangle &= \frac{df}{d\lambda} \\ &= \frac{\partial f}{\partial x^\alpha} \frac{dx^\alpha}{d\lambda} \\ &= (\tilde{d}f)_\alpha V^\alpha. \end{aligned} \quad (5.10)$$

Le produit intérieur est donc une opération *bilinéaire* associant un scalaire à un couple forme-vecteur :

$$\langle \tilde{\Omega}, \vec{V} \rangle = \Omega_\alpha V^\alpha. \quad (5.11)$$

La définition (5.10) implique que ce scalaire est indépendant des coordonnées. L'utilisation des relations (5.4) et (5.9) prouve d'ailleurs directement son invariance : $\Omega_\alpha V^\alpha = \underline{\Omega}_\alpha \underline{V}^\alpha$. Du reste, le produit intérieur se prête à une interprétation géométrique simple. En effet, $(df/d\lambda)_P$ mesure la tendance à la variation par unité de λ de la fonction f en un point P donné. Par exemple, $(df/d\lambda)_P = -3$ signifie que l'extrapolation linéaire de f (c'est-à-dire son gradient $\tilde{d}f$) a pour valeur $f(P) - 3$ lorsque $\lambda = 1$ le long de la droite tangente à la courbe $\mathcal{C}(\lambda)$ en P . En d'autres termes, en $\lambda = 1$ cette droite atteint l'hypersurface H_{-3} du gradient. Or, d'après la note 1, le tronçon de droite allant de $\lambda = 0$ à $\lambda = 1$ correspond précisément à la flèche symbolisant le vecteur $(d/d\lambda)_P$. Le produit intérieur coïncide ainsi avec le nombre d'hypersurfaces H_k du gradient (au besoin interpolées si le produit intérieur n'est pas un entier relatif) transpercées par cette flèche.

Dans le cas particulier des vecteurs et des formes de base, on a

$$\begin{aligned} \langle \tilde{\omega}^\alpha, \vec{e}_\beta \rangle &= \langle \tilde{d}x^\alpha, \partial/\partial x^\beta \rangle \\ &= \frac{\partial x^\alpha}{\partial x^\beta} \\ &= \delta_\beta^\alpha. \end{aligned} \quad (5.12)$$

2. Géométriquement il s'agit bien de formes différentielles mais, dans une acception algébrique (éléments d'un espace vectoriel), on peut aussi les qualifier de « vecteurs de base ».

La base des $\tilde{\omega}^\alpha$ et celle des \vec{e}_β sont dites *duales* l'une de l'autre. Observons que

$$V^\alpha = \langle \tilde{\omega}^\alpha, \vec{V} \rangle, \quad \Omega_\alpha = \langle \tilde{\Omega}, \vec{e}_\alpha \rangle. \quad (5.13)$$

5.2.5 Tenseurs

Le produit intérieur nous autorise à envisager les vecteurs comme des entités agissant linéairement sur des formes différentielles, et vice-versa. On acte cela en disant qu'un vecteur est un *tenseur d'ordre 1* et de variance $(1, 0)$, tandis qu'une forme est un tenseur d'ordre 1 et de variance $(0, 1)$. Cette approche peut être étendue à des objets plus élaborés que nous allons décrire au travers de l'exemple, facile à transposer, d'un tenseur \mathbf{T} d'ordre 3 et de variance $(1, 2)$, dit *une fois contravariant et deux fois covariant*. Un tel tenseur est un élément d'un espace vectoriel abstrait à n^3 dimensions défini en un point donné de la variété différentielle et sous-tendu par des vecteurs de base notés $\vec{e}_\alpha \otimes \tilde{\omega}^\beta \otimes \tilde{\omega}^\gamma$:

$$\mathbf{T} = T^\alpha_{\beta\gamma} \vec{e}_\alpha \otimes \tilde{\omega}^\beta \otimes \tilde{\omega}^\gamma. \quad (5.14)$$

Le symbole \otimes s'appelle *produit tensoriel* et les quantités $T^\alpha_{\beta\gamma}$ constituent les *composantes* du tenseur dans la base choisie, elle-même relative à des coordonnées \vec{x} .

Chaque vecteur de base peut être imaginé comme une machine agissant *linéairement* sur une forme et deux vecteurs (disposés dans un ordre précis) pour produire un nombre réel selon la règle

$$\begin{aligned} (\vec{e}_\alpha \otimes \tilde{\omega}^\beta \otimes \tilde{\omega}^\gamma) (\tilde{\Omega}, \vec{V}, \vec{W}) &= \langle \tilde{\Omega}, \vec{e}_\alpha \rangle \langle \tilde{\omega}^\beta, \vec{V} \rangle \langle \tilde{\omega}^\gamma, \vec{W} \rangle \\ &= \Omega_\alpha V^\beta W^\gamma. \end{aligned} \quad (5.15)$$

L'action du tenseur \mathbf{T} est alors simplement définie par

$$\mathbf{T}(\tilde{\Omega}, \vec{V}, \vec{W}) = T^\alpha_{\beta\gamma} \Omega_\alpha V^\beta W^\gamma. \quad (5.16)$$

Remarquons que

$$T^\alpha_{\beta\gamma} = \mathbf{T}(\tilde{\omega}^\alpha, \vec{e}_\beta, \vec{e}_\gamma), \quad (5.17)$$

dans l'esprit des relations (5.13). En résumé, le tenseur agit linéairement sur les vecteurs et les formes concernés :

$$\mathbf{T}(\tilde{\Omega}, \vec{V}, \vec{W}) = \Omega_\alpha V^\beta W^\gamma \mathbf{T}(\tilde{\omega}^\alpha, \vec{e}_\beta, \vec{e}_\gamma). \quad (5.18)$$

Formellement, on requiert du produit tensoriel les propriétés suivantes, compatibles avec la définition opérationnelle (5.15). Elles s'ajoutent et se combinent aux axiomes habituels des espaces vectoriels.

Propriété 1. Distributivité par rapport à l'addition des espaces vectoriels. Par exemple,

$$\begin{aligned} \tilde{\Omega} \otimes (\vec{V} + \vec{W}) &= \tilde{\Omega} \otimes \vec{V} + \tilde{\Omega} \otimes \vec{W} \\ (\tilde{\Omega} + \tilde{\Sigma}) \otimes \vec{V} &= \tilde{\Omega} \otimes \vec{V} + \tilde{\Sigma} \otimes \vec{V}. \end{aligned} \quad (5.19)$$

Propriété 2. Associativité de la multiplication par un scalaire λ . Par exemple,

$$(\lambda \vec{V}) \otimes \tilde{\Omega} = \vec{V} \otimes (\lambda \tilde{\Omega}) = \lambda (\vec{V} \otimes \tilde{\Omega}). \quad (5.20)$$

Propriété 3. Associativité. Par exemple,

$$(\vec{V} \otimes \tilde{\Omega}) \otimes \tilde{\Sigma} = \vec{V} \otimes (\tilde{\Omega} \otimes \tilde{\Sigma}) = \vec{V} \otimes \tilde{\Omega} \otimes \tilde{\Sigma}. \quad (5.21)$$

Insistons sur le fait que *le produit tensoriel* n'est pas *commutatif*. Ainsi,

$$(\vec{e}_\alpha \otimes \vec{e}_\beta)(\tilde{\Omega}, \tilde{\Sigma}) = \Omega_\alpha \Sigma_\beta \neq (\vec{e}_\beta \otimes \vec{e}_\alpha)(\tilde{\Omega}, \tilde{\Sigma}) = \Omega_\beta \Sigma_\alpha.$$

A titre d'exercice, on s'entraînera à justifier des relations du genre

$$(\Omega_\alpha \tilde{\omega}^\alpha) \otimes (T^{\mu\nu} \vec{e}_\mu \otimes \vec{e}_\nu) = \Omega_\alpha T^{\mu\nu} \tilde{\omega}^\alpha \otimes \vec{e}_\mu \otimes \vec{e}_\nu,$$

ou encore

$$(\vec{V} \otimes \tilde{\Omega})(\tilde{\Sigma}, \vec{W}) = \langle \tilde{\Sigma}, \vec{V} \rangle \langle \tilde{\Omega}, \vec{W} \rangle.$$

En particulier, les relations (5.3) et (5.8) donnant les vecteurs de base de l'espace tangent et de l'espace cotangent relatifs à un autre système de coordonnées \underline{x} impliquent

$$\begin{aligned} T^{\alpha}_{\beta\gamma} \vec{e}_\alpha \otimes \tilde{\omega}^\beta \otimes \tilde{\omega}^\gamma &= \frac{\partial \underline{x}^\mu}{\partial x^\alpha} \frac{\partial x^\beta}{\partial \underline{x}^\nu} \frac{\partial x^\gamma}{\partial \underline{x}^\xi} T^{\alpha}_{\beta\gamma} \vec{e}_\mu \otimes \tilde{\omega}^\nu \otimes \tilde{\omega}^\xi. \\ &= \underline{T}^{\mu}_{\nu\xi} \vec{e}_\mu \otimes \tilde{\omega}^\nu \otimes \tilde{\omega}^\xi. \end{aligned}$$

On obtient ainsi le lien entre les composantes du tenseur dans chacun des deux systèmes de coordonnées :

$$\underline{T}^{\mu}_{\nu\xi} = \frac{\partial \underline{x}^\mu}{\partial x^\alpha} \frac{\partial x^\beta}{\partial \underline{x}^\nu} \frac{\partial x^\gamma}{\partial \underline{x}^\xi} T^{\alpha}_{\beta\gamma}, \quad (5.22)$$

équivalente à

$$T^{\alpha}_{\beta\gamma} = \frac{\partial x^\alpha}{\partial \underline{x}^\mu} \frac{\partial \underline{x}^\nu}{\partial x^\beta} \frac{\partial \underline{x}^\xi}{\partial x^\gamma} \underline{T}^{\mu}_{\nu\xi}. \quad (5.23)$$

C'est la *loi tensorielle*. Comme le montre la conjugaison des lois de transformation (5.3) et (5.8), de la linéarité (5.18) et de la formulation (5.17) des composantes du tenseur dans le système de coordonnées initial, la loi tensorielle ne fait, en somme, que traduire le résultat naturel

$$\underline{T}^{\mu}_{\nu\xi} = \mathbf{T}(\tilde{\omega}^\mu, \vec{e}_\nu, \vec{e}_\xi).$$

Du coup, la relation (5.16) exprimant l'action du tenseur s'avère entièrement covariante. En effet, on peut écrire le scalaire $\mathbf{T}(\tilde{\Omega}, \vec{V}, \vec{W})$ indifféremment sous la forme

$$\mathbf{T}(\Omega_\alpha \tilde{\omega}^\alpha, V^\beta \vec{e}_\beta, W^\gamma \vec{e}_\gamma) = T^{\alpha}_{\beta\gamma} \Omega_\alpha V^\beta W^\gamma,$$

ou bien

$$\mathbf{T}(\underline{\Omega}_\alpha \tilde{\omega}^\alpha, \underline{V}^\beta \vec{e}_\beta, \underline{W}^\gamma \vec{e}_\gamma) = \underline{T}^{\alpha}_{\beta\gamma} \underline{\Omega}_\alpha \underline{V}^\beta \underline{W}^\gamma.$$

Au demeurant, cette covariance procède directement de la loi tensorielle :

$$\begin{aligned} \mathbf{T}(\tilde{\Omega}, \vec{V}, \vec{W}) &= T^{\alpha}_{\beta\gamma} \Omega_\alpha V^\beta W^\gamma \\ &= \frac{\partial x^\alpha}{\partial \underline{x}^\mu} \frac{\partial \underline{x}^\nu}{\partial x^\beta} \frac{\partial \underline{x}^\xi}{\partial x^\gamma} \underline{T}^{\mu}_{\nu\xi} \frac{\partial \underline{x}^\rho}{\partial x^\alpha} \frac{\partial x^\beta}{\partial \underline{x}^\sigma} \frac{\partial x^\gamma}{\partial \underline{x}^\tau} \underline{\Omega}_\rho \underline{V}^\sigma \underline{W}^\tau \\ &= \underline{T}^{\mu}_{\nu\xi} \underline{\Omega}_\mu \underline{V}^\nu \underline{W}^\xi. \end{aligned}$$

Terminons en définissant la *contraction sur* (par exemple) les indices α et γ d'un tenseur $\mathbf{R} = R^{\alpha}_{\beta\gamma\delta} \vec{e}_\alpha \otimes \tilde{\omega}^\beta \otimes \tilde{\omega}^\gamma \otimes \tilde{\omega}^\delta$: c'est le nouveau tenseur

$$\begin{aligned} \mathbf{R}' &= R^{\alpha}_{\beta\gamma\delta} \langle \tilde{\omega}^\gamma, \vec{e}_\alpha \rangle \tilde{\omega}^\beta \otimes \tilde{\omega}^\delta \\ &= R^{\alpha}_{\beta\alpha\delta} \tilde{\omega}^\beta \otimes \tilde{\omega}^\delta. \end{aligned} \quad (5.24)$$

L'opération de contraction convertit ainsi un tenseur de variance (p, q) en un tenseur de variance $(p-1, q-1)$. Il est facile de voir que la relation $R'_{\beta\delta} = R^{\alpha}_{\beta\alpha\delta}$ est covariante : dans les coordonnées \underline{x} , on obtient $\underline{R}'_{\beta\delta} = \underline{R}^{\alpha}_{\beta\alpha\delta}$.

5.3 Notion de transport parallèle

On dote une variété différentielle d'une *structure affine* en se donnant un critère permettant de comparer les éléments d'un champ tensoriel définis en des points différents de cette variété. On utilise également le terme de *connexion affine*, suggestif car il s'agit précisément de « connecter » les informations relatives à ces éléments. Cela revient à définir un *transport parallèle*.

Sur une variété différentielle rapportée à des coordonnées x^α de vecteurs de base \vec{e}_α , soit une courbe $\mathcal{C}(\lambda)$ de vecteur tangent

$$\vec{U} = \frac{dx^\alpha}{d\lambda} \vec{e}_\alpha. \quad (5.25)$$

Au point $P(\lambda_0)$, le taux de variation d'un champ tensoriel \mathbf{T} selon la courbe \mathcal{C} , appelé sa *dérivée covariante dans la direction \vec{U}* et noté $\nabla_{\vec{U}} \mathbf{T}$, ou encore $\nabla_\lambda \mathbf{T}$, s'obtient en comparant le résultat du transport parallèle du champ tensoriel depuis $P(\lambda_0 + \epsilon)$ jusqu'en $P(\lambda_0)$, avec le champ au point $P(\lambda_0)$, et en passant à la limite :

$$\left(\nabla_{\vec{U}} \mathbf{T} \right)_{P(\lambda_0)} = \lim_{\epsilon \rightarrow 0} \frac{\mathbf{T}[P(\lambda_0 + \epsilon)]_{\parallel \rightarrow P(\lambda_0)} - \mathbf{T}[P(\lambda_0)]}{\epsilon}. \quad (5.26)$$

Cette relation définit en fait le transport parallèle. En pratique en effet, celui-ci résulte de

$$\mathbf{T}[P(\lambda_0 + \epsilon)]_{\parallel \rightarrow P(\lambda_0)} = \mathbf{T}[P(\lambda_0)] + \epsilon \left(\nabla_{\vec{U}} \mathbf{T} \right)_{P(\lambda_0)} + O(2). \quad (5.27)$$

Il reste toutefois à préciser la définition de la dérivée covariante : à son tour, celle-ci est guidée par la notion intuitive de parallélisme dans le cas particulier de l'espace \mathbb{R}^n de la géométrie élémentaire.

5.4 Les coordonnées cartésiennes dans \mathbb{R}^n

Cet espace, que nous qualifierons d'ores et déjà d'« euclidien » et dont les propriétés seront précisées au paragraphe 5.18, admet des systèmes de coordonnées x^α , dites *cartésiennes*, de vecteurs de base \vec{e}_α , dans lesquelles on décrète que *les composantes de la dérivée covariante d'un champ tensoriel selon la courbe $\mathcal{C}(\lambda)$ sont égales à la dérivée ordinaire, par rapport à λ , des dites composantes de ce champ*. Ainsi, un champ tensoriel est transporté parallèlement selon une courbe si et seulement si ses composantes sont constantes le long de cette courbe. Dans le cas particulier d'un champ vectoriel, on a donc

$$\left[\left(\nabla_{\vec{U}} \vec{V} \right)_{P(\lambda_0)} \right]^\beta = \left[\frac{dV^\beta(\lambda)}{d\lambda} \right]_{\lambda_0}$$

et on peut écrire

$$\left(\nabla_{\vec{U}} \vec{V} \right)^\beta = \frac{dx^\alpha}{d\lambda} \frac{\partial V^\beta}{\partial x^\alpha} = U^\alpha \frac{\partial V^\beta}{\partial x^\alpha}. \quad (5.28)$$

Notamment,

$$\left(\nabla_{\vec{e}_\alpha} \vec{V} \right)^\beta = \frac{\partial V^\beta}{\partial x^\alpha}.$$

Ainsi, l'opérateur de dérivation covariante est linéaire par rapport à la direction \vec{U} :

$$\nabla_{\vec{U}} \vec{V} = U^\alpha \nabla_{\vec{e}_\alpha} \vec{V}. \quad (5.29)$$

Compte tenu de (5.28), les vecteurs de base \vec{e}_β sont transportés parallèlement dans n'importe quelle direction :

$$\nabla_{\vec{U}} \vec{e}_\beta = 0. \quad (5.30)$$

Il est facile de généraliser ces propos à des champs tensoriels quelconques.

5.5 Les coordonnées générales dans \mathbb{R}^n

Appelons $\{x^\alpha\}$ un système de coordonnées cartésiennes parmi ceux qui ont été évoqués dans la section 5.4, et $\{\underline{x}^\alpha\}$ un système de coordonnées générales. Notons leurs vecteurs de base, respectivement, \vec{e}_α et $\underline{\vec{e}}_\alpha$. On a

$$\vec{e}_\alpha = \frac{\partial \underline{x}^\beta}{\partial x^\alpha} \underline{\vec{e}}_\beta, \quad \underline{\vec{e}}_\alpha = \frac{\partial x^\beta}{\partial \underline{x}^\alpha} \vec{e}_\beta.$$

Posons

$$\nabla_{\vec{e}_\alpha} = \nabla_\alpha \quad \text{et} \quad \nabla_{\underline{\vec{e}}_\alpha} = \nabla_{\underline{\alpha}}.$$

Pour chaque α , la ligne de coordonnée \underline{x}^α est une courbe particulière de paramètre $\lambda = \underline{x}^\alpha$ et de vecteur tangent $\underline{\vec{e}}_\alpha$. Compte tenu des résultats de la section 5.4, on peut écrire

$$\nabla_{\underline{\alpha}} = \frac{\partial x^\beta}{\partial \underline{x}^\alpha} \nabla_\beta.$$

Donc, plus généralement,

$$\nabla_{\vec{U}} = U^\beta \nabla_\beta = \frac{\partial x^\beta}{\partial \underline{x}^\alpha} \underline{U}^\alpha \nabla_\beta = \underline{U}^\alpha \nabla_{\underline{\alpha}}.$$

5.5.1 Champs de vecteurs

On a :

$$\begin{aligned} \nabla_{\underline{\alpha}} \vec{V} &= \nabla_{\underline{\alpha}} (V^\mu \underline{\vec{e}}_\mu) \\ &= \nabla_{\underline{\alpha}} \left(V^\mu \frac{\partial x^\gamma}{\partial \underline{x}^\mu} \vec{e}_\gamma \right) \\ &= \frac{\partial}{\partial \underline{x}^\alpha} \left(V^\mu \frac{\partial x^\gamma}{\partial \underline{x}^\mu} \right) \vec{e}_\gamma \\ &= \frac{\partial V^\mu}{\partial \underline{x}^\alpha} \frac{\partial x^\gamma}{\partial \underline{x}^\mu} \vec{e}_\gamma + V^\mu \frac{\partial^2 x^\gamma}{\partial \underline{x}^\alpha \partial \underline{x}^\mu} \vec{e}_\gamma \\ &= \frac{\partial V^\mu}{\partial \underline{x}^\alpha} \underline{\vec{e}}_\mu + V^\rho \frac{\partial^2 x^\gamma}{\partial \underline{x}^\alpha \partial \underline{x}^\rho} \vec{e}_\gamma \\ &= \left(\frac{\partial V^\mu}{\partial \underline{x}^\alpha} + V^\rho \frac{\partial^2 x^\gamma}{\partial \underline{x}^\alpha \partial \underline{x}^\rho} \frac{\partial \underline{x}^\mu}{\partial x^\gamma} \right) \underline{\vec{e}}_\mu \\ &= \left(\frac{\partial V^\mu}{\partial \underline{x}^\alpha} + V^\rho \Gamma_{\rho\alpha}^\mu \right) \underline{\vec{e}}_\mu \end{aligned}$$

où les quantités

$$\Gamma_{\rho\alpha}^\mu = \frac{\partial^2 x^\gamma}{\partial \underline{x}^\alpha \partial \underline{x}^\rho} \frac{\partial \underline{x}^\mu}{\partial x^\gamma}, \quad (5.31)$$

symétriques sur leurs deux indices inférieurs, portent le nom de *symboles de connexion affine*. On a donc démontré que

$$\nabla_{\underline{\alpha}} \left(\underline{V}^{\mu} \underline{\tilde{e}}_{\mu} \right) = \left(\frac{\partial \underline{V}^{\mu}}{\partial \underline{x}^{\alpha}} + \Gamma_{\rho\alpha}^{\mu} \underline{V}^{\rho} \right) \underline{\tilde{e}}_{\mu}. \quad (5.32)$$

En particulier,

$$\nabla_{\underline{\alpha}} \underline{\tilde{e}}_{\mu} = \Gamma_{\mu\alpha}^{\beta} \underline{\tilde{e}}_{\beta}. \quad (5.33)$$

Exercice 5.1 Calculer les symboles de connexion affine pour les coordonnées elliptiques $(\underline{x}^1, \underline{x}^2) = (\xi, \eta)$ dans le plan, reliées aux coordonnées cartésiennes $(x^1, x^2) = (x, y)$ par les relations

$$x = \cosh \xi \cos \eta, \quad y = \sinh \xi \sin \eta \quad (5.34)$$

où $\xi \in]0, +\infty[$, $\eta \in]0, 2\pi[$.

Solution. On calcule facilement

$$\begin{pmatrix} \partial x / \partial \xi & \partial x / \partial \eta \\ \partial y / \partial \xi & \partial y / \partial \eta \end{pmatrix} = \begin{pmatrix} \sinh \xi \cos \eta & -\cosh \xi \sin \eta \\ \cosh \xi \sin \eta & \sinh \xi \cos \eta \end{pmatrix}$$

$$\begin{aligned} \begin{pmatrix} \partial \xi / \partial x & \partial \xi / \partial y \\ \partial \eta / \partial x & \partial \eta / \partial y \end{pmatrix} &= \begin{pmatrix} \sinh \xi \cos \eta & -\cosh \xi \sin \eta \\ \cosh \xi \sin \eta & \sinh \xi \cos \eta \end{pmatrix}^{-1} \\ &= \frac{1}{\sinh^2 \xi \cos^2 \eta + \cosh^2 \xi \sin^2 \eta} \begin{pmatrix} \sinh \xi \cos \eta & \cosh \xi \sin \eta \\ -\cosh \xi \sin \eta & \sinh \xi \cos \eta \end{pmatrix}. \end{aligned}$$

Par (5.31), on en déduit :

$$\underline{\Gamma}_{11}^1 = \frac{\sinh \xi \cosh \xi}{\sinh^2 \xi \cos^2 \eta + \cosh^2 \xi \sin^2 \eta}, \quad \underline{\Gamma}_{11}^2 = \frac{-\sin \eta \cos \eta}{\sinh^2 \xi \cos^2 \eta + \cosh^2 \xi \sin^2 \eta}, \quad \text{etc.}$$

5.5.2 Champs de formes

Raisonnant comme dans la section 5.5.1, il vient

$$\begin{aligned} \nabla_{\underline{\alpha}} \tilde{\Omega} &= \nabla_{\underline{\alpha}} \left(\underline{\Omega}_{\mu} \tilde{\omega}^{\mu} \right) \\ &= \nabla_{\underline{\alpha}} \left(\underline{\Omega}_{\mu} \frac{\partial \underline{x}^{\mu}}{\partial x^{\gamma}} \tilde{\omega}^{\gamma} \right) \\ &= \frac{\partial}{\partial \underline{x}^{\alpha}} \left(\underline{\Omega}_{\mu} \frac{\partial \underline{x}^{\mu}}{\partial x^{\gamma}} \right) \tilde{\omega}^{\gamma} \\ &= \frac{\partial \underline{\Omega}_{\mu}}{\partial \underline{x}^{\alpha}} \frac{\partial \underline{x}^{\mu}}{\partial x^{\gamma}} \tilde{\omega}^{\gamma} + \underline{\Omega}_{\mu} \frac{\partial}{\partial \underline{x}^{\alpha}} \frac{\partial \underline{x}^{\mu}}{\partial x^{\gamma}} \tilde{\omega}^{\gamma} \\ &= \frac{\partial \underline{\Omega}_{\mu}}{\partial \underline{x}^{\alpha}} \tilde{\omega}^{\mu} + \underline{\Omega}_{\rho} \frac{\partial}{\partial \underline{x}^{\alpha}} \frac{\partial \underline{x}^{\rho}}{\partial x^{\gamma}} \tilde{\omega}^{\gamma} \\ &= \left(\frac{\partial \underline{\Omega}_{\mu}}{\partial \underline{x}^{\alpha}} + \underline{\Omega}_{\rho} \frac{\partial}{\partial \underline{x}^{\alpha}} \frac{\partial \underline{x}^{\rho}}{\partial x^{\gamma}} \frac{\partial x^{\gamma}}{\partial \underline{x}^{\mu}} \right) \tilde{\omega}^{\mu}. \end{aligned}$$

Or,

$$\frac{\partial \underline{x}^{\rho}}{\partial x^{\gamma}} \frac{\partial x^{\gamma}}{\partial \underline{x}^{\mu}} = \delta_{\mu}^{\rho}.$$

En dérivant cette relation par rapport à \underline{x}^α , on obtient

$$\frac{\partial}{\partial \underline{x}^\alpha} \frac{\partial \underline{x}^\rho}{\partial x^\gamma} \frac{\partial x^\gamma}{\partial \underline{x}^\mu} + \frac{\partial \underline{x}^\rho}{\partial x^\gamma} \frac{\partial^2 x^\gamma}{\partial \underline{x}^\alpha \partial \underline{x}^\mu} = 0,$$

donc

$$\begin{aligned} \frac{\partial}{\partial \underline{x}^\alpha} \frac{\partial \underline{x}^\rho}{\partial x^\gamma} \frac{\partial x^\gamma}{\partial \underline{x}^\mu} &= - \frac{\partial \underline{x}^\rho}{\partial x^\gamma} \frac{\partial^2 x^\gamma}{\partial \underline{x}^\alpha \partial \underline{x}^\mu} \\ &= -\underline{\Gamma}_{\mu\alpha}^\rho, \end{aligned}$$

en vertu de la définition (5.31). Ainsi,

$$\nabla_{\underline{\alpha}} \left(\underline{\Omega}_\mu \tilde{\omega}^\mu \right) = \left(\frac{\partial \underline{\Omega}_\mu}{\partial \underline{x}^\alpha} - \underline{\Gamma}_{\mu\alpha}^\rho \underline{\Omega}_\rho \right) \tilde{\omega}^\mu. \quad (5.35)$$

En particulier,

$$\nabla_{\underline{\alpha}} \tilde{\omega}^\mu = -\underline{\Gamma}_{\beta\alpha}^\mu \tilde{\omega}^\beta. \quad (5.36)$$

5.5.3 Champs tensoriels quelconques

La dérivée covariante d'un champ tensoriel quelconque s'obtient à partir de la distributivité du produit tensoriel sur les combinaisons linéaires de vecteurs ou de formes, de la définition imposée en coordonnées cartésiennes (section 5.4) et des résultats obtenus dans les sections 5.5.1 et 5.5.2. Considérons par exemple un champ tensoriel \mathbf{T} une fois contravariant et deux fois covariant. On a :

$$\begin{aligned} \nabla_{\underline{\alpha}} \mathbf{T} &= \nabla_{\underline{\alpha}} \left(\underline{T}_{\nu\xi}^\mu \tilde{e}_\mu \otimes \tilde{\omega}^\nu \otimes \tilde{\omega}^\xi \right) \\ &= \nabla_{\underline{\alpha}} \left(\underline{T}_{\nu\xi}^\mu \frac{\partial x^\rho}{\partial \underline{x}^\mu} \frac{\partial \underline{x}^\nu}{\partial x^\sigma} \frac{\partial \underline{x}^\xi}{\partial x^\tau} \tilde{e}_\rho \otimes \tilde{\omega}^\sigma \otimes \tilde{\omega}^\tau \right) \\ &= \frac{\partial}{\partial \underline{x}^\alpha} \left(\underline{T}_{\nu\xi}^\mu \frac{\partial x^\rho}{\partial \underline{x}^\mu} \frac{\partial \underline{x}^\nu}{\partial x^\sigma} \frac{\partial \underline{x}^\xi}{\partial x^\tau} \right) \tilde{e}_\rho \otimes \tilde{\omega}^\sigma \otimes \tilde{\omega}^\tau \\ &= \dots \end{aligned}$$

Il vient finalement

$$\nabla_{\underline{\alpha}} \left(\underline{T}_{\nu\xi}^\mu \tilde{e}_\mu \otimes \tilde{\omega}^\nu \otimes \tilde{\omega}^\xi \right) = \left(\frac{\partial \underline{T}_{\nu\xi}^\mu}{\partial \underline{x}^\alpha} + \underline{\Gamma}_{\lambda\alpha}^\mu \underline{T}_{\nu\xi}^\lambda - \underline{\Gamma}_{\nu\alpha}^\lambda \underline{T}_{\lambda\xi}^\mu - \underline{\Gamma}_{\xi\alpha}^\lambda \underline{T}_{\nu\lambda}^\mu \right) \tilde{e}_\mu \otimes \tilde{\omega}^\nu \otimes \tilde{\omega}^\xi. \quad (5.37)$$

Nous allégerons l'écriture en notant dorénavant $\{x^\alpha\}$ un système de coordonnées générales.

Exercice 5.2 Démontrer que la dérivée covariante obéit à la règle de Leibniz pour le produit intérieur :

$$\nabla_\alpha \langle \tilde{\Omega}, \tilde{V} \rangle = \langle \nabla_\alpha \tilde{\Omega}, \tilde{V} \rangle + \langle \tilde{\Omega}, \nabla_\alpha \tilde{V} \rangle. \quad (5.38)$$

Suggestion. Le membre de gauche est la dérivée covariante du scalaire $\Omega_\mu V^\mu$, égale à sa dérivée ordinaire puisque les scalaires ne dépendent pas de la base.

Exercice 5.3 Montrer que la dérivée covariante vérifie la règle de Leibniz pour le produit tensoriel, par exemple :

$$\begin{aligned} \nabla_\alpha (T_{\nu\xi}^\mu \tilde{e}_\mu \otimes \tilde{\omega}^\nu \otimes \tilde{\omega}^\xi) &= \frac{\partial T_{\nu\xi}^\mu}{\partial x^\alpha} \tilde{e}_\mu \otimes \tilde{\omega}^\nu \otimes \tilde{\omega}^\xi + T_{\nu\xi}^\mu (\nabla_\alpha \tilde{e}_\mu) \otimes \tilde{\omega}^\nu \otimes \tilde{\omega}^\xi \\ &+ T_{\nu\xi}^\mu \tilde{e}_\mu \otimes (\nabla_\alpha \tilde{\omega}^\nu) \otimes \tilde{\omega}^\xi + T_{\nu\xi}^\mu \tilde{e}_\mu \otimes \tilde{\omega}^\nu \otimes (\nabla_\alpha \tilde{\omega}^\xi). \end{aligned} \quad (5.39)$$

5.6 Propriétés axiomatiques de la dérivée covariante

Soit une variété différentielle quelconque rapportée à des coordonnées x^α de vecteurs de base \vec{e}_α et de formes duales $\tilde{\omega}^\alpha$. La dérivée covariante a pour vocation de connecter l'information relative aux espaces tangents définis en des points voisins. Ses propriétés doivent être compatibles avec celles du cas particulier de l'espace \mathbb{R}^n , étudiées dans la section 5.5. Il est dès lors naturel de *systematiser les propriétés imposées à la dérivée covariante sur une variété différentielle quelconque en exigeant qu'elles soient formellement identiques à celles établies de façon heuristique dans l'espace euclidien rapporté à des coordonnées générales*. La similitude entre ces deux situations est, d'un point de vue physique, en étroite relation avec le principe d'équivalence.

Propriété 1. La dérivée covariante est linéaire par rapport à la direction de dérivation et additive par rapport aux tenseurs auxquels elle s'applique ; par exemple, dans le cas de champs vectoriels \vec{V} et \vec{W} :

$$\nabla_{\vec{U}} \vec{V} = U^\alpha \nabla_{\vec{e}_\alpha} \vec{V} \quad (5.40)$$

$$\nabla_{\vec{U}} (\vec{V} + \vec{W}) = \nabla_{\vec{U}} \vec{V} + \nabla_{\vec{U}} \vec{W}. \quad (5.41)$$

Propriété 2. La dérivée covariante d'un champ scalaire f est identique à la dérivée ordinaire :

$$\nabla_{\vec{U}} f = U^\alpha \frac{\partial f}{\partial x^\alpha}. \quad (5.42)$$

Propriété 3. La dérivée covariante vérifie la règle de Leibniz pour le produit tensoriel, par exemple :

$$\nabla_{\vec{U}} (\vec{V} \otimes \vec{W}) = (\nabla_{\vec{U}} \vec{V}) \otimes \vec{W} + \vec{V} \otimes (\nabla_{\vec{U}} \vec{W}) \quad (5.43)$$

$$\nabla_{\vec{U}} (f \vec{V}) = (\nabla_{\vec{U}} f) \vec{V} + f (\nabla_{\vec{U}} \vec{V}). \quad (5.44)$$

En pratique, la dérivée covariante est déterminée par la donnée des *symboles de connexion affine* :

$$\nabla_\mu \vec{e}_\alpha = \Gamma_{\alpha\mu}^\beta \vec{e}_\beta, \quad (5.45)$$

Etablissons leur loi de transformation.

$$\begin{aligned} \Gamma_{\xi\nu}^\rho \vec{e}_\rho &= \nabla_{\underline{x}} \vec{e}_\xi \\ &= \nabla_{\underline{x}} \left(\frac{\partial x^\sigma}{\partial \underline{x}^\xi} \vec{e}_\sigma \right) \\ &= \frac{\partial^2 x^\sigma}{\partial \underline{x}^\nu \partial \underline{x}^\xi} \vec{e}_\sigma + \frac{\partial x^\sigma}{\partial \underline{x}^\xi} \nabla_{\underline{x}} \vec{e}_\sigma \\ &= \frac{\partial^2 x^\sigma}{\partial \underline{x}^\nu \partial \underline{x}^\xi} \vec{e}_\sigma + \frac{\partial x^\sigma}{\partial \underline{x}^\xi} \frac{\partial x^\varphi}{\partial \underline{x}^\nu} \nabla_\varphi \vec{e}_\sigma \\ &= \left(\frac{\partial^2 x^\pi}{\partial \underline{x}^\nu \partial \underline{x}^\xi} + \frac{\partial x^\sigma}{\partial \underline{x}^\xi} \frac{\partial x^\varphi}{\partial \underline{x}^\nu} \Gamma_{\sigma\varphi}^\pi \right) \vec{e}_\pi \\ &= \left(\frac{\partial^2 x^\pi}{\partial \underline{x}^\nu \partial \underline{x}^\xi} \frac{\partial \underline{x}^\rho}{\partial x^\pi} + \frac{\partial x^\sigma}{\partial \underline{x}^\xi} \frac{\partial x^\varphi}{\partial \underline{x}^\nu} \frac{\partial \underline{x}^\rho}{\partial x^\pi} \Gamma_{\sigma\varphi}^\pi \right) \vec{e}_\rho. \end{aligned}$$

[La troisième ligne résulte de (5.44) et (5.42). Le deuxième terme de la quatrième ligne résulte de (5.40).] Donc

$$\underline{\Gamma}_{\xi\nu}^{\rho} = \frac{\partial^2 x^{\pi}}{\partial \underline{x}^{\xi} \partial \underline{x}^{\nu}} \frac{\partial \underline{x}^{\rho}}{\partial x^{\pi}} + \frac{\partial x^{\sigma}}{\partial \underline{x}^{\xi}} \frac{\partial x^{\varphi}}{\partial \underline{x}^{\nu}} \frac{\partial \underline{x}^{\rho}}{\partial x^{\pi}} \Gamma_{\sigma\varphi}^{\pi}. \quad (5.46)$$

Ainsi,

- les symboles de connexion affine ne constituent pas les composantes d'un tenseur, sauf pour les transformations *linéaires* de coordonnées ;
- par contre, les quantités

$$\mathcal{T}_{\mu\nu}^{\xi} = \Gamma_{\mu\nu}^{\xi} - \Gamma_{\nu\mu}^{\xi} \quad (5.47)$$

constituent les composantes d'un tenseur \mathcal{T} , appelé *tenseur de torsion*.

Pour dériver un champ vectoriel \vec{V} quelconque on utilise les propriétés (5.41), (5.42) et (5.44), et on obtient

$$(\nabla_{\beta} \vec{V})^{\alpha} = \frac{\partial V^{\alpha}}{\partial x^{\beta}} + \Gamma_{\mu\beta}^{\alpha} V^{\mu}. \quad (5.48)$$

On raccourcit habituellement l'écriture en notant

$$(\nabla_{\beta} \vec{V})^{\alpha} = V^{\alpha}_{;\beta}.$$

L'objet $\nabla \vec{V}$, défini par ses composantes

$$\begin{aligned} (\nabla \vec{V})^{\alpha}_{\beta} &= \nabla \vec{V}(\tilde{\omega}^{\alpha}, \vec{e}_{\beta}) \\ &= (\nabla_{\beta} \vec{V})^{\alpha} \end{aligned}$$

constitue un tenseur d'ordre deux, une fois contravariant et une fois covariant. En effet,

$$\begin{aligned} (\nabla_{\beta} \vec{V})^{\alpha} &= \frac{\partial \underline{x}^{\nu}}{\partial x^{\beta}} (\nabla_{\underline{\nu}} \vec{V})^{\alpha} \\ &= \frac{\partial \underline{x}^{\nu}}{\partial x^{\beta}} \frac{\partial x^{\alpha}}{\partial \underline{x}^{\rho}} (\nabla_{\underline{\nu}} \vec{V})^{\rho} \end{aligned}$$

[La première ligne résulte de la linéarité (5.40)]. On utilise souvent la notation suggestive

$$\nabla \vec{V}(\cdot, \vec{U}) = \nabla_{\vec{U}} \vec{V}. \quad (5.49)$$

Propriété 4. La dérivée covariante commute avec la contraction.

Or, la contraction du tenseur $\tilde{\omega}^{\beta} \otimes \vec{e}_{\alpha}$ est égale au produit intérieur $\langle \tilde{\omega}^{\beta}, \vec{e}_{\alpha} \rangle$. Compte tenu des propriétés 3 et 4, on a donc

$$\nabla_{\mu} \langle \tilde{\omega}^{\beta}, \vec{e}_{\alpha} \rangle = \langle \nabla_{\mu} \tilde{\omega}^{\beta}, \vec{e}_{\alpha} \rangle + \langle \tilde{\omega}^{\beta}, \nabla_{\mu} \vec{e}_{\alpha} \rangle.$$

Vu la définition $\langle \tilde{\omega}^{\beta}, \vec{e}_{\alpha} \rangle = \delta_{\alpha}^{\beta}$ et la propriété 2, on obtient

$$\langle \nabla_{\mu} \tilde{\omega}^{\beta}, \vec{e}_{\alpha} \rangle = - \langle \tilde{\omega}^{\beta}, \nabla_{\mu} \vec{e}_{\alpha} \rangle.$$

La définition (5.45) permet de conclure :

$$\nabla_{\mu} \tilde{\omega}^{\beta} = -\Gamma_{\alpha\mu}^{\beta} \tilde{\omega}^{\alpha}. \quad (5.50)$$

La dérivée covariante d'un champ quelconque de formes $\tilde{\Omega}$ résulte des propriétés (5.41), (5.42) et (5.44) :

$$\left(\nabla_{\mu} \tilde{\Omega}\right)_{\beta} = \frac{\partial \Omega_{\beta}}{\partial x^{\mu}} - \Gamma_{\beta\mu}^{\alpha} \Omega_{\alpha}. \quad (5.51)$$

On utilisera fréquemment la notation

$$\left(\nabla_{\alpha} \tilde{\Omega}\right)_{\beta} = \Omega_{\beta;\alpha}.$$

La dérivée covariante d'un champ tensoriel quelconque découle également des propriétés 1 à 3. Par exemple, pour le champ tensoriel $\mathbf{T} = T_{\alpha\beta} \tilde{\omega}^{\alpha} \otimes \tilde{\omega}^{\beta}$, on trouve

$$\left(\nabla_{\mu} \mathbf{T}\right)_{\alpha\beta} = \frac{\partial T_{\alpha\beta}}{\partial x^{\mu}} - \Gamma_{\alpha\mu}^{\xi} T_{\xi\beta} - \Gamma_{\beta\mu}^{\xi} T_{\alpha\xi}. \quad (5.52)$$

Propriété 5. Le tenseur de torsion est nul; en d'autres termes, les symboles de connexion affine sont symétriques sur leurs deux indices inférieurs.

Exercice 5.4 Démontrer la relation

$$\nabla_{\mu} \left[\mathbf{T}(\vec{V}, \vec{W}) \right] = \left(\nabla_{\mu} \mathbf{T}\right) (\vec{V}, \vec{W}) + \mathbf{T}(\nabla_{\mu} \vec{V}, \vec{W}) + \mathbf{T}(\vec{V}, \nabla_{\mu} \vec{W}). \quad (5.53)$$

5.7 Géodésiques dans un espace affine

Soit, sur une variété affine, une courbe $\mathcal{C}(\lambda)$ de vecteur tangent \vec{U} donné par (5.25). On dit qu'elle est *géodésique* si \vec{U} est transporté parallèlement dans sa propre direction :

$$\nabla_{\vec{U}} \vec{U} = 0. \quad (5.54)$$

Insistons sur le fait qu'une géodésique n'est pas caractérisée seulement par un lieu géométrique de points : il faut également préciser le paramétrage. Or, l'équation (5.54) n'est pas nécessairement préservée lors d'un changement de paramètre. Imaginons en effet un autre étiquetage $\mu = f(\lambda)$. Le nouveau vecteur tangent prend la valeur

$$\begin{aligned} (U')^{\alpha} &= \frac{dx^{\alpha}}{d\mu} \\ &= \left(\frac{d\mu}{d\lambda}\right)^{-1} \frac{dx^{\alpha}}{d\lambda} \\ &= \left(\frac{d\mu}{d\lambda}\right)^{-1} U^{\alpha} \end{aligned}$$

et on a

$$\begin{aligned} \nabla_{\vec{U}'} \vec{U}' &= \left(\frac{d\mu}{d\lambda}\right)^{-1} \nabla_{\vec{U}} \left[\left(\frac{d\mu}{d\lambda}\right)^{-1} \vec{U} \right] \\ &= - \left(\frac{d\mu}{d\lambda}\right)^{-3} \frac{d^2\mu}{d\lambda^2} \vec{U} + \left(\frac{d\mu}{d\lambda}\right)^{-2} \nabla_{\vec{U}} \vec{U} \\ &= - \left(\frac{d\mu}{d\lambda}\right)^{-3} \frac{d^2\mu}{d\lambda^2} \vec{U}. \end{aligned}$$

Cette quantité n'est nulle que si la relation entre μ et λ est *linéaire*.

Souvent, par abus de langage, on donne le nom de *géodésique* au lieu des points, quel que soit le paramétrage choisi ; et on appelle *affine* un paramètre λ pour lequel l'équation (5.54) est vérifiée. Dès lors, tous les paramètres affines sont reliés entre eux par des transformations linéaires.

En composantes, la relation (5.54) s'écrit

$$\begin{aligned} 0 &= U^\alpha (\nabla_\alpha \vec{U})^\mu \\ &= \frac{dx^\alpha}{d\lambda} \left(\frac{\partial U^\mu}{\partial x^\alpha} + \Gamma_{\beta\alpha}^\mu U^\beta \right) \\ &= \frac{dU^\mu}{d\lambda} + \Gamma_{\beta\alpha}^\mu U^\beta U^\alpha, \end{aligned}$$

donc

$$0 = \frac{d^2 x^\mu}{d\lambda^2} + \Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda}. \quad (5.55)$$

Remarquons que *la connaissance des géodésiques est équivalente à la connaissance des symboles de connexion affine (supposés symétriques, propriété 5).*

5.8 Les variétés riemanniennes

Si on leur attribue la même origine, les paramètres affines d'une géodésique sont tous proportionnels entre eux. Cette propriété permet d'interpréter l'abscisse curviligne comme une sorte de *distance* mesurée, le long de la géodésique, relativement à l'étalonnage correspondant au paramètre affine choisi. Ainsi, en exploitant la définition du parallélisme il est possible de graduer une droite de l'espace euclidien en reportant bout à bout son vecteur tangent, ou de découper un segment en un nombre arbitraire de « parties égales » (théorème de Thalès). Cependant, on ne dispose d'aucun critère pour comparer les « distances » établies de cette façon selon deux géodésiques différentes : la structure affine seule ne permet pas de construire un cercle ou de faire de la trigonométrie. Il faudrait donc intégrer dans le formalisme un outil habilité à modéliser les mesures de longueurs et d'angles nécessitées par la géométrie pratique. En d'autres termes, il s'agit d'introduire sur la variété différentielle une entité mathématique destinée à y jouer le rôle de la règle et du rapporteur.

Une variété différentielle est dite *riemannienne* si, en tout point, on la munit d'un *tenseur métrique*

$$\mathbf{g} = g_{\alpha\beta} \tilde{\omega}^\alpha \otimes \tilde{\omega}^\beta$$

défini symétrique ($g_{\alpha\beta} = g_{\beta\alpha}$) et non dégénéré (la matrice G constituée des composantes $g_{\alpha\beta}$ est inversible).³ Il agit linéairement sur deux vecteurs $\vec{X} = X^\alpha \vec{e}_\alpha$ et $\vec{Y} = Y^\alpha \vec{e}_\alpha$ de l'espace tangent au point considéré, pour produire le scalaire

$$\mathbf{g}(\vec{X}, \vec{Y}) = g_{\alpha\beta} X^\alpha Y^\beta \quad (5.56)$$

également désigné par le symbole $\vec{X} \cdot \vec{Y}$ et appelé *produit scalaire* des vecteurs \vec{X} et \vec{Y} . Si deux points infiniment proches ont pour coordonnées respectives \vec{x} et $\vec{x} + d\vec{x}$, le carré de leur distance est qualifié d'*élément de longueur* et on note

$$ds^2 = \mathbf{g}(d\vec{x}, d\vec{x}).$$

3. D'ordinaire on ajoute la *condition de positivité* : $\forall \vec{X} \neq 0, \mathbf{g}(\vec{X}, \vec{X}) > 0$. Si elle n'est pas satisfaite (ce qui est le cas en théorie de la relativité), on parle usuellement de *pseudo-métrique* et de *variété pseudo-riemannienne*. Pour ne pas alourdir le texte nous omettons le préfixe.

Par exemple, l'élément de longueur du plan euclidien rapporté aux coordonnées elliptiques (5.34) s'écrit

$$\begin{aligned} ds^2 &= dx^2 + dy^2 \\ &= (\sinh^2 \xi \cos^2 \eta + \cosh^2 \xi \sin^2 \eta) (d\xi^2 + d\eta^2). \end{aligned}$$

On dote l'espace cotangent d'un produit scalaire à l'aide du *tenseur métrique inverse*

$$\underline{g}^{-1} = g^{\alpha\beta} \vec{e}_\alpha \otimes \vec{e}_\beta.$$

Ses composantes $g^{\alpha\beta}$ sont définies dans tout système de coordonnées comme étant les éléments de la matrice G^{-1} inverse de la matrice G . Montrons qu'elles satisfont bien la loi tensorielle

$$\underline{g}^{\mu\nu} = \frac{\partial \underline{x}^\mu}{\partial x^\alpha} \frac{\partial \underline{x}^\nu}{\partial x^\beta} g^{\alpha\beta}. \quad (5.57)$$

Pour cela, nommons J la matrice jacobienne d'éléments $J_{\beta\nu} = \partial x^\beta / \partial \underline{x}^\nu$. La matrice inverse J^{-1} vérifie alors $(J^{-1})_{\mu\alpha} = \partial \underline{x}^\mu / \partial x^\alpha$. Sous forme matricielle, la loi tensorielle pour le tenseur métrique

$$\underline{g}_{\mu\nu} = \frac{\partial x^\alpha}{\partial \underline{x}^\mu} \frac{\partial x^\beta}{\partial \underline{x}^\nu} g_{\alpha\beta} \quad (5.58)$$

prend la forme

$$\underline{G}_{\mu\nu} = \sum_\alpha \sum_\beta \tilde{J}_{\mu\alpha} G_{\alpha\beta} J_{\beta\nu},$$

c'est-à-dire $\underline{G} = \tilde{J} G J$. En inversant les deux membres on trouve $\underline{G}^{-1} = J^{-1} G^{-1} \tilde{J}^{-1}$, ou encore

$$(\underline{G}^{-1})_{\mu\nu} = \sum_\alpha \sum_\beta (J^{-1})_{\mu\alpha} (G^{-1})_{\alpha\beta} (\tilde{J}^{-1})_{\beta\nu},$$

égalité équivalente à la relation (5.57) qu'il fallait prouver.

La métrique permet de jeter un pont intéressant entre les éléments de l'espace tangent et ceux de l'espace cotangent. En effet, à tout vecteur $\vec{V} = V^\alpha \vec{e}_\alpha$ on peut associer sans ambiguïté une forme $\tilde{V} = V_\alpha \tilde{\omega}^\alpha$ par la condition

$$\forall \vec{W} : \langle \tilde{V}, \vec{W} \rangle = \underline{g}(\vec{V}, \vec{W}). \quad (5.59)$$

On utilise habituellement la notation symbolique

$$\tilde{V} = \underline{g}(\vec{V}, \cdot), \quad (5.60)$$

suggestive car le membre de droite présente une case libre prête à accueillir un vecteur afin d'engendrer linéairement un scalaire, ce qui correspond exactement au concept de forme différentielle. Posant $\vec{W} = \vec{e}_\alpha$ dans (5.59), on déduit les composantes de \tilde{V} :

$$V_\alpha = g_{\alpha\beta} V^\beta. \quad (5.61)$$

Cette relation peut être inversée, ce qui assure que le lien entre vecteurs et formes est biunivoque :

$$V^\alpha = g^{\alpha\beta} V_\beta. \quad (5.62)$$

Cela revient à associer un vecteur $\vec{\Omega}$ à une forme $\tilde{\Omega}$ en posant

$$\vec{\Omega} = \underline{g}^{-1}(\cdot, \tilde{\Omega}). \quad (5.63)$$

Ainsi, le *vecteur gradient* d'un champ scalaire Φ a pour composantes

$$(\vec{\nabla}\Phi)^\alpha = g^{\alpha\beta} (\tilde{d}\Phi)_\beta. \quad (5.64)$$

Cette procédure s'étend aisément à des tenseurs de variance quelconque. Par exemple, à un tenseur \mathbf{R} de variance $(1, 3)$ on associe un tenseur \mathbf{R}' de variance $(0, 4)$ par la relation

$$\mathbf{R}'(\vec{V}, \vec{X}, \vec{Y}, \vec{Z}) = \mathbf{R}(\tilde{V}, \vec{X}, \vec{Y}, \vec{Z})$$

où $\tilde{V} = \mathbf{g}(\vec{V}, \cdot)$, conformément à (5.60). Il s'ensuit

$$R'_{\alpha\beta\gamma\delta} = g_{\alpha\mu} R^\mu_{\beta\gamma\delta}.$$

Par commodité, on omet en général le « prime » et on écrit $R_{\alpha\beta\gamma\delta} = g_{\alpha\mu} R^\mu_{\beta\gamma\delta}$. La métrique joue ainsi le rôle d'un *ascenseur d'indices*.

Exercice 5.5 Montrer que $\vec{\nabla}\Phi$ est orthogonal à l'hypersurface tangente à l'équipotentielle du champ Φ au point considéré.

Preuve. Pour n'importe quel vecteur \vec{T} appartenant à cette hypersurface on a, par définition, $\langle \tilde{d}\Phi, \vec{T} \rangle = 0$. Donc, en vertu de (5.59),

$$\begin{aligned} \mathbf{g}(\vec{\nabla}\Phi, \vec{T}) &= \langle \tilde{d}\Phi, \vec{T} \rangle \\ &= 0. \end{aligned}$$

5.9 Coordonnées localement cartésiennes

Soit un tenseur métrique défini sur une variété riemannienne de dimension n , au voisinage d'un point P . Notons $g_{\lambda\mu}(\vec{x})$ ses composantes dans un système de coordonnées $\{x^\rho\}$. Ses composantes dans un autre système de coordonnées $\{\underline{x}^\rho\}$ sont fournies par la loi tensorielle :

$$\underline{g}_{\alpha\beta}(\underline{\vec{x}}) = \frac{\partial x^\lambda}{\partial \underline{x}^\alpha} \frac{\partial x^\mu}{\partial \underline{x}^\beta} g_{\lambda\mu}[\vec{x}(\underline{\vec{x}})] \quad (5.65)$$

Posons

$$A_\alpha^\lambda = \frac{\partial x^\lambda}{\partial \underline{x}^\alpha}(P), \quad B_{\alpha\beta}^\lambda = \frac{\partial^2 x^\lambda}{\partial \underline{x}^\alpha \partial \underline{x}^\beta}(P), \quad C_{\alpha\beta\gamma}^\lambda = \frac{\partial^3 x^\lambda}{\partial \underline{x}^\alpha \partial \underline{x}^\beta \partial \underline{x}^\gamma}(P). \quad (5.66)$$

Il y a n^2 quantités A_α^λ et $n D_2^n$ quantités⁴ $B_{\alpha\beta}^\lambda$. Quant aux $C_{\alpha\beta\gamma}^\lambda$, il y en a $n D_3^n$, c'est-à-dire $n^2(n+1)(n+2)/6$. Ces nombres déterminent les premiers termes du développement en série du changement de variables au voisinage de P : en notant \vec{x}_P et $\underline{\vec{x}}_P$, respectivement, les coordonnées de P dans chacun des deux systèmes, on a

$$\frac{\partial x^\lambda}{\partial \underline{x}^\alpha}(\underline{\vec{x}}) = A_\alpha^\lambda + (\underline{x}^\beta - \underline{x}_P^\beta) B_{\alpha\beta}^\lambda + \frac{1}{2} (\underline{x}^\beta - \underline{x}_P^\beta) (\underline{x}^\gamma - \underline{x}_P^\gamma) C_{\alpha\beta\gamma}^\lambda + \dots \quad (5.67)$$

4. Le symbole D_p^n désigne le nombre de combinaisons avec répétitions de n objets pris p à p (nombre de façons de tirer p balles dans n cibles). Il est facile de démontrer que

$$D_p^n = C_p^{n+p-1} = C_{n-1}^{n+p-1}.$$

Il est possible de trouver des nombres A_α^λ et $B_{\alpha\beta}^\lambda$ engendrant des *coordonnées localement cartésiennes* $\{\underline{x}^p\}$ (qui, en référence à la physique, pourraient s'appeler *localement inertielles*) dans lesquelles le tenseur métrique prend la forme diagonale canonique⁵

$$\underline{g}_{\alpha\beta}(P) = \text{diag}(\pm 1, \pm 1, \dots, \pm 1) \quad (5.68)$$

et où s'annulent ses dérivées premières :

$$\frac{\partial \underline{g}_{\alpha\beta}}{\partial \underline{x}^\gamma}(P) = 0. \quad (5.69)$$

Pour prouver cela, considérons la loi tensorielle (5.65) et ses dérivées. On obtient

$$\underline{g}_{\alpha\beta}(P) = A_\alpha^\lambda A_\beta^\mu g_{\lambda\mu}(P) \quad (5.70)$$

$$\underline{g}_{\alpha\beta,\gamma}(P) = \left(B_{\alpha\gamma}^\lambda A_\beta^\mu + B_{\beta\gamma}^\mu A_\alpha^\lambda \right) g_{\lambda\mu}(P) + A_\alpha^\lambda A_\beta^\mu A_\gamma^\nu g_{\lambda\mu,\nu}(P) \quad (5.71)$$

$$\begin{aligned} \underline{g}_{\alpha\beta,\gamma\delta}(P) &= \left(A_\beta^\mu C_{\alpha\gamma\delta}^\lambda + B_{\alpha\gamma}^\lambda B_{\beta\delta}^\mu + A_\alpha^\lambda C_{\beta\gamma\delta}^\mu + B_{\beta\gamma}^\mu B_{\alpha\delta}^\lambda \right) g_{\lambda\mu}(P) \\ &+ \left(A_\beta^\mu A_\delta^\nu B_{\alpha\gamma}^\lambda + A_\alpha^\lambda A_\delta^\nu B_{\beta\gamma}^\mu + A_\beta^\mu A_\gamma^\nu B_{\alpha\delta}^\lambda + A_\alpha^\lambda A_\gamma^\nu B_{\beta\delta}^\mu + A_\alpha^\lambda A_\beta^\mu B_{\gamma\delta}^\nu \right) g_{\lambda\mu,\nu}(P) \\ &+ A_\alpha^\lambda A_\beta^\mu A_\gamma^\nu A_\delta^\xi g_{\lambda\mu,\nu\xi}(P) \end{aligned} \quad (5.72)$$

où, pour alléger l'écriture, on a posé

$$\underline{g}_{\alpha\beta,\gamma} = \frac{\partial \underline{g}_{\alpha\beta}}{\partial \underline{x}^\gamma}(\vec{x}) \quad \text{et} \quad \underline{g}_{\alpha\beta,\gamma\delta} = \frac{\partial^2 \underline{g}_{\alpha\beta}}{\partial \underline{x}^\gamma \partial \underline{x}^\delta}(\vec{x})$$

(les dérivées secondes n'interviendront pas dans cette démonstration-ci mais leur calcul sera utilisé dans la section 5.16).

Les $n(n+1)/2$ conditions (5.68) peuvent être satisfaites par un choix judicieux des coefficients A_α^λ . Il subsiste $n(n-1)/2$ degrés de liberté, correspondant à la dimension du groupe de transformations qui préserve la forme diagonale canonique du tenseur métrique. Quant aux équations (5.69), elles constituent un système linéaire de $n^2(n+1)/2$ équations pour le même nombre d'inconnues $B_{\alpha\beta}^\lambda$. Génériquement il admet donc une solution.

5.10 Les connexions riemanniennes

La définition du parallélisme utilisée en physique n'est pas arbitraire et se trouve liée à la métrique. Emmanuel Kant avait déjà remarqué que l'intuition associe une propriété *métrique* au concept *affine* de ligne droite dans l'espace euclidien :⁶

« Que la ligne droite soit entre deux points la plus courte, c'est une proposition synthétique. Car mon concept de *droit* ne contient rien qui se rapporte à la quantité, mais seulement une qualité. Le concept de plus court vient donc entièrement s'ajouter, et ne peut être tiré par aucune analyse du concept de la ligne droite. Il faut donc s'aider de l'intuition, au moyen de laquelle seulement la synthèse est possible. »

5. D'après la loi tensorielle et le théorème d'inertie de J.J. Sylvester, le nombre de signes + (resp. de signes -) apparaissant dans (5.68) est égal au nombre de valeurs propres strictement positives (resp. négatives) de la matrice $\{g_{\lambda\mu}(P)\}$.

6. E. Kant, *Critique de la raison pure*, édition publiée sous la direction de F. Alquié, traduit de l'allemand par A. J.-L. Delamarre et F. Marty à partir de la traduction de J. Berni, folio essais numéro 145, Gallimard, 1980, p. 76.

Dans les traités de géométrie euclidienne élémentaire, on décide implicitement que, dans un système de coordonnées où les composantes du tenseur métrique sont constantes, un champ vectoriel est transporté parallèlement si et seulement si ses composantes sont constantes. Cela implique l'annulation de tous les symboles de connexion affine dans ces coordonnées. Ainsi, on impose un lien entre la métrique et le parallélisme. Il peut être généralisé comme suit.

Théorème 5.1 *Sachant qu'en un point P donné d'une variété riemannienne il est possible de choisir un système de coordonnées localement cartésiennes dans lesquelles s'annulent les dérivées premières du tenseur métrique (voir section 5.9), on définit complètement une connexion, c'est-à-dire un critère de parallélisme, en décrétant que dans ce système de coordonnées les symboles de connexion affine sont nuls en P . Autrement dit, dans ce système de coordonnées un tenseur est transporté parallèlement au voisinage de P si et seulement si les dérivées premières de ses composantes sont nulles en P .*

Preuve. Remarquons d'abord que le tenseur de torsion (5.47) est nul et la connexion, symétrique. D'autre part, en vertu de (5.52) on a

$$\nabla_{\mu} \mathbf{g} = 0. \quad (5.73)$$

Dans des coordonnées quelconques, cette relation tensorielle prend la forme

$$g_{\alpha\beta,\mu} = \Gamma_{\alpha\mu}^{\xi} g_{\xi\beta} + \Gamma_{\beta\mu}^{\xi} g_{\alpha\xi}. \quad (5.74)$$

En permutant les indices, on obtient

$$\begin{aligned} g_{\mu\alpha,\beta} &= \Gamma_{\mu\beta}^{\xi} g_{\xi\alpha} + \Gamma_{\alpha\beta}^{\xi} g_{\mu\xi} \\ &= \Gamma_{\beta\mu}^{\xi} g_{\alpha\xi} + \Gamma_{\alpha\beta}^{\xi} g_{\mu\xi} \end{aligned} \quad (5.75)$$

$$\begin{aligned} g_{\beta\mu,\alpha} &= \Gamma_{\beta\alpha}^{\xi} g_{\xi\mu} + \Gamma_{\mu\alpha}^{\xi} g_{\beta\xi} \\ &= \Gamma_{\alpha\beta}^{\xi} g_{\mu\xi} + \Gamma_{\alpha\mu}^{\xi} g_{\xi\beta}. \end{aligned} \quad (5.76)$$

La combinaison linéaire (5.74)+(5.75)-(5.76) permet d'exprimer la connexion en fonction de la métrique :

$$\Gamma_{\beta\mu}^{\delta} = \frac{1}{2} g^{\delta\alpha} (g_{\alpha\beta,\mu} + g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha}). \quad (5.77)$$

Elle est dite *riemannienne*, ou encore : *de Levi-Civita*. C'est celle-là que nous utiliserons désormais en présence d'une métrique. Les symboles (5.77) ont reçu le nom de E.B. Christoffel (1869). \square

Dans le cadre des connexions riemanniennes on peut expliciter les coefficients $B_{\alpha\beta}^{\lambda}$ de la section 5.9. En effet, dans le système des coordonnées barrées l'annulation des dérivées premières du tenseur métrique équivaut alors à celle des symboles de Christoffel. Or, d'après la loi de transformation (5.46) de ceux-ci, on a

$$\underline{\Gamma}_{\alpha\beta}^{\rho} = \left(\frac{\partial^2 x^{\lambda}}{\partial \underline{x}^{\alpha} \partial \underline{x}^{\beta}} + \frac{\partial x^{\epsilon}}{\partial \underline{x}^{\alpha}} \frac{\partial x^{\eta}}{\partial \underline{x}^{\beta}} \Gamma_{\epsilon\eta}^{\lambda}[\vec{x}(\vec{\underline{x}})] \right) \frac{\partial \underline{x}^{\rho}}{\partial x^{\lambda}}[\vec{x}(\vec{\underline{x}})].$$

Le membre de gauche étant nul au point P , on en déduit successivement

$$\begin{aligned} \frac{\partial^2 x^{\lambda}}{\partial \underline{x}^{\alpha} \partial \underline{x}^{\beta}}(P) &= -\frac{\partial x^{\epsilon}}{\partial \underline{x}^{\alpha}}(P) \frac{\partial x^{\eta}}{\partial \underline{x}^{\beta}}(P) \Gamma_{\epsilon\eta}^{\lambda}(P) \\ B_{\alpha\beta}^{\lambda} &= -A_{\alpha}^{\epsilon} A_{\beta}^{\eta} \Gamma_{\epsilon\eta}^{\lambda}(P). \end{aligned} \quad (5.78)$$

Il est facile de vérifier que les nombres (5.78) annulent effectivement le membre de gauche de l'équation (5.71).

Exercice 5.6 Calculer les coefficients A_α^λ et $B_{\alpha\beta}^\lambda$ pour la métrique de Lobatchevski dans les coordonnées du demi-plan de Poincaré (4.31) :

$$ds^2 = \frac{1}{y^2} (dx^2 + dy^2), \quad y > 0.$$

Solution. Cette métrique est définie positive. Les équations (5.68) imposent

$$\begin{aligned} A_1^1 A_1^1 + A_1^2 A_1^2 &= (y_P)^2 \\ A_2^1 A_2^1 + A_2^2 A_2^2 &= (y_P)^2 \\ A_1^1 A_2^1 + A_1^2 A_2^2 &= 0. \end{aligned}$$

Les solutions dépendent d'un paramètre continu $\theta \in [0, 2\pi[$ et d'un paramètre discret égal à ± 1 :

$$A_1^1 = y_P \cos \theta, \quad A_1^2 = y_P \sin \theta, \quad A_2^1 = \mp y_P \sin \theta, \quad A_2^2 = \pm y_P \cos \theta.$$

Le seuls symboles de Christoffel non nuls sont

$$\Gamma_{12}^1 = \Gamma_{21}^1 = -\Gamma_{11}^2 = \Gamma_{22}^2 = -\frac{1}{y_P}.$$

Donc, compte tenu de (5.78) et de la valeur obtenue pour les A_α^λ :

$$B_{11}^1 = \pm B_{12}^2 = -B_{22}^1 = y_P \sin(2\theta), \quad B_{22}^2 = \pm B_{12}^1 = -B_{11}^2 = y_P \cos(2\theta).$$

5.11 Métrique et connexion sur une surface immergée dans \mathbb{R}^3

Nous allons compléter les propos de la section précédente en examinant le cas particulier d'une surface \mathcal{S} immergée dans l'espace euclidien à trois dimensions doté de la dérivée covariante habituelle définie en coordonnées cartésiennes par (5.28) [et qui, pour des raisons de clarté, sera exprimée ici à l'aide du symbole ∂ , comme dans le chapitre 4].

Plaçons-nous en un point P de cette surface, et supposons-y définie la normale unitaire \vec{N} . Tout vecteur \vec{X} de \mathbb{R}^3 peut être décomposé en une composante orthogonale à la surface et une composante $\vec{X}_\mathcal{S}$ tangente à la surface :

$$\vec{X} = (\vec{X} \cdot \vec{N}) \vec{N} + \vec{X}_\mathcal{S}$$

où le point exprime le produit scalaire dans \mathbb{R}^3 .

On définit très naturellement le *tenseur métrique* $g_\mathcal{S}$ sur la surface immergée (agissant sur des couples de vecteurs définis dans le plan tangent à la surface au point considéré) à partir du tenseur métrique dans \mathbb{R}^3 :

$$g_\mathcal{S}(\vec{X}_\mathcal{S}, \vec{Y}_\mathcal{S}) = g(\vec{X}_\mathcal{S}, \vec{Y}_\mathcal{S}). \quad (5.79)$$

D'autre part, sur \mathcal{S} , on définit la *dérivée covariante du champ vectoriel* $\vec{X}_\mathcal{S}$ (défini dans le plan tangent à la surface) par rapport à la coordonnée x^μ (elle-même définie sur \mathcal{S}), par

$$\begin{aligned} \nabla_\mu \vec{X}_\mathcal{S} &= (\partial_\mu \vec{X}_\mathcal{S})_\mathcal{S} \\ &= \partial_\mu \vec{X}_\mathcal{S} - (\partial_\mu \vec{X}_\mathcal{S} \cdot \vec{N}) \vec{N}. \end{aligned} \quad (5.80)$$

Théorème 5.2 *Des définitions (5.80) et (5.79) il résulte que la connexion sur la surface est riemannienne.*

Démonstration. Vu (5.53), on a

$$\partial_\mu \left[\mathbf{g}(\vec{X}_S, \vec{Y}_S) \right] = (\partial_\mu \mathbf{g})(\vec{X}_S, \vec{Y}_S) + \mathbf{g}(\partial_\mu \vec{X}_S, \vec{Y}_S) + \mathbf{g}(\vec{X}_S, \partial_\mu \vec{Y}_S).$$

Le premier terme du membre de droite s'annule car la connexion définie dans \mathbb{R}^3 est riemannienne. D'autre part, de (5.80) il résulte

$$\partial_\mu \vec{X}_S = \nabla_\mu \vec{X}_S + (\partial_\mu \vec{X}_S \cdot \vec{N}) \vec{N}$$

(considérations analogues pour $\partial_\mu \vec{Y}_S$), donc

$$\begin{aligned} \mathbf{g}(\partial_\mu \vec{X}_S, \vec{Y}_S) &= \mathbf{g}(\nabla_\mu \vec{X}_S, \vec{Y}_S) + 0 \\ \mathbf{g}(\vec{X}_S, \partial_\mu \vec{Y}_S) &= \mathbf{g}(\vec{X}_S, \nabla_\mu \vec{Y}_S) + 0. \end{aligned}$$

Ainsi,

$$\partial_\mu \left[\mathbf{g}(\vec{X}_S, \vec{Y}_S) \right] = \mathbf{g}(\nabla_\mu \vec{X}_S, \vec{Y}_S) + \mathbf{g}(\vec{X}_S, \nabla_\mu \vec{Y}_S). \quad (5.81)$$

Compte tenu de (5.79) on peut, dans cette équation, substituer \mathbf{g}_S à \mathbf{g} . Enfin, la quantité $\mathbf{g}_S(\vec{X}_S, \vec{Y}_S)$ étant un scalaire, on peut remplacer ∂_μ par ∇_μ dans le membre de gauche.

On obtient donc finalement

$$\nabla_\mu \left[\mathbf{g}_S(\vec{X}_S, \vec{Y}_S) \right] = \mathbf{g}_S(\nabla_\mu \vec{X}_S, \vec{Y}_S) + \mathbf{g}_S(\vec{X}_S, \nabla_\mu \vec{Y}_S), \quad (5.82)$$

relation que l'on peut noter sous la forme compacte

$$\nabla_\mu \mathbf{g}_S = 0,$$

exprimant que la connexion sur la surface, définie par (5.80), est riemannienne. \square

On en déduit le lien entre le tenseur métrique et les symboles de Christoffel. En effet, en composantes l'équation (5.82) s'exprime comme suit (tous les termes se rapportant à la surface, nous omettrons l'indice S) :

$$\left(g_{\alpha\beta} X^\alpha Y^\beta \right)_{,\mu} = g_{\alpha\beta} \left(X^\alpha_{,\mu} + \Gamma_{\xi\mu}^\alpha X^\xi \right) Y^\beta + g_{\alpha\beta} X^\alpha \left(Y^\beta_{,\mu} + \Gamma_{\xi\mu}^\beta Y^\xi \right)$$

c'est-à-dire

$$g_{\alpha\beta,\mu} = \Gamma_{\alpha\mu}^\xi g_{\xi\beta} + \Gamma_{\beta\mu}^\xi g_{\alpha\xi}$$

impliquant, comme nous l'avons vu, la relation (5.77).

Exercice 5.7 *Calculer la métrique sur un tore à deux dimensions (figure 4.5) à partir de celle de \mathbb{R}^3 . En déduire les symboles de Christoffel relatifs aux coordonnées choisies.*

Solution.

$$x = (a + b \cos \psi) \cos \varphi, \quad y = (a + b \cos \psi) \sin \varphi, \quad z = b \sin \psi$$

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + dz^2 \\ &= (a + b \cos \psi)^2 d\varphi^2 + b^2 d\psi^2. \end{aligned} \quad (5.83)$$

On vérifie facilement que les symboles de Christoffel, obtenus intrinsèquement à l'aide de la relation (5.77), sont conformes au résultat (4.25).

5.12 Géodésiques dans un espace riemannien

Une courbe est dite *temporelle*, *nulle* ou *spatiale* en un point selon que la norme au carré de son vecteur tangent y est négative, nulle ou positive. Or, on a le

Théorème 5.3 *Si la connexion est riemannienne, alors le transport parallèle conserve le produit scalaire; en d'autres termes, si la connexion est riemannienne, alors le produit scalaire $\mathbf{g}(\vec{V}, \vec{W})$ de deux vecteurs \vec{V} et \vec{W} , transportés parallèlement dans une direction \vec{U} , est constant dans cette direction.*

Preuve. La relation générale (5.53) permet d'écrire

$$\begin{aligned}\nabla_{\vec{U}} [\mathbf{g}(\vec{V}, \vec{W})] &= (\nabla_{\vec{U}} \mathbf{g})(\vec{V}, \vec{W}) + \mathbf{g}(\nabla_{\vec{U}} \vec{V}, \vec{W}) + \mathbf{g}(\vec{V}, \nabla_{\vec{U}} \vec{W}) \\ &= 0 + 0 + 0.\end{aligned}$$

□

En découlent les corollaires suivants.

- Si une géodésique est temporelle, nulle ou spatiale en un de ses points, elle l'est en tous ses points. Pour s'en convaincre, il suffit, en notant \vec{U} son vecteur tangent dans un paramétrage affine, de poser $\vec{V} = \vec{W} = \vec{U}$ dans le théorème ci-dessus.
- Soit une géodésique de paramètre affine λ . La longueur d'arc, mesurée à partir de λ_0 respectivement par la relation

$$\sigma(\lambda) = \int_{\lambda_0}^{\lambda} \sqrt{-g_{\mu\nu}[\vec{x}(\lambda')] \frac{dx^\mu}{d\lambda'} \frac{dx^\nu}{d\lambda'}} d\lambda'$$

pour une géodésique temporelle (remarquons que $\Delta\sigma = c\Delta\tau$, où τ est le temps propre), et

$$s(\lambda) = \int_{\lambda_0}^{\lambda} \sqrt{g_{\mu\nu}[\vec{x}(\lambda')] \frac{dx^\mu}{d\lambda'} \frac{dx^\nu}{d\lambda'}} d\lambda'$$

pour une géodésique spatiale, constitue également un paramètre affine. En effet, l'intégrand est constant et donc, la relation entre σ et λ est linéaire, de même qu'entre s et λ .

Nous allons montrer que les géodésiques non nulles sont des courbes de longueur extrémale entre deux points donnés. Pour cela, on utilise le formalisme des principes variationnels. En mécanique classique, l'intégrale d'action est extrémale sur la trajectoire d'une particule entre deux événements fixés de coordonnées généralisées $q^i(t_1)$ et $q^i(t_2)$:

$$\delta \int_{t_1}^{t_2} L(q^i, \dot{q}^i, t) dt = 0, \quad (5.84)$$

où L est le lagrangien. En effet, la relation (5.84) est équivalente aux équations de Lagrange

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0.$$

Pour appliquer ce principe sur la variété riemannienne, il faut prendre une précaution supplémentaire car, dans le voisinage (topologique) de chaque courbe temporelle, on peut trouver des courbes nulles ou spatiales. On formule donc le principe de moindre action en décrétant que, entre deux points donnés P_1 et P_2 , les géodésiques temporelles (resp.

spatiales) ont une longueur extrémale *parmi celles de toutes les courbes temporelles (resp. spatiales) situées dans leur voisinage* :

$$\delta \int_{P_1}^{P_2} d\sigma = 0 \quad \text{ou} \quad \delta \int_{P_1}^{P_2} ds = 0.$$

Dès lors, le principe de moindre action s'écrit (le point désigne la dérivée par rapport à λ)

$$\delta \int_{\lambda_1}^{\lambda_2} L(x^\alpha, \dot{x}^\alpha, \lambda) d\lambda = 0, \quad (5.85)$$

où $x^\alpha(\lambda_1)$ et $x^\alpha(\lambda_2)$ sont fixés et où le lagrangien est donné par

$$L(x^\alpha, \dot{x}^\alpha, \lambda) = \sqrt{|g_{\mu\nu}[\vec{x}(\lambda)] \dot{x}^\mu \dot{x}^\nu|} \quad (5.86)$$

(en fait il ne dépend pas explicitement de la variable indépendante λ ; d'autre part, il est défini à une constante multiplicative près, sans importance ici). Il génère les équations d'Euler-Lagrange

$$\frac{d}{d\lambda} \frac{\partial L}{\partial \dot{x}^\alpha} - \frac{\partial L}{\partial x^\alpha} = 0. \quad (5.87)$$

Celles-ci sont équivalentes aux équations des géodésiques (5.55). Montrons-le pour les géodésiques temporelles :

$$L = \sqrt{-g_{\mu\nu}[\vec{x}(\lambda)] \dot{x}^\mu \dot{x}^\nu} = \dot{\sigma},$$

donc

$$\begin{aligned} \frac{\partial L}{\partial \dot{x}^\alpha} &= \frac{-g_{\mu\alpha} \dot{x}^\mu}{\dot{\sigma}} \\ \frac{d}{d\lambda} \frac{\partial L}{\partial \dot{x}^\alpha} &= \frac{-1}{\dot{\sigma}} (g_{\mu\alpha, \xi} \dot{x}^\xi \dot{x}^\mu + g_{\mu\alpha} \ddot{x}^\mu) + \frac{\ddot{\sigma}}{\dot{\sigma}^2} g_{\mu\alpha} \dot{x}^\mu \\ \frac{\partial L}{\partial x^\alpha} &= \frac{-g_{\mu\nu, \alpha} \dot{x}^\mu \dot{x}^\nu}{2\dot{\sigma}} \end{aligned}$$

et les équations (5.87) prennent la forme

$$\begin{aligned} \frac{\ddot{\sigma}}{\dot{\sigma}} g_{\mu\alpha} \dot{x}^\mu &= g_{\mu\alpha} \ddot{x}^\mu + g_{\mu\alpha, \xi} \dot{x}^\xi \dot{x}^\mu - \frac{1}{2} g_{\mu\nu, \alpha} \dot{x}^\mu \dot{x}^\nu \\ &= g_{\mu\alpha} \ddot{x}^\mu + \frac{1}{2} (g_{\mu\alpha, \xi} + g_{\xi\alpha, \mu} - g_{\mu\xi, \alpha}) \dot{x}^\mu \dot{x}^\xi, \end{aligned}$$

c'est-à-dire, compte tenu de (5.77),

$$\frac{\ddot{\sigma}}{\dot{\sigma}} \dot{x}^\rho = \ddot{x}^\rho + \Gamma_{\mu\xi}^\rho \dot{x}^\mu \dot{x}^\xi. \quad (5.88)$$

Si λ est un paramètre affine, alors il est relié linéairement au paramètre affine σ , donc $\ddot{\sigma} = 0$ et le membre de gauche est nul, ce qui achève la démonstration.

5.13 L'ascenseur d'Einstein

Théorème 5.4 *Sur une variété différentielle affine sans torsion, il est toujours possible de choisir un système de coordonnées dans lequel tous les symboles de connexion affine s'annulent en tout point P_0 d'une géodésique Γ donnée.*

Commentaire. Le long d'une géodésique, la variété peut être assimilée à une portion d'espace euclidien (jusqu'au premier ordre inclus). Cette propriété se prête à une interprétation géométrique simple dans le cas des surfaces. Par exemple, le long de l'équateur d'une sphère on peut dérouler un ruban adhésif suffisamment étroit sans devoir le froisser. Cela n'est pas possible pour un parallèle quelconque.

Démonstration. Une première coordonnée x^0 est naturellement identifiée à un paramètre affine graduant la géodésique donnée.

Construisons alors, tout le long de Γ , une base de n vecteurs linéairement indépendants, constituée du vecteur $\vec{e}_0 = \partial/\partial x^0$ tangent à la géodésique et de $n - 1$ autres vecteurs \vec{e}_i , transportés eux aussi parallèlement dans la direction de \vec{e}_0 . En tout point P_0 de Γ , on a donc

$$\begin{aligned} 0 &= \nabla_{\vec{e}_0} \vec{e}_\alpha \\ &= \nabla_0 \vec{e}_\alpha \\ &= \Gamma_{0\alpha}^\beta(P_0) \vec{e}_\beta. \end{aligned}$$

Cela implique

$$\Gamma_{0\alpha}^\beta(P_0) = 0 \quad \forall \alpha, \beta. \quad (5.89)$$

Soit, émanant du point P_0 d'abscisse x^0 sur Γ , les géodésiques de paramètre affine λ ($\lambda = 0$ sur Γ) et de vecteur tangent

$$\begin{aligned} \vec{v} &= \frac{d}{d\lambda} \\ &= \sum_{i=1}^{n-1} v^i \vec{e}_i. \end{aligned}$$

Tout point P donné suffisamment proche de Γ appartient à une et une seule de ces géodésiques pour un certain x^0 : on définit ses coordonnées par les n nombres

$$(x^0, x^1 = \lambda v^1, x^2 = \lambda v^2, \dots, x^{n-1} = \lambda v^{n-1}).$$

Dans l'équation de cette géodésique

$$\frac{d^2 x^\alpha}{d\lambda^2} + \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{d\lambda} \frac{dx^\gamma}{d\lambda} = 0$$

le premier terme est nul, soit parce que la coordonnée est constante ($\alpha = 0$), soit parce qu'elle est linéaire en λ ($\alpha = 1, \dots, n - 1$). Ainsi,

$$\begin{aligned} 0 &= \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{d\lambda} \frac{dx^\gamma}{d\lambda} \\ &= \Gamma_{ij}^\alpha \frac{dx^i}{d\lambda} \frac{dx^j}{d\lambda} \\ &= \Gamma_{ij}^\alpha v^i v^j. \end{aligned}$$

Toutes ces géodésiques se croisent en P_0 . Donc $\Gamma_{ij}^\alpha(P_0) v^i v^j = 0$. Comme le vecteur \vec{v} est quelconque, et la connexion symétrique, on en déduit

$$\Gamma_{ij}^\alpha(P_0) = 0 \quad \forall \alpha, i, j. \quad (5.90)$$

Les équations (5.89) et (5.90) montrent que, dans ce système de coordonnées,

$$\Gamma_{\beta\gamma}^{\alpha} = 0 \quad \forall \alpha, \beta, \gamma \quad (5.91)$$

le long de la géodésique Γ .

□

La signification physique de ce théorème est particulièrement claire si la variété différentielle est riemannienne. Adoptons une connexion riemannienne et une métrique de signature $(-, +, +, +)$, comme le prescrit la relativité générale, et supposons que la géodésique Γ soit temporelle. Choisissons une base orthonormée en un de ses points. D'après les propriétés du transport parallèle, elle reste orthonormée en tout point de Γ . On réalise ainsi les *coordonnées normales de Fermi*⁷ dans lesquelles on a, tout le long de Γ :

$$\begin{aligned} (g_{\mu\nu}) &= \text{diag}(-1, 1, 1, 1) \\ g_{\mu\nu,\xi} &= 0. \end{aligned}$$

La géodésique temporelle peut être interprétée comme la ligne d'univers d'une particule matérielle en chute libre, et dans ces coordonnées le tenseur métrique simule celui de la relativité restreinte jusqu'à l'ordre 1 de dérivation inclus. C'est la raison pour laquelle on les désigne parfois sous l'appellation pittoresque d'*ascenseur d'Einstein*.

5.14 Commutateur de deux champs de vecteurs

5.14.1 Définition algébrique

Le *commutateur* de deux champs vectoriels \vec{X} et \vec{Y} est le champ vectoriel

$$\begin{aligned} [\vec{X}, \vec{Y}] &= [X^{\alpha} \partial_{\alpha}, Y^{\beta} \partial_{\beta}] \\ &= X^{\alpha} (\partial_{\alpha} Y^{\beta}) \partial_{\beta} + X^{\alpha} Y^{\beta} \partial_{\alpha} \partial_{\beta} - Y^{\beta} (\partial_{\beta} X^{\alpha}) \partial_{\alpha} - Y^{\beta} X^{\alpha} \partial_{\beta} \partial_{\alpha} \\ &= (X^{\alpha} \partial_{\alpha} Y^{\mu} - Y^{\alpha} \partial_{\alpha} X^{\mu}) \partial_{\mu}. \end{aligned} \quad (5.92)$$

Ce concept relève de la topologie différentielle. A titre d'exercice, on vérifiera que les n quantités $X^{\alpha} \partial_{\alpha} Y^{\mu} - Y^{\alpha} \partial_{\alpha} X^{\mu}$ se transforment bien comme les composantes d'un vecteur.

Si les champs \vec{X} et \vec{Y} sont tangents à des lignes de coordonnées, autrement dit s'il existe des coordonnées $\{x^{\mu}\}$ telles que $\vec{X} = \partial/\partial x^{\alpha}$ et $\vec{Y} = \partial/\partial x^{\beta}$, alors leur commutateur est nul :

$$[\partial/\partial x^{\alpha}, \partial/\partial x^{\beta}] = 0. \quad (5.93)$$

Pour une connexion affine symétrique (sans torsion), on voit facilement que

$$\begin{aligned} [\vec{X}, \vec{Y}]^{\mu} &= X^{\alpha} (\partial_{\alpha} Y^{\mu} + \Gamma_{\xi\alpha}^{\mu} Y^{\xi}) - Y^{\alpha} (\partial_{\alpha} X^{\mu} + \Gamma_{\xi\alpha}^{\mu} X^{\xi}) \\ &= X^{\alpha} (\nabla_{\alpha} \vec{Y})^{\mu} - Y^{\alpha} (\nabla_{\alpha} \vec{X})^{\mu}, \end{aligned}$$

c'est-à-dire

$$[\vec{X}, \vec{Y}] = \nabla_{\vec{X}} \vec{Y} - \nabla_{\vec{Y}} \vec{X}. \quad (5.94)$$

7. C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation*, Freeman, 1973, pp. 327-332.

5.14.2 Interprétation géométrique

Notons $\mathcal{X}(x)$ les courbes intégrales (section 5.2) d'un champ vectoriel $\vec{X}(\vec{x}) = d/dx$. De même, désignons par $\mathcal{Y}(y)$ les courbes intégrales d'un autre champ vectoriel $\vec{Y}(\vec{x}) = d/dy$.

Soit un point P de coordonnées $x^\mu(P)$ où passe une courbe $\mathcal{X}(x)$ de vecteur tangent $\vec{X}(P)$ et une courbe $\mathcal{Y}(y)$ de vecteur tangent $\vec{Y}(P)$, avec $P = \mathcal{X}(0) = \mathcal{Y}(0)$. Posons $P_1 = \mathcal{X}(\epsilon_1)$. Par P_1 passe la courbe $\mathcal{Y}'(y)$ de vecteur tangent $\vec{Y}(P_1)$. Supposant que $P_1 = \mathcal{Y}'(0)$, définissons $P_3 = \mathcal{Y}'(\epsilon_2)$. Posons ensuite $P_2 = \mathcal{Y}(\epsilon_2)$. Par P_2 passe la courbe $\mathcal{X}'(x)$ de vecteur tangent $\vec{X}(P_2)$. Supposant que $P_2 = \mathcal{X}'(0)$, définissons enfin $P_4 = \mathcal{X}'(\epsilon_1)$ (figure 5.3). Remarquons d'ores et déjà qu'en dimension $n > 2$, les courbes \mathcal{X}' et \mathcal{Y}' ne se croisent pas nécessairement.

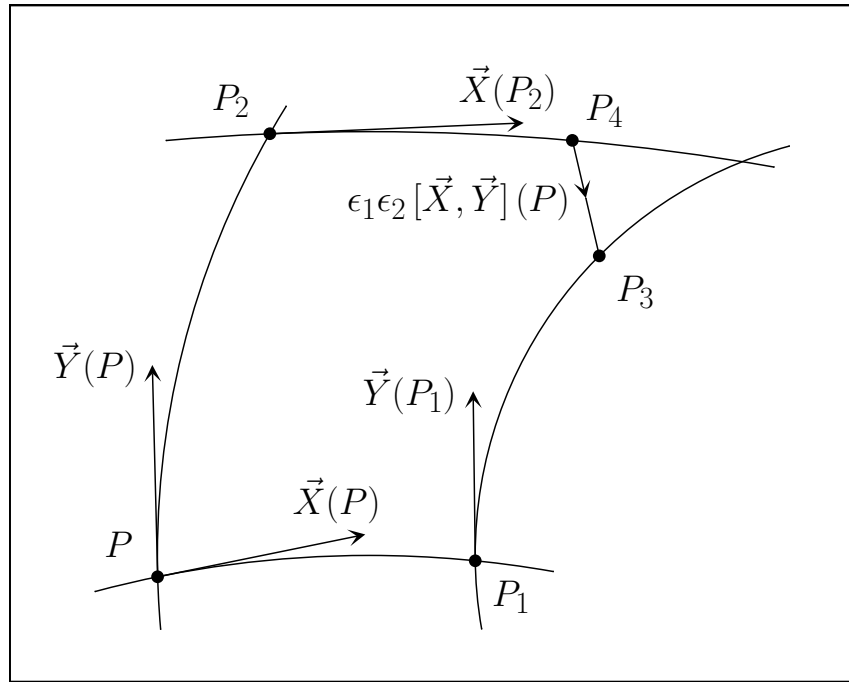


FIGURE 5.3: *Interprétation géométrique du commutateur dans l'espace \mathbf{R}^n localement homéomorphe à la variété (les vecteurs tangents ne sont pas à l'échelle). Quatre courbes intégrales ont été représentées : en bas, $\mathcal{X}(x)$; en haut, $\mathcal{X}'(x)$; à gauche, $\mathcal{Y}(y)$; à droite, $\mathcal{Y}'(y)$. On a $P = \mathcal{X}(0) = \mathcal{Y}(0)$, $P_1 = \mathcal{X}(\epsilon_1) = \mathcal{Y}'(0)$, $P_3 = \mathcal{Y}'(\epsilon_2)$, $P_2 = \mathcal{Y}(\epsilon_2) = \mathcal{X}'(0)$, $P_4 = \mathcal{X}'(\epsilon_1)$. Le vecteur intuitif reliant P_4 à P_3 est alors donné par $\epsilon_1 \epsilon_2 [\vec{X}, \vec{Y}](P) + O(3)$.*

En utilisant le résultat général (5.2), on calcule

$$x^\mu(P_1) = x^\mu(P) + \epsilon_1 X^\mu(P) + \frac{1}{2}(\epsilon_1)^2 \left(\frac{dX^\mu}{dx} \right) (P) + O(3),$$

$$x^\mu(P_2) = x^\mu(P) + \epsilon_2 Y^\mu(P) + \frac{1}{2}(\epsilon_2)^2 \left(\frac{dY^\mu}{dy} \right) (P) + O(3),$$

$$x^\mu(P_3) = x^\mu(P_1) + \epsilon_2 Y^\mu(P_1) + \frac{1}{2}(\epsilon_2)^2 \left(\frac{dY^\mu}{dy} \right) (P_1) + O(3),$$

$$x^\mu(P_4) = x^\mu(P_2) + \epsilon_1 X^\mu(P_2) + \frac{1}{2}(\epsilon_1)^2 \left(\frac{dX^\mu}{dx} \right) (P_2) + O(3).$$

Or,

$$Y^\mu(P_1) = Y^\mu(P) + \epsilon_1 \left(\frac{dY^\mu}{dx} \right) (P) + O(2)$$

$$X^\mu(P_2) = X^\mu(P) + \epsilon_2 \left(\frac{dX^\mu}{dy} \right) (P) + O(2).$$

D'autre part,

$$\left(\frac{dY^\mu}{dy} \right) (P_1) = \left(\frac{dY^\mu}{dy} \right) (P) + O(1)$$

$$\left(\frac{dX^\mu}{dx} \right) (P_2) = \left(\frac{dX^\mu}{dx} \right) (P) + O(1).$$

De tout cela on déduit

$$\begin{aligned} x^\mu(P_3) - x^\mu(P_4) &= \epsilon_1 \epsilon_2 \left(\frac{dY^\mu}{dx} \right) (P) - \epsilon_1 \epsilon_2 \left(\frac{dX^\mu}{dy} \right) (P) + O(3) \\ &= \epsilon_1 \epsilon_2 (X^\alpha \partial_\alpha Y^\mu - Y^\alpha \partial_\alpha X^\mu) (P) + O(3) \\ &= \epsilon_1 \epsilon_2 [\vec{X}, \vec{Y}]^\mu (P) + O(3), \end{aligned}$$

c'est-à-dire

$$\overrightarrow{P_4 P_3} = \epsilon_1 \epsilon_2 [\vec{X}, \vec{Y}] (P) + O(3). \quad (5.95)$$

5.15 L'opérateur de courbure

L'opérateur de courbure $\mathcal{R}(\vec{X}, \vec{Y})$ est défini par

$$\mathcal{R}(\vec{X}, \vec{Y}) = [\nabla_{\vec{X}}, \nabla_{\vec{Y}}] - \nabla_{[\vec{X}, \vec{Y}]}. \quad (5.96)$$

Il s'applique à des tenseurs d'ordre quelconque.

Théorème 5.5 *L'opérateur de courbure $\mathcal{R}(\vec{X}, \vec{Y})$ est linéaire par rapport à chacun de ses arguments \vec{X} et \vec{Y} .*

Notons d'abord que

$$\begin{aligned} [\nabla_{\vec{X}}, \nabla_{\vec{Y}}] &= X^\alpha \nabla_\alpha (Y^\beta \nabla_\beta) - Y^\alpha \nabla_\alpha (X^\beta \nabla_\beta) \\ &= X^\alpha \partial_\alpha Y^\beta \nabla_\beta + X^\alpha Y^\beta \nabla_\alpha \nabla_\beta - Y^\alpha \partial_\alpha X^\beta \nabla_\beta - Y^\alpha X^\beta \nabla_\alpha \nabla_\beta \\ \nabla_{[\vec{X}, \vec{Y}]} &= (X^\alpha \partial_\alpha Y^\beta - Y^\alpha \partial_\alpha X^\beta) \nabla_\beta \\ [\nabla_{\vec{X}}, \nabla_{\vec{Y}}] - \nabla_{[\vec{X}, \vec{Y}]} &= X^\alpha Y^\beta \nabla_\alpha \nabla_\beta - Y^\alpha X^\beta \nabla_\alpha \nabla_\beta. \end{aligned}$$

Par conséquent,

$$\mathcal{R}(\vec{X}, \vec{Y}) = X^\alpha Y^\beta (\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha). \quad (5.97)$$

La linéarité par rapport à \vec{X} et \vec{Y} apparaît clairement.

□

Lemme 5.1 *Si f est un champ scalaire, alors $\mathcal{R}(\vec{X}, \vec{Y})f = 0$.*

En effet, la dérivée covariante d'un champ scalaire n'est autre que sa dérivée partielle. Donc, d'après (5.97),

$$\begin{aligned}\mathcal{R}(\vec{X}, \vec{Y})f &= X^\alpha Y^\beta (\partial_\alpha \partial_\beta - \partial_\beta \partial_\alpha) f \\ &= 0.\end{aligned}$$

□

Théorème 5.6 *L'opérateur de courbure $\mathcal{R}(\vec{X}, \vec{Y})$ est linéaire par rapport à l'argument \vec{V} auquel il s'applique.*

En effet, d'une part

$$\begin{aligned}([\nabla_{\vec{X}}, \nabla_{\vec{Y}}] - \nabla_{[\vec{X}, \vec{Y}]})(f\vec{V}) &= f([\nabla_{\vec{X}}, \nabla_{\vec{Y}}] - \nabla_{[\vec{X}, \vec{Y}]})\vec{V} + ([\nabla_{\vec{X}}, \nabla_{\vec{Y}}] - \nabla_{[\vec{X}, \vec{Y}]})f \vec{V} \\ &= f([\nabla_{\vec{X}}, \nabla_{\vec{Y}}] - \nabla_{[\vec{X}, \vec{Y}]})\vec{V} + 0\end{aligned}$$

en vertu du lemme 5.1 ; et, d'autre part,

$$([\nabla_{\vec{X}}, \nabla_{\vec{Y}}] - \nabla_{[\vec{X}, \vec{Y}]})(\vec{V} + \vec{W}) = ([\nabla_{\vec{X}}, \nabla_{\vec{Y}}] - \nabla_{[\vec{X}, \vec{Y}]})\vec{V} + ([\nabla_{\vec{X}}, \nabla_{\vec{Y}}] - \nabla_{[\vec{X}, \vec{Y}]})\vec{W}.$$

□

La propriété suivante établit le lien entre l'action de l'opérateur de courbure sur des vecteurs et son action sur des formes :

Théorème 5.7

$$\langle \mathcal{R}(\vec{X}, \vec{Y})\tilde{\Omega}, \vec{V} \rangle = - \langle \tilde{\Omega}, \mathcal{R}(\vec{X}, \vec{Y})\vec{V} \rangle .$$

En effet, de la règle de Leibniz

$$\nabla_{\vec{Y}} \langle \tilde{\Omega}, \vec{V} \rangle = \langle \nabla_{\vec{Y}} \tilde{\Omega}, \vec{V} \rangle + \langle \tilde{\Omega}, \nabla_{\vec{Y}} \vec{V} \rangle$$

il résulte successivement

$$\begin{aligned}\nabla_{[\vec{X}, \vec{Y}]} \langle \tilde{\Omega}, \vec{V} \rangle &= \langle \nabla_{[\vec{X}, \vec{Y}]} \tilde{\Omega}, \vec{V} \rangle + \langle \tilde{\Omega}, \nabla_{[\vec{X}, \vec{Y}]} \vec{V} \rangle \\ \mathcal{R}(\vec{X}, \vec{Y}) \langle \tilde{\Omega}, \vec{V} \rangle &= \langle \mathcal{R}(\vec{X}, \vec{Y})\tilde{\Omega}, \vec{V} \rangle + \langle \tilde{\Omega}, \mathcal{R}(\vec{X}, \vec{Y})\vec{V} \rangle .\end{aligned}$$

Or, le membre de gauche est nul en vertu du lemme 5.1, car on applique l'opérateur de courbure au scalaire $\langle \tilde{\Omega}, \vec{V} \rangle$. Le théorème est démontré.

□

5.16 Approche intuitive du tenseur de Riemann-Christoffel

La relation (5.72) exprime les dérivées secondes du tenseur métrique en fonction des nombres $A_\alpha^\lambda, B_{\alpha\beta}^\lambda, C_{\alpha\beta\gamma}^\lambda$ associés à un changement de variables $\vec{x} \rightarrow \vec{x}$ et définis en (5.66). Tâchons de voir ce qu'il est possible d'affirmer au sujet de ces dérivées secondes, en supposant que les A_α^λ et les $B_{\alpha\beta}^\lambda$ ont déjà été fixés pour réduire le tenseur métrique à la forme canonique et annuler ses dérivées premières.

Apparemment, les relations (5.72) constituent un système de $n^2(n+1)^2/4$ équations linéaires pour les $n^2(n+1)(n+2)/6$ inconnues $C_{\alpha\beta\gamma}^\lambda$. Il y aurait donc

$$\frac{1}{4}n^2(n+1)^2 - \frac{1}{6}n^2(n+1)(n+2) = \frac{1}{12}n^2(n^2-1)$$

équations de trop et le système semble insoluble quels que soient les $\underline{g}_{\alpha\beta,\gamma\delta}(P)$, ce qui est paradoxal puisque les coefficients $C_{\alpha\beta\gamma}^\lambda$ devraient résulter sans ambiguïté du changement de variables générant les $\underline{g}_{\alpha\beta,\gamma\delta}(P)$: le système *doit* être compatible.

En fait, *il existe* $n^2(n^2-1)/12$ relations linéaires entre les lignes du système (5.72), ce qui porte son rang à la valeur $n^2(n+1)(n+2)/6$, égal au nombre d'inconnues $C_{\alpha\beta\gamma}^\lambda$. Pour justifier cela, posons, dans un système de coordonnées quelconques,

$$R_{\lambda\mu\nu\xi} = \frac{1}{2}(g_{\lambda\xi,\mu\nu} + g_{\mu\nu,\lambda\xi} - g_{\mu\xi,\lambda\nu} - g_{\lambda\nu,\mu\xi}) + g_{\epsilon\eta}(\Gamma_{\mu\nu}^\epsilon\Gamma_{\lambda\xi}^\eta - \Gamma_{\mu\xi}^\epsilon\Gamma_{\lambda\nu}^\eta). \quad (5.98)$$

Comme nous le verrons dans la section 5.21.2, ces quantités sont en nombre $n^2(n^2-1)/12$. D'autre part, par définition, dans le système des coordonnées $\{\underline{x}^\rho\}$ le deuxième terme de (5.98) s'annule au point P . Ainsi,

$$\underline{R}_{\alpha\beta\gamma\delta}(P) = \frac{1}{2}[\underline{g}_{\alpha\delta,\beta\gamma}(P) + \underline{g}_{\beta\gamma,\alpha\delta}(P) - \underline{g}_{\beta\delta,\alpha\gamma}(P) - \underline{g}_{\alpha\gamma,\beta\delta}(P)].$$

En développant le membre de droite à l'aide des relations (5.72), on constate que les inconnues $C_{\alpha\beta\gamma}^\lambda$ n'apparaissent pas dans ces combinaisons linéaires. Autrement dit, il y a $n^2(n^2-1)/12$ relations entre les lignes du système compatible (5.72). Elles fournissent l'expression des $\underline{R}_{\alpha\beta\gamma\delta}$ en fonction du tenseur métrique et des symboles de Christoffel dans le système $\{x^\rho\}$.

Voici le calcul (dans ce qui suit, les fonctions sont implicitement évaluées en P) :

$$\begin{aligned} \underline{R}_{\alpha\beta\gamma\delta} &= \frac{1}{2}g_{\lambda\mu}(\underbrace{C_{\beta\gamma\alpha}^\lambda A_\delta^\mu + C_{\beta\gamma\delta}^\mu A_\alpha^\lambda + \underline{B_{\beta\alpha}^\lambda B_{\gamma\delta}^\mu} + B_{\gamma\alpha}^\mu B_{\beta\delta}^\lambda}_{\text{}} \\ &\quad + \underline{C_{\alpha\delta\beta}^\lambda A_\gamma^\mu} + \underline{C_{\alpha\delta\gamma}^\mu A_\beta^\lambda} + \underline{B_{\alpha\beta}^\lambda B_{\delta\gamma}^\mu} + B_{\delta\beta}^\mu B_{\alpha\gamma}^\lambda \\ &\quad - \underline{C_{\alpha\gamma\beta}^\lambda A_\delta^\mu} - \underline{C_{\alpha\gamma\delta}^\mu A_\beta^\lambda} - \underline{B_{\alpha\beta}^\lambda B_{\gamma\delta}^\mu} - B_{\gamma\beta}^\mu B_{\alpha\delta}^\lambda \\ &\quad - \underline{C_{\beta\delta\alpha}^\lambda A_\gamma^\mu} - \underline{C_{\beta\delta\gamma}^\mu A_\alpha^\lambda} - \underline{B_{\beta\alpha}^\lambda B_{\delta\gamma}^\mu} - B_{\delta\alpha}^\mu B_{\beta\gamma}^\lambda) \\ &\quad + \frac{1}{2}g_{\lambda\mu,\nu}[(\underline{B_{\beta\alpha}^\lambda A_\delta^\mu} + B_{\beta\delta}^\mu A_\alpha^\lambda)A_\gamma^\nu + (B_{\gamma\alpha}^\lambda A_\delta^\mu + \underline{B_{\gamma\delta}^\mu A_\alpha^\lambda})A_\beta^\nu + A_\alpha^\lambda A_\delta^\mu B_{\beta\gamma}^\nu \\ &\quad + (\underline{B_{\alpha\beta}^\lambda A_\gamma^\mu} + B_{\alpha\gamma}^\mu A_\beta^\lambda)A_\delta^\nu + (B_{\delta\beta}^\lambda A_\gamma^\mu + \underline{B_{\delta\gamma}^\mu A_\beta^\lambda})A_\alpha^\nu + A_\beta^\lambda A_\gamma^\mu B_{\alpha\delta}^\nu \\ &\quad - (B_{\alpha\beta}^\lambda A_\delta^\mu + B_{\alpha\delta}^\mu A_\beta^\lambda)A_\gamma^\nu - (B_{\gamma\beta}^\lambda A_\delta^\mu + \underline{B_{\gamma\delta}^\mu A_\beta^\lambda})A_\alpha^\nu - A_\beta^\lambda A_\delta^\mu B_{\alpha\gamma}^\nu \\ &\quad - (\underline{B_{\beta\alpha}^\lambda A_\gamma^\mu} + B_{\beta\gamma}^\mu A_\alpha^\lambda)A_\delta^\nu - (B_{\delta\alpha}^\lambda A_\gamma^\mu + \underline{B_{\delta\gamma}^\mu A_\alpha^\lambda})A_\beta^\nu - A_\alpha^\lambda A_\gamma^\mu B_{\beta\delta}^\nu] \\ &\quad + \frac{1}{2}g_{\lambda\mu,\nu\xi}(A_\alpha^\lambda A_\delta^\mu A_\beta^\nu A_\gamma^\xi + A_\beta^\lambda A_\gamma^\mu A_\alpha^\nu A_\delta^\xi - A_\beta^\lambda A_\delta^\mu A_\alpha^\nu A_\gamma^\xi - A_\alpha^\lambda A_\gamma^\mu A_\beta^\nu A_\delta^\xi) \end{aligned}$$

$$\begin{aligned}
&= A_\alpha^\lambda A_\beta^\mu A_\gamma^\nu A_\delta^\xi \frac{1}{2} (g_{\lambda\xi, \mu\nu} + g_{\mu\nu, \lambda\xi} - g_{\mu\xi, \lambda\nu} - g_{\lambda\nu, \mu\xi}) + g_{\lambda\mu} (B_{\alpha\gamma}^\lambda B_{\beta\delta}^\mu - B_{\alpha\delta}^\lambda B_{\beta\gamma}^\mu) \\
&\quad + \frac{1}{2} (A_\alpha^\lambda A_\gamma^\nu B_{\beta\delta}^\epsilon g_{\lambda\epsilon, \nu} + A_\delta^\xi A_\beta^\mu B_{\gamma\alpha}^\epsilon g_{\epsilon\xi, \mu} + A_\alpha^\lambda A_\delta^\xi B_{\beta\gamma}^\epsilon g_{\lambda\xi, \epsilon} \\
&\quad + A_\beta^\mu A_\delta^\xi B_{\alpha\gamma}^\epsilon g_{\mu\epsilon, \xi} + A_\gamma^\nu A_\alpha^\lambda B_{\delta\beta}^\epsilon g_{\epsilon\nu, \lambda} + A_\beta^\mu A_\gamma^\nu B_{\alpha\delta}^\epsilon g_{\mu\nu, \epsilon} \\
&\quad - A_\beta^\mu A_\gamma^\nu B_{\alpha\delta}^\epsilon g_{\mu\epsilon, \nu} - A_\delta^\xi A_\alpha^\lambda B_{\gamma\beta}^\epsilon g_{\epsilon\xi, \lambda} - A_\beta^\mu A_\delta^\xi B_{\alpha\gamma}^\epsilon g_{\mu\xi, \epsilon} \\
&\quad - A_\alpha^\lambda A_\delta^\xi B_{\beta\gamma}^\epsilon g_{\lambda\epsilon, \xi} - A_\gamma^\nu A_\beta^\mu B_{\delta\alpha}^\epsilon g_{\epsilon\nu, \mu} - A_\alpha^\lambda A_\gamma^\nu B_{\beta\delta}^\epsilon g_{\lambda\nu, \epsilon}).
\end{aligned}$$

Les quatre dernières lignes peuvent être regroupées sous la forme

$$\begin{aligned}
&\frac{1}{2} [A_\alpha^\lambda A_\gamma^\nu (B_{\beta\delta}^\epsilon g_{\lambda\epsilon, \nu} + B_{\beta\delta}^\epsilon g_{\epsilon\nu, \lambda} - B_{\beta\delta}^\epsilon g_{\lambda\nu, \epsilon}) + A_\beta^\mu A_\delta^\xi (B_{\gamma\alpha}^\epsilon g_{\epsilon\xi, \mu} + B_{\alpha\gamma}^\epsilon g_{\mu\epsilon, \xi} - B_{\alpha\gamma}^\epsilon g_{\mu\xi, \epsilon}) \\
&\quad + A_\alpha^\lambda A_\delta^\xi (B_{\beta\gamma}^\epsilon g_{\lambda\xi, \epsilon} - B_{\gamma\beta}^\epsilon g_{\epsilon\xi, \lambda} - B_{\beta\gamma}^\epsilon g_{\lambda\epsilon, \xi}) + A_\beta^\mu A_\gamma^\nu (B_{\alpha\delta}^\epsilon g_{\mu\nu, \epsilon} - B_{\alpha\delta}^\epsilon g_{\mu\epsilon, \nu} - B_{\delta\alpha}^\epsilon g_{\epsilon\nu, \mu})] \\
&= A_\alpha^\lambda A_\gamma^\nu B_{\beta\delta}^\epsilon g_{\epsilon\eta} \Gamma_{\lambda\nu}^\eta + A_\beta^\mu A_\delta^\xi B_{\alpha\gamma}^\epsilon g_{\epsilon\eta} \Gamma_{\mu\xi}^\eta - A_\alpha^\lambda A_\delta^\xi B_{\beta\gamma}^\epsilon g_{\epsilon\eta} \Gamma_{\lambda\xi}^\eta - A_\beta^\mu A_\gamma^\nu B_{\alpha\delta}^\epsilon g_{\epsilon\eta} \Gamma_{\mu\nu}^\eta,
\end{aligned}$$

la dernière égalité résultant du lien entre les symboles de Christoffel et le tenseur métrique lorsque la connexion est riemannienne. Enfin, remplaçons les $B_{\alpha\beta}^\lambda$ par leur valeur tirée de (5.78). Il en résulte

$$\begin{aligned}
\mathcal{R}_{\alpha\beta\gamma\delta} &= A_\alpha^\lambda A_\beta^\mu A_\gamma^\nu A_\delta^\xi \left[\frac{1}{2} (g_{\lambda\xi, \mu\nu} + g_{\mu\nu, \lambda\xi} - g_{\mu\xi, \lambda\nu} - g_{\lambda\nu, \mu\xi}) + g_{\epsilon\eta} (\underbrace{\Gamma_{\lambda\nu}^\epsilon \Gamma_{\mu\xi}^\eta}_{\text{sym}} - \underbrace{\Gamma_{\lambda\xi}^\epsilon \Gamma_{\mu\nu}^\eta}_{\text{sym}}) \right. \\
&\quad \left. + g_{\epsilon\eta} (-\Gamma_{\mu\xi}^\epsilon \Gamma_{\lambda\nu}^\eta - \underbrace{\Gamma_{\lambda\nu}^\epsilon \Gamma_{\mu\xi}^\eta}_{\text{sym}} + \Gamma_{\mu\nu}^\epsilon \Gamma_{\lambda\xi}^\eta + \underbrace{\Gamma_{\lambda\xi}^\epsilon \Gamma_{\mu\nu}^\eta}_{\text{sym}}) \right] \\
&= A_\alpha^\lambda A_\beta^\mu A_\gamma^\nu A_\delta^\xi \left[\frac{1}{2} (g_{\lambda\xi, \mu\nu} + g_{\mu\nu, \lambda\xi} - g_{\mu\xi, \lambda\nu} - g_{\lambda\nu, \mu\xi}) + g_{\epsilon\eta} (\Gamma_{\mu\nu}^\epsilon \Gamma_{\lambda\xi}^\eta - \Gamma_{\mu\xi}^\epsilon \Gamma_{\lambda\nu}^\eta) \right] \\
&= A_\alpha^\lambda A_\beta^\mu A_\gamma^\nu A_\delta^\xi R_{\lambda\mu\nu\xi}. \tag{5.99}
\end{aligned}$$

Ce résultat prouve directement le caractère tensoriel de l'expression (5.98). Ainsi, *les conditions de compatibilité du système (5.72) constituent la loi de transformation d'un tenseur du quatrième ordre. On l'appelle tenseur de Riemann-Christoffel, ou encore, tenseur de courbure.* Ses composantes s'expriment à l'aide de $n^2(n^2 - 1)/12$ combinaisons linéaires « irréductibles » des dérivées secondes du tenseur métrique : cela signifie qu'on ne peut les annuler par un changement de coordonnées, sauf bien sûr si le tenseur de courbure est nul.

Ces considérations peuvent se généraliser aux dérivées d'ordre k du tenseur métrique. Il y en a $D_2^n D_k^n$. Au point P , ces quantités sont régies par les dérivées k ièmes de la matrice jacobienne, en nombre égal à $n D_{k+1}^n$. Pour ces dernières on dispose donc d'un système linéaire compatible de dimension $D_2^n D_k^n$, admettant

$$N(n, k) = D_2^n D_k^n - n D_{k+1}^n = \frac{n(k-1)(n+k-1)(n+k-2) \cdots n(n-1)}{2(k+1)!} \tag{5.100}$$

contraintes. Celles-ci expriment la loi tensorielle pour les $N(n, k)$ dérivées covariantes d'ordre $k - 2$

$$(\nabla_{\omega_1} \nabla_{\omega_2} \cdots \nabla_{\omega_{k-2}} \mathbf{R})_{\lambda\mu\nu\xi}(P)$$

du tenseur de courbure en dimension n . Remarquons que

$$N(4, k) = (k-1)(k+2)(k+3)$$

$$N(n, 3) = \frac{1}{24} n^2 (n^2 - 1) (n + 2).$$

5.17 Le tenseur de Riemann-Christoffel

5.17.1 Définition affine

Soit l'entité mathématique produisant un champ scalaire à partir d'un champ de formes et de trois champs de vecteurs, selon la définition que voici :

$$\mathbf{R}(\tilde{\Omega}, \vec{V}, \vec{X}, \vec{Y}) = \langle \tilde{\Omega}, \mathcal{R}(\vec{X}, \vec{Y})\vec{V} \rangle. \quad (5.101)$$

De la définition il résulte que l'action sur les formes est linéaire ; il en va de même pour l'action sur les vecteurs, en vertu des théorèmes 5.5 et 5.6. Cet objet est donc un *tenseur à quatre indices*, appelé *tenseur de Riemann-Christoffel* ou encore, *tenseur de courbure*. On a ainsi

$$\begin{aligned} \mathbf{R}(\tilde{\Omega}, \vec{V}, \vec{X}, \vec{Y}) &= \Omega_\alpha V^\beta X^\gamma Y^\delta \mathbf{R}(\tilde{\omega}^\alpha, \vec{e}_\beta, \vec{e}_\gamma, \vec{e}_\delta) \\ &= \Omega_\alpha V^\beta X^\gamma Y^\delta R^\alpha_{\beta\gamma\delta}. \end{aligned}$$

Il est facile de calculer ses composantes dans une base $\{\vec{e}_\alpha\}$ *naturelle*, c'est-à-dire telle que $[\vec{e}_\alpha, \vec{e}_\beta] = 0$:

$$\begin{aligned} R^\alpha_{\beta\gamma\delta} &= \mathbf{R}(\tilde{\omega}^\alpha, \vec{e}_\beta, \vec{e}_\gamma, \vec{e}_\delta) \\ &= \langle \tilde{\omega}^\alpha, \mathcal{R}(\vec{e}_\gamma, \vec{e}_\delta)\vec{e}_\beta \rangle \\ &= \langle \tilde{\omega}^\alpha, \nabla_\gamma \nabla_\delta \vec{e}_\beta - \nabla_\delta \nabla_\gamma \vec{e}_\beta \rangle \\ &= \langle \tilde{\omega}^\alpha, \nabla_\gamma (\Gamma_{\delta\beta}^\xi \vec{e}_\xi) - \nabla_\delta (\Gamma_{\gamma\beta}^\xi \vec{e}_\xi) \rangle \\ &= \langle \tilde{\omega}^\alpha, \Gamma_{\delta\beta,\gamma}^\xi \vec{e}_\xi + \Gamma_{\delta\beta}^\xi \Gamma_{\gamma\xi}^\mu \vec{e}_\mu - \Gamma_{\gamma\beta,\delta}^\xi \vec{e}_\xi - \Gamma_{\gamma\beta}^\xi \Gamma_{\delta\xi}^\mu \vec{e}_\mu \rangle \\ &= \Gamma_{\beta\delta,\gamma}^\alpha + \Gamma_{\gamma\xi}^\alpha \Gamma_{\beta\delta}^\xi - \Gamma_{\beta\gamma,\delta}^\alpha - \Gamma_{\delta\xi}^\alpha \Gamma_{\beta\gamma}^\xi. \end{aligned} \quad (5.102)$$

5.17.2 Symétries

Le tenseur de Riemann possède les propriétés suivantes :

$$\mathbf{R}(\cdot, \vec{V}, \vec{X}, \vec{Y}) = -\mathbf{R}(\cdot, \vec{V}, \vec{Y}, \vec{X}) \quad (5.103)$$

$$\mathbf{R}(\cdot, \vec{V}, \vec{X}, \vec{Y}) + \mathbf{R}(\cdot, \vec{X}, \vec{Y}, \vec{V}) + \mathbf{R}(\cdot, \vec{Y}, \vec{V}, \vec{X}) = 0. \quad (5.104)$$

La relation (5.103) résulte naturellement des définitions (5.101) et (5.96). D'autre part, la composante α du membre de gauche de (5.104) vaut

$$V^\beta X^\gamma Y^\delta \left(R^\alpha_{\beta\gamma\delta} + R^\alpha_{\gamma\delta\beta} + R^\alpha_{\delta\beta\gamma} \right)$$

où, compte tenu de (5.102), le terme entre parenthèses peut s'écrire

$$\begin{aligned} &\Gamma_{\beta\delta,\gamma}^\alpha + \Gamma_{\gamma\xi}^\alpha \Gamma_{\beta\delta}^\xi - \Gamma_{\beta\gamma,\delta}^\alpha - \Gamma_{\delta\xi}^\alpha \Gamma_{\beta\gamma}^\xi \\ &+ \Gamma_{\gamma\beta,\delta}^\alpha + \Gamma_{\delta\xi}^\alpha \Gamma_{\gamma\beta}^\xi - \Gamma_{\gamma\delta,\beta}^\alpha - \Gamma_{\beta\xi}^\alpha \Gamma_{\gamma\delta}^\xi \\ &+ \Gamma_{\delta\gamma,\beta}^\alpha + \Gamma_{\beta\xi}^\alpha \Gamma_{\delta\gamma}^\xi - \Gamma_{\delta\beta,\gamma}^\alpha - \Gamma_{\gamma\xi}^\alpha \Gamma_{\delta\beta}^\xi = 0. \end{aligned}$$

En composantes, les relations (5.103) et (5.104) prennent respectivement la forme

$$R^\alpha_{\beta\gamma\delta} = -R^\alpha_{\beta\delta\gamma} \quad (5.105)$$

$$R^\alpha_{[\beta\gamma\delta]} = 0, \quad (5.106)$$

où les crochets désignent l'opérateur d'antisymétrisation.

5.17.3 Interprétation géométrique

Nous suivons ici la démonstration de R. Wald⁸. Soit, incluse dans une variété affine à n dimensions, une surface dotée d'un système de coordonnées $\{t, s\}$. En un point P , donnons-nous un vecteur quelconque \vec{V} (pas nécessairement tangent à la surface) et transportons-le parallèlement le long du circuit $PP_1P_2P_3P$, comme le montre la figure 5.4. Soit également un champ de formes $\tilde{\Omega}$, qui va nous servir d'intermédiaire de calcul.

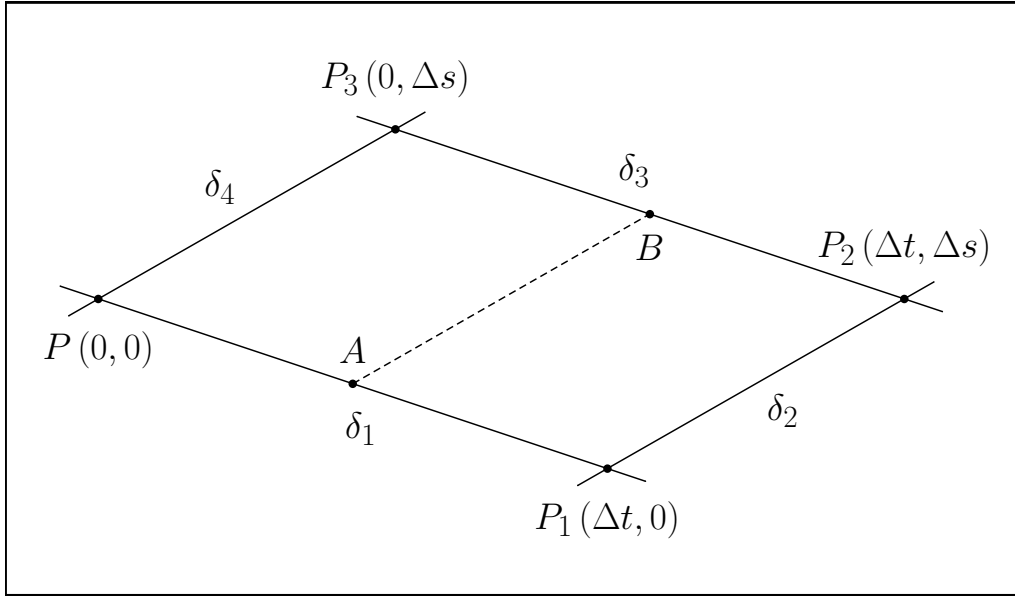


FIGURE 5.4: Transport parallèle d'un vecteur \vec{V} le long du circuit $PP_1P_2P_3P$. Les vecteurs \vec{T} et \vec{S} sont respectivement tangents aux lignes de coordonnées t et s : $\vec{T} = \partial/\partial t$, $\vec{S} = \partial/\partial s$. Les points A et B ont respectivement pour coordonnées $(\Delta t/2, 0)$ et $(\Delta t/2, \Delta s)$.

Appelant δ_1 la variation d'une quantité le long du chemin PP_1 , on calcule

$$\begin{aligned}
 \delta_1 \langle \tilde{\Omega}, \vec{V} \rangle &= \Delta t \left[\frac{d}{dt} \langle \tilde{\Omega}, \vec{V} \rangle \right]_P + \frac{1}{2} (\Delta t)^2 \left[\frac{d^2}{dt^2} \langle \tilde{\Omega}, \vec{V} \rangle \right]_P + O(3) \\
 &= \Delta t \left[\frac{d}{dt} \langle \tilde{\Omega}, \vec{V} \rangle \right]_A + O(3) \\
 &= \Delta t \left[\nabla_{\vec{T}} \langle \tilde{\Omega}, \vec{V} \rangle \right]_A + O(3) \\
 &= \Delta t \langle \nabla_{\vec{T}} \tilde{\Omega}, \vec{V} \rangle_A + \Delta t \langle \tilde{\Omega}, \nabla_{\vec{T}} \vec{V} \rangle_A + O(3).
 \end{aligned}$$

Or, la quantité $\nabla_{\vec{T}} \vec{V}$ est nulle, puisque \vec{V} est transporté parallèlement.

Raisonnant de même pour la variation correspondante le long du chemin P_2P_3 , notée δ_3 , on obtient

$$\delta_1 \langle \tilde{\Omega}, \vec{V} \rangle + \delta_3 \langle \tilde{\Omega}, \vec{V} \rangle = \Delta t \langle \nabla_{\vec{T}} \tilde{\Omega}, \vec{V} \rangle_A - \Delta t \langle \nabla_{\vec{T}} \tilde{\Omega}, \vec{V} \rangle_B + O(3), \quad (5.107)$$

c'est-à-dire une quantité du deuxième ordre. Reprenant ces considérations pour $\delta_2 + \delta_4$, on arrive à la conclusion selon laquelle *la variation d'un vecteur transporté parallèlement le long d'une courbe fermée du premier ordre, est une quantité du deuxième ordre.*

8. R. Wald, *General Relativity*, The University of Chicago Press, 1984, pp. 36-38.

En vertu de cette remarque, le vecteur \vec{V}_B résultant du transport parallèle de \vec{V} le long du « véritable » circuit AP_1P_2B est identique, aux termes du deuxième ordre près, au vecteur résultant du transport parallèle de \vec{V} le long du « raccourci » AB :

$$\vec{V}_B = \vec{V}_A \parallel_{\rightarrow B} + O(2). \quad (5.108)$$

D'autre part, on a⁹

$$(\nabla_{\vec{T}} \tilde{\Omega})_B = (\nabla_{\vec{T}} \tilde{\Omega})_A \parallel_{\rightarrow B} + \Delta s \left(\nabla_{\vec{S}} \nabla_{\vec{T}} \tilde{\Omega} \right)_B + O(2) \quad (5.109)$$

où $(\nabla_{\vec{T}} \tilde{\Omega})_A \parallel_{\rightarrow B}$ désigne le résultat du transport parallèle de la forme $\nabla_{\vec{T}} \tilde{\Omega}$ depuis A jusqu'à B , le long du tronçon AB . Les équations (5.108) et (5.109) permettent ainsi d'écrire

$$\langle \nabla_{\vec{T}} \tilde{\Omega}, \vec{V} \rangle_B = \langle \nabla_{\vec{T}} \tilde{\Omega}, \vec{V} \rangle_A \parallel_{\rightarrow B} + \Delta s \langle \left(\nabla_{\vec{S}} \nabla_{\vec{T}} \tilde{\Omega} \right)_B, \vec{V}_A \parallel_{\rightarrow B} \rangle + O(2). \quad (5.110)$$

Or, d'après (5.38), si une forme et un vecteur sont transportés parallèlement le long d'une courbe, le nombre résultant de l'application de la forme au vecteur est le même en tout point de la courbe :

$$\langle \nabla_{\vec{T}} \tilde{\Omega}, \vec{V} \rangle_A \parallel_{\rightarrow B} = \langle \nabla_{\vec{T}} \tilde{\Omega}, \vec{V} \rangle_A. \quad (5.111)$$

L'injection de (5.111) dans (5.110) donne

$$\langle \nabla_{\vec{T}} \tilde{\Omega}, \vec{V} \rangle_B = \langle \nabla_{\vec{T}} \tilde{\Omega}, \vec{V} \rangle_A + \Delta s \langle \left(\nabla_{\vec{S}} \nabla_{\vec{T}} \tilde{\Omega} \right)_B, \vec{V}_A \parallel_{\rightarrow B} \rangle + O(2). \quad (5.112)$$

Ce résultat (5.112) permet d'écrire l'équation (5.107) sous la forme

$$\begin{aligned} \delta_1 + \delta_3 &= -\Delta t \Delta s \langle \left(\nabla_{\vec{S}} \nabla_{\vec{T}} \tilde{\Omega} \right)_B, \vec{V}_A \parallel_{\rightarrow B} \rangle + O(3) \\ &= -\Delta t \Delta s \left[\langle \nabla_{\vec{S}} \nabla_{\vec{T}} \tilde{\Omega}, \vec{V} \rangle_P + O(1) \right] + O(3) \\ &= -\Delta t \Delta s \langle \nabla_{\vec{S}} \nabla_{\vec{T}} \tilde{\Omega}, \vec{V} \rangle_P + O(3). \end{aligned} \quad (5.113)$$

Procédant de même pour les tronçons δ_2 et δ_4 , il vient enfin

$$\begin{aligned} \delta_1 + \delta_2 + \delta_3 + \delta_4 &= \delta \langle \tilde{\Omega}, \vec{V} \rangle \\ &= \Delta t \Delta s \langle (\nabla_{\vec{T}} \nabla_{\vec{S}} - \nabla_{\vec{S}} \nabla_{\vec{T}}) \tilde{\Omega}, \vec{V} \rangle_P + O(3). \end{aligned} \quad (5.114)$$

D'autre part, $[\vec{T}, \vec{S}] = 0$ car les vecteurs \vec{T} et \vec{S} sont respectivement tangents aux lignes de coordonnées t et s . Donc $\mathcal{R}(\vec{T}, \vec{S}) = \nabla_{\vec{T}} \nabla_{\vec{S}} - \nabla_{\vec{S}} \nabla_{\vec{T}}$. Par (5.114) et le théorème 5.7, on a, en posant $\delta \vec{V} = \vec{V}_P \parallel_{\rightarrow P_1 \rightarrow P_2 \rightarrow P_3 \rightarrow P} - \vec{V}_P$:

$$\begin{aligned} \langle \tilde{\Omega}, \delta \vec{V} \rangle &= \delta \langle \tilde{\Omega}, \vec{V} \rangle \\ &= \Delta t \Delta s \langle \mathcal{R}(\vec{T}, \vec{S}) \tilde{\Omega}, \vec{V} \rangle_P + O(3) \\ &= -\Delta t \Delta s \langle \tilde{\Omega}, \mathcal{R}(\vec{T}, \vec{S}) \vec{V} \rangle_P + O(3). \end{aligned}$$

9. Rappelons une relation générale relative à la dérivée covariante. Etant donnés un champ de formes $\tilde{\Sigma}$ et une courbe $\mathcal{C}(\lambda)$ de vecteur tangent \vec{U} , on a [cf. la relation (5.27)]

$$\tilde{\Sigma}[P(\lambda_0)] = \tilde{\Sigma}[P(\lambda_0 + \epsilon)] \parallel_{\rightarrow P(\lambda_0)} - \epsilon \left(\nabla_{\vec{U}} \tilde{\Sigma} \right)_{P(\lambda_0)} + O(2).$$

Posant $\lambda = s$, $\lambda_0 = \Delta s$, $\epsilon = -\Delta s$ et $\tilde{\Sigma} = \nabla_{\vec{T}} \tilde{\Omega}$, on obtient $\vec{U} = \vec{S}$, $P(\lambda_0) = B$, $P(\lambda_0 + \epsilon) = A$ et on déduit la relation (5.109).

Le champ de formes $\tilde{\Omega}$ étant arbitraire, on en déduit la signification géométrique du tenseur de Riemann :

$$\delta\vec{V} = -\Delta t \Delta s \mathcal{R}(\vec{T}, \vec{S})\vec{V} + O(3)$$

$$\boxed{\delta\vec{V} = -\Delta t \Delta s \mathbf{R}(\cdot, \vec{V}, \vec{T}, \vec{S}) + O(3)}. \quad (5.115)$$

Remarquons que $\vec{V}_{P \parallel \rightarrow P_1 \rightarrow P_2 \rightarrow P_3 \rightarrow P} = \vec{V}_{P \parallel \rightarrow P_1 \rightarrow P_2}$. Or, le membre de gauche n'est autre que $(\vec{V} + \delta\vec{V})_{P \parallel \rightarrow P_3 \rightarrow P_2} = \vec{V}_{P \parallel \rightarrow P_3 \rightarrow P_2} + \delta\vec{V}_{P \parallel \rightarrow P_3 \rightarrow P_2}$. Mais

$$(\delta\vec{V}_{P \parallel \rightarrow P_3 \rightarrow P_2})^\mu = \delta V^\mu + O(3).$$

Par conséquent, les composantes δV^μ représentent également la différence entre les composantes du vecteur \vec{V} transporté parallèlement de P en P_2 le long du circuit $P P_1 P_2$ et celles du vecteur \vec{V} transporté parallèlement de P en P_2 le long du circuit $P P_3 P_2$.

5.18 Les espaces euclidiens

Une variété différentielle V^n , dotée d'une structure affine, est qualifiée d'*euclidienne* si le tenseur de Riemann-Christoffel s'y annule.

Théorème 5.8 *Dans une variété euclidienne simplement connexe le parallélisme est global, autrement dit, le résultat du transport parallèle est indépendant du chemin utilisé.*

Démonstration. Dans un ouvert simplement connexe, deux courbes d'extrémités identiques peuvent être appliquées l'une sur l'autre par une transformation continue (*homotopie*). Pour démontrer le théorème, il suffit donc de s'assurer que le transport parallèle d'un vecteur le long de deux chemins C_1 et C_2 arbitrairement proches et de mêmes extrémités A et B , fournit des résultats identiques (figure 5.5).

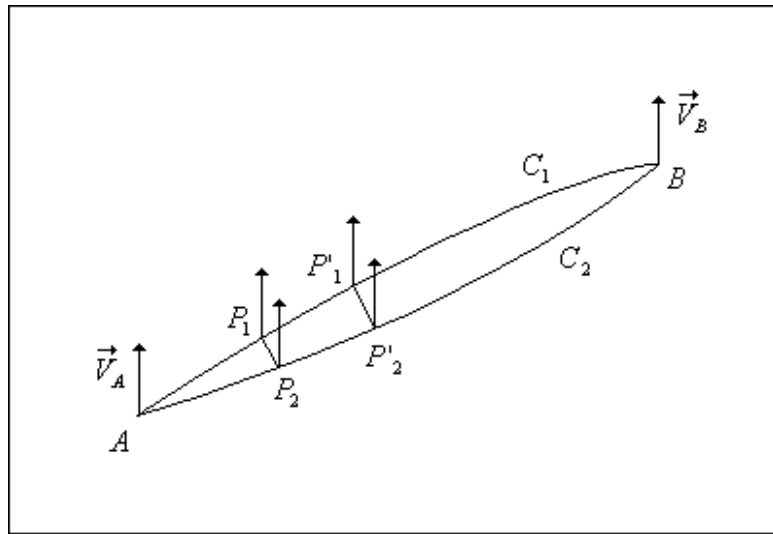


FIGURE 5.5: *Transport parallèle d'un vecteur le long de deux courbes voisines C_1 et C_2 de mêmes extrémités A et B .*

Appelons \vec{V}_1 le vecteur obtenu par transport parallèle, en suivant \mathcal{C}_1 , de \vec{V}_A depuis A jusqu'à P_1 infiniment proche de A . De même, \vec{V}_2 résulte du transport parallèle, en suivant \mathcal{C}_2 , de \vec{V}_A depuis A jusqu'à P_2 infiniment proche de A . En vertu de l'annulation du tenseur de Riemann il y a parallélisme local et, par conséquent, *le résultat du transport parallèle de \vec{V}_1 le long du segment infinitésimal P_1P_2 n'est autre que \vec{V}_2 .*

On procède ainsi de proche en proche en découpant en petits quadrilatères $P'_1P_1P_2P'_2$ l'espace compris entre les deux courbes : si \vec{V}'_1 est le vecteur obtenu par transport parallèle de \vec{V}_1 le long du segment infinitésimal $P_1P'_1$ et si \vec{V}'_2 est obtenu par transport parallèle de \vec{V}_2 le long du segment infinitésimal $P_2P'_2$, alors \vec{V}'_2 résulte du transport parallèle de \vec{V}'_1 le long du segment infinitésimal $P'_1P'_2$.

Les vecteurs respectivement transportés le long de \mathcal{C}_1 et \mathcal{C}_2 restent donc « connectés » l'un à l'autre, même loin de A , par transport parallèle le long d'un segment infinitésimal qui, finalement, se réduit à un point en B ; ce qui signifie que le vecteur \vec{V}_B obtenu en B est unique. □

Lemme 5.2 (Frobenius) *Dans V^n simplement connexe, un ensemble de n champs de vecteurs linéairement indépendants $\{\vec{e}_\mu\}$ constitue une base naturelle relative à un système de coordonnées $\{\underline{x}^\mu\}$, autrement dit on peut écrire $\vec{e}_\mu = \partial/\partial \underline{x}^\mu$, si et seulement si $[\vec{e}_\mu, \vec{e}_\nu] = 0 \quad \forall \mu, \nu = 1, \dots, n$.*

Démonstration. La condition nécessaire est évidente. Pour justifier la réciproque, considérons une base naturelle de vecteurs $\vec{e}_\alpha = \partial/\partial x^\alpha$ et rapportons-y les champs vectoriels :

$$\vec{e}_\mu = \Lambda_\mu^\alpha \vec{e}_\alpha.$$

Procédons de même avec les champs de formes duales :

$$\tilde{\omega}^\nu = \Psi_\beta^\nu \tilde{\omega}^\beta.$$

Donc, par construction,

$$\langle \tilde{\omega}^\nu, \vec{e}_\mu \rangle = \Psi_\alpha^\nu \Lambda_\mu^\alpha = \delta_\mu^\nu. \quad (5.116)$$

Cela étant, d'après l'hypothèse, on a

$$\langle \tilde{\omega}^\xi, [\vec{e}_\mu, \vec{e}_\nu] \rangle = 0$$

pour tout ξ, μ, ν . Développons cette relation :

$$\begin{aligned} 0 &= \left\langle \Psi_\gamma^\xi \tilde{\omega}^\gamma, [\Lambda_\mu^\alpha \vec{e}_\alpha, \Lambda_\nu^\beta \vec{e}_\beta] \right\rangle \\ &= \left\langle \Psi_\gamma^\xi \tilde{\omega}^\gamma, \Lambda_\mu^\alpha \Lambda_\nu^\beta [\vec{e}_\alpha, \vec{e}_\beta] + \Lambda_\mu^\alpha \frac{\partial \Lambda_\nu^\beta}{\partial x^\alpha} \vec{e}_\beta - \Lambda_\nu^\beta \frac{\partial \Lambda_\mu^\alpha}{\partial x^\beta} \vec{e}_\alpha \right\rangle \\ &= 0 + \Psi_\beta^\xi \Lambda_\mu^\alpha \frac{\partial \Lambda_\nu^\beta}{\partial x^\alpha} - \Psi_\alpha^\xi \Lambda_\nu^\beta \frac{\partial \Lambda_\mu^\alpha}{\partial x^\beta}. \end{aligned}$$

Vu la relation (5.116), ce résultat peut s'écrire sous la forme

$$\begin{aligned} 0 &= -\frac{\partial \Psi_\beta^\xi}{\partial x^\alpha} \Lambda_\mu^\alpha \Lambda_\nu^\beta + \frac{\partial \Psi_\alpha^\xi}{\partial x^\beta} \Lambda_\nu^\beta \Lambda_\mu^\alpha \\ &= \Lambda_\mu^\alpha \Lambda_\nu^\beta \left(\frac{\partial \Psi_\alpha^\xi}{\partial x^\beta} - \frac{\partial \Psi_\beta^\xi}{\partial x^\alpha} \right). \end{aligned}$$

Il en résulte

$$\frac{\partial \Psi_\alpha^\xi}{\partial x^\beta} = \frac{\partial \Psi_\beta^\xi}{\partial x^\alpha} \quad \forall \xi, \alpha, \beta. \quad (5.117)$$

D'après la théorie de la primitivation dans un espace à n dimensions simplement connexe, les conditions (5.117) prouvent l'existence de n fonctions $\underline{x}^\xi(\vec{x})$ telles que

$$\Psi_\alpha^\xi = \frac{\partial \underline{x}^\xi}{\partial x^\alpha}. \quad (5.118)$$

On en déduit

$$\tilde{\omega}^\nu = \frac{\partial \underline{x}^\nu}{\partial x^\beta} \tilde{\omega}^\beta = \tilde{d}\underline{x}^\nu \quad \text{et} \quad \tilde{e}_\mu = \frac{\partial x^\alpha}{\partial \underline{x}^\mu} \vec{e}_\alpha = \frac{\partial}{\partial \underline{x}^\mu}, \quad (5.119)$$

ce qu'il fallait démontrer. \square

Théorème 5.9 *Une variété V^n simplement connexe est euclidienne si et seulement si on peut y définir des coordonnées où s'annulent tous les symboles de connexion affine.*

Démonstration. La condition suffisante est triviale puisque l'annulation des symboles de connexion implique celle du tenseur de courbure.

Passons à la condition nécessaire. Remarquons d'abord que le transport parallèle préserve l'indépendance linéaire des vecteurs. Il permet donc de propager une base $\{\vec{e}_\mu\}$ sur V^n à partir de la donnée de n vecteurs linéairement indépendants en un point initial. La relation

$$\nabla_{\vec{e}_\beta} \vec{e}_\alpha = 0 \quad (5.120)$$

engendre ainsi une base en tout point de la variété affine. La cohérence de la démarche (univocité du résultat) est garantie par le fait que le parallélisme est global dans V^n , vu le théorème 5.8. De plus,

$$\begin{aligned} [\vec{e}_\alpha, \vec{e}_\beta] &= \nabla_{\vec{e}_\alpha} \vec{e}_\beta - \nabla_{\vec{e}_\beta} \vec{e}_\alpha \\ &= 0 - 0, \end{aligned}$$

ce qui assure, en raison du lemme 5.2, que l'on détermine bien, de cette manière, un système de coordonnées. Comme

$$\nabla_{\vec{e}_\beta} \vec{e}_\alpha = \Gamma_{\alpha\beta}^\gamma \vec{e}_\gamma,$$

on a donc, en vertu de (5.120),

$$\Gamma_{\alpha\beta}^\gamma(P) = 0$$

en tout point P de V^n .¹⁰ \square

5.19 Equation de déviation géodésique

Sur une variété affine, soit une courbe $\mathcal{C}(s)$ interceptant un champ de géodésiques $\gamma_s(t)$ de vecteur tangent $\vec{T}_s(t)$, où t est un paramètre affine tel que $t = 0$ sur $\mathcal{C}(s)$. Considérons les courbes $\mathcal{C}'(s)$, $\mathcal{C}''(s)$ etc., reliant les points des géodésiques $\gamma_s(t)$ correspondant à des t identiques. On détermine ainsi une surface dotée de lignes de coordonnées $\{t, s\}$, de vecteurs tangents respectifs $\vec{T} = \partial/\partial t$ et $\vec{S} = \partial/\partial s$ (figure 5.6).

10. Si la variété et la connexion sont riemanniennes, il en résulte la constance du tenseur métrique qui, au demeurant, peut être ramené à une forme canonique. Par exemple, en relativité générale

$$(g_{\mu\nu})(P) = \text{diag}(-1, 1, 1, 1).$$

L'espace-temps est alors rapporté à un *référentiel inertiel* et les coordonnées sont qualifiées de *cartésiennes*.

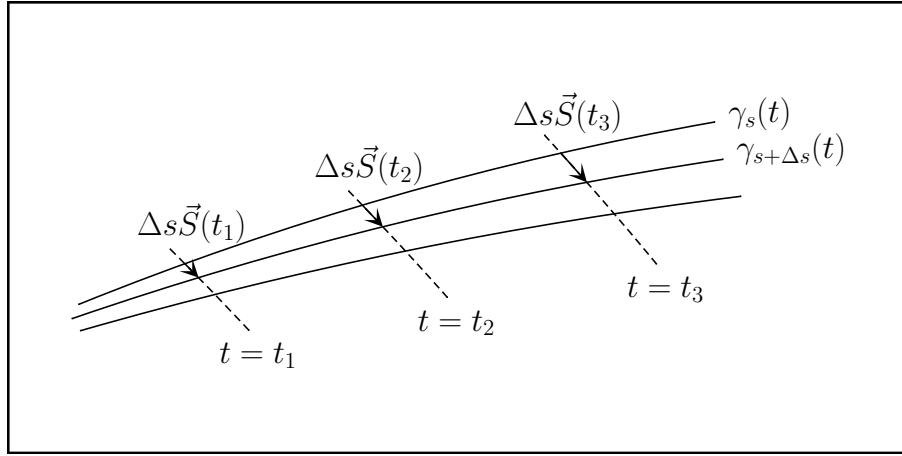


FIGURE 5.6: Evolution de l'écart $\Delta_s \vec{S}(t)$ entre deux géodésiques voisines $\gamma_s(t)$ et $\gamma_{s+\Delta_s}(t)$.

Au premier ordre, $\Delta_s \vec{S}(t)$ est le vecteur reliant le point $\gamma_s(t)$ au point $\gamma_{s+\Delta_s}(t)$. Le tenseur de Riemann permet d'exprimer simplement son « accélération », c'est-à-dire la dérivée covariante deuxième de $\Delta_s \vec{S}$ par rapport au vecteur \vec{T} tangent aux géodésiques :

$$\vec{a} = \nabla_{\vec{T}} \nabla_{\vec{T}} (\Delta_s \vec{S}).$$

Dans un espace euclidien, il est facile de démontrer la constance de la « vitesse » de $\Delta_s \vec{S}$, c'est-à-dire de la quantité

$$\vec{V} = \nabla_{\vec{T}} (\Delta_s \vec{S})$$

(le théorème de Thalès en constitue une application particulière). En d'autres termes, on peut d'ores et déjà prédire que l'accélération géodésique est nulle dans un espace euclidien.

Pour une variété affine quelconque on a, Δ_s étant constant :

$$\begin{aligned} \vec{a} &= \Delta_s \nabla_{\vec{T}} \nabla_{\vec{T}} \vec{S} \\ &= \Delta_s \nabla_{\vec{T}} \nabla_{\vec{S}} \vec{T} \end{aligned}$$

en vertu de (5.94) et du fait que $[\vec{T}, \vec{S}] = 0$. Compte tenu des définitions (5.101) et (5.96), il s'ensuit

$$\vec{a} = \Delta_s \nabla_{\vec{S}} \nabla_{\vec{T}} \vec{T} + \Delta_s \mathbf{R}(\cdot, \vec{T}, \vec{T}, \vec{S}).$$

Or, le premier terme est nul puisque les courbes $\gamma_s(t)$ sont géodésiques. Il vient donc

$$\boxed{\vec{a} = \Delta_s \mathbf{R}(\cdot, \vec{T}, \vec{T}, \vec{S}).} \quad (5.121)$$

Si la variété est riemannienne et que t désigne la longueur d'arc, on a $\vec{T} \cdot \vec{T} = \pm 1$. Le produit scalaire $\vec{T} \cdot \vec{S}$ est alors constant le long des géodésiques. En effet,

$$\frac{d(\vec{T} \cdot \vec{S})}{dt} = \nabla_{\vec{T}} (\vec{T} \cdot \vec{S}) = \vec{T} \cdot \nabla_{\vec{T}} \vec{S} = \vec{T} \cdot \nabla_{\vec{S}} \vec{T} = \frac{1}{2} \nabla_{\vec{S}} (\vec{T} \cdot \vec{T}) = 0.$$

Par conséquent, si la déviation \vec{S} est orthogonale à \vec{T} en $t = 0$, alors cette propriété subsiste $\forall t$. Ainsi, la composante parallèle à \vec{T} d'une déviation ne joue pas de rôle fondamental. C'est la raison pour laquelle, en général, on la suppose nulle.

5.20 Les identités de Bianchi

5.20.1 Aspects algébriques

Lemme 5.3 *Le commutateur vérifie l'identité de Jacobi*

$$\left[\left[\vec{X}, \vec{Y} \right], \vec{Z} \right] + \left[\left[\vec{Y}, \vec{Z} \right], \vec{X} \right] + \left[\left[\vec{Z}, \vec{X} \right], \vec{Y} \right] = 0. \quad (5.122)$$

Pour le prouver, il suffit de développer les calculs. Posant $\vec{C} = [\vec{X}, \vec{Y}]$, il vient

$$\begin{aligned} \left[\left[\vec{X}, \vec{Y} \right], \vec{Z} \right]^\alpha &= C^\gamma Z^\alpha_{,\gamma} - Z^\gamma C^\alpha_{,\gamma} \\ &= \left(X^\mu Y^\gamma_{,\mu} - Y^\mu X^\gamma_{,\mu} \right) Z^\alpha_{,\gamma} - Z^\gamma \left(X^\mu Y^\alpha_{,\mu} - Y^\mu X^\alpha_{,\mu} \right)_{,\gamma} \\ &= X^\mu Y^\gamma_{,\mu} Z^\alpha_{,\gamma} - Y^\mu X^\gamma_{,\mu} Z^\alpha_{,\gamma} - Z^\gamma X^\mu_{,\gamma} Y^\alpha_{,\mu} - Z^\gamma X^\mu Y^\alpha_{,\mu\gamma} + Z^\gamma Y^\mu_{,\gamma} X^\alpha_{,\mu} + Z^\gamma Y^\mu X^\alpha_{,\mu\gamma}. \end{aligned}$$

En effectuant la permutation circulaire sur \vec{X} , \vec{Y} et \vec{Z} , on trouve identiquement zéro. \square

Théorème 5.10 (identités de L. Bianchi) *L'opérateur de dérivation covariante vérifie l'identité de Jacobi*

$$\left[\left[\nabla_{\vec{X}}, \nabla_{\vec{Y}} \right], \nabla_{\vec{Z}} \right] + \left[\left[\nabla_{\vec{Y}}, \nabla_{\vec{Z}} \right], \nabla_{\vec{X}} \right] + \left[\left[\nabla_{\vec{Z}}, \nabla_{\vec{X}} \right], \nabla_{\vec{Y}} \right] = 0. \quad (5.123)$$

Démonstration. Transformons d'abord le premier terme de (5.123) :

$$\begin{aligned} \left[\left[\nabla_{\vec{X}}, \nabla_{\vec{Y}} \right], \nabla_{\vec{Z}} \right] \vec{V} &= \left[\mathcal{R}(\vec{X}, \vec{Y}) + \nabla_{[\vec{X}, \vec{Y}]}, \nabla_{\vec{Z}} \right] \vec{V} \\ &= \left[\mathcal{R}(\vec{X}, \vec{Y}), \nabla_{\vec{Z}} \right] \vec{V} + \mathcal{R}([\vec{X}, \vec{Y}], \vec{Z}) \vec{V} + \nabla_{[[\vec{X}, \vec{Y}], \vec{Z}]} \vec{V} \end{aligned} \quad (5.124)$$

Lorsque l'on effectue la permutation circulaire sur \vec{X} , \vec{Y} et \vec{Z} dans le membre de droite de (5.124), le troisième terme disparaît en vertu du lemme 5.3. Quant au premier terme, il vaut

$$\mathbf{R}(\cdot, \nabla_{\vec{Z}} \vec{V}, \vec{X}, \vec{Y}) - \nabla_{\vec{Z}} \mathbf{R}(\cdot, \vec{V}, \vec{X}, \vec{Y})$$

dont la composante α , dans une base donnée, est égale à

$$\begin{aligned} &\mathbf{R}(\tilde{\omega}^\alpha, \nabla_{\vec{Z}} \vec{V}, \vec{X}, \vec{Y}) - \langle \tilde{\omega}^\alpha, \nabla_{\vec{Z}} \mathbf{R}(\cdot, \vec{V}, \vec{X}, \vec{Y}) \rangle \\ &= \mathbf{R}(\tilde{\omega}^\alpha, \nabla_{\vec{Z}} \vec{V}, \vec{X}, \vec{Y}) - \nabla_{\vec{Z}} \langle \tilde{\omega}^\alpha, \mathbf{R}(\cdot, \vec{V}, \vec{X}, \vec{Y}) \rangle + \langle \nabla_{\vec{Z}} \tilde{\omega}^\alpha, \mathbf{R}(\cdot, \vec{V}, \vec{X}, \vec{Y}) \rangle \\ &= \mathbf{R}(\tilde{\omega}^\alpha, \nabla_{\vec{Z}} \vec{V}, \vec{X}, \vec{Y}) - \nabla_{\vec{Z}} \mathbf{R}(\tilde{\omega}^\alpha, \vec{V}, \vec{X}, \vec{Y}) + \mathbf{R}(\nabla_{\vec{Z}} \tilde{\omega}^\alpha, \vec{V}, \vec{X}, \vec{Y}) \\ &= Z^\mu \left(V^\beta_{,\mu} + \Gamma_{\mu\xi}^\beta V^\xi \right) X^\gamma Y^\delta R^\alpha_{\beta\gamma\delta} - Z^\mu \left(V^\beta X^\gamma Y^\delta R^\alpha_{\beta\gamma\delta} \right)_{,\mu} - Z^\mu \Gamma_{\mu\xi}^\alpha V^\beta X^\gamma Y^\delta R^\xi_{\beta\gamma\delta} \\ &= Z^\mu V^\beta X^\gamma Y^\delta \left(\Gamma_{\mu\beta}^\xi R^\alpha_{\xi\gamma\delta} - \Gamma_{\mu\xi}^\alpha R^\xi_{\beta\gamma\delta} - R^\alpha_{\beta\gamma\delta,\mu} \right) - Z^\mu V^\beta \left(X^\gamma Y^\delta \right)_{,\mu} R^\alpha_{\beta\gamma\delta} \\ &= Z^\mu V^\beta X^\gamma Y^\delta \left[\left(-\nabla_\mu \mathbf{R} \right)^\alpha_{\beta\gamma\delta} - \Gamma_{\mu\gamma}^\xi R^\alpha_{\beta\xi\delta} - \Gamma_{\mu\delta}^\xi R^\alpha_{\beta\gamma\xi} \right] - Z^\mu V^\beta \left(X^\gamma Y^\delta \right)_{,\mu} R^\alpha_{\beta\gamma\delta}. \end{aligned}$$

Le membre de gauche de (5.123) prend alors la forme

$$Z^\mu V^\beta X^\gamma Y^\delta \left[\left(-\nabla_\mu \mathbf{R} \right)^\alpha_{\beta\gamma\delta} - \Gamma_{\mu\gamma}^\xi R^\alpha_{\beta\xi\delta} - \Gamma_{\mu\delta}^\xi R^\alpha_{\beta\gamma\xi} \right] - \underbrace{Z^\mu V^\beta \left(X^\gamma Y^\delta \right)_{,\mu} R^\alpha_{\beta\gamma\delta}}$$

$$\begin{aligned}
& + X^\mu V^\beta Y^\gamma Z^\delta \left[(-\nabla_\mu \mathbf{R})^\alpha_{\beta\gamma\delta} - \underline{\underline{\Gamma_{\mu\gamma}^\xi R^\alpha_{\beta\xi\delta}}} - \underline{\underline{\Gamma_{\mu\delta}^\xi R^\alpha_{\beta\gamma\xi}}} \right] - \underline{\underline{X^\mu V^\beta (Y^\gamma Z^\delta)_{,\mu} R^\alpha_{\beta\gamma\delta}}} \\
& + Y^\mu V^\beta Z^\gamma X^\delta \left[(-\nabla_\mu \mathbf{R})^\alpha_{\beta\gamma\delta} - \underline{\underline{\Gamma_{\mu\gamma}^\xi R^\alpha_{\beta\xi\delta}}} - \underline{\underline{\Gamma_{\mu\delta}^\xi R^\alpha_{\beta\gamma\xi}}} \right] - \underline{\underline{Y^\mu V^\beta (Z^\gamma X^\delta)_{,\mu} R^\alpha_{\beta\gamma\delta}}} \\
& + V^\beta R^\alpha_{\beta\gamma\delta} \left[\left(\underline{\underline{X^\mu Y^\gamma_{,\mu}}} - \underline{\underline{Y^\mu X^\gamma_{,\mu}}} \right) Z^\delta + \left(\underline{\underline{Y^\mu Z^\gamma_{,\mu}}} - \underline{\underline{Z^\mu Y^\gamma_{,\mu}}} \right) X^\delta + \left(\underline{\underline{Z^\mu X^\gamma_{,\mu}}} - \underline{\underline{X^\mu Z^\gamma_{,\mu}}} \right) Y^\delta \right]
\end{aligned}$$

(les termes soulignés de façon identique s'annulent entre eux).

La thèse (5.123) se réduit ainsi à l'identité

$$V^\beta X^\gamma Y^\delta Z^\mu \left[(\nabla_\mu \mathbf{R})^\alpha_{\beta\gamma\delta} + (\nabla_\gamma \mathbf{R})^\alpha_{\beta\delta\mu} + (\nabla_\delta \mathbf{R})^\alpha_{\beta\mu\gamma} \right] = 0. \quad (5.125)$$

Pour démontrer cette relation tensorielle en un point donné, utilisons un système de coordonnées où tous les symboles de connexion affine s'annulent en ce point (l'existence d'un tel système est prouvée dans l'annexe B). On a alors

$$\begin{aligned}
(\nabla_\mu \mathbf{R})^\alpha_{\beta\gamma\delta} &= \Gamma_{\beta\delta,\gamma\mu}^\alpha - \Gamma_{\beta\gamma,\delta\mu}^\alpha \\
(\nabla_\gamma \mathbf{R})^\alpha_{\beta\delta\mu} &= \Gamma_{\beta\mu,\delta\gamma}^\alpha - \Gamma_{\beta\delta,\mu\gamma}^\alpha \\
(\nabla_\delta \mathbf{R})^\alpha_{\beta\mu\gamma} &= \Gamma_{\beta\gamma,\mu\delta}^\alpha - \Gamma_{\beta\mu,\gamma\delta}^\alpha
\end{aligned}$$

et la somme de ces trois quantités est bien égale à zéro.

□

5.20.2 Interprétation géométrique

Le paragraphe 5.17.3 exprimait le résultat du transport parallèle d'un vecteur le long d'un parallélogramme infinitésimal [relation (5.115)]. Les identités de Bianchi généralisent ce résultat pour un parallépipède infinitésimal dans l'espace à trois dimensions : un vecteur transporté parallèlement, depuis un sommet, le long de *toutes* les arêtes de ce parallépipède revient à son point de départ sans changement. Du point de vue du transport parallèle, le trajet total est donc comme nul.¹¹

Soit un parallépipède infinitésimal de sommet O , origine d'un système de coordonnées normales de Riemann (x, y, z) (annexe B) de vecteurs de base $(\vec{e}_x, \vec{e}_y, \vec{e}_z)$ tangents à ses arêtes. Attribuons au sommet P_6 , opposé à O , les coordonnées $(\Delta x, \Delta y, \Delta z)$. Commençons par calculer le résultat du transport parallèle d'un vecteur le long du trajet $O P_1 P_2 P_3 O P_4 P_5 P_6 P_7 P_4 O$ (figure 5.7).

D'après la relation (5.115), le résultat du transport parallèle d'un vecteur initial de coordonnées V^α au terme du parcours $O P_1 P_2 P_3 O$, est le vecteur de coordonnées

$$V^\alpha - R^\alpha_{\beta zy}(O) V^\beta \Delta y \Delta z + O(3)(\Delta y, \Delta z).$$

Qu'obtient-on à l'issue du tronçon $O P_4$? Par hypothèse, dans notre système de coordonnées tous les symboles de Christoffel sont nuls en O . Un transport parallèle le long dudit

11. Lors d'un tel trajet, chaque arête est parcourue une fois dans chaque sens. On peut démontrer (ou, au moins, se persuader en le testant sur quelques exemples) qu'il n'existe pas de chemin parcourant l'ensemble des douze arêtes sans emprunter deux fois chacune d'entre elles, une fois dans chaque sens.

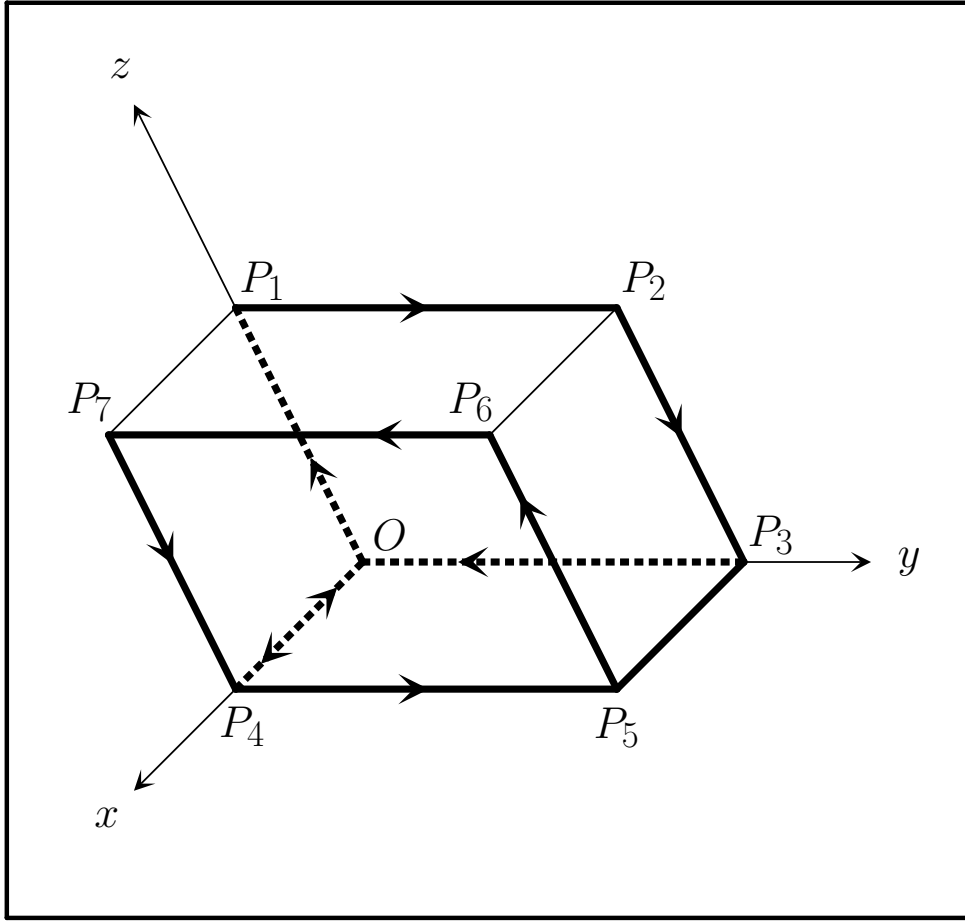


FIGURE 5.7: *Interprétation géométrique des identités de Bianchi. En gras, on a indiqué un tiers du circuit total, celui qui est relatif aux faces $x = 0$ et $x = \Delta x$. Chaque face est parcourue dans le sens tel que le vecteur de rotation associé pointe vers l'extérieur.*

tronçon, c'est-à-dire depuis le point O de coordonnées $x = 0, y = 0, z = 0$ jusqu'au point P_4 de coordonnées $x = \Delta x, y = 0, z = 0$, ne modifie donc les composantes que d'une quantité d'ordre $(\Delta x)^2$. Il en résulte le vecteur

$$V^\alpha - R^\alpha_{\beta zy}(O) V^\beta \Delta y \Delta z + O(3)(\Delta y, \Delta z) + O(\Delta x)^2 + O(4).$$

Toujours d'après (5.115), au terme du tronçon $P_4 P_5 P_6 P_7 P_4$ on obtient le vecteur de composantes

$$\begin{aligned} V^\alpha &- R^\alpha_{\beta zy}(O) V^\beta \Delta y \Delta z + O(3)(\Delta y, \Delta z) + O(\Delta x)^2 \\ &- R^\alpha_{\beta yz}(P_4) V^\beta \Delta y \Delta z - O(3)(\Delta y, \Delta z) + O(4). \end{aligned}$$

Les termes d'ordre 3 étant construits de façon identique à partir des composantes du vecteur initial V^μ et des quantités Δy et Δz , ils s'annulent. En d'autres termes, le résultat est

$$V^\alpha - R^\alpha_{\beta zy}(O) V^\beta \Delta y \Delta z + O(\Delta x)^2 + R^\alpha_{\beta yz}(P_4) V^\beta \Delta y \Delta z + O(4),$$

c'est-à-dire

$$V^\alpha + \frac{\partial R^\alpha_{\beta zy}}{\partial x}(O) V^\beta \Delta x \Delta y \Delta z + O(\Delta x)^2 + O(4).$$

Enfin, au terme du tronçon $P_4 O$, on obtient

$$V^\alpha + \frac{\partial R^\alpha_{\beta zy}}{\partial x}(O) V^\beta \Delta x \Delta y \Delta z + O(\Delta x)^2 + O(4) - O(\Delta x)^2.$$

Les termes d'ordre 2 se simplifient car ils sont construits de façon identique à partir des composantes initiales V^μ et de $(\Delta x)^2$.

En résumé, à l'issue du parcours $O P_1 P_2 P_3 O P_4 P_5 P_6 P_7 P_4 O$, les composantes du vecteur initial sont modifiées d'une quantité

$$R^\alpha_{\beta zy,x}(O) V^\beta \Delta x \Delta y \Delta z + O(4).$$

Poursuivons le chemin parcourant toutes les arêtes en raisonnant à l'identique pour les deux circuits, respectivement le long des faces $y = 0$ et $y = \Delta y$ puis $z = 0$ et $z = \Delta z$. On obtient le résultat final :¹²

$$\delta V^\alpha = \left[R^\alpha_{\beta zy,x}(O) + R^\alpha_{\beta xz,y}(O) + R^\alpha_{\beta yx,z}(O) \right] V^\beta \Delta x \Delta y \Delta z + O(4). \quad (5.126)$$

On peut remplacer les dérivations partielles par des dérivations covariantes, puisque ces deux opérations sont identiques dans un système de coordonnées normales. La relation (5.125) garantit l'annulation des quantités δV^α dans l'équation (5.126).

5.21 Courbure des variétés riemanniennes

5.21.1 Nouvelles propriétés de symétrie

Par abus de langage, le tenseur

$$R_{\alpha\beta\gamma\delta} = g_{\alpha\mu} R^\mu_{\beta\gamma\delta}$$

sera toujours appelé *tenseur de Riemann*. En utilisant la relation (5.102), en exprimant les symboles de Christoffel dérivés en fonction du tenseur métrique et en tenant compte du fait que

$$\begin{aligned} g_{\alpha\mu} g^{\mu\xi}_{,\gamma} &= -g_{\alpha\mu,\gamma} g^{\mu\xi} \\ &= -\left(\Gamma^\rho_{\alpha\gamma} g_{\rho\mu} + \Gamma^\rho_{\mu\gamma} g_{\alpha\rho} \right) g^{\mu\xi}, \end{aligned}$$

il est assez facile de voir que

$$R_{\alpha\beta\gamma\delta} = \frac{1}{2} (g_{\alpha\delta,\beta\gamma} + g_{\beta\gamma,\alpha\delta} - g_{\beta\delta,\alpha\gamma} - g_{\alpha\gamma,\beta\delta}) + g_{\epsilon\eta} \left(\Gamma^\epsilon_{\beta\gamma} \Gamma^\eta_{\alpha\delta} - \Gamma^\epsilon_{\beta\delta} \Gamma^\eta_{\alpha\gamma} \right). \quad (5.127)$$

On en déduit

$$R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta}. \quad (5.128)$$

12. A strictement parler, les arêtes $O P_1$, $O P_3$ et $O P_4$ sont alors parcourues *deux* fois dans chaque sens ; mais cela n'est pas essentiel, comme on vient de le voir pour le tronçon $O P_4$ dont la contribution est nulle lors du premier circuit.

Bien entendu, les relations affines (5.105) et (5.106) restent valables. Il en découle une importante propriété de symétrie « par blocs » :

$$\begin{aligned}
R_{\alpha\beta\gamma\delta} &= -R_{\alpha\gamma\delta\beta} - R_{\alpha\delta\beta\gamma} \\
&= R_{\gamma\alpha\delta\beta} + R_{\delta\alpha\beta\gamma} \\
&= -R_{\gamma\delta\beta\alpha} - R_{\gamma\beta\alpha\delta} - R_{\delta\beta\gamma\alpha} - R_{\delta\gamma\alpha\beta} \\
&= R_{\gamma\delta\alpha\beta} - R_{\beta\gamma\delta\alpha} - R_{\beta\delta\alpha\gamma} + R_{\gamma\delta\alpha\beta} \\
&= R_{\gamma\delta\alpha\beta} + R_{\beta\alpha\gamma\delta} + R_{\gamma\delta\alpha\beta} \\
&= 2R_{\gamma\delta\alpha\beta} - R_{\alpha\beta\gamma\delta} \\
R_{\alpha\beta\gamma\delta} &= R_{\gamma\delta\alpha\beta}. \tag{5.129}
\end{aligned}$$

Il est dès lors facile de vérifier que l'opérateur d'antisymétrisation appliqué à *n'importe quel triplet d'indices* engendre un tenseur identiquement nul. Par exemple,

$$\begin{aligned}
R_{[\alpha\beta\gamma]\delta} &= \frac{1}{3} (R_{\alpha\beta\gamma\delta} + R_{\beta\gamma\alpha\delta} + R_{\gamma\alpha\beta\delta}) \\
&= \frac{1}{3} (R_{\alpha\beta\gamma\delta} + R_{\alpha\delta\beta\gamma} + R_{\alpha\gamma\delta\beta}) \\
&= R_{\alpha[\beta\gamma\delta]} \\
&= 0.
\end{aligned}$$

5.21.2 Nombre de composantes indépendantes du tenseur de courbure

A n dimensions, combien de composantes $R_{\alpha\beta\gamma\delta}$ indépendantes subsiste-t-il? Les relations (5.128) et (5.105) ne laissent subsister que $[n(n-1)/2]^2$ quantités indépendantes, soumises aux nC_3^n contraintes (5.106). Il y a donc

$$\left[\frac{1}{2} n(n-1) \right]^2 - nC_3^n = \frac{1}{12} n^2(n^2 - 1). \tag{5.130}$$

composantes indépendantes.

5.21.3 La matrice de Petrov

Les composantes non nulles $R_{\alpha\beta\gamma\delta}$ peuvent être logées dans une matrice P_{AB} (*notation de A.Z. Petrov*), où A et B représentent respectivement les couples d'indices (α, β) , $\alpha \neq \beta$ et (γ, δ) , $\gamma \neq \delta$, sur lesquels le tenseur de Riemann est antisymétrique en vertu des relations (5.128) et (5.105). La matrice P constitue donc un tableau carré à $n(n-1)/2$ dimensions. De plus, en vertu de (5.129) elle est symétrique, ce qui ne lui laisse que

$$N = \frac{1}{2} \frac{n(n-1)}{2} \left[\frac{n(n-1)}{2} + 1 \right] \tag{5.131}$$

composantes indépendantes.

Le nombre de relations (5.106) $R_{\alpha[\beta\gamma\delta]} = 0$ qui n'ont pas encore été exploitées est facile à évaluer. En effet, les relations déjà prises en compte impliquent

$$\begin{aligned}
R_{\alpha[\beta\gamma\delta]} &= \frac{1}{4} (R_{\alpha[\beta\gamma\delta]} - R_{\beta[\alpha\gamma\delta]} + R_{\gamma[\alpha\beta\delta]} - R_{\delta[\alpha\beta\gamma]}) \\
&= R_{\alpha[\beta\gamma\delta]}.
\end{aligned}$$

Les composantes rescapées, en nombre (5.131), sont ainsi soumises aux

$$C_4^n = \frac{n(n-1)(n-2)(n-3)}{24} \quad (5.132)$$

contraintes linéaires $R_{[\alpha\beta\gamma\delta]} = 0$. On retrouve ainsi les

$$N - C_4^n = \frac{1}{12} n^2(n^2 - 1)$$

composantes $R_{\alpha\beta\gamma\delta}$ indépendantes.

5.21.4 Tenseurs associés

Le *tenseur de Ricci*, noté **Ricci** et de composantes $R_{\beta\delta}$, est la contraction du tenseur de Riemann sur les premier et troisième indice :

$$R_{\beta\delta} = R^\alpha_{\beta\alpha\delta}. \quad (5.133)$$

Symétrique, il admet $n(n+1)/2$ composantes indépendantes.

La *courbure scalaire* est la contraction du tenseur de Ricci :

$$R = R^\alpha_{\alpha}. \quad (5.134)$$

Enfin, les quantités

$$G_{\mu\nu} = R_{\mu\nu} - \frac{R}{2} g_{\mu\nu} \quad (5.135)$$

sont les composantes du *tenseur d'Einstein*, noté **G**.

5.21.5 Les identités de Bianchi contractées

Contractons les équations (5.125)

$$(\nabla_\mu \mathbf{R})^\alpha_{\beta\gamma\delta} + (\nabla_\gamma \mathbf{R})^\alpha_{\beta\delta\mu} + (\nabla_\delta \mathbf{R})^\alpha_{\beta\mu\gamma} = 0$$

sur les indices α et γ . On obtient

$$(\nabla_\mu \mathbf{R})^\alpha_{\beta\alpha\delta} + (\nabla_\alpha \mathbf{R})^\alpha_{\beta\delta\mu} + (\nabla_\delta \mathbf{R})^\alpha_{\beta\mu\alpha} = 0.$$

Or, l'opérateur de dérivation covariante commute avec la contraction (voir p. 87). L'égalité précédente peut donc s'écrire

$$(\nabla_\mu \mathbf{Ricci})_{\beta\delta} + (\nabla_\alpha \mathbf{R})^\alpha_{\beta\delta\mu} - (\nabla_\delta \mathbf{Ricci})_{\beta\mu} = 0.$$

Multipliant maintenant par $g^{\beta\delta}$, on obtient

$$\nabla_\mu R - (\nabla_\alpha \mathbf{Ricci})^\alpha_{\mu} - (\nabla_\delta \mathbf{Ricci})^\delta_{\mu} = 0.$$

Ce résultat peut s'écrire plus simplement sous les formes successives

$$\begin{aligned} (\nabla_\alpha \mathbf{Ricci})^\alpha_{\mu} - \frac{1}{2} \nabla_\mu R &= 0 \\ (\nabla_\alpha \mathbf{Ricci})^{\alpha\nu} - \frac{1}{2} g^{\mu\nu} \nabla_\mu R &= 0 \\ \left[\nabla_\mu \left(\mathbf{Ricci} - \frac{R}{2} \mathbf{g} \right) \right]^{\mu\nu} &= 0 \\ (\nabla_\mu \mathbf{G})^{\mu\nu} &= 0. \end{aligned}$$

Nous avons démontré l'annulation de la divergence covariante du tenseur d'Einstein :

$$\left(R^{\mu\nu} - \frac{R}{2} g^{\mu\nu} \right)_{;\nu} = 0. \quad (5.136)$$

5.21.6 Lien avec la courbure de Gauss

Voici la relation entre le tenseur de Riemann-Christoffel et la courbure de Gauss.¹³ Soit deux vecteurs linéairement indépendants \vec{a} et \vec{b} définis en un point P d'une variété riemannienne à n dimensions dotée d'une structure métrique $g_{\mu\nu}$. La surface définie en ce point par toutes les géodésiques de direction $\alpha\vec{a} + \beta\vec{b}$ possède une courbure gaussienne donnée par

$$K(\vec{a}, \vec{b}, P) = \frac{R_{\lambda\mu\nu\xi}(P) a^\lambda b^\mu a^\nu b^\xi}{[g_{\lambda\nu}(P)g_{\mu\xi}(P) - g_{\lambda\xi}(P)g_{\mu\nu}(P)] a^\lambda b^\mu a^\nu b^\xi}.$$

Le tenseur de courbure $R_{\lambda\mu\nu\xi}$ possède les mêmes symétries que l'expression $g_{\lambda\nu}g_{\mu\xi} - g_{\lambda\xi}g_{\mu\nu}$ et admet donc le même nombre (5.130) de composantes indépendantes.¹⁴

Pour $n = 3$, la connaissance du tenseur de Riemann est équivalente à celle du tenseur de Ricci car ils possèdent tous les deux le même nombre de composantes indépendantes, 6.

Lorsque $n = 2$, l'information contenue dans l'unique composante non nulle se trouve essentiellement dans la courbure scalaire, reliée très simplement à la courbure de Gauss K par la relation

$$R = 2K.$$

En effet, dans l'espace euclidien à 3 dimensions, une surface peut être approximée, jusqu'au deuxième ordre inclus, par la quadrique d'équation

$$z = \frac{1}{2}(K_1 x^2 + K_2 y^2)$$

où K_1 et K_2 sont les deux courbures principales (chapitre 4). On peut donc décrire localement la surface à l'aide de la métrique

$$ds^2 = [1 + (K_1)^2 x^2] dx^2 + [1 + (K_2)^2 y^2] dy^2 + 2K_1 K_2 xy dx dy.$$

Au voisinage de l'origine $x = y = 0$, on calcule

$$(g_{\mu\nu}) = \begin{pmatrix} 1 + K_1^2 x^2 & K_1 K_2 xy \\ K_1 K_2 xy & 1 + K_2^2 y^2 \end{pmatrix},$$

$$(g^{\mu\nu}) = \begin{pmatrix} 1 - K_1^2 x^2 & -K_1 K_2 xy \\ -K_1 K_2 xy & 1 - K_2^2 y^2 \end{pmatrix} + O(4),$$

$$\Gamma_{11}^1 = K_1^2 x + O(3), \quad \Gamma_{12}^1 = 0, \quad \Gamma_{22}^1 = K_1 K_2 x + O(3),$$

$$\Gamma_{11}^2 = K_1 K_2 y + O(3), \quad \Gamma_{12}^2 = 0, \quad \Gamma_{22}^2 = K_2^2 y + O(3),$$

$$R_{212}^1 = K_1 K_2 + O(2) = K + O(2),$$

$$R_{121}^2 = K + O(2),$$

$$R_{11} = K + O(2), \quad R_{12} = O(2), \quad R_{22} = K + O(2),$$

$$R = 2K + O(2).$$

13. Voir par exemple S. Weinberg, *Gravitation and Cosmology*, John Wiley & Sons, 1972, pp. 147-148.

14. A partir des n vecteurs linéairement indépendants en P , on ne peut construire que $n(n-1)/2$ surfaces géodésiques. L'annulation de la courbure gaussienne de chacune de ces surfaces ne suffit donc pas à rendre l'espace euclidien si $n \geq 3$.

5.22 Exercices

Exercice 5.8 Dans l'espace euclidien à trois dimensions, considérons les coordonnées du parabolöide de révolution (ξ, η, φ) reliées aux coordonnées cartésiennes (x, y, z) par les relations

$$\begin{aligned}x &= \xi \eta \cos \varphi \\y &= \xi \eta \sin \varphi \\z &= \frac{\xi^2 - \eta^2}{2}.\end{aligned}$$

Appelons $(\vec{e}_\xi, \vec{e}_\eta, \vec{e}_\varphi)$ les vecteurs de base des coordonnées du parabolöide de révolution. Sur la surface $\xi = \xi_0$, soit le vecteur $\vec{W} = \vec{e}_\eta$ défini au point P de coordonnées (η_0, φ_0) , et le vecteur \vec{W}' résultant du transport parallèle de \vec{W} jusqu'au point P' infiniment proche de coordonnées $(\eta_0 + d\eta, \varphi_0)$.

- Représenter la situation sur un croquis.
- Expliquer la méthode à utiliser pour trouver \vec{W}' .
- Calculer les symboles de Christoffel nécessaires.
- En déduire \vec{W}' .

Solution. La surface $\xi = \xi_0$ a pour équation cartésienne

$$2z = \xi_0^2 \left(1 - \frac{x^2 + y^2}{\xi_0^4} \right).$$

Il s'agit d'un parabolöide de révolution autour de l'axe des z , tourné vers le bas et dont le sommet se trouve en $z = \xi_0^2/2$. La surface de coordonnée η_0 est également un parabolöide de révolution, mais tourné vers le haut et de sommet situé en $z = -\eta_0^2/2$. La coordonnée φ a la signification de l'angle polaire habituel.

Appelons \vec{V} le champ vectoriel transporté parallèlement :

$$\vec{V}(P) = \vec{W}, \quad \vec{V}(P)_{\parallel \rightarrow P'} = \vec{W}'.$$

On a (attention, pour la commodité de l'écriture les indices η et φ représentent les composantes selon les coordonnées correspondantes)

$$\begin{aligned}0 &= (\nabla_\eta \vec{V})^\alpha \\ &= V^\alpha_{,\eta} + \Gamma_{\beta\eta}^\alpha V^\beta\end{aligned}$$

donc

$$V^\eta_{,\eta} = -\Gamma_{\eta\eta}^\eta \quad \text{et} \quad V^\varphi_{,\eta} = -\Gamma_{\eta\eta}^\varphi.$$

Sur le parabolöide ξ_0 , on calcule

$$\begin{aligned}ds^2 &= (\xi_0^2 + \eta^2) d\eta^2 + \xi_0^2 \eta^2 d\varphi^2, \\ \Gamma_{\eta\eta}^\eta &= \frac{\eta}{\xi_0^2 + \eta^2}, \quad \Gamma_{\eta\eta}^\varphi = 0.\end{aligned}$$

Par conséquent,

$$\begin{aligned}(\vec{W}')^\eta &= 1 - (\Gamma_{\eta\eta}^\eta)_P d\eta = 1 - \frac{\eta_0}{\xi_0^2 + \eta_0^2} d\eta \\ (\vec{W}')^\varphi &= -(\Gamma_{\eta\eta}^\varphi)_P d\eta = 0.\end{aligned}$$

Exercice 5.9 (le pendule de Foucault) *Sur la sphère de métrique*

$$ds^2 = \mathcal{R}^2 (d\theta^2 + \sin^2 \theta d\varphi^2),$$

étudier le transport parallèle d'un vecteur \vec{S} le long du cercle d'équations paramétriques

$$\theta(s) = \theta_0, \quad \varphi(s) = s/(\mathcal{R} \sin \theta_0),$$

où s désigne la longueur d'arc et où θ_0 est une constante.

Solution. On a

$$\begin{aligned} 0 &= \nabla_{\vec{U}} \vec{S} \\ &= U^\theta \nabla_\theta \vec{S} + U^\varphi \nabla_\varphi \vec{S}, \end{aligned}$$

avec

$$U^\theta = 0, \quad U^\varphi = 1/(\mathcal{R} \sin \theta_0).$$

Les seuls symboles de Christoffel utiles et non nuls sont

$$\Gamma_{\varphi\varphi}^\theta = -\sin \theta \cos \theta, \quad \Gamma_{\theta\varphi}^\varphi = \cotg \theta.$$

On en arrive ainsi aux deux équations différentielles

$$\begin{aligned} 0 &= S^\theta_{,\varphi} - \sin \theta_0 \cos \theta_0 S^\varphi \\ 0 &= S^\varphi_{,\varphi} + \cotg \theta_0 S^\theta. \end{aligned}$$

Par conséquent,

$$S^\theta_{,\varphi\varphi} + \cos^2 \theta_0 S^\theta = 0.$$

Introduisant la latitude $\lambda \in]-\pi/2, \pi/2[$ par la relation

$$\cos \theta_0 = \sin \lambda$$

(impliquant $\sin \theta_0 = \cos \lambda$), on obtient la solution

$$S^\theta(\varphi) = A \sin(\pm\varphi \sin \lambda + B), \quad S^\varphi(\varphi) = \frac{\pm A}{\cos \lambda} \cos(\pm\varphi \sin \lambda + B)$$

où les signes positif et négatif sont relatifs à l'hémisphère boréal et à l'hémisphère austral, respectivement. On vérifie facilement que la norme de $\vec{S}(\varphi)$ est préservée tout au long du transport, comme il se doit, et vaut $\mathcal{R}|A|$. Si l'on suppose que \vec{S} est unitaire et que $\vec{S}(0)$ pointe dans le sens de \vec{e}_θ , alors $A = 1/\mathcal{R}$ et $B = \pi/2$. En définitive,

$$S^\theta(\varphi) = \frac{1}{\mathcal{R}} \cos(\varphi \sin \lambda), \quad S^\varphi(\varphi) = \frac{-1}{\mathcal{R} \cos \lambda} \sin(\varphi \sin \lambda).$$

Le vecteur \vec{S} précesse ainsi, dans la base locale $(\vec{e}_\theta, \vec{e}_\varphi)$, d'un angle $\alpha(\varphi)$ donné par

$$\cos \alpha(\varphi) = \frac{\vec{S}(\varphi) \cdot \vec{e}_\theta}{|\vec{S}||\vec{e}_\theta|} = \cos(\varphi \sin \lambda) \quad \text{et} \quad \sin \alpha(\varphi) = \frac{\vec{S}(\varphi) \cdot \vec{e}_\varphi}{|\vec{S}||\vec{e}_\varphi|} = -\sin(\varphi \sin \lambda),$$

donc $\alpha(\varphi) = -\varphi \sin \lambda$. Par conséquent, au terme de chaque révolution il a tourné d'un angle

$$\alpha(2\pi) = -2\pi \sin \lambda.$$

Exercice 5.10 Soit la sphère S^2 de rayon \mathcal{R} . Calculer son aire, les tenseurs de Riemann, de Ricci et la courbure scalaire.

Solution. Le passage aux coordonnées sphériques (r, θ, φ) fournit immédiatement

$$ds^2 = \mathcal{R}^2(d\theta^2 + \sin^2 \theta d\varphi^2).$$

Dans un espace riemannien de dimension 2 rapporté à des coordonnées générales (x^1, x^2) et à métrique (g_{ij}) définie positive, comme c'est le cas ici, la mesure d'intégration est donnée par $\sqrt{g} dx^1 dx^2$. En effet, le jacobien du changement de variables permettant de passer des coordonnées cartésiennes locales (dans lesquelles le tenseur métrique est représenté par la matrice identité) aux coordonnées générales est égal à \sqrt{g} .

On a ici

$$\sqrt{g} = \mathcal{R}^2 \sin \theta,$$

d'où l'aire de la sphère S^2 :

$$A = \int_0^\pi d\theta \int_0^{2\pi} \sqrt{g} d\varphi = 4\pi \mathcal{R}^2.$$

Posons $\theta = x^1$, $\varphi = x^2$. Les seuls symboles de Christoffel non nuls sont

$$\Gamma_{22}^1 = -\sin \theta \cos \theta, \quad \Gamma_{12}^2 = \Gamma_{21}^2 = \cotg \theta.$$

Comme l'espace est à deux dimensions, il ne subsiste qu'une composante indépendante pour le tenseur de courbure, mettons R_{1212} . On obtient

$$\begin{aligned} R_{1212} &= \mathcal{R}^2 \sin^2 \theta \\ R_{11} &= 1, \quad R_{12} = 0, \quad R_{22} = \sin^2 \theta. \end{aligned}$$

Remarquons que

$$R_{\alpha\beta\gamma\delta} = \frac{1}{\mathcal{R}^2} (g_{\alpha\gamma} g_{\beta\delta} - g_{\alpha\delta} g_{\beta\gamma}), \quad R_{\alpha\beta} = \frac{1}{\mathcal{R}^2} g_{\alpha\beta}, \quad R = \frac{2}{\mathcal{R}^2}. \quad (5.137)$$

On résume les relations (5.137) en disant que la sphère est un *espace à symétrie maximale de courbure positive*. En dimension n , les espaces à symétrie maximale admettent un groupe d'isométries à $n(n+1)/2$ paramètres (donc trois paramètres pour des espaces à deux dimensions). Leur courbure scalaire est constante.

Exercice 5.11 Soit l'hyperboloïde à deux nappes de dimension 2, d'équation

$$x^2 + y^2 - z^2 = -\mathcal{R}^2.$$

On peut également définir cette surface à l'aide des équations paramétriques

$$x = \mathcal{R} \sinh \theta \cos \varphi, \quad y = \mathcal{R} \sinh \theta \sin \varphi, \quad z = \mathcal{R} \cosh \theta.$$

Si on l'immerge dans un espace pseudo-euclidien de métrique

$$ds^2 = dx^2 + dy^2 - dz^2,$$

on en déduit l'expression de son élément de longueur :

$$ds^2 = \mathcal{R}^2 (d\theta^2 + \sinh^2 \theta d\varphi^2).$$

Montrer que cette surface est un espace à symétrie maximale de courbure négative :

$$R_{\alpha\beta\gamma\delta} = \frac{-1}{\mathcal{R}^2} (g_{\alpha\gamma} g_{\beta\delta} - g_{\alpha\delta} g_{\beta\gamma}), \quad R_{\alpha\beta} = \frac{-1}{\mathcal{R}^2} g_{\alpha\beta}, \quad R = \frac{-2}{\mathcal{R}^2}. \quad (5.138)$$

Exercice 5.12 *Montrer que la pseudosphère de Beltrami est également un espace à symétrie maximale de courbure négative.*

Solution. Rappelons l'expression (4.30) de l'élément de longueur sur cette surface :

$$ds^2 = \frac{\mathcal{R}^2}{r^2} dr^2 + r^2 d\varphi^2.$$

Un simple calcul permet d'en déduire les relations (5.138).

Pour des \mathcal{R} identiques, la courbure scalaire de la pseudosphère est la même que celle de l'hyperboloïde de l'exercice 5.11. Or, on peut démontrer que les espaces à symétrie maximale sont uniques pour une courbure scalaire donnée, c'est-à-dire que deux espaces à symétrie maximale de même courbure scalaire peuvent être reliés par un changement de coordonnées transformant la métrique du premier en la métrique du second¹⁵.

Exercice 5.13 *Soit le tore à deux dimensions, de métrique (5.83) :*

$$ds^2 = (a + b \cos \psi)^2 d\varphi^2 + b^2 d\psi^2.$$

Calculer son aire. Exprimer ensuite les tenseurs de Riemann, de Ricci et la courbure scalaire sur cette surface.

Solution. De l'information $\sqrt{g} = b(a + b \cos \psi)$ on déduit l'aire du tore :

$$A = \int_0^{2\pi} d\varphi \int_0^{2\pi} \sqrt{g} d\psi = 4\pi^2 ab.$$

Posons ensuite $\varphi = x^1$, $\psi = x^2$. Les seuls symboles de Christoffel non nuls sont

$$\Gamma_{12}^1 = \Gamma_{21}^1 = \frac{-b \sin \psi}{a + b \cos \psi}, \quad \Gamma_{11}^2 = \frac{\sin \psi (a + b \cos \psi)}{b}.$$

Ainsi,

$$\begin{aligned} R_{1212} &= b \cos \psi (a + b \cos \psi) \\ R_{11} &= \frac{\cos \psi (a + b \cos \psi)}{b}, \quad R_{12} = 0, \quad R_{22} = \frac{b \cos \psi}{a + b \cos \psi} \\ R &= \frac{2 \cos \psi}{b(a + b \cos \psi)}. \end{aligned}$$

Exercice 5.14 *Montrer que, dans tout espace riemannien de dimension 2, on a*

$$R_{\alpha\beta\gamma\delta} = (g_{\alpha\gamma} g_{\beta\delta} - g_{\alpha\delta} g_{\beta\gamma}) \frac{R_{1212}}{g}, \quad R_{\alpha\beta} = g_{\alpha\beta} \frac{R_{1212}}{g}, \quad R = 2 \frac{R_{1212}}{g}.$$

Le tenseur d'Einstein (5.135) s'y annule donc identiquement. On en déduit également qu'un espace de dimension 2 est à symétrie maximale si et seulement si sa courbure scalaire est constante.

15. Voir par exemple S. Weinberg, *Gravitation and Cosmology*, John Wiley & Sons, 1972, p. 381. Plus précisément, le théorème demande également que le nombre de valeurs propres du tenseur métrique de chaque signe soit identique pour les deux métriques. En l'occurrence, c'est bien le cas pour les métriques des exercices 5.11 et 5.12, lesquelles sont définies positives. Le changement de coordonnées qui les relie est donné dans la section 4.12.2.

Chapitre 6

Le tenseur d'énergie-impulsion

6.1 Le statut de l'énergie volumique en relativité restreinte

Il est raisonnable d'attendre qu'une théorie relativiste de la gravitation se ramène, comme cas limite, à l'équation fondamentale de la théorie newtonienne : l'équation de Poisson $\nabla^2\Phi = 4\pi G\mu$, dans laquelle Φ désigne le potentiel gravitationnel, G la constante de Cavendish et μ la masse volumique. Or, lors d'un changement de référentiel, cette dernière se transforme comme un scalaire puisque la masse et le volume spatial sont des invariants galiléens.

Quelles nouveautés l'invariance lorentzienne introduit-elle ? Nous savons déjà qu'il faut abandonner le concept de masse au profit de celui d'énergie. En relativité restreinte l'énergie constitue la composante temporelle d'un quadrivecteur, ce qui clarifie considérablement son statut. Cependant, le volume spatial n'est plus un invariant scalaire. Pour fixer les idées, raisonnons sur le cas particulier d'un fluide de poussière dont toutes les particules possèdent la même masse propre m_0 . Leur énergie volumique dans le référentiel inertiel comobile vaut $n_0 m_0 c^2$, où n_0 est le nombre de particules par unité de volume spatial propre. Dans un repère inertiel quelconque de vitesse spatiale V par rapport au référentiel comobile, l'énergie volumique est donnée par $n m c^2$, où

$$n = \frac{n_0}{\sqrt{1 - V^2/c^2}} \quad \text{et} \quad m = \frac{m_0}{\sqrt{1 - V^2/c^2}}.$$

L'énergie volumique se transforme ainsi en faisant apparaître l'élément de matrice jacobienne $\sqrt{1 - V^2/c^2}$ non pas *une*, mais *deux* fois. On peut d'ores et déjà y détecter la signature caractéristique d'un *tenseur d'ordre deux*. Dans ce chapitre, nous allons montrer d'une façon générale qu'effectivement, l'énergie volumique représente la composante deux fois temporelle d'un tenseur à deux indices, le *tenseur d'énergie-impulsion*.

6.2 Définitions

Dans l'espace-temps de Minkowski rapporté à des coordonnées cartésiennes $\{x^\alpha\} = \{ct, x, y, z\}$, on définit le *gradient* d'une famille d'hypersurfaces d'équation $\phi = \text{constante}$, par la relation

$$\tilde{d}\phi = \frac{\partial\phi}{\partial x^\mu} \tilde{d}x^\mu.$$

Notant $(\eta^{\mu\nu})$ l'inverse du tenseur métrique $(\eta_{\mu\nu}) = \text{diag}(-1, 1, 1, 1)$, on dit que l'une de ces hypersurfaces est localement

- spatiale si

$$\eta^{\mu\nu} \frac{\partial\phi}{\partial x^\mu} \frac{\partial\phi}{\partial x^\nu} < 0;$$

- temporelle si

$$\eta^{\mu\nu} \frac{\partial\phi}{\partial x^\mu} \frac{\partial\phi}{\partial x^\nu} > 0.$$

On appelle *norme* du gradient la quantité

$$|\tilde{d}\phi| = \sqrt{\left| \eta^{\mu\nu} \frac{\partial\phi}{\partial x^\mu} \frac{\partial\phi}{\partial x^\nu} \right|}.$$

Si $|\tilde{d}\phi| \neq 0$, la *normale unitaire* à l'hypersurface est la forme définie par

$$\tilde{n} = \tilde{d}\phi / |\tilde{d}\phi|.$$

En un point donné, le *flux* d'une grandeur physique au travers de l'hypersurface spatiale d'équation $\phi = ct$ est alors défini comme la quantité de cette grandeur par unité de volume spatial. Le flux d'une grandeur physique au travers de l'hypersurface temporelle d'équation $\phi = x^i$ est la quantité de cette grandeur traversant, dans le sens des x^i croissants, l'unité de surface spatiale orthogonale à x^i par unité de ct . Les grandeurs physiques utilisées dans ce contexte seront l'énergie et la quantité de mouvement transportées par les particules matérielles.¹

Définissons à présent des quantités $T^{\alpha\beta}$ égales respectivement au flux d'énergie (si $\alpha = 0$) ou au flux de c fois la i ème composante de la quantité de mouvement (si $\alpha = i$), au travers de l'hypersurface $x^\beta = \text{constante}$. T^{00} représente donc l'énergie volumique; T^{i0} vaut c fois la i ème composante de la quantité de mouvement volumique; T^{0j} est la quantité d'énergie s'écoulant, par unité de ct , au travers de l'unité de surface $x^j = k$, où k est une constante, dans le sens des x^j croissants; et T^{ij} désigne la quantité de c fois la i ème composante de la quantité de mouvement s'écoulant, par unité de ct , au travers de l'unité de surface $x^j = k$ dans le sens des x^j croissants, c'est-à-dire la i ème composante de la force exercée sur l'unité de surface $x^j = k$, par le milieu « arrière » ($x^j < k$) sur le milieu « avant » ($x^j > k$).

Il est facile de généraliser ce discours au cas d'une hypersurface Φ quelconque, spatiale ou temporelle, de normale unitaire $\tilde{n} = n_\alpha \tilde{d}x^\alpha$ et dont l'hyperplan tangent au point considéré admet pour équation $n_\alpha x^\alpha = \text{constante}$: le flux d'une grandeur est la quantité de cette grandeur par unité de volume dans Φ et qui s'écoule « dans le sens de \tilde{n} », c'est-à-dire dans le sens des ϕ croissants. Comme le montre la figure 6.1, les flux $(F_\Phi)^\alpha$ d'énergie ($\alpha = 0$) ou de c fois la quantité de mouvement ($\alpha = i$) au travers de Φ sont alors donnés par²

$$(F_\Phi)^\alpha = T^{\alpha\beta} n_\beta. \quad (6.1)$$

1. A ce stade, nous n'envisageons pas l'énergie et la quantité de mouvement associées aux champs classiques. Le cas du champ électromagnétique sera traité dans la section 6.6. Concernant le champ gravitationnel, il est toujours possible de l'annuler par un changement de référentiel, en vertu du principe d'équivalence : on peut donc prédire qu'il serait incohérent de tenter de décrire son énergie et sa quantité de mouvement à l'aide d'un tenseur. Nous verrons comment l'adoption d'un espace-temps courbe permet de trancher la question.

2. L'égalité (6.1) permet notamment d'exprimer qu'un flux dans le sens des x^i décroissants est l'opposé dudit flux dans le sens des x^i croissants. En particulier, la force exercée par le milieu « avant » sur le milieu « arrière » est l'opposée de la force exercée par le milieu « arrière » sur le milieu « avant ».

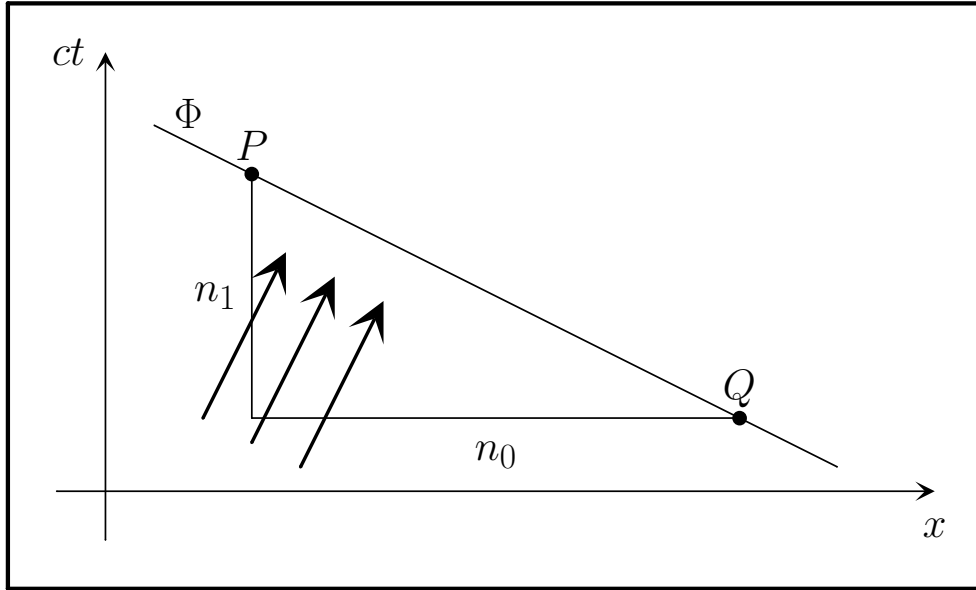


FIGURE 6.1: Flux au travers d'une hypersurface spatiale $\Phi \equiv n_0 ct + n_1 x = \text{constante}$ ($n_0 > n_1 > 0$). Le segment PQ , de longueur $\sqrt{(n_0)^2 - (n_1)^2} = 1$, détermine une des arêtes d'un cube unitaire (les deux autres arêtes, de longueur $\Delta y = \Delta z = 1$, sont orthogonales au plan de la page). Les lignes fléchées illustrent les trajectoires spatio-temporelles des particules matérielles véhiculant de l'énergie et de la quantité de mouvement. Le flux d'une quantité au travers de Φ est ainsi associé au nombre de lignes percutant l'unité d'hypersurface oblique. On peut l'interpréter comme la somme de deux termes : le flux au travers de l'hypersurface spatiale (horizontale) de côtés $\Delta x = n_0, \Delta y = \Delta z = 1$, plus le flux au travers de l'hypersurface temporelle (verticale) de côtés $\Delta(ct) = n_1, \Delta y = \Delta z = 1$. Ainsi, $(F_\Phi)^\alpha = T^{\alpha 0} n_0 + T^{\alpha 1} n_1$. Pour une hypersurface quelconque, on obtiendrait $(F_\Phi)^\alpha = T^{\alpha\beta} n_\beta$.

Cette relation révèle en outre que les $T^{\alpha\beta}$ se transforment comme les composantes d'un tenseur d'ordre deux. En effet, le membre de gauche constitue, par définition, un quadrivecteur énergie-quantité de mouvement ; et à droite, les n_β caractérisent une forme quelconque. On l'appelle *tenseur d'énergie-impulsion* \mathbf{T} . Le caractère tensoriel résulte du fait que la définition générale des flux est purement géométrique, c'est-à-dire indépendante des coordonnées. On observera les deux propriétés suivantes :

- La « colonne » β n'est autre que le vecteur $\vec{F}_\Phi = \mathbf{T}(\cdot, \tilde{d}x^\beta)$ où $\Phi \equiv x^\beta = \text{constante}$. D'une manière générale, le vecteur flux au travers d'une hypersurface Φ quelconque de normale unitaire \tilde{n} peut s'écrire sous la forme $\vec{F}_\Phi = \mathbf{T}(\cdot, \tilde{n})$.
- La « ligne » α constitue également un vecteur, à savoir $\mathbf{T}(\tilde{d}x^\alpha, \cdot)$.

6.3 Le tenseur d'énergie-impulsion des fluides

6.3.1 Symétrie du tenseur d'énergie-impulsion

Montrons tout d'abord que $T^{i0} = T^{0i}$. En effet, T^{i0} est égal à c fois la i ème composante de la quantité de mouvement volumique, donc c fois la masse volumique fois la i ème composante de la vitesse, donc l'énergie volumique fois la i ème composante de la vitesse

divisée par c , donc la quantité d'énergie traversant l'unité de surface spatiale x^i par unité de c .temps, c'est-à-dire T^{0i} .

Pour démontrer la symétrie de la partie spatiale du tenseur d'énergie-impulsion, on peut procéder comme suit³. Considérons un élément infinitésimal de fluide, de forme cubique. Des axes émanant de son centre O , orthogonaux deux à deux et en translation, définissent un référentiel non inertiel dans lequel le cube possède un point fixe. Nous allons supposer que le mouvement du cube autour de son centre de masse O est régi par les équations d'Euler, exprimant la conservation du moment cinétique d'un solide passant par un point fixe. Dans le système des axes principaux d'inertie (d'origine O et perpendiculaires aux différentes faces), elles s'écrivent

$$I_1 \dot{\omega}_1 = (I_2 - I_3) \omega_2 \omega_3 + \mathcal{M}_1 \quad (6.2)$$

$$I_2 \dot{\omega}_2 = (I_3 - I_1) \omega_3 \omega_1 + \mathcal{M}_2 \quad (6.3)$$

$$I_3 \dot{\omega}_3 = (I_1 - I_2) \omega_1 \omega_2 + \mathcal{M}_3 \quad (6.4)$$

Dans ces relations, $\vec{\omega}$ désigne le vecteur instantané de rotation ; $\vec{\mathcal{M}} = \oint_S \overrightarrow{OP} \wedge \vec{f} dS$ est le moment par rapport à O des forces extérieures agissant sur le solide (seules interviennent les forces surfaciques car, d'une part, nous négligeons les forces volumiques et, d'autre part, le moment de la force fictive d'entraînement est nul) ; les quantités I_1, I_2, I_3 , représentent les moments principaux d'inertie, par exemple, si μ est la masse volumique :

$$I_3 = \int_V \mu(\overrightarrow{OP})^2 \text{distance}^2(P, \text{axe des } z) dV .$$

Si V et V' désignent respectivement un cube de côté l et un cube de côté 1, on a

$$I_1 = I_2 = I_3 = l^5 \int_{V'} \mu(\overrightarrow{OP'})^2 \text{distance}^2(P', \text{axe des } z) dV' = \frac{1}{6} \mu(O) l^5 + O(l^5) = \alpha l^5$$

où α est une constante. Les équations (6.2)-(6.4) prennent ainsi la forme

$$\dot{\vec{\omega}} = \frac{1}{\alpha l^5} \vec{\mathcal{M}} . \quad (6.5)$$

D'autre part, si les nombres n_l désignent les composantes de la normale unitaire (à trois dimensions) à la surface du cube et dirigée vers l'extérieur de ce dernier, on a⁴

$$\begin{aligned} \mathcal{M}_k &= \left[\oint_S \overrightarrow{OP} \wedge \vec{f} dS \right]_k = \oint_S \varepsilon_{ijk} x_i f_j dS \\ &= - \oint_S \varepsilon_{ijk} x_i T_{jl} n_l dS \\ &= - \int_V \varepsilon_{ijk} (x_i T_{jl})_{,l} dV \\ &= - \int_V \varepsilon_{ijk} (\delta_{il} T_{jl} + x_i T_{jl,l}) dV \\ &= - \int_V \varepsilon_{ijk} (T_{ji} + x_i T_{jl,l}) dV . \end{aligned}$$

3. Cette démonstration, inspirée de B. Schutz, *A first course in general relativity*, Cambridge University Press, 2009, chapitre 4, présente l'avantage de rester intuitive.

4. Ici comme dans la suite, lorsque nous raisonnerons en coordonnées cartésiennes nous n'accorderons guère d'importance à la position des indices spatiaux. Sauf mention explicite du contraire, la convention de sommation d'Einstein sera supposée d'application pour tout indice répété, quelle que soit sa variance.

Par exemple,

$$\mathcal{M}_3 = \int_V (T_{12} - T_{21}) \, dV + \int_V (y T_{1l,l} - x T_{2l,l}) \, dV. \quad (6.6)$$

Développons T_{ij} en série de puissances :

$$T_{ij}(x, y, z) = T_{ij}(O) + x a_{ij} + y b_{ij} + z c_{ij} + x^2 d_{ij} + y^2 e_{ij} + z^2 f_{ij} + xy g_{ij} + xz h_{ij} + yz k_{ij} + O(3).$$

Ainsi,

$$\mathcal{M}_3 = [T_{12}(O) - T_{21}(O)] l^3 + O(l^5). \quad (6.7)$$

Pour fixer les idées, examinons la troisième composante de l'équation (6.5) :

$$\dot{\omega}_3 = \frac{1}{\alpha l^2} [T_{12}(O) - T_{21}(O)] + O(l^0). \quad (6.8)$$

Si $T_{12}(O)$ était différent de $T_{21}(O)$, l'accélération de rotation d'un élément du fluide croîtrait sans borne lorsque la dimension caractéristique de ce dernier tend vers zéro. L'inspection des deux premières composantes de l'équation (6.5) mène à des conclusions analogues si $T_{23}(O) \neq T_{32}(O)$ ou $T_{13}(O) \neq T_{31}(O)$. Un tel résultat étant inadmissible, on en déduit la symétrie de la partie spatiale du tenseur d'énergie-impulsion au point O dans le référentiel des axes principaux d'inertie, donc en tout point et dans tout système de référence.

6.3.2 Conservation de l'énergie et de la quantité de mouvement

L'utilisation du tenseur d'énergie-impulsion permet une formulation très naturelle des principes de conservation de l'énergie et de la quantité de mouvement.

Une modification de la quantité d'énergie ou de mouvement contenue dans un volume spatial V ne peut résulter que du flux de l'entité correspondante au travers des parois S : en adoptant une normale unitaire qui pointe constamment vers l'extérieur, on peut écrire

$$\frac{\partial}{c \partial t} \int_V T^{\mu 0}(ct, x, y, z) \, dV + \oint_S T^{\mu j}(ct, x, y, z) n_j \, dS = 0, \quad (6.9)$$

où $\mu = 0$ si l'on traite l'énergie et $\mu = 1, 2$ ou 3 si l'on s'intéresse aux diverses composantes de la quantité de mouvement. Or, par le théorème de Gauss,

$$\oint_S T^{\mu j}(ct, x, y, z) n_j \, dS = \int_V \frac{\partial T^{\mu j}(ct, x, y, z)}{\partial x^j} \, dV.$$

L'équation (6.9) devient donc

$$\int_V \frac{\partial T^{\mu \nu}(ct, x, y, z)}{\partial x^\nu} \, dV = 0$$

quel que soit le volume V . Il s'ensuit, en tout point,

$$\boxed{\frac{\partial T^{\mu \nu}}{\partial x^\nu} = 0.} \quad (6.10)$$

Lue dans l'espace-temps, la relation (6.9) exprime en fait que

$$\oint_S T^{\mu \nu}(ct, x, y, z) n_\nu \, dS = 0,$$

où \mathcal{S} désigne la frontière de l'hypervolume correspondant à l'évolution du volume tridimensionnel V entre deux hypersurfaces de coordonnées respectives x^0 et $x^0 + dx^0$.

6.3.3 Conservation du nombre de particules

Si n_0 représente le nombre de particules par unité de volume dans le référentiel localement comobile et si les U^μ désignent les composantes de la quadrivitesse de ce dernier par rapport à un référentiel inertiel arbitraire, on définit le quadrivecteur *flux de particules* par la relation

$$N^\mu = \frac{n_0}{c} U^\mu.$$

Si l'on appelle \vec{V} la tri-vitesse du référentiel localement comobile, la quantité

$$N^0 = \frac{n_0}{\sqrt{1 - V^2/c^2}}$$

est le nombre de particules par unité de volume, mesuré dans le référentiel arbitraire, tandis que chacun des

$$N^i = \frac{n_0}{\sqrt{1 - V^2/c^2}} \frac{V^i}{c}$$

désigne le nombre de particules traversant l'unité de surface orthogonale à l'axe x^i par unité de c . temps. Or, dans un grand nombre de situations physiques il n'y a ni création, ni annihilation de particules. Le raisonnement de la section précédente peut alors être repris et fournit la relation

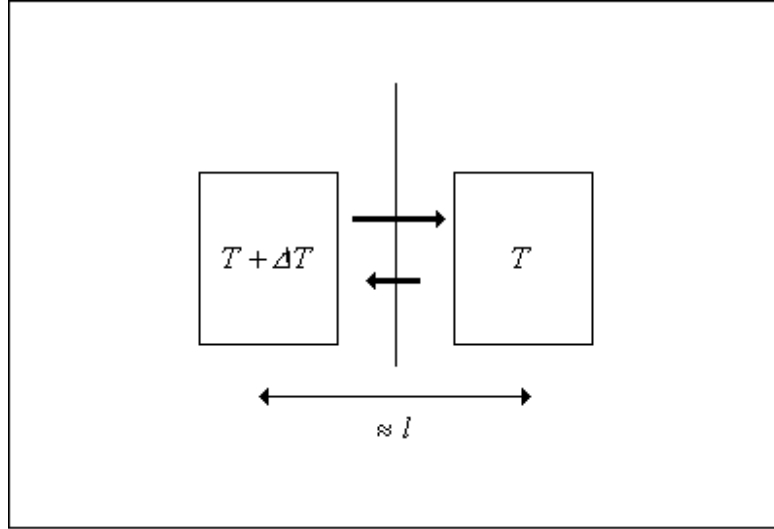
$$\frac{\partial(n_0 U^\nu)}{\partial x^\nu} = 0. \quad (6.11)$$

6.4 Les fluides parfaits

6.4.1 Expression du tenseur d'énergie-impulsion

Soit un fluide en équilibre thermodynamique local : par exemple, dans chaque élément du système les particules vérifient la statistique de Maxwell-Boltzmann pour une certaine température. Si celle-ci varie de point en point (donc en l'absence d'équilibre thermique global), il se produit un *transport d'énergie* appelé *conduction thermique*. Ce phénomène n'est compréhensible que si, dans le modèle adopté pour décrire le fluide, les particules matérielles, dotées d'une certaine extension spatiale, peuvent entrer en collision les unes avec les autres ; en d'autres termes, si leur libre parcours moyen l n'est pas infini.

Voici, succinctement, l'explication du phénomène. Raisonnons dans le référentiel comobile où, par définition, la vitesse vectorielle moyenne des particules matérielles est égale à zéro. Imaginons une paroi virtuelle séparant deux éléments voisins du fluide, immobiles l'un par rapport à l'autre et de températures respectives $T + \Delta T$ et T , avec $0 < \Delta T \approx |\vec{\nabla} T| l \ll T$ (figure 6.2). Le flux moyen de particules traversant la paroi est identique (en valeur absolue) dans les deux sens, autrement dit le flux moyen total est nul. Cependant, une particule traversant la paroi depuis la gauche vers la droite aura, statistiquement, une énergie plus importante que les particules de l'élément où elle arrive et qu'elle va donc « réchauffer » par collisions. Inversement, une particule venant de la droite aura tendance à « refroidir » son milieu d'accueil. La mécanique statistique permet de calculer le flux résultant d'énergie au travers de la paroi. Ce dernier, dont la fonction est, en fait, de tendre à établir l'équilibre thermique global, s'avère proportionnel au gradient de température. Qu'il nous suffise de comprendre que, dans notre exemple, ce flux est bien non nul et dirigé vers la droite.

FIGURE 6.2: *La conduction thermique.*

L'identité entre le flux d'énergie et c fois la quantité de mouvement volumique, déjà mentionnée dans la section (6.3.1), implique la propriété, inconnue de la mécanique newtonienne, selon laquelle *en présence de conduction, la quantité de mouvement volumique est non nulle dans le référentiel comobile de l'élément de fluide considéré*. On peut concevoir cette quantité de mouvement comme étant celle de l'« onde thermique » (donc d'origine microscopique) traversant le référentiel macroscopiquement comobile.⁵

D'autre part, la *viscosité* d'un milieu résulte de la possibilité de *transporter de la quantité de mouvement* d'une région à l'autre. Cela implique que le flux de quantité de mouvement interne ne soit pas orthogonal à la surface virtuelle considérée.

On appelle *fluide parfait* un fluide *sans conduction de chaleur* et *sans viscosité*. Compte tenu des paragraphes précédents, cela signifie respectivement que, dans le référentiel localement comobile,

$$T^{0i} = 0 \quad \text{et} \quad (T^{ij}) = \begin{pmatrix} P_1 & & \\ & P_2 & \\ & & P_3 \end{pmatrix}.$$

Cette dernière propriété doit cependant être indépendante du choix des axes spatiaux. Or, lors d'un changement R des axes spatiaux, la partie spatiale T_s du tenseur se transforme

5. Il est facile de comprendre, à partir d'un modèle simple, qu'un élément de volume comobile puisse contenir une quantité de mouvement non nulle. Imaginons un parallépipède rectangle dont un côté est parallèle à un axe, appelons-le « horizontal ». A l'intérieur de cette boîte vont et viennent, horizontalement, des particules de masse propre m_0 , en quantité n_0 par unité de volume. Supposons qu'elles se meuvent de la gauche vers la droite avec une vitesse de composante $2V > 0$, et de la droite vers la gauche avec une vitesse $-V$ (les changements de vitesses résultent de collisions élastiques avec les particules de boîtes contiguës semblables). Leur vitesse vectorielle moyenne est bien nulle. Mais la quantité de mouvement volumique est donnée par

$$\frac{n_0}{3} \left(\frac{2m_0V}{\sqrt{1-4V^2/c^2}} - \frac{2m_0V}{\sqrt{1-V^2/c^2}} \right) > 0.$$

Il s'agit donc d'un effet relativiste, conséquence de l'équivalence de la masse et de l'énergie comme nous l'avons déjà signalé : chez Newton la masse ne dépend pas de la vitesse (il n'y a pas de racines carrées aux dénominateurs) et aucune quantité de mouvement n'accompagne le flux d'énergie.

selon une loi du type $\tilde{R}T_s R$, et ne peut rester diagonale que si $P_1 = P_2 = P_3 = P$. Ainsi, dans le référentiel comobile, on a

$$(T^{\mu\nu}) = \begin{pmatrix} \rho & & & \\ & P & & \\ & & P & \\ & & & P \end{pmatrix},$$

où le paramètre ρ symbolise l'énergie volumique tandis que P n'est autre que la pression. Il est facile de voir qu'on peut l'écrire de manière tensorielle sous la forme

$$T^{\mu\nu} = \frac{1}{c^2} (\rho + P) U^\mu U^\nu + P g^{\mu\nu}, \quad (6.12)$$

expression que nous utiliserons dans la suite puisqu'elle est valable dans n'importe quel référentiel. Insistons néanmoins sur le fait que les quantités ρ et P demeurent définies dans le référentiel comobile.

6.4.2 Equation d'état

Les paramètres intervenant dans l'expression du tenseur énergie-impulsion ne sont pas tous indépendants. On appelle *équation d'état* une relation entre la pression et l'énergie volumique. Dans la suite, nous utiliserons l'*équation barotrope*

$$P = (\gamma - 1) (\rho - n_0 m_0 c^2) \quad (6.13)$$

où m_0 représente la masse propre des particules et où l'indice adiabatique γ vaut

- 1 pour un fluide de poussière (pression nulle) ;
- 5/3 pour un gaz parfait monoatomique et non relativiste (la pression vaut alors les deux tiers de l'énergie interne volumique) ;
- 4/3 pour un gaz parfait monoatomique et ultra-relativiste, c'est-à-dire tel que $\rho \gg n_0 m_0 c^2$ (la pression vaut alors le tiers de l'énergie volumique). Ce dernier cas s'applique notamment au gaz de photons, pour lequel $m_0 = 0$.

6.4.3 Equations de conservation pour les fluides parfaits non relativistes

Soit un fluide parfait d'énergie volumique ρ et de pression P . Son tenseur d'énergie-impulsion, mesuré dans un référentiel inertiel rapporté à des coordonnées cartésiennes, est donné par (6.12) :

$$T^{\mu\nu} = \frac{1}{c^2} (\rho + P) U^\mu U^\nu + P \eta^{\mu\nu}.$$

Pour passer à la limite newtonienne, on fait les trois approximations suivantes.

1. La vitesse macroscopique de l'élément de fluide est très inférieure à celle de la lumière : $V^i \ll c$, d'où $U^0 \approx c$ et $U^i \approx V^i$.
2. De même pour les vitesses internes. Cela permet de définir univoquement la masse propre du système comme somme des masses propres de ses constituants élémentaires. (Rappelons que, en général, la notion de masse propre d'un système relativiste n'est pas très fondée car cette dernière n'est pas égale à la somme des masses propres des constituants du système, d'ailleurs non conservée lors de processus faisant intervenir des collisions inélastiques ; notamment, s'il y a apport d'énergie au système et donc variation de la température.) On peut alors écrire sans ambiguïté $\rho \approx \mu c^2$, μ désignant la masse volumique propre.

3. La pression est négligeable par rapport à l'énergie volumique : $P \ll \mu c^2$. Dans les situations courantes, cette hypothèse s'avère pertinente (dans le cas de l'atmosphère terrestre par exemple, $P \approx 10^5$ Pa et $\mu \approx 1$ gramme par litre, donc $P/(\mu c^2) \approx 10^{-12} \ll 1$).

On a ainsi

$$T^{00} = \frac{1}{c^2} (\rho + P) U^0 U^0 - P \approx \mu c^2 \quad (6.14)$$

$$T^{0i} = \frac{1}{c^2} (\rho + P) U^0 U^i \approx \mu c V^i \quad (6.15)$$

$$T^{ij} = \frac{1}{c^2} (\rho + P) U^i U^j + P \delta^{ij} \approx \mu V^i V^j + P \delta^{ij}. \quad (6.16)$$

La composante temporelle de l'équation (6.10) devient ainsi

$$\frac{\partial \mu}{\partial t} + \vec{\nabla} \cdot (\mu \vec{V}) = 0 \quad (\text{équation de continuité}) \quad (6.17)$$

tandis que ses composantes spatiales prennent la forme

$$\frac{\partial(\mu V^i)}{\partial t} + \frac{\partial}{\partial x^j} (\mu V^i V^j) + \frac{\partial P}{\partial x^i} = 0$$

soit, en tenant compte de (6.17),

$$\frac{\partial \vec{V}}{\partial t} + (\vec{V} \cdot \vec{\nabla}) \vec{V} = -\frac{1}{\mu} \vec{\nabla} P \quad (\text{équation d'Euler}). \quad (6.18)$$

6.4.4 Considérations thermodynamiques

L'équation (6.10) se prête à quelques manipulations conduisant à des résultats intéressants. Dans un système de coordonnées cartésiennes, on a :

$$0 = T^{\mu\nu}_{,\nu} \quad (6.19)$$

$$\begin{aligned} &= \left[\frac{1}{c^2} (\rho + P) U^\mu U^\nu + P \eta^{\mu\nu} \right]_{,\nu} \\ &= \frac{1}{c^2} [(\rho + P) U^\nu]_{,\nu} U^\mu + \frac{1}{c^2} (\rho + P) U^\nu U^\mu_{,\nu} + P_{,\nu} \eta^{\mu\nu}. \end{aligned} \quad (6.20)$$

Multiplions cette équation par U_μ , en sommant sur μ . Si l'on note que, d'une part, $U^\mu U_\mu = -c^2$ et que, d'autre part, $U^\mu_{,\nu} U_\mu = 0$, on obtient

$$\begin{aligned} -[(\rho + P) U^\nu]_{,\nu} + 0 + P_{,\nu} \eta^{\mu\nu} U_\mu &= 0 \\ [(\rho + P) U^\nu]_{,\nu} &= P_{,\nu} U^\nu. \end{aligned} \quad (6.21)$$

S'il y a conservation du nombre de particules, alors l'équation (6.11) est vérifiée :

$$(n_0 U^\nu)_{,\nu} = 0.$$

Dès lors, dans l'équation (6.21) réécrite sous la forme

$$\left(\frac{\rho + P}{n_0} \right)_{,\nu} n_0 U^\nu + \frac{\rho + P}{n_0} (n_0 U^\nu)_{,\nu} = P_{,\nu} U^\nu, \quad ,$$

le deuxième terme s'annule. D'autre part, pour tout champ scalaire α ,

$$\begin{aligned}\alpha_{,\nu} U^\nu &= \frac{\partial \alpha}{\partial x^\nu} \frac{dx^\nu}{d\tau} \\ &= \frac{d\alpha}{d\tau}\end{aligned}$$

où τ est le temps propre. Ainsi, on a

$$\begin{aligned}n_0 \frac{d}{d\tau} \left(\frac{\rho + P}{n_0} \right) &= \frac{dP}{d\tau} \\ \frac{d\rho}{d\tau} - \frac{\rho + P}{n_0} \frac{dn_0}{d\tau} &= 0.\end{aligned}\tag{6.22}$$

Or, un élément de fluide contenant N particules vérifie le premier principe de la thermodynamique :

$$T dS = dE + P dV$$

où T est la *température* de l'élément de fluide, S son *entropie*, $V = N/n_0$ son *volume propre* et $E = \rho V$ son *énergie interne*. Ainsi,

$$\begin{aligned}T dS &= d\rho V + (\rho + P) dV \\ &= d\rho \frac{N}{n_0} + (\rho + P) \frac{dN n_0 - N dn_0}{(n_0)^2}.\end{aligned}$$

Le nombre de particules étant supposé constant dans l'élément de fluide, on a $dN = 0$ et

$$\begin{aligned}T dS &= \frac{N}{n_0} \left[d\rho - \frac{\rho + P}{n_0} dn_0 \right] \\ T \frac{dS}{d\tau} &= \frac{N}{n_0} \left[\frac{d\rho}{d\tau} - \frac{\rho + P}{n_0} \frac{dn_0}{d\tau} \right].\end{aligned}$$

Or, cette quantité est nulle en vertu de l'équation (6.22). Ainsi, *un fluide parfait conservant le nombre de particules conserve son entropie : on parle de fluide adiabatique*.

6.4.5 Fluides parfaits dans un champ de gravitation

Injectons (6.21) dans (6.20) :

$$P_{,\nu} U^\nu U^\mu + (\rho + P) U^\nu U^\mu_{,\nu} + c^2 P_{,\nu} \eta^{\mu\nu} = 0.\tag{6.23}$$

Pour un fluide de poussières, $P = 0$ et l'équation devient

$$U^\nu U^\mu_{,\nu} = 0,$$

c'est-à-dire

$$\frac{dU^\mu}{d\tau} = 0.\tag{6.24}$$

Ainsi qu'il se doit, cette relation admet comme solution la ligne droite, c'est-à-dire la courbe transportant son vecteur tangent parallèlement à lui-même.

En présence d'un champ de gravitation, il est clair qu'un tel résultat est incorrect. En d'autres termes, la relation (6.19), écrite dans l'espace-temps de Minkowski rapporté aux

coordonnées cartésiennes, doit être modifiée si l'on veut prendre en compte l'interaction gravitationnelle.

Face à ce constat, deux attitudes sont alors possibles :

- *Incorporer des forces volumiques aux équations du mouvement dans l'espace-temps de Minkowski.* C'est le point de vue adopté par la mécanique newtonienne.
- *Maintenir autant que possible l'équation (6.19) et inclure le champ de gravitation dans la structure de l'espace-temps.*⁶ Ce renversement de perspective, profondément original, constitue une des idées maîtresses de la relativité générale.

6.5 Les fluides visqueux

Afin d'ouvrir notre étude à la théorie relativiste de la viscosité, mentionnons sommairement le cas particulier des fluides isotropes en *écoulement indivergentiel* (sans expansion des trajectoires spatio-temporelles des particules), c'est-à-dire tels que

$$U^\alpha_{;\alpha} = 0.$$

On peut leur attribuer un tenseur d'énergie-impulsion de composantes⁷

$$\begin{aligned} T_{\mu\nu}^{(\text{visc})} &= -\eta P_\mu^\alpha P_\nu^\beta (U_{\alpha;\beta} + U_{\beta;\alpha}) \\ &= -\eta \left(U_{\mu;\nu} + U_{\nu;\mu} + \frac{1}{c^2} U^\alpha U_{\mu;\alpha} U_\nu + \frac{1}{c^2} U^\alpha U_{\nu;\alpha} U_\mu \right), \end{aligned} \quad (6.25)$$

lequel doit, bien sûr, s'ajouter au tenseur d'énergie-impulsion des fluides parfaits établi en (6.12). Dans la relation (6.25), P_μ^α désigne l'opérateur de projection sur la section spatiale locale orthogonale à la quadrivitesse (section 3.2) :

$$P_\mu^\alpha = \delta_\mu^\alpha + \frac{1}{c^2} U^\alpha U_\mu,$$

et η est un paramètre positif dépendant du fluide, fonction de la température et éventuellement de la pression, appelé *coefficient de viscosité laminaire*.

A la limite non relativiste, on retrouve ainsi les *liquides newtoniens isotropes* (un *liquide* est un fluide à écoulement incompressible :

$$\frac{\partial \mu}{\partial t} + \vec{V} \cdot \vec{\nabla} \mu = 0,$$

ce qui implique, vu l'équation de continuité, l'annulation de la divergence de son champ de vitesses). Dans un référentiel inertiel en coordonnées cartésiennes, leur tenseur d'énergie-impulsion est donné par

$$T_{\mu 0}^{(\text{visc})} = 0, \quad T_{ij}^{(\text{visc})} = -\eta \left(\frac{\partial V_i}{\partial x_j} + \frac{\partial V_j}{\partial x_i} \right).$$

6. En fait, c'est l'expression de (6.19) sous forme covariante qui est maintenue (remplacement de la virgule par le point-virgule), conformément au principe d'équivalence : en présence d'un champ gravitationnel, les lois de la physique admettent la même formulation mathématique que ces lois en l'absence de gravitation, rapportées à des référentiels non inertiels.

7. A. Lichnerowicz, *Théories relativistes de la gravitation et de l'électromagnétisme*, Masson, 1955, p. 91 ; C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation*, Freeman, 1973, p. 567.

On parle d'*écoulement de Couette* lorsque ce liquide, supposé à pression constante (on néglige donc notamment l'influence de la gravitation), est soumis à un écoulement stationnaire entre deux plaques horizontales : l'une immobile à l'altitude $z = 0$ et l'autre dotée d'une vitesse V_h à l'altitude $z = h$, et tel qu'entre les deux plaques la vitesse du liquide soit donnée par la loi

$$(\vec{V}) = \left(\frac{zV_h}{h}, 0, 0 \right).$$

Les composantes non nulles du tenseur de viscosité sont alors

$$T_{13}^{(\text{visc})} = T_{31}^{(\text{visc})} = -\eta \frac{V_h}{h}.$$

Il est facile d'en déduire les forces agissant sur un élément de liquide, par exemple un petit cube dont les arêtes seraient dirigées selon les trois axes cartésiens.

6.6 Le tenseur d'énergie-impulsion du champ électromagnétique

6.6.1 Lois de conservation

Considérons⁸ le tenseur $T^{(\text{ch})}$, de composantes

$$T^{(\text{ch})\mu\nu} = \frac{1}{\mu_0} \left(F^{\mu\xi} F_{\nu\xi} - \frac{1}{4} \eta^{\mu\nu} F_{\rho\xi} F^{\rho\xi} \right). \quad (6.26)$$

Il possède les dimensions physiques d'une énergie volumique, c'est-à-dire d'une pression. On a

$$(T^{(\text{ch})\mu\nu}) = \begin{pmatrix} W & S_1/c & S_2/c & S_3/c \\ S_1/c & \sigma_{11} & \sigma_{12} & \sigma_{13} \\ S_2/c & \sigma_{21} & \sigma_{22} & \sigma_{23} \\ S_3/c & \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \quad (6.27)$$

où

$$W = \frac{\varepsilon_0}{2} (E^2 + c^2 B^2) \quad (6.28)$$

$$\vec{S} = \frac{1}{\mu_0} \vec{E} \wedge \vec{B} \quad (\text{vecteur de Poynting}) \quad (6.29)$$

$$\sigma_{ij} = \varepsilon_0 \left[- (E_i E_j + c^2 B_i B_j) + \frac{1}{2} (E^2 + c^2 B^2) \delta_{ij} \right] \quad (\text{tenseur de Maxwell}). \quad (6.30)$$

Calculons la quadridivergence $T^{(\text{ch})\mu}_{\nu,\mu}$:

$$\begin{aligned} T^{(\text{ch})\mu}_{\nu,\mu} &= \frac{1}{\mu_0} (F^{\mu\xi} F_{\nu\xi} - \frac{1}{4} \delta_{\nu}^{\mu} F_{\rho\xi} F^{\rho\xi})_{,\mu} \\ &= \frac{1}{\mu_0} (F^{\mu\xi}_{,\mu} F_{\nu\xi} + F^{\mu\xi} F_{\nu\xi,\mu} - \frac{1}{4} \delta_{\nu}^{\mu} F_{\rho\xi,\mu} F^{\rho\xi} - \frac{1}{4} \delta_{\nu}^{\mu} F_{\rho\xi} F^{\rho\xi}_{,\mu}) \\ &= \frac{1}{\mu_0} (F^{\mu\xi}_{,\mu} F_{\nu\xi} + F^{\mu\xi} F_{\nu\xi,\mu} - \frac{1}{2} F^{\rho\xi} F_{\rho\xi,\nu}). \end{aligned}$$

8. La signification des symboles utilisés dans cette section est exposée en annexe A.1 p. 250.

Compte tenu des équations de Maxwell (A.3) et (A.4)

$$\begin{aligned} F_{\rho\xi,\nu} &= -F_{\xi\nu,\rho} - F_{\nu\rho,\xi} \\ F^{\mu\xi}_{,\mu} &= -\mu_0 j^\xi, \end{aligned}$$

on ramène l'expression à

$$\begin{aligned} T^{(\text{ch})\mu}_{\nu,\mu} &= \frac{1}{\mu_0} (-\mu_0 j^\xi F_{\nu\xi} + F^{\mu\xi} F_{\nu\xi,\mu} + \frac{1}{2} F^{\rho\xi} F_{\xi\nu,\rho} + \frac{1}{2} F^{\rho\xi} F_{\nu\rho,\xi}) \\ &= -F_{\nu\xi} j^\xi + \frac{1}{2\mu_0} F^{\mu\xi} (2F_{\nu\xi,\mu} + F_{\xi\nu,\mu} + F_{\nu\mu,\xi}). \end{aligned}$$

Or, le terme entre parenthèses vaut

$$F_{\nu\xi,\mu} + F_{\xi\nu,\mu} + F_{\nu\xi,\mu} + F_{\nu\mu,\xi}.$$

Vu l'antisymétrie du tenseur de Faraday, la somme des deux premiers termes est nulle. La somme des deux derniers termes est, quant à elle, symétrique sur les indices μ et ξ et s'annule par multiplication contractée avec $F^{\mu\xi}$. On obtient ainsi l'important résultat

$$T^{(\text{ch})\mu}_{\nu,\mu} = -F_{\nu\xi} j^\xi. \quad (6.31)$$

Exprimons le membre de droite en termes des quantités communes :

$$(F^\nu_\xi j^\xi) = \left[\frac{1}{c} \vec{j} \cdot \vec{E}, (\rho \vec{E} + \vec{j} \wedge \vec{B}) \right]. \quad (6.32)$$

Compte tenu de (A.8)-(A.9), la composante temporelle de (6.32) est égale à $1/c$ fois le taux de variation temporelle, par unité de volume donné, de l'énergie cinétique des particules matérielles (incluant leur énergie au repos); tandis que les composantes spatiales représentent le taux de variation temporelle, par unité de volume donné, de leur quantité de mouvement. Dès lors, vu les relations (6.27) et (6.28) à (6.30), l'expression (6.31) se décompose de la manière suivante.

La composante temporelle $T^{(\text{ch})\mu}_{0,\mu} = -F_{0\xi} j^\xi$ peut s'écrire

$$\frac{\partial W}{\partial t} + \vec{\nabla} \cdot \vec{S} = -\vec{j} \cdot \vec{E}. \quad (6.33)$$

Cette équation signifie que, dans un volume donné, la variation d'énergie du champ électromagnétique contrebalance la variation d'énergie cinétique de la matière. La quantité $T^{(\text{ch})00} = W$ représente l'énergie volumique du champ. En vertu du théorème de la divergence de Gauss, la quantité S_i désigne la quantité d'énergie passant par unité de temps à travers l'unité de surface orthogonale à l'axe x^i dans le sens des x^i croissants. Les composantes $T^{(\text{ch})0i} = S_i/c$ représentent donc la quantité d'énergie passant par unité de c temps à travers l'unité de surface orthogonale à l'axe x^i dans le sens des x^i croissants.

Quant aux composantes spatiales $T^{(\text{ch})\mu}_{i,\mu} = -F_{i\xi} j^\xi$, elles prennent la forme

$$\frac{1}{c^2} \frac{\partial S_i}{\partial t} + \sum_{k=1}^3 \frac{\partial \sigma_{ki}}{\partial x^k} = -(\rho \vec{E} + \vec{j} \wedge \vec{B})_i \quad (6.34)$$

exprimant que, dans un volume donné, la variation de quantité de mouvement du champ compense celle de la matière. Par conséquent, les quantités $T^{(\text{ch})i0}/c = S_i/c^2$ sont les

composantes de la *quantité de mouvement volumique du champ*. Quant aux quantités $T^{(\text{ch})ij} = \sigma_{ij}$, toujours en vertu du théorème de Gauss, elles constituent la *ième composante de la quantité de mouvement du champ traversant par unité de temps l'unité de surface orthogonale à l'axe x^j dans le sens des x^j croissants*.

De tout cela il résulte que $T^{(\text{ch})\mu\nu}$ représente le flux d'énergie ($\mu = 0$) ou de c fois la quantité de mouvement ($\mu = 1, 2, 3$) du champ au travers de l'hypersurface $x^\nu = \text{constante}$. $\mathbf{T}^{(\text{ch})}$ est donc le *tenseur d'énergie-impulsion du champ électromagnétique*.

Des deux interprétations de S^i résultant de (6.33) et (6.34), on conclut également que le flux d'énergie du champ électromagnétique au travers du plan $x^j = \text{constante}$, est égal au produit par c de la j ème composante de sa quantité de mouvement volumique.

En résumé, considérons un parallélépipède infinitésimal à faces respectivement orthogonales aux trois axes spatiaux x^1, x^2, x^3 , et de côtés égaux à A, B, C . Cette boîte contient une quantité d'énergie électromagnétique égale à $ABC T^{(\text{ch})00}$ et une quantité de i ème composante de la quantité de mouvement égale à $ABC T^{(\text{ch})i0}/c$. Par unité de c .temps il s'écoule, par chacune des deux faces orthogonales à x^1 et dans le sens des x^1 croissants, une quantité d'énergie égale à $BC T^{(\text{ch})01}$ et une quantité de i ème composante de la quantité de mouvement égale à $BC T^{(\text{ch})i1}/c$, où le champ est évalué sur la face correspondante. Il est facile de transposer ces propos aux autres faces.

6.6.2 Fluide de particules chargées de type « poussière »

Dans un référentiel localement comobile avec cette poussière, appelons μ_0 la masse volumique du fluide, ρ_0 sa charge électrique volumique, n_0 le nombre de particules par unité de volume, m_0 la masse au repos de chaque particule et q la charge électrique de chaque particule. Remarquons que $\mu_0/m_0 = \rho_0/q = n_0$. Ces poussières sont supposées sans interaction mutuelle : leur énergie d'interaction, précisément, est logée dans le champ.

Soit $T^{(\text{p})\mu\nu} = \mu_0 U^\mu U^\nu = n_0 m_0 U^\mu U^\nu$ le tenseur d'énergie-impulsion du fluide en question (on peut considérer que l'on a autant de fluides que de rapports différents q/m_0). On a

$$\begin{aligned} T^{(\text{p})\mu}_{\nu,\mu} &= \eta_{\nu\xi} T^{(\text{p})\mu\xi}_{,\mu} \\ &= \eta_{\nu\xi} m_0 (n_0 U^\mu U^\xi)_{,\mu} \\ &= \eta_{\nu\xi} m_0 (n_0 U^\mu)_{,\mu} U^\xi + \eta_{\nu\xi} m_0 n_0 U^\mu (U^\xi)_{,\mu}. \end{aligned}$$

Or, l'équation de continuité (A.5) entraîne

$$(n_0 U^\mu)_{,\mu} = 0,$$

exprimant la conservation du nombre de particules. Par conséquent,

$$\begin{aligned} T^{(\text{p})\mu}_{\nu,\mu} &= \eta_{\nu\xi} m_0 n_0 U^\mu (U^\xi)_{,\mu} \\ &= \eta_{\nu\xi} m_0 n_0 \frac{dx^\mu}{d\tau} \frac{\partial U^\xi}{\partial x^\mu} \\ &= \eta_{\nu\xi} m_0 n_0 \frac{dU^\xi}{d\tau}. \end{aligned}$$

D'autre part, l'équation de mouvement (A.6)-(A.7) donne

$$\eta_{\nu\xi} m_0 \frac{dU^\xi}{d\tau} = q F_{\nu\lambda} U^\lambda.$$

Puisque $\rho_0/q = n_0$, on peut l'écrire sous la forme continue

$$\begin{aligned} \eta_{\nu\xi} m_0 n_0 \frac{dU^\xi}{d\tau} &= \rho_0 F_{\nu\lambda} U^\lambda \\ &= F_{\nu\lambda} j^\lambda. \end{aligned}$$

Par suite,

$$T^{(\text{p})\mu}_{\nu,\mu} = F_{\nu\lambda} j^\lambda. \quad (6.35)$$

En conclusion, des équations (6.31) et (6.35) il découle

$$\boxed{T^{(\text{ch})\mu}_{\nu,\mu} + T^{(\text{p})\mu}_{\nu,\mu} = 0} \quad (6.36)$$

qui exprime la *conservation de l'énergie et de la quantité de mouvement totales (champ + particules)*.

Chapitre 7

Les équations d'Einstein

7.1 Retour sur le principe d'équivalence

La question se pose de trouver les principes permettant de conduire à l'écriture des lois de la physique en présence d'un champ de gravitation.¹

Dans son acception forte, le *principe d'équivalence* s'exprime comme suit : *il est toujours possible de trouver un référentiel tel que, dans une région suffisamment petite de l'espace-temps, les lois de la physique soient celles de la relativité restreinte*. En d'autres mots, même en présence d'un champ de gravitation on peut trouver un système de coordonnées locales dans lequel ces lois sont celles de l'espace-temps euclidien.

L'utilisation d'un espace-temps non euclidien pour décrire l'interaction gravitationnelle, déjà annoncée dans le chapitre 2, s'avère cependant tout à fait compatible avec ce principe et même, suggérée par lui. En effet, en n'importe quel point donné d'une variété riemannienne il est toujours possible de trouver un système de coordonnées où s'annulent tous les symboles de Christoffel (voir les coordonnées normales de Riemann dans l'annexe B). On dispose donc de la latitude d'annuler localement les dérivées premières du tenseur métrique. De plus, on peut toujours s'arranger pour réduire ses composantes à la forme canonique $\text{diag}(-1, 1, 1, 1)$ de la relativité restreinte si l'on impose une métrique lorentzienne de signature $(-+++)$. Or, comme on l'a noté dans les chapitres 2 et 3, il paraît fondé d'attribuer au tenseur métrique la signification de potentiels gravitationnels. Les effets du champ gravifique peuvent donc toujours être annulés, *jusqu'au premier ordre de dérivation inclus*, par une transformation de coordonnées, c'est-à-dire, en langage physique, par un changement de référentiel. Ainsi, de même qu'un champ gravitationnel inhomogène peut être annulé mais seulement localement, un espace courbe peut localement être assimilé à son espace tangent, euclidien mais différent en chaque point.

Bien entendu, ces réflexions n'ont pas valeur de démonstration : elles reflètent l'harmonie des rapports qu'entretiennent le principe d'équivalence et la notion de variété riemannienne, mais il n'existe pas d'argument pour prouver que l'espace-temps de Minkowski doit *nécessairement* être remplacé par un espace courbe. Si on l'admet cependant, force est de constater que l'on ne peut pas mieux « simuler » la relativité restreinte en un point, qu'à

1. Voir également

S.M. Carroll, *Lecture Notes on General Relativity*, <http://preposterousuniverse.com/grnotes/>, 1997, chap. 4 ;

M. Beuthe, *La quête d'une nouvelle théorie de la gravitation*, dans *initiation à la cosmologie*, J.-M. Liétard et L. Willems éditeurs, Bruxelles, 1998, paragraphes 1.5, 1.6 et 2.1 ;

B. Schutz, *A first course in general relativity*, Cambridge University Press, 2009, pp. 171-175.

travers l'utilisation des coordonnées normales de Riemann.

Dès lors, dans ces coordonnées les équations de la physique doivent être *exactement* celles de la relativité restreinte. Par exemple, les équations de Maxwell (A.4) demeurent inchangées. Le principe d'équivalence y proscriit l'intervention de termes géométriques relatifs à la courbure, comme le tenseur de Riemann-Christoffel et ses dérivés, et interdit des formulations du genre

$$(F^{\mu\nu} + \alpha R^{\mu\nu})_{,\nu} = \mu_0 j^\mu$$

si la constante α diffère de zéro.

D'une façon générale, considérons une équation E de la relativité restreinte, tensorielle pour le groupe de Lorentz. Dans le système des coordonnées normales de Riemann, tous les symboles de Christoffel sont nuls et donc, les dérivées partielles ordinaires sont identiques à des dérivées covariantes. Pour obtenir l'équation tensorielle valide dans n'importe quel système de coordonnées et se réduisant à E dans les coordonnées normales, il suffit donc (du moins pour les équations du premier ordre) de remplacer dans E les dérivées ordinaires par des dérivées covariantes.

Ainsi conçu dans le cadre d'un espace-temps courbe, le principe d'équivalence assure une transition très naturelle entre le principe de covariance restreinte (aux transformations de Lorentz) et un principe de covariance générale, en fournissant une règle simple permettant d'écrire les lois de la physique en présence d'un champ de gravitation : *ce sont les équations de la relativité restreinte en coordonnées générales*. Par exemple, en relativité générale la loi (6.10) de conservation de l'énergie et de la quantité de mouvement prend la forme

$$T^{\mu\nu}_{;\nu} = 0. \quad (7.1)$$

Dans le jargon des habitués, on parle de la « règle de passage de la virgule au point-virgule ».

Pour cette raison, le tenseur métrique doit satisfaire les conditions $g_{\mu\nu;\xi} = 0$. La connexion est donc riemannienne et le lien (5.77) vérifié entre les symboles de Christoffel et la métrique, comme la cohérence du raisonnement l'exigeait.

L'équation de mouvement d'une particule ponctuelle soumise uniquement à son inertie et à la gravitation constitue une autre application importante de cette règle. En relativité restreinte, une particule libre se meut le long d'une droite, c'est-à-dire une *géodésique de l'espace-temps de Minkowski* :

$$\frac{d^2 x^\mu}{d\tau^2} = 0$$

où τ est le temps propre. Cette équation peut s'écrire en fonction du quadrivecteur vitesse de composantes $U^\mu = dx^\mu/d\tau$, sous la forme

$$0 = \frac{dU^\mu}{d\tau} = U^\alpha \frac{\partial U^\mu}{\partial x^\alpha},$$

soit, de façon covariante :

$$U^\alpha (\nabla_\alpha \vec{U})^\mu = 0$$

ou encore

$$\begin{aligned} U^\alpha \left(\frac{\partial U^\mu}{\partial x^\alpha} + \Gamma_{\alpha\beta}^\mu U^\beta \right) &= 0 \\ \frac{d^2 x^\mu}{d\tau^2} + \Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} &= 0. \end{aligned} \quad (7.2)$$

Telle est l'équation de mouvement d'une particule en relativité générale. La trajectoire demeure *géodésique*, mais cette fois sur la variété riemannienne.

7.2 « Die Lochbetrachtung »

Le *principe de covariance générale* est la formulation mathématique de la requête d'équivalence formelle de tous les référentiels. Il demande que *les lois de la physique soient tensorielles pour le groupe des transformations quelconques de coordonnées*.

Nous avons montré comment le principe d'équivalence permettait de souscrire à cette exigence en ce qui concerne les lois issues de la relativité restreinte. Il reste à trouver une formulation totalement covariante pour l'équation relativiste du champ de gravitation lui-même, généralisant l'équation de Poisson de la théorie newtonienne :

$$\nabla^2 \Phi = 4\pi G \mu,$$

où Φ est le potentiel gravitationnel et μ la masse volumique.

D'après le chapitre 6, l'être géométrique le plus simple faisant intervenir μ est un tenseur d'ordre deux, le tenseur d'énergie-impulsion \mathbf{T} . D'autre part, dans l'esprit de la relativité générale c'est le tenseur métrique \mathbf{g} qui doit assurer le rôle de Φ . Dès 1912 Einstein avait bien intégré ces divers aspects du problème. Il sollicita alors la collaboration de son ami mathématicien Marcel Grossmann, afin de trouver des équations du type

$$\mathbf{G}(\mathbf{g}) = \kappa \mathbf{T} \quad (7.3)$$

où \mathbf{G} est un tenseur d'ordre deux, fonction du tenseur métrique (et probablement de ses dérivées jusqu'à l'ordre deux, comme dans l'équation de Poisson), et κ une constante de couplage entre la géométrie et la matière.

Cependant, les auteurs se convainquirent bientôt que leur projet était voué à l'échec, au terme d'un raisonnement assez subtil qu'ils appelèrent « Lochbetrachtung », c'est-à-dire « l'argument du trou ». Nous allons l'exposer car il permet de mieux comprendre les enjeux mathématiques et épistémologiques d'une théorie covariante de la gravitation.

Supposons que le tenseur d'énergie-impulsion s'annule dans une certaine région L de l'espace-temps (c'est le « trou »). A ces sources $T_{\mu\nu}(\vec{x})$ correspond une solution $g_{\mu\nu}(\vec{x})$ des équations du champ. Choisissons alors de nouvelles coordonnées $\underline{\vec{x}}$, définies par le difféomorphisme (passif)

$$\vec{x} = \vec{f}(\underline{\vec{x}}) \quad (7.4)$$

et telles que $\underline{\vec{x}} \neq \vec{x}$ dans L et $\underline{\vec{x}} = \vec{x}$ hors de L .² (A l'intérieur de L , le domaine de variation de $\underline{\vec{x}}$ est identique à celui de \vec{x} .) On a

$$\underline{T}_{\mu\nu}(\underline{\vec{x}}) = \frac{\partial f^\alpha}{\partial \underline{x}^\mu} \frac{\partial f^\beta}{\partial \underline{x}^\nu} T_{\alpha\beta}[\vec{f}(\underline{\vec{x}})]. \quad (7.5)$$

A l'intérieur du trou les coordonnées changent mais $T_{\alpha\beta}(\vec{x}) = 0$; et à l'extérieur où $T_{\alpha\beta}(\vec{x}) \neq 0$, les coordonnées ne sont pas modifiées. Les relations (7.5) imposent donc

$$\underline{T}_{\mu\nu}(\underline{\vec{x}}) = T_{\mu\nu}(\underline{\vec{x}}) \text{ partout.} \quad (7.6)$$

Si l'on postule des équations tensorielles du type (7.3), à $\underline{T}_{\mu\nu}(\underline{\vec{x}})$ correspond la solution

$$\underline{g}_{\mu\nu}(\underline{\vec{x}}) = \frac{\partial f^\alpha}{\partial \underline{x}^\mu} \frac{\partial f^\beta}{\partial \underline{x}^\nu} g_{\alpha\beta}[\vec{f}(\underline{\vec{x}})] \quad (7.7)$$

2. Pour simplifier, on supposera que la fonction \vec{f} détermine un difféomorphisme de classe C^∞ , c'est-à-dire une bijection infiniment continûment dérivable dans les deux sens.

tandis qu'à $T_{\mu\nu}(\vec{x})$ correspond la solution $g_{\mu\nu}(\vec{x})$. Or, à l'intérieur du trou, le domaine de variation de \vec{x} est le même que celui de \vec{x} . Si nous rebaptisons \vec{x} en \vec{x} dans les équations (7.6) et (7.7), nous pouvons donc affirmer qu'aux sources $T_{\mu\nu}(\vec{x}) = \underline{T}_{\mu\nu}(\vec{x})$ correspond, en plus de la solution $g_{\mu\nu}(\vec{x})$, la solution

$$\begin{aligned} g_{\mu\nu}(\vec{x}) & \text{ hors de L} \\ \underline{g}_{\mu\nu}(\vec{x}) &= \frac{\partial f^\alpha(\vec{x})}{\partial x^\mu} \frac{\partial f^\beta(\vec{x})}{\partial x^\nu} g_{\alpha\beta}[\vec{f}(\vec{x})] \neq g_{\mu\nu}(\vec{x}) \quad \text{dans L.} \end{aligned} \quad (7.8)$$

Enfin, on peut s'arranger pour que le trou occupe un domaine de l'espace-temps situé tout entier dans le futur d'une hypersurface spatiale $t = 0$ contenant les conditions initiales pour le tenseur métrique. La conclusion est implacable : *on a découvert deux solutions différentes pour une même distribution de sources et des conditions initiales identiques*. La causalité est violée ! Ainsi, l'hypothèse d'équations tensorielles pour le groupe des transformations quelconques des coordonnées semble conduire à une absurdité.

Pourtant, ce raisonnement est parfaitement correct d'un point de vue mathématique. Mais il faut réviser son *interprétation*. L'élément fondamental de la théorie des variétés différentielles est le concept de *point*. Ce dernier jouit d'une existence autonome et on peut lui attribuer des coordonnées par le biais de tel ou tel homéomorphisme local avec \mathbb{R}^n . Il est dès lors tentant de supposer que la notion de point mathématique traduit sans ambiguïté celle d'*événement physique*. La démarche de la relativité générale consisterait alors en ceci : étant donné un point-événement (« là »), on lui attribue des coordonnées et, ensuite, on résout des équations pour connaître la valeur du tenseur métrique dans le système de coordonnées choisi en ce point. Si cette analyse était correcte, alors effectivement l'argument du trou révélerait une absurdité : en tel événement (« là »), des équations généralement covariantes autoriseraient deux solutions différentes.

Il apparaît cependant que cette lecture est erronée. *Les coordonnées seules n'ont aucun sens physique, on ne peut associer un point mathématique univoque à un événement physique donné*. Certes, résoudre des équations généralement covariantes consiste à trouver une solution $g_{\mu\nu}(\vec{x})$; mais les coordonnées \vec{x} ne sont pas données *a priori*, elles font partie intégrante de la solution. C'est la métrique qui confère un sens physique aux points de la variété. Arthur Eddington a résumé la situation en une formule choc : « Space is not a lot of points close together, it is a lot of distances interlocked ».

Tâchons de préciser cette explication en inspectant l'élément de longueur dans L. Le lien entre les deux solutions $\underline{g}_{\mu\nu}(\vec{x})$ et $g_{\mu\nu}(\vec{x})$ y apparaît clairement :

$$\begin{aligned} ds^2 &= \underline{g}_{\mu\nu}(\vec{x}) dx^\mu dx^\nu \\ &= \frac{\partial f^\alpha(\vec{x})}{\partial x^\mu} \frac{\partial f^\beta(\vec{x})}{\partial x^\nu} g_{\alpha\beta}[\vec{f}(\vec{x})] dx^\mu dx^\nu \\ &= g_{\alpha\beta}[\vec{f}(\vec{x})] df^\alpha(\vec{x}) df^\beta(\vec{x}). \end{aligned} \quad (7.9)$$

Considérons alors le difféomorphisme (7.4), non plus comme un changement *passif* de coordonnées, mais comme une transformation *active* appliquant L sur lui-même.³ Le point de coordonnées \vec{x} , associé à la solution $\underline{g}_{\mu\nu}(\vec{x})$, est appliqué sur le point de coordonnées $\vec{f}(\vec{x})$ associé à la solution $g_{\mu\nu}(\vec{x})$. Ils constituent le *même* événement physique car, d'une

3. Voir par exemple la section 4.1 de l'article de M. Gaul et C. Rovelli, *Loop Quantum Gravity and the Meaning of Diffeomorphism Invariance*, ArXiv : gr-qc/9910079v2, 20 décembre 1999.

part, les seuls éléments de réalité physique sont les coïncidences spatio-temporelles (l'intersection des lignes d'univers de deux particules), et celles-ci sont bien préservées par le difféomorphisme ; et, d'autre part, les ds^2 sont identiques (figure 7.1).

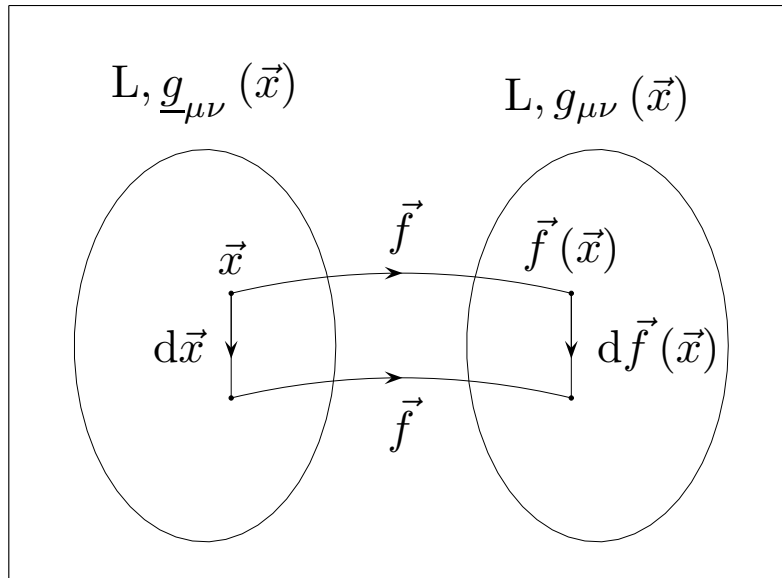


FIGURE 7.1: Le difféomorphisme f applique le trou L sur lui-même en préservant le ds^2 , ainsi que le montrent les équations (7.9). On dit que l'espace-temps de gauche est isométriquement équivalent à celui de droite. Ils décrivent tous les deux le même champ de gravitation mais l'expression mathématique de la métrique n'est pas univoque. La relativité générale est une théorie de jauge : étant donné une distribution de sources et les conditions initiales, des équations tensorielles admettent plusieurs solutions différentes mathématiquement mais physiquement équivalentes. Cette particularité de la théorie sera étudiée plus amplement dans le cadre de l'approximation des champs faibles (section 7.6).

Il faudra attendre le dernier trimestre de 1915 pour qu'Einstein réalise son erreur. La voie était à nouveau libre pour la recherche d'une théorie covariante de la gravitation : un mois plus tard, il obtenait les équations du champ. Dans un texte de vulgarisation écrit des années plus tard, il a jugé opportun d'insister sur la signification profonde de la métrique et l'immense fossé conceptuel qui sépare la relativité restreinte de la relativité générale :⁴

« Si l'on suppose le champ de gravitation, c'est-à-dire les g_{ik} , éliminé, il ne reste pas un espace du type [Minkowski], mais absolument *rien*, pas même un « espace topologique » [...] Un espace vide, c'est-à-dire un espace sans champ n'existe pas. »

Citons enfin un excellent commentaire de John Stachel :

« Even relativists have not yet fully adopted the point of view, "no metric, no nothing". If you look at the way the general theory of relativity is formulated mathematically in even the most careful treatises, for example, you see this clearly. They start out by introducing a global manifold (the points of which are usually identified forthwith with events [...]), and then put such structures on this manifold as the metric tensor field. Is that the way that any

4. A. Einstein, *La relativité et le problème de l'espace* (1953), dans *La relativité*, traduit de l'allemand par M. Solovine, Gauthier-Villars, 1998, pp. 177-178.

one of us actually goes about solving the field equations of general relativity? Of course not. One first solves them on a generic patch, and *then* one tries to maximally extend the local solution (using some criteria for acceptable extensions) from that patch to a global manifold, which is not known ahead of time. Before solving the field equations, one generally doesn't know the global manifold on which the solution will turn out to be maximally extended. So we are pulling a swindle when we tell students, as our definitions imply, that you first pick the manifold and then solve the field equations on it. »⁵

7.3 La démarche heuristique

Que faut-il mettre dans le membre de gauche de l'équation (7.3)? Le tenseur de Ricci et les quantités associées sont des objets géométriques fonctions du tenseur métrique et ce, à l'instar du laplacien, jusqu'à l'ordre deux de dérivation. D'autre part, la similitude entre (5.136) et (7.1) apparaît comme un critère de sélection. Une *équation reliant le tenseur d'Einstein et le tenseur d'énergie-impulsion* semble dès lors tout indiquée :

$$\boxed{R^{\mu\nu} - \frac{R}{2} g^{\mu\nu} = \kappa T^{\mu\nu}}, \quad (7.10)$$

la constante de couplage κ restant à déterminer. L'importance qu'il convient d'accorder au tenseur d'Einstein dans ce contexte est légitimée par un théorème dû à Elie Cartan. Ce dernier a démontré⁶ que le tenseur \mathbf{X} de rang deux le plus général, qui soit

- symétrique ;
- construit entièrement à partir du tenseur métrique \mathbf{g} et de ses dérivées d'ordre un et deux ;
- quasi-linéaire, c'est-à-dire linéaire en les dérivées secondes du tenseur métrique ;
- de quadridivergence covariante nulle ;

est une combinaison linéaire du tenseur d'Einstein et du tenseur métrique :

$$\mathbf{X} = a \mathbf{G} + b \mathbf{g}$$

où a et b sont des constantes arbitraires.⁷

En adaptant la constante de couplage entre la géométrie et la matière, on peut poser $\mathbf{X} = \mathbf{G} + \Lambda \mathbf{g}$ sans perte de généralité. Historiquement, ce sont néanmoins les équations (7.10), donc sans le paramètre Λ , qu'après bien des tâtonnements A. Einstein a soumises le 25 novembre 1915 aux *Sitzungsberichte* de la *Preussische Akademie der Wissenschaften* (publication le 2 décembre). On les appelle *équations d'Einstein*, ou encore *équations du champ de gravitation*. Il faut saluer l'intuition profonde qui a permis à leur auteur de reconnaître leur pertinence car, en fait, à cette époque il ne connaissait pas encore les identités de Bianchi!

Dans le vide, $T^{\mu\nu} = 0$. Les relations (7.10), contractées sur leurs deux indices, imposent alors $R = 0$ et les équations d'Einstein se réduisent à l'annulation du tenseur de Ricci :

$$R^{\mu\nu} = 0. \quad (7.11)$$

5. J. Stachel, *Einstein from 'B' to 'Z'*, Einstein Studies, Volume 9, Birkhäuser, 2002, p. 423.

6. E. Cartan, « Sur les équations de la gravitation d'Einstein », *Journal de mathématiques pures et appliquées*, 9^e série, tome 1, pp. 141-204 (1922).

7. En dimension 4 on peut même lever les hypothèses de symétrie et de quasi-linéarité : D. Lovelock, *Journal of Mathematical Physics* **12**, pp. 498-501 (1971) ; *Journal of Mathematical Physics* **13**, pp. 874-876 (1972).

Les équations d'Einstein ne sont pas linéaires. En voici l'explication : le champ de gravitation engendré par une distribution de matière possède lui-même de l'énergie et de la quantité de mouvement, autrement dit, il constitue lui-même la source d'un champ de gravitation.⁸ Dans le langage de la théorie quantique des champs, on dirait que les bosons vecteurs de l'interaction gravitationnelle (les hypothétiques *gravitons*) interagissent entre eux, étant eux-mêmes soumis à la gravitation universelle : en somme, la non-linéarité de la relativité générale est une conséquence du principe d'équivalence. La situation est tout autre en électromagnétisme, théorie régie par les équations de Maxwell, linéaires, et où le champ ne transporte pas de charges électriques.

Cependant, les équations du champ possèdent la propriété de *quasi-linéarité*, c'est-à-dire qu'elles sont linéaires en les termes d'ordre maximal, 2. Comme nous le verrons dans un cours plus avancé, il s'avère possible d'extraire du tenseur d'Einstein ces termes en les dérivées secondes du tenseur métrique, et de regrouper les termes complémentaires dans une expression appelée *pseudo-tenseur d'énergie-impulsion du champ gravitationnel*. Assez naturellement, on le range habituellement parmi les sources, à côté du tenseur d'énergie-impulsion traditionnel.

7.4 Equations du champ et équations du mouvement

Dans la théorie de Newton, pour établir la loi de conservation de l'énergie mécanique totale, cinétique plus potentielle, il faut exploiter l'équation de Poisson *et* les équations du mouvement. Or, les équations du champ relativiste engendrent automatiquement des lois de conservation. On peut ainsi s'attendre à ce qu'elles contiennent une information sur le mouvement des particules. Nous avons déjà remarqué cette propriété dans le cas du fluide de poussières [voir la relation (6.24)]. En 1927, Einstein et Jakob Grommer ont constaté que les équations du mouvement des particules d'épreuve, envisagées comme singularités de la variété riemannienne, résultent en fait des équations du champ, ce qui rend superflu le postulat des géodésiques. Nous allons le montrer dans cette section.

Récrivons d'abord les lois de conservation (7.1) :

$$\begin{aligned} 0 &= T^{\mu\nu}{}_{;\nu} \\ &= T^{\mu\nu}{}_{,\nu} + \Gamma_{\alpha\nu}^{\mu} T^{\alpha\nu} + \Gamma_{\alpha\nu}^{\nu} T^{\mu\alpha} \\ &= \frac{1}{\sqrt{-g}} (\sqrt{-g} T^{\mu\nu})_{,\nu} + \Gamma_{\alpha\nu}^{\mu} T^{\alpha\nu}, \end{aligned} \quad (7.12)$$

compte tenu de la relation

$$\Gamma_{\alpha\nu}^{\nu} = \frac{g_{,\alpha}}{2g}.$$

Considérons une particule test de masse propre m_0 qui se déplace, au fil de son temps propre τ , sur la ligne d'univers $\mathcal{L} \equiv \vec{x} = \vec{x}(\tau)$, et notons \mathcal{V} un petit tronçon de tube entourant la trajectoire sur une longueur $c\Delta\tau$. On peut définir son énergie volumique propre ρ par la relation

$$\int_{\mathcal{V}} \rho(\vec{x}) \sqrt{-g(\vec{x})} d\vec{x} = m_0 c^2 c\Delta\tau. \quad (7.13)$$

8. En conséquence de cela, il est possible de construire des modèles d'univers vides de matière (c'est-à-dire, dont le tenseur de Ricci est nul partout) possédant un champ de gravitation non nul (c'est-à-dire, à courbure non nulle). Cette situation se rencontre dans le contexte des modèles cosmologiques spatialement homogènes et anisotropes, appelés *modèles de Bianchi*.

En effet, comme le membre de gauche est un scalaire, on peut le calculer dans un système de coordonnées cartésiennes localement comobiles $(c\tau, x, y, z)$, et on obtient ainsi $\int \rho c d\tau dx dy dz$, égal au membre de droite. D'autre part,

$$\begin{aligned} c\Delta\tau &= \int_{\mathcal{L}} c d\tau \int_{\mathcal{V}} \delta[\vec{x} - \vec{x}(\tau)] d\vec{x} \\ &= \int_{\mathcal{V}} \left\{ \int_{\mathcal{L}} \delta[\vec{x} - \vec{x}(\tau)] c d\tau \right\} d\vec{x} \end{aligned} \quad (7.14)$$

où δ est la distribution de Dirac dans l'espace à quatre dimensions. De la comparaison de (7.13) et (7.14), on tire

$$\rho(\vec{x}) \sqrt{-g(\vec{x})} = m_0 c^2 \int_{\mathcal{L}} \delta[\vec{x} - \vec{x}(\tau)] c d\tau. \quad (7.15)$$

Dès lors, le tenseur d'énergie-impulsion de la particule s'écrit

$$\begin{aligned} T^{\mu\nu}(\vec{x}) &= \rho(\vec{x}) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \\ &= \frac{m_0 c^3}{\sqrt{-g(\vec{x})}} \int_{\mathcal{L}} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \delta[\vec{x} - \vec{x}(\tau)] d\tau. \end{aligned} \quad (7.16)$$

En injectant (7.16) dans (7.12), on obtient successivement

$$\begin{aligned} 0 &= \int_{\mathcal{L}} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \frac{\partial}{\partial x^\nu} \delta[\vec{x} - \vec{x}(\tau)] d\tau + \Gamma_{\alpha\nu}^\mu \int_{\mathcal{L}} \frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau} \delta[\vec{x} - \vec{x}(\tau)] d\tau \\ &= - \int_{\mathcal{L}} \frac{dx^\mu}{d\tau} \frac{d}{d\tau} \delta[\vec{x} - \vec{x}(\tau)] d\tau + \Gamma_{\alpha\nu}^\mu \int_{\mathcal{L}} \frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau} \delta[\vec{x} - \vec{x}(\tau)] d\tau \\ &= - \int_{\mathcal{L}} \frac{d}{d\tau} \left\{ \frac{dx^\mu}{d\tau} \delta[\vec{x} - \vec{x}(\tau)] \right\} d\tau + \int_{\mathcal{L}} \left(\frac{d^2 x^\mu}{d\tau^2} + \Gamma_{\alpha\nu}^\mu \frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau} \right) \delta[\vec{x} - \vec{x}(\tau)] d\tau \\ &= \left\{ - \frac{dx^\mu}{d\tau} \delta[\vec{x} - \vec{x}(\tau)] \right\}_{\tau_1}^{\tau_2} + \int_{\mathcal{L}} \left(\frac{d^2 x^\mu}{d\tau^2} + \Gamma_{\alpha\nu}^\mu \frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau} \right) \delta[\vec{x} - \vec{x}(\tau)] d\tau \end{aligned}$$

où τ_1 et τ_2 désignent les temps propres aux extrémités de la trajectoire considérée. Il en résulte, presque partout,

$$\int_{\mathcal{L}} \left(\frac{d^2 x^\mu}{d\tau^2} + \Gamma_{\alpha\nu}^\mu \frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau} \right) \delta[\vec{x} - \vec{x}(\tau)] d\tau = 0,$$

ce qui implique l'annulation de l'intégrand le long de la trajectoire, qui dès lors vérifie l'équation des géodésiques

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma_{\alpha\nu}^\mu \frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau} = 0.$$

En 1938, Einstein, Banesh Hoffmann et Leopold Infeld ont développé une méthode approchée pour étudier le mouvement d'un système de corps ponctuels graves. (Le problème est complexe car, non seulement chaque corps influence le mouvement des autres, mais en plus les équations du champ ne sont pas linéaires. Ainsi, on ne connaît pas de solution exacte au problème des deux corps en relativité générale.) Cette méthode est connue sous le nom d'*approximation post-newtonienne*. Nous en toucherons quelques mots dans la section 7.8.

7.5 La méthode de Hilbert

Le 31 mars 1916, la *Gesellschaft der Wissenschaften* de Göttingen publiait un article signé David Hilbert, et dans lequel les équations (7.10) étaient méthodiquement déduites d'un principe variationnel, donc par une voie différente de celle d'Einstein.⁹ Ce résultat autonome contribue à nous rappeler l'importance des principes variationnels, qui semblent au cœur de nos tentatives de modélisation mathématique de la nature. Nous allons nous limiter ici à donner une idée de la méthode de Hilbert en dérivant les équations (7.10) dans le vide.

7.5.1 L'action du champ gravitationnel

La densité lagrangienne la plus simple que l'on puisse concevoir est la courbure scalaire R . Cependant, $\int R d\vec{x}$ n'est pas un scalaire : c'est la quantité $\int R \sqrt{-g} d\vec{x}$ qui peut prétendre à ce statut d'invariance. Aussi retiendra-t-on comme densité lagrangienne de Hilbert la quantité

$$R \sqrt{-g}.$$

Théorème 7.1 (décomposition de la densité lagrangienne) *La densité lagrangienne de Hilbert peut s'écrire comme somme de deux termes :*

$$R \sqrt{-g} = G \sqrt{-g} + (\sqrt{-g} w^\alpha)_{,\alpha} \quad (7.17)$$

où

$$G = g^{\mu\nu} (\Gamma_{\mu\alpha}^\beta \Gamma_{\nu\beta}^\alpha - \Gamma_{\mu\nu}^\alpha \Gamma_{\alpha\beta}^\beta) \quad \text{et} \quad w^\alpha = g^{\mu\nu} \Gamma_{\mu\nu}^\alpha - g^{\mu\alpha} \Gamma_{\mu\nu}^\nu.$$

Commentaire. Ce théorème est d'un grand intérêt car il permet de transformer comme suit l'intégrale d'action :

$$\int_V R \sqrt{-g} d\vec{x} = \int_V G \sqrt{-g} d\vec{x} + \int_V (\sqrt{-g} w^\alpha)_{,\alpha} d\vec{x}.$$

Le deuxième terme du membre de droite peut être transformé par le théorème de Gauss et sa variation s'annule si l'on postule, comme c'est l'usage, une variation du champ nulle sur la frontière du domaine V . On aboutit ainsi à la conclusion selon laquelle la densité lagrangienne de Hilbert $R \sqrt{-g}$ est équivalente à la densité lagrangienne $G \sqrt{-g}$, dans la mesure où elle fournit les mêmes équations du champ :

$$\delta \int_V R \sqrt{-g} d\vec{x} = \delta \int_V G \sqrt{-g} d\vec{x}.$$

9. Les deux scientifiques étaient en contact depuis octobre 1912 au moins, et au début de l'été 1915 Einstein avait donné à l'Université de Göttingen une série de séminaires portant sur ses recherches en relativité générale. L'article de Hilbert a été soumis le 20 novembre 1915, donc cinq jours avant celui d'Einstein, mais il a subi ensuite d'importantes refontes sans que la date de soumission ne fût modifiée. Ces circonstances ont engendré une querelle de priorité. Voir par exemple A. Pais, « *Subtle is the Lord...* », *The Science and the Life of Albert Einstein*, Oxford University Press, 1982, chapitre 14 ; J. Stachel, *op. cit.*, pp. 353-364.

Or, la quantité G ne dépend que de $g^{\mu\nu}$ et de ses dérivées premières $g^{\mu\nu}_{,\lambda}$. D'après un résultat classique de calcul des variations, l'équation $\delta \int_V G \sqrt{-g} \, d\vec{x} = 0$ est alors équivalente aux équations d'Euler-Lagrange

$$\frac{\partial}{\partial x^\lambda} \frac{\partial(G\sqrt{-g})}{\partial(g^{\mu\nu}_{,\lambda})} - \frac{\partial(G\sqrt{-g})}{\partial g^{\mu\nu}} = 0. \quad (7.18)$$

Cela confirme que les équations déduites de la variation de l'action $\delta \int_V R\sqrt{-g} \, d\vec{x} = 0$ sont du deuxième ordre en les dérivées du tenseur métrique.

Pour démontrer le théorème, il faut d'abord prouver deux lemmes.

Lemme 7.1

$$g^{\rho\sigma} \Gamma_{\rho\sigma}^\mu = -\frac{1}{\sqrt{-g}} (\sqrt{-g} g^{\mu\alpha})_{,\alpha}.$$

Preuve. Cette relation découle de

$$\begin{aligned} g^{\rho\sigma} \Gamma_{\rho\sigma}^\mu &= \frac{1}{2} g^{\rho\sigma} g^{\mu\alpha} (g_{\alpha\rho,\sigma} + g_{\alpha\sigma,\rho} - g_{\rho\sigma,\alpha}) \\ &= g^{\rho\sigma} g^{\mu\alpha} g_{\alpha\rho,\sigma} - \frac{1}{2} g^{\mu\alpha} g^{\rho\sigma} g_{\rho\sigma,\alpha}. \end{aligned}$$

Or, $(g^{\mu\alpha} g_{\alpha\rho})_{,\sigma} = 0$ implique $g^{\mu\alpha} g_{\alpha\rho,\sigma} = -g^{\mu\alpha}_{,\sigma} g_{\alpha\rho}$ et, d'autre part, $g_{,\alpha} = g g^{\rho\sigma} g_{\rho\sigma,\alpha}$, d'où

$$\begin{aligned} g^{\rho\sigma} \Gamma_{\rho\sigma}^\mu &= -g^{\rho\sigma} g^{\mu\alpha}_{,\sigma} g_{\alpha\rho} - \frac{1}{2g} g_{,\alpha} g^{\mu\alpha} \\ &= -g^{\mu\alpha}_{,\alpha} - \frac{1}{2g} g_{,\alpha} g^{\mu\alpha} \\ &= -\frac{1}{\sqrt{-g}} (\sqrt{-g} g^{\mu\alpha})_{,\alpha}. \end{aligned}$$

□

Lemme 7.2

$$-\Gamma_{\mu\nu}^\alpha (\sqrt{-g} g^{\mu\nu})_{,\alpha} + \Gamma_{\mu\nu}^\nu (\sqrt{-g} g^{\mu\alpha})_{,\alpha} = 2 g^{\mu\nu} \sqrt{-g} (\Gamma_{\mu\alpha}^\beta \Gamma_{\nu\beta}^\alpha - \Gamma_{\mu\nu}^\alpha \Gamma_{\alpha\beta}^\beta). \quad (7.19)$$

Preuve. Il faut utiliser les trois identités

$$g^{\rho\sigma} \Gamma_{\rho\sigma}^\mu = -\frac{1}{\sqrt{-g}} (\sqrt{-g} g^{\mu\alpha})_{,\alpha} \quad (7.20)$$

$$(\sqrt{-g})_{,\alpha} = \sqrt{-g} \Gamma_{\alpha\rho}^\rho \quad (7.21)$$

$$g^{\mu\nu}_{,\alpha} = -\Gamma_{\rho\alpha}^\mu g^{\rho\nu} - \Gamma_{\rho\alpha}^\nu g^{\mu\rho}. \quad (7.22)$$

La relation (7.20) découle du lemme 7.1. Quant à la relation (7.22), elle résulte de

$$0 = g^{\mu\nu}_{,\alpha} = g^{\mu\nu}_{,\alpha} + \Gamma_{\rho\alpha}^\mu g^{\rho\nu} + \Gamma_{\rho\alpha}^\nu g^{\mu\rho}.$$

En utilisant les identités (7.20) à (7.22) pour traiter le membre de gauche de (7.19), on obtient

$$\sqrt{-g} (-\Gamma_{\mu\nu}^\nu g^{\rho\sigma} \Gamma_{\rho\sigma}^\mu - \Gamma_{\mu\nu}^\alpha \Gamma_{\alpha\rho}^\rho g^{\mu\nu} + \Gamma_{\mu\nu}^\alpha \Gamma_{\rho\alpha}^\mu g^{\rho\nu} + \Gamma_{\mu\nu}^\alpha \Gamma_{\rho\alpha}^\nu g^{\mu\rho}).$$

Les deux premiers termes sont identiques, ainsi que les deux derniers. Ainsi,

$$\begin{aligned} -\Gamma_{\mu\nu}^\alpha (\sqrt{-g} g^{\mu\nu})_{,\alpha} + \Gamma_{\mu\nu}^\nu (\sqrt{-g} g^{\mu\alpha})_{,\alpha} &= 2 g^{\rho\nu} \sqrt{-g} (\Gamma_{\rho\alpha}^\mu \Gamma_{\mu\nu}^\alpha - \Gamma_{\rho\nu}^\mu \Gamma_{\mu\beta}^\beta) \\ &= 2 g^{\mu\nu} \sqrt{-g} (\Gamma_{\mu\alpha}^\beta \Gamma_{\nu\beta}^\alpha - \Gamma_{\mu\nu}^\alpha \Gamma_{\alpha\beta}^\beta). \end{aligned} \quad \square$$

Nous pouvons enfin démontrer le théorème 7.1. On a

$$\begin{aligned}
\sqrt{-g} R &= \sqrt{-g} g^{\mu\nu} R_{\mu\nu} \\
&= \sqrt{-g} (g^{\mu\nu} \Gamma_{\mu\nu,\alpha}^\alpha - g^{\mu\nu} \Gamma_{\mu\alpha,\nu}^\alpha + g^{\mu\nu} \Gamma_{\mu\nu}^\alpha \Gamma_{\alpha\beta}^\beta - g^{\mu\nu} \Gamma_{\mu\alpha}^\beta \Gamma_{\nu\beta}^\alpha) \\
&= [\sqrt{-g} (g^{\mu\nu} \Gamma_{\mu\nu}^\alpha - g^{\mu\alpha} \Gamma_{\mu\nu}^\nu)],_\alpha - \Gamma_{\mu\nu}^\alpha (\sqrt{-g} g^{\mu\nu}),_\alpha + \Gamma_{\mu\nu}^\nu (\sqrt{-g} g^{\mu\alpha}),_\alpha \\
&\quad - (\Gamma_{\mu\alpha}^\beta \Gamma_{\nu\beta}^\alpha - \Gamma_{\mu\nu}^\alpha \Gamma_{\alpha\beta}^\beta) g^{\mu\nu} \sqrt{-g}.
\end{aligned}$$

Or, d'après le lemme 7.2, la somme du deuxième et du troisième terme est égale à

$$2 g^{\mu\nu} \sqrt{-g} (\Gamma_{\mu\alpha}^\beta \Gamma_{\nu\beta}^\alpha - \Gamma_{\mu\nu}^\alpha \Gamma_{\alpha\beta}^\beta).$$

On a donc bien, comme annoncé en (7.17),

$$\begin{aligned}
R\sqrt{-g} &= g^{\mu\nu} \sqrt{-g} (\Gamma_{\mu\alpha}^\beta \Gamma_{\nu\beta}^\alpha - \Gamma_{\mu\nu}^\alpha \Gamma_{\alpha\beta}^\beta) + [\sqrt{-g} (g^{\mu\nu} \Gamma_{\mu\nu}^\alpha - g^{\mu\alpha} \Gamma_{\mu\nu}^\nu)],_\alpha \\
&= G\sqrt{-g} + (\sqrt{-g} w^\alpha),_\alpha.
\end{aligned}$$

□

7.5.2 Obtention des équations d'Einstein à l'aide d'un principe variationnel

Généralement, dans la littérature¹⁰, on montre, en effectuant un calcul *ad hoc*, que la variation de la densité lagrangienne $R\sqrt{-g}$ fournit les équations d'Einstein dans le vide

$$R_{\mu\nu} - \frac{R}{2} g_{\mu\nu} = 0. \quad (7.23)$$

Dans l'annexe C nous utilisons un procédé techniquement plus lourd mais offrant l'avantage, nous semble-t-il, d'une économie de pensée considérable : nous montrons que les équations d'Euler-Lagrange (7.18) appliquées à la densité lagrangienne effective

$$G\sqrt{-g} = g^{\mu\nu} (\Gamma_{\mu\alpha}^\beta \Gamma_{\nu\beta}^\alpha - \Gamma_{\mu\nu}^\alpha \Gamma_{\alpha\beta}^\beta) \sqrt{-g} \quad (7.24)$$

sont bien les équations (7.23).

7.6 L'approximation des champs faibles

7.6.1 Généralités

Si le champ de gravitation est suffisamment faible, on peut traiter l'espace-temps comme une perturbation au premier ordre de l'espace-temps euclidien. Négliger les termes d'ordre 2 ou plus dans les équations d'Einstein revient à linéariser ces dernières, donc à simplifier considérablement leur écriture et, par conséquent, leur résolution. Outre qu'un certain nombre de situations physiques justifient une telle procédure (par exemple, l'étude de la propagation des ondes gravitationnelles loin de leurs sources), l'approximation des champs faibles offre l'avantage de permettre le calcul de la constante de couplage entre la géométrie et la matière.

10. Voir par exemple

L. Landau et E. Lifchitz, *Physique théorique, Tome 2, Théorie des champs*, Mir, 1989, § 95 ;
R. d'Inverno, *Introducing Einstein's Relativity*, Oxford University Press, 1992, chapitre 11.

La covariance générale rend cependant assez délicate la formulation mathématique de ces idées. En effet, une quantité « petite » dans un système de coordonnées peut être rendue aussi « grande » que l'on veut par un changement de variables approprié. On dira donc qu'un champ est faible s'il existe un système de coordonnées (dites « cartésiennes ») tel que

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad \text{avec} \quad |h_{\mu\nu}| \ll 1. \quad (7.25)$$

Nous supposons en outre que les dérivées des quantités $h_{\mu\nu}$ sont également d'ordre 1 :

$$|h_{\mu\nu,\alpha}| = O(1), \quad |h_{\mu\nu,\alpha\beta}| = O(1).$$

Bien entendu, la décomposition (7.25) n'est pas covariante : seul $(g_{\mu\nu})$ possède le statut de tenseur. Cependant, elle est préservée par les transformations de Lorentz

$$\underline{x}^\alpha = \Lambda^\alpha_\beta x^\beta$$

et les quantités $h_{\mu\nu}$ se transforment *tensoriellement* pour ce sous-groupe du groupe des transformations générales.

Théorème 7.2 *Les transformations infinitésimales de coordonnées*

$$\underline{x}^\alpha = x^\alpha + \xi^\alpha(x^\beta), \quad |\xi^\alpha_{,\beta}| \ll 1, \quad (7.26)$$

où les quantités $\xi^\alpha_{,\beta}$ sont d'ordre 1, préservent également la décomposition (7.25).

En effet,

$$\underline{g}_{\mu\nu} = \frac{\partial x^\alpha}{\partial \underline{x}^\mu} \frac{\partial x^\beta}{\partial \underline{x}^\nu} g_{\alpha\beta}.$$

Or,

$$\frac{\partial \underline{x}^\alpha}{\partial x^\mu} = \delta^\alpha_\mu + \xi^\alpha_{,\mu}$$

implique

$$\left(\frac{\partial \underline{x}^\alpha}{\partial \underline{x}^\mu} \right) = \left(\frac{\partial \underline{x}^\alpha}{\partial x^\mu} \right)^{-1} = (\delta^\alpha_\mu - \xi^\alpha_{,\mu}) + O(2).$$

Donc, si l'on note $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}^{old}$,

$$\begin{aligned} \underline{g}_{\mu\nu} &= (\delta^\alpha_\mu - \xi^\alpha_{,\mu})(\delta^\beta_\nu - \xi^\beta_{,\nu}) g_{\alpha\beta} + O(2) \\ &= g_{\mu\nu} - \xi^\alpha_{,\mu} g_{\alpha\nu} - \xi^\beta_{,\nu} g_{\mu\beta} + O(2) \\ &= \eta_{\mu\nu} + h_{\mu\nu}^{old} - \eta_{\alpha\nu} \xi^\alpha_{,\mu} - \eta_{\mu\beta} \xi^\beta_{,\nu} + O(2) \\ &= \eta_{\mu\nu} + h_{\mu\nu}^{new} + O(2) \end{aligned}$$

où l'on a posé

$$\xi_\mu = \eta_{\mu\alpha} \xi^\alpha \quad (7.27)$$

$$h_{\mu\nu}^{new} = h_{\mu\nu}^{old} - \xi_{\mu,\nu} - \xi_{\nu,\mu}. \quad (7.28)$$

□

Remarquons que les quantités $h_{\mu\nu}^{new}$ et $h_{\mu\nu}^{old}$ ne sont pas reliées par la loi tensorielle :

$$\begin{aligned} h_{\mu\nu}^{new} \neq h_{\mu\nu}^{old} &= \frac{\partial x^\alpha}{\partial \underline{x}^\mu} \frac{\partial x^\beta}{\partial \underline{x}^\nu} h_{\alpha\beta}^{old} = (\delta_\mu^\alpha - \xi^{\alpha, \mu}) (\delta_\nu^\beta - \xi^{\beta, \nu}) h_{\alpha\beta}^{old} + O(2) \\ &= h_{\mu\nu}^{old} + O(2). \end{aligned}$$

Exemple : le changement de variables (dans cet exemple, on ne somme pas sur les indices répétés)

$$\underline{x}^\mu = x^\mu + A^\mu \sin x^\mu, \quad |A^\mu| \ll 1,$$

fournit

$$\begin{aligned} \frac{\partial \underline{x}^\mu}{\partial x^\alpha} &= (1 + A^\mu \cos x^\mu) \delta_\alpha^\mu \\ \frac{\partial x^\alpha}{\partial \underline{x}^\mu} &= (1 - A^\alpha \cos x^\alpha)_{x(\underline{x})} \delta_\mu^\alpha + O(2) \\ &= (1 - A^\alpha \cos \underline{x}^\alpha) \delta_\mu^\alpha + O(2) \end{aligned}$$

donc

$$\begin{aligned} h_{\mu\nu}^{new} &= h_{\mu\nu}^{old} - 2 A_\mu \cos \underline{x}^\mu \quad \text{si } \mu = \nu \\ &= h_{\mu\nu}^{old} \quad \text{si } \mu \neq \nu. \end{aligned}$$

En résumé, du point de vue mathématique tout se passe comme si le champ de gravitation faible était décrit à l'aide d'un champ tensoriel $h_{\mu\nu}$ défini dans l'espace-temps absolu de Minkowski, de tenseur métrique $\eta_{\mu\nu}$. Dans cette optique, les seuls changements de coordonnées admissibles sont les transformations de Lorentz, pour lesquelles le champ $h_{\mu\nu}$ se transforme, effectivement, comme un tenseur.¹¹ De plus, les champs tensoriels *mathématiquement différents* $h_{\mu\nu}^{old}$ et $h_{\mu\nu}^{new}$ décrivent la *même situation physique* en vertu du principe de covariance générale et du fait qu'ils se déduisent l'un de l'autre par un changement de coordonnées [du type (7.26)].

Cette interprétation rappelle une caractéristique importante de l'électromagnétisme. En effet, cette théorie se fonde également sur la donnée d'un champ A^μ (*vectorel*, cette fois) défini dans l'espace-temps de Minkowski. Or, il se peut que deux champs vectoriels mathématiquement différents, $A^{old \mu}$ et $A^{new \mu}$, décrivent la même situation physique. Cette propriété, appelée *invariance de jauge*, est exposée en annexe A.2. L'intérêt de cette notion est considérable, car le fait de disposer d'une certaine latitude mathématique permet de simplifier les équations sans rien changer à leur contenu physique.

Le passage d'un système de coordonnées x^μ à un système \underline{x}^μ à l'aide d'une transformation infinitésimale du type (7.26) sera donc, dans la suite, interprété comme un *changement de jauge* $h_{\mu\nu}^{old} \rightarrow h_{\mu\nu}^{new}$ pour les potentiels gravitationnels au point de coordonnées x^μ dans l'espace-temps de Minkowski.

11. Ce passage par la relativité restreinte est une transition nécessaire si l'on veut, à terme, réinterpréter le champ de gravitation dans le cadre de l'espace absolu de Newton. En réalité, l'espace-temps n'est pas absolu (son tenseur métrique $g_{\mu\nu}$ étant relatif à la matière) et, comme nous l'avons déjà signalé, $\eta_{\mu\nu}$ n'est pas un tenseur, non plus que $h_{\mu\nu}$. Dans la décomposition (7.25) on relit la gravitation comme un champ indépendant défini dans l'espace-temps, alors qu'en fait elle est extraite de sa structure géométrique (c'est la signification du signe +). Méthodologiquement, le recours à des coordonnées « cartésiennes » est requis si l'on veut, au travers des $h_{\mu\nu}$, décrire le champ de gravitation newtonien « pur », c'est-à-dire débarrassé des forces fictives d'inertie. On perçoit combien toutes ces contorsions enfreignent la philosophie du principe d'équivalence.

Il est dès lors naturel d'élever et abaisser les indices à l'aide des $\eta_{\mu\nu}$, ainsi que nous l'avons déjà suggéré dans l'équation (7.27). Par exemple, nous poserons

$$h^{\alpha\beta} = \eta^{\alpha\mu} \eta^{\beta\nu} h_{\mu\nu}.$$

Il ne faut pas, cependant, se laisser abuser par l'interprétation exposée dans les paragraphes précédents et qui repose sur une sorte de leurre formel : insistons sur ce point, les quantités $\eta_{\mu\nu}$ et $h_{\mu\nu}$ ne réalisent pas de véritables champs tensoriels au sens où l'entend la covariance générale, ni même la covariance restreinte aux coordonnées « cartésiennes » telles que nous les avons définies ci-dessus. Seule la métrique $g_{\mu\nu}$ peut prétendre à ce statut de tenseur. Ainsi, on *n'a pas* $g^{\mu\nu} = \eta^{\mu\nu} + h^{\mu\nu}$. En effet, si $\eta = \text{diag}(-1, 1, 1, 1)$ et si h est une matrice constituée d'éléments d'ordre 1, alors

$$(\eta + h)^{-1} = \eta - \eta h \eta + O(2),$$

ce qui signifie que, en fait,

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} + O(2).$$

7.6.2 Les équations d'Einstein linéarisées

Il est aisé de calculer les composantes du tenseur de Riemann-Christoffel en se limitant aux termes du premier ordre :

$$R_{\lambda\mu\nu\xi} = \frac{1}{2} (h_{\lambda\xi, \mu\nu} + h_{\mu\nu, \lambda\xi} - h_{\mu\xi, \lambda\nu} - h_{\lambda\nu, \mu\xi}) + O(2). \quad (7.29)$$

On en déduit le tenseur de Ricci

$$\begin{aligned} R_{\mu\nu} &= \eta^{\lambda\xi} R_{\lambda\mu\xi\nu} + O(2) \\ &= -\eta^{\lambda\xi} R_{\lambda\mu\nu\xi} + O(2) \\ &= -\frac{1}{2} \left(h^{\lambda}_{\lambda, \mu\nu} + h_{\mu\nu, \lambda}^{\lambda} - h^{\lambda}_{\mu, \lambda\nu} - h^{\lambda}_{\nu, \lambda\mu} \right) + O(2), \end{aligned}$$

la courbure scalaire

$$\begin{aligned} R &= \eta^{\alpha\beta} R_{\alpha\beta} + O(2) \\ &= -h^{\lambda}_{\lambda, \alpha}{}^{\alpha} + h^{\lambda\alpha}_{, \lambda\alpha} + O(2) \end{aligned}$$

et le tenseur d'Einstein

$$\begin{aligned} G_{\mu\nu} &= R_{\mu\nu} - \frac{R}{2} \eta_{\mu\nu} + O(2) \\ &= -\frac{1}{2} \left(h^{\lambda}_{\lambda, \mu\nu} + h_{\mu\nu, \lambda}^{\lambda} - h^{\lambda}_{\mu, \lambda\nu} - h^{\lambda}_{\nu, \lambda\mu} - \eta_{\mu\nu} h^{\lambda}_{\lambda, \alpha}{}^{\alpha} + \eta_{\mu\nu} h^{\lambda\alpha}_{, \lambda\alpha} \right) + O(2). \end{aligned}$$

Si l'on pose

$$h = h^{\lambda}_{\lambda} \quad (7.30)$$

$$\tilde{h}_{\mu\nu} = h_{\mu\nu} - \frac{h}{2} \eta_{\mu\nu}, \quad (7.31)$$

on peut écrire

$$G_{\mu\nu} = -\frac{1}{2} \left(\tilde{h}_{\mu\nu, \lambda}^{\lambda} - \tilde{h}_{\lambda\mu, \nu}^{\lambda} - \tilde{h}_{\lambda\nu, \mu}^{\lambda} + \eta_{\mu\nu} \tilde{h}_{\lambda\alpha, \lambda\alpha}^{\lambda\alpha} \right) + O(2). \quad (7.32)$$

Les quantités $R_{\lambda\mu\nu\xi}$ et ses associées sont tensorielles pour des transformations quelconques de coordonnées (contrairement aux $h_{\mu\nu}$), notamment pour les changements de coordonnées du type (7.26) : $R_{\lambda\mu\nu\xi}^{new} = \underline{R}_{\lambda\mu\nu\xi}^{old}$. Comme, de plus, $\underline{R}_{\lambda\mu\nu\xi}^{old} = R_{\lambda\mu\nu\xi}^{old} + O(2)$, on obtient donc $R_{\lambda\mu\nu\xi}^{new} = R_{\lambda\mu\nu\xi}^{old} + O(2)$: ces quantités sont invariantes de jauge et on peut les écrire, indifféremment, en fonction de $h_{\mu\nu}^{old}$ ou de $h_{\mu\nu}^{new}$. On peut le vérifier directement en injectant la transformation (7.28) dans les expressions ci-dessus. Pour le tenseur de Riemann-Christoffel par exemple, on trouve

$$\begin{aligned} R_{\lambda\mu\nu\xi}^{new} &= R_{\lambda\mu\nu\xi}^{old} - \frac{1}{2} [(\xi_{\lambda,\xi} + \xi_{\xi,\lambda})_{,\mu\nu} + (\xi_{\mu,\nu} + \xi_{\nu,\mu})_{,\lambda\xi} - (\xi_{\mu,\xi} + \xi_{\xi,\mu})_{,\lambda\nu} - (\xi_{\lambda,\nu} + \xi_{\nu,\lambda})_{,\mu\xi}] \\ &= R_{\lambda\mu\nu\xi}^{old}. \end{aligned}$$

Théorème 7.3 (jauge de L. Lorenz en relativité générale) *Il est possible de trouver une jauge telle que*

$$\tilde{h}^{new\ \mu\nu}_{,\nu} = 0$$

et, par conséquent, dans laquelle

$$G_{\mu\nu} = -\frac{1}{2} \square \tilde{h}_{\mu\nu}$$

où \square désigne l'opérateur de J. d'Alembert.

Démonstration. Voyons d'abord comment se transforment les quantités $\tilde{h}_{\mu\nu}$ lors du changement de coordonnées (7.26). On a

$$\begin{aligned} \tilde{h}_{\mu\nu}^{new} &= h_{\mu\nu}^{new} - \frac{h^{new}}{2} \eta_{\mu\nu} \\ &= h_{\mu\nu}^{new} - \frac{1}{2} \eta^{\alpha\beta} h_{\alpha\beta}^{new} \eta_{\mu\nu} \\ &= h_{\mu\nu}^{old} - \xi_{\mu,\nu} - \xi_{\nu,\mu} - \frac{1}{2} \eta^{\alpha\beta} [h_{\alpha\beta}^{old} - \xi_{\alpha,\beta} - \xi_{\beta,\alpha}] \eta_{\mu\nu} \\ &= \tilde{h}_{\mu\nu}^{old} - \xi_{\mu,\nu} - \xi_{\nu,\mu} + \frac{1}{2} (\xi^{\beta}_{,\beta} + \xi^{\alpha}_{,\alpha}) \eta_{\mu\nu} \\ &= \tilde{h}_{\mu\nu}^{old} - \xi_{\mu,\nu} - \xi_{\nu,\mu} + \xi^{\lambda}_{,\lambda} \eta_{\mu\nu}. \end{aligned}$$

Ainsi,

$$\begin{aligned} \tilde{h}^{new\ \mu\nu}_{,\nu} &= \eta^{\mu\alpha} \eta^{\nu\beta} \tilde{h}^{new}_{\alpha\beta,\nu} \\ &= \eta^{\mu\alpha} \eta^{\nu\beta} \left(\tilde{h}_{\alpha\beta}^{old} - \xi_{\alpha,\beta} - \xi_{\beta,\alpha} + \xi^{\lambda}_{,\lambda} \eta_{\alpha\beta} \right)_{,\nu} \\ &= \tilde{h}^{old\ \mu\nu}_{,\nu} - \xi^{\mu}_{,\nu} - \xi^{\nu}_{,\mu} + \eta^{\mu\nu} \xi^{\lambda}_{,\lambda\nu} \\ &\quad \text{(les termes soulignés en ondulé se suppriment)} \\ &= \tilde{h}^{old\ \mu\nu}_{,\nu} - \square \xi^{\mu}. \end{aligned}$$

Pour annuler cette quantité, il faut donc trouver des ξ^{μ} tels que

$$\square \xi^{\mu} = \tilde{h}^{old\ \mu\nu}_{,\nu}. \quad (7.33)$$

Cette équation hyperbolique admet toujours une solution si le membre de droite est suffisamment régulier (problème de Cauchy dans la théorie des équations aux dérivées partielles). Dans la jauge ainsi obtenue (également appelée *jauge de De Donder*, ou encore *jauge harmonique*), les trois derniers termes de l'équation (7.32) s'annulent. Le théorème est démontré. \square

Les équations d'Einstein $G^{\mu\nu} = \kappa T^{\mu\nu}$ s'écrivent ainsi, dans l'approximation linéaire :

$$\boxed{\square \bar{h}^{\mu\nu} = -2\kappa T^{\mu\nu}}. \quad (7.34)$$

7.6.3 La limite newtonienne

La théorie linéarisée présente en elle-même un certain intérêt : en plus de l'obtention de solutions radiatives, déjà mentionnées, elle permet d'établir, de façon approchée, l'effet Einstein (décalage en fréquence de la lumière se propageant dans un champ gravitationnel) et le calcul de la déviation des rayons lumineux au voisinage du soleil¹².

La moindre de ses applications n'est cependant pas le calcul de la constante de couplage κ entre la géométrie et la matière. Le principe de la méthode est simple : si la théorie de la relativité générale est correcte, on doit, moyennant quelques approximations, pouvoir retrouver l'équation de D. Poisson pour le potentiel gravifique newtonien, à partir des équations d'Einstein (7.34). La valeur de κ s'en déduit par comparaison.

Passons en revue ces différentes approximations.

- On raisonne donc dans le cadre de la théorie des champs faibles (7.25).
- On utilise les approximations pour les fluides parfaits non relativistes envisagées dans la section 6.4.3. Rappelons les conséquences des équations (6.14) à (6.16) :

$$\begin{aligned} |T^{00}| &\gg |T^{0i}| \\ |T^{00}| &\gg |T^{ij}|. \end{aligned}$$

- On se limite aux solutions quasi-statiques (voir annexe A.3); en d'autres termes, les dérivées temporelles sont supposées négligeables par rapport aux dérivées spatiales dans les équations (7.34).

Si μ désigne la masse volumique de la matière, on a alors

$$\nabla^2 \bar{h}^{00} \approx -2\kappa \mu c^2 \quad (7.35)$$

$$|\nabla^2 \bar{h}^{\mu\nu}| \ll |\nabla^2 \bar{h}^{00}| \quad \text{si } (\mu, \nu) \neq (0, 0) \quad (7.36)$$

donc

$$\nabla^2 (\bar{h}^{\mu\nu}) \approx 0 \quad \text{si } (\mu, \nu) \neq (0, 0).$$

La théorie des équations aux dérivées partielles enseigne que, si le laplacien d'une quantité nulle à l'infini (ce qui est bien le cas des $\bar{h}^{\mu\nu}$) est partout égal à zéro, alors cette quantité est identiquement nulle. Donc

$$\bar{h}^{\mu\nu} \approx 0 \quad \text{si } (\mu, \nu) \neq (0, 0).$$

12. Elle ne fournit néanmoins aucune information quant aux éventuelles corrections relativistes de la trajectoire des planètes car, comme nous allons le montrer plus loin, l'équation des géodésiques écrite pour une particule matérielle dans le cadre de l'approximation des champs faibles conduit *exactement* à l'équation de Newton. Il faut donc passer à une approximation d'un ordre supérieur ou envisager des solutions exactes des équations de la relativité générale. Voir par exemple C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation*, Freeman, 1973, pp. 445-447.

Posant

$$\hbar^{00} = -4\Phi/c^2, \quad (7.37)$$

il en résulte

$$(\hbar^{\mu\nu}) \approx \begin{pmatrix} -4\Phi/c^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Par conséquent,

$$\begin{aligned} (h^{\mu\nu}) &= \left(\hbar^{\mu\nu} - \frac{\hbar}{2} \eta^{\mu\nu} \right) \\ &\approx \begin{pmatrix} -2\Phi/c^2 & 0 & 0 & 0 \\ 0 & -2\Phi/c^2 & 0 & 0 \\ 0 & 0 & -2\Phi/c^2 & 0 \\ 0 & 0 & 0 & -2\Phi/c^2 \end{pmatrix} \\ &\approx (h_{\mu\nu}). \end{aligned}$$

Ainsi, il est possible de rendre compte de la théorie newtonienne de la gravitation à l'aide du formalisme géométrique de la relativité générale, via la métrique approchée

$$ds^2 \approx - \left(1 + \frac{2\Phi}{c^2} \right) c^2 dt^2 + \left(1 - \frac{2\Phi}{c^2} \right) (dx^2 + dy^2 + dz^2), \quad |\Phi/c^2| \ll 1. \quad (7.38)$$

Compte tenu de (7.35) et (7.37), les équations du champ se réduisent à

$$\nabla^2 \Phi \approx \frac{\kappa c^4}{2} \mu. \quad (7.39)$$

7.6.4 Calcul de la constante de couplage

Nous allons étudier les trajectoires des particules matérielles libres dans le cadre de la métrique (7.38). Prenons pour paramètre affine le temps propre τ et pour coordonnées $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$. On a l'équation des géodésiques

$$\frac{d^2 x^\alpha}{d\tau^2} + \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{d\tau} \frac{dx^\gamma}{d\tau} = 0 \quad (7.40)$$

avec les valeurs suivantes, calculées au premier ordre, pour les symboles de Christoffel (pas de sommation sur les indices répétés) :

$$\Gamma_{00}^0 = \frac{1}{c^2} \frac{\partial \Phi}{\partial x^0}, \quad \Gamma_{0i}^0 = \frac{1}{c^2} \frac{\partial \Phi}{\partial x^i}, \quad \Gamma_{ii}^0 = -\frac{1}{c^2} \frac{\partial \Phi}{\partial x^0}, \quad \Gamma_{ij}^0 = 0 \quad (i \neq j) \quad (7.41)$$

$$\Gamma_{00}^i = \frac{1}{c^2} \frac{\partial \Phi}{\partial x^i}, \quad \Gamma_{0i}^i = -\frac{1}{c^2} \frac{\partial \Phi}{\partial x^0}, \quad \Gamma_{0j}^i = 0 \quad (i \neq j)$$

$$\Gamma_{ki}^i = -\frac{1}{c^2} \frac{\partial \Phi}{\partial x^k} \quad (i = k \text{ ou } i \neq k), \quad \Gamma_{kk}^i = \frac{1}{c^2} \frac{\partial \Phi}{\partial x^i} \quad (i \neq k), \quad \Gamma_{jk}^i = 0 \quad (i \neq j, i \neq k, j \neq k).$$

Dans le contexte des solutions quasi-statiques, on peut continuer à négliger les dérivées temporelles devant les dérivées spatiales, ce qui implique

$$\begin{aligned} |\Gamma_{ij}^0| \approx |\Gamma_{0j}^i| \approx |\Gamma_{jk}^i| &\ll |\Gamma_{00}^0| \approx |\Gamma_{ii}^0| \approx |\Gamma_{0i}^i| \\ &\ll |\Gamma_{0i}^0| \approx |\Gamma_{00}^i| \approx |\Gamma_{ki}^i| \approx |\Gamma_{kk}^i|. \end{aligned} \quad (7.42)$$

Pour retrouver la limite newtonienne, il est plus naturel d'écrire les équations des géodésiques en fonction du temps-coordonnée x^0 (qui n'est pas un paramètre affine) : puisque

$$\begin{aligned}\frac{d}{d\tau} &= \left(\frac{d\tau}{dx^0}\right)^{-1} \frac{d}{dx^0} \\ \frac{d^2}{d\tau^2} &= \left(\frac{d\tau}{dx^0}\right)^{-2} \frac{d^2}{(dx^0)^2} - \left(\frac{d\tau}{dx^0}\right)^{-3} \frac{d^2\tau}{(dx^0)^2} \frac{d}{dx^0},\end{aligned}$$

on a

$$\frac{d^2x^\alpha}{(dx^0)^2} + \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{dx^0} \frac{dx^\gamma}{dx^0} = \left(\frac{d\tau}{dx^0}\right)^{-1} \frac{d^2\tau}{(dx^0)^2} \frac{dx^\alpha}{dx^0}. \quad (7.43)$$

De l'équation $\alpha = 0$, on tire

$$\left(\frac{d\tau}{dx^0}\right)^{-1} \frac{d^2\tau}{(dx^0)^2} = \Gamma_{\beta\gamma}^0 \frac{dx^\beta}{dx^0} \frac{dx^\gamma}{dx^0}. \quad (7.44)$$

En injectant cette expression dans (7.43) pour $\alpha = i$, on obtient

$$\frac{d^2x^i}{(dx^0)^2} + \left(\Gamma_{\beta\gamma}^i - \Gamma_{\beta\gamma}^0 \frac{dx^i}{dx^0}\right) \frac{dx^\beta}{dx^0} \frac{dx^\gamma}{dx^0} = 0. \quad (7.45)$$

Or,

$$\begin{aligned}\left|\frac{dx^j}{dx^0} \frac{dx^k}{dx^0}\right| &= \left|\frac{V^j}{c} \frac{V^k}{c}\right| \ll \left|\frac{dx^0}{dx^0} \frac{dx^k}{dx^0}\right| = \left|\frac{V^k}{c}\right| \\ &\ll \left|\frac{dx^0}{dx^0} \frac{dx^0}{dx^0}\right| = 1.\end{aligned}$$

Le terme dominant dans la sommation sur β et sur γ est donc celui pour lequel $\beta = \gamma = 0$, étant entendu qu'il contient bien lui-même des symboles de Christoffel dominants, en l'occurrence les Γ_{00}^i , comme le montrent les inégalités (7.42). Comme

$$|\Gamma_{00}^0| \ll \left|\Gamma_{00}^i = \frac{1}{c^2} \frac{\partial\Phi}{\partial x^i}\right|, \quad (7.46)$$

l'équation i se réduit à

$$\frac{d^2x^i}{(dx^0)^2} + \frac{1}{c^2} \frac{\partial\Phi}{\partial x^i} = 0,$$

c'est-à-dire

$$\frac{d^2\vec{x}}{dt^2} = -\vec{\nabla}\Phi. \quad (7.47)$$

C'est l'équation de Newton pour une particule astreinte à se déplacer dans un champ de potentiel gravifique Φ . Telle est donc la signification de cette quantité introduite en (7.37).

Dès lors, il suffit de comparer (7.39) avec l'équation de Poisson

$$\nabla^2\Phi = 4\pi G\mu \quad (7.48)$$

pour en déduire la valeur de la constante de couplage

$$\boxed{\kappa = \frac{8\pi G}{c^4}} \quad (7.49)$$

et donc les équations d'Einstein (7.10) en présence de matière :

$$\boxed{R^{\mu\nu} - \frac{R}{2}g^{\mu\nu} = \frac{8\pi G}{c^4}T^{\mu\nu}}. \quad (7.50)$$

7.6.5 Commentaire

L'équation du mouvement (7.47) peut être lue de deux façons conceptuellement très différentes. Dans l'approche newtonienne, t désigne le temps absolu et l'espace euclidien est rapporté aux coordonnées (x, y, z) d'un référentiel inertiel. Mais d'après la relativité générale, (7.38) caractérise un espace-temps courbe dans un système de coordonnées (ct, x, y, z) forcément non inertiel. Remarquons tout d'abord que, comme la métrique reste orthogonale et isotrope spatialement, il suffit, pour se représenter intuitivement les sections spatiales, de les imaginer sous-tendues par les lignes de coordonnées inertielles de Newton dans lesquelles l'élément de longueur physique n'est pas $(dx^2 + dy^2 + dz^2)$ mais se trouve dilaté ou raccourci selon le facteur $(1 - 2\Phi/c^2)$.

Génériquement ces sections spatiales sont courbes. Pourtant, cette propriété n'intervient pas dans l'équation du mouvement. En effet, celle-ci est dominée par la quantité Γ_{00}^i , laquelle ne dépend que de g_{00} . Autrement dit, à la limite newtonienne on peut rendre compte du mouvement des particules matérielles à l'aide d'une métrique effective spatialement euclidienne :¹³

$$ds^2 = - \left(1 + \frac{2\Phi}{c^2} \right) c^2 dt^2 + (dx^2 + dy^2 + dz^2). \quad (7.51)$$

Or, justement, on se rappelle que les premières tentatives d'Einstein pour décrire l'interaction gravitationnelle reposaient sur la métrique de la fusée accélérée (chapitre 2), qui est un cas particulier de (7.51) ; car si dans (2.20) on pose $\tau = t$ et $\xi = x$, et si d'autre part on se limite aux termes du premier ordre en $1/c^2$, la métrique se ramène à

$$ds^2 = - \left(1 + \frac{2gx}{c^2} \right) c^2 dt^2 + (dx^2 + dy^2 + dz^2).$$

Le résultat (7.38) étonna donc beaucoup Einstein. Il s'en ouvrit auprès de son ami Michele Besso dans une lettre datée du 10 décembre 1915 :¹⁴

« Tu seras surpris par la présence des $g_{11} \dots g_{33}$. »,

puis dans une lettre du 21 décembre 1915 :

« [...] le plus remarquable est le fait que la théorie newtonienne *du champ* ne soit pas correcte même à l'ordre 1 (du fait de la présence de $g_{11} - g_{33}$). C'est uniquement parce que les $g_{11} \dots g_{33}$ ne figurent pas dans la première approximation des équations du mouvement d'un point que la théorie de Newton est si simple. »

On peut donc résumer les deux approches de la façon suivante. Selon Newton, le temps est absolu mais il règne, dans l'espace euclidien, un champ de gravitation $\Phi(t, x, y, z)$ dont l'effet consiste à dévier les particules de leur tendance à suivre un mouvement rectiligne et uniforme. Selon Einstein, le temps physique indiqué par les horloges dépend de leur position, par le biais de la fonction Φ : il en résulte l'espace-temps effectif (7.51), qui est courbe¹⁵. Le mouvement d'une particule matérielle libre est une géodésique déterminée par l'exigence de maximiser le temps propre de la particule entre deux événements donnés, comme nous le verrons dans la section 7.6.7. Il est instructif, et capital pour une bonne compréhension du principe d'équivalence, d'observer qu'en relativité générale le champ $\vec{g} = -\vec{\nabla} \Phi$ est identifié à un triplet de symboles de Christoffel !

13. Toutefois cette métrique simplifiée n'explique pas le mouvement des photons, comme nous le verrons dans la section 7.6.8.

14. *Albert Einstein, œuvres choisies 2*, textes choisis et présentés par F. Balibar, O. Darrigol, B. Jech et J. Stachel, Seuil, 1993, p. 172, notes 20 et 21.

15. Sauf dans le cas particulier où $\Phi = gx$.

7.6.6 Fluide parfait dans un champ de gravitation newtonien

Il est maintenant facile de retrouver l'équation d'Euler pour un fluide parfait non relativiste dans un champ de gravitation newtonien. La loi de conservation de la quantité de mouvement s'écrit

$$\begin{aligned} 0 &= T^{i\mu}{}_{;\mu} \\ &= T^{i\mu}{}_{,\mu} + \Gamma_{\alpha\mu}^i T^{\alpha\mu} + \Gamma_{\alpha\mu}^\mu T^{i\alpha}. \end{aligned}$$

Pour la métrique (7.38), le terme dominant parmi les $\Gamma_{\alpha\mu}^i T^{\alpha\mu}$ est

$$\Gamma_{00}^i T^{00} = \mu \frac{\partial \Phi}{\partial x^i}.$$

Il domine également les $\Gamma_{\alpha\mu}^\mu T^{i\alpha}$. On trouve ainsi

$$0 = \frac{\partial (\mu V^i)}{\partial t} + (\mu V^i V^j)_{,j} + P_{,i} + \mu \frac{\partial \Phi}{\partial x^i},$$

c'est-à-dire, en tenant compte de l'équation de continuité (6.17) (laquelle n'est pas modifiée à l'ordre newtonien),

$$\mu \frac{\partial \vec{V}}{\partial t} + \mu (\vec{V} \cdot \vec{\nabla}) \vec{V} + \vec{\nabla} P + \mu \vec{\nabla} \Phi = 0.$$

Si le mouvement est stationnaire (les dérivées partielles par rapport au temps s'annulent) et si le fluide est homogène ($\vec{\nabla} \mu = 0$), cette équation se réduit à

$$(\vec{V} \cdot \vec{\nabla}) (\mu \vec{V}) + \vec{\nabla} (P + \mu \Phi) = 0.$$

Remarquons que, dans ces conditions, μ est une constante spatio-temporelle. En effectuant la multiplication scalaire par \vec{V} , on obtient l'intégrale première d'énergie

$$\vec{V} \cdot \vec{\nabla} \left(\frac{1}{2} \mu V^2 + P + \mu \Phi \right) = 0.$$

La quantité $\mu V^2/2 + P + \mu \Phi$ est constante le long d'une ligne de courant : c'est la *loi de Bernoulli*.

7.6.7 L'intégrale d'action

Soit une particule matérielle de temps propre τ . La relation $ds^2 = -c^2 d\tau^2$ permet d'écrire (7.38) sous la forme

$$d\tau^2 = dt^2 \left[\left(1 + \frac{2\Phi}{c^2} \right) - \left(1 - \frac{2\Phi}{c^2} \right) \frac{dx^2 + dy^2 + dz^2}{c^2 dt^2} \right],$$

soit, en laissant τ s'écouler dans le même sens que t :

$$d\tau = dt \left[1 + \frac{\Phi}{c^2} - \frac{V^2}{2c^2} + O(1/c^4) \right]. \quad (7.52)$$

Ce résultat apporte une perspective nouvelle sur le formalisme de la gravitation classique, en permettant de réinterpréter l'intégrale d'action S à la limite newtonienne. Pour mémoire, le lagrangien relativiste d'une particule libre (c'est-à-dire, soumise uniquement à son inertie-gravitation) de masse propre m_0 est donné dans la section 3.5 :

$$L = -m_0 c^2 d\tau/d\lambda.$$

En utilisant comme paramètre λ la coordonnée temporelle t , on obtient donc

$$\begin{aligned} S &= -m_0 c^2 \int_{\tau(t_1)}^{\tau(t_2)} d\tau \\ &= -m_0 c^2 \int_{t_1}^{t_2} \left(1 + \frac{\Phi}{c^2} - \frac{V^2}{2c^2} \right) dt. \end{aligned}$$

Comme les équations du mouvement rendent extrémale l'intégrale d'action avec t_1 et t_2 fixés, on a

$$0 = \delta \int_{t_1}^{t_2} \left(\frac{1}{2} m_0 V^2 - m_0 \Phi \right) dt. \quad (7.53)$$

On reconnaît le lagrangien de la mécanique newtonienne, différence entre l'énergie cinétique et l'énergie potentielle de gravitation. Le temps propre écoulé entre les deux extrémités du mouvement est maximal si l'action est minimale.¹⁶

Il est instructif de considérer un exemple concret dans le champ de pesanteur homogène de mesure g . Soit, par exemple, le déplacement vertical, le long du demi-axe $z \geq 0$, d'une particule entre les événements $(t = 0, z = 0)$ et $(t = 2, z = 0)$.¹⁷ Pour simplifier, on se limitera, dans l'espace-temps (t, z) , aux trajectoires paraboliques de sommet $(t = 1, z = h)$, c'est-à-dire $z(t) = ht(2-t)$. Il est alors facile d'effectuer le calcul explicite du temps propre écoulé pour chacune de ces courbes : il s'avère maximal lorsque $h = g/2$, c'est-à-dire lorsque la trajectoire vérifie les équations du mouvement.

L'équation (7.52) fournit également un éclairage complémentaire dans le cas particulier des champs stationnaires où $\partial\Phi/\partial t = 0$. Comme le précise l'équation (3.20), on a

$$L(x^\alpha, \dot{x}^\alpha) = -m_0 c \sqrt{-g_{\mu\nu}[x^\alpha(\lambda)] \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}}.$$

Si donc, par hypothèse, la coordonnée temporelle t n'apparaît pas explicitement dans la métrique (7.38), les équations de Lagrange engendrent l'intégrale première

$$\left(1 + \frac{2\Phi}{c^2} \right) \frac{dt}{d\tau} = K, \quad (7.54)$$

où K est une constante le long de la trajectoire. Introduisant (7.52) dans (7.54), on trouve

$$1 + \frac{\Phi}{c^2} + \frac{V^2}{2c^2} + O(1/c^4) = K,$$

relation équivalente à l'équation de conservation de l'énergie

$$\frac{1}{2} m_0 V^2 + m_0 \Phi + m_0 c^2 = \text{constante}. \quad (7.55)$$

16. Cependant, l'extremum n'est pas toujours un minimum : voir par exemple L. Landau et E. Lifchitz, *Physique théorique, Tome 1, Mécanique*, Mir, 1982, § 2 ; C.W. Misner, K.S. Thorne, J.A. Wheeler, *op. cit.*, pp. 318-320.

17. R. Signore, *L'histoire de la chute des corps*, Vuibert, 2008.

Exercice 7.1 En utilisant la relation (7.52) dans (7.44) et en recourant aux valeurs approchées (7.41) des symboles de Christoffel, trouver la généralisation de l'équation de conservation de l'énergie (7.55) en présence d'un champ Φ non stationnaire.

Solution. On obtient

$$\frac{d}{dt} \left(\frac{V^2}{2} + \Phi \right) = \frac{\partial \Phi}{\partial t}.$$

7.6.8 Déviation des rayons lumineux par un corps sphérique

La métrique linéarisée (7.38) restitue le mouvement newtonien d'une particule test matérielle, sans lui apporter de correction relativiste due à la gravitation.¹⁸ Mais nous allons montrer qu'elle prédit un comportement différent pour la lumière, en annonçant au voisinage d'un astre sphérique une déviation de la trajectoire *double* de la valeur newtonienne.

Soit, en coordonnées sphériques (r, θ, φ) , le champ faible newtonien, solution de l'équation de Poisson (7.48), engendré par une masse M à symétrie sphérique :

$$\Phi(r) = \frac{-GM}{r}.$$

Dans le plan $\theta = \pi/2$, la métrique (7.38) prend la forme

$$ds^2 = - \left(1 - \frac{2GM}{rc^2} \right) c^2 dt^2 + \left(1 + \frac{2GM}{rc^2} \right) (dr^2 + r^2 d\varphi^2).$$

On raisonnera partout au premier ordre en GM/rc^2 . (Au voisinage du disque solaire, cette quantité est de l'ordre de $2 \times 10^{-6} \ll 1$.)

Choisissons un paramètre affine τ et notons par un point les dérivées par rapport à τ . Les trois relations suivantes déterminent complètement le mouvement de la lumière.

- La longueur de la trajectoire spatio-temporelle est nulle :

$$0 = - \left(1 - \frac{2GM}{rc^2} \right) c^2 \dot{t}^2 + \left(1 + \frac{2GM}{rc^2} \right) (\dot{r}^2 + r^2 \dot{\varphi}^2). \quad (7.56)$$

- Le lagrangien étant indépendant de t , on a l'intégrale première

$$\left(1 - \frac{2GM}{rc^2} \right) \dot{t} = l, \quad (7.57)$$

où l est une constante.

- Le lagrangien étant indépendant de φ , on a l'intégrale première

$$r^2 \left(1 + \frac{2GM}{rc^2} \right) \dot{\varphi} = h, \quad (7.58)$$

où h est une constante.

En injectant (7.57) et (7.58) dans (7.56), on trouve successivement, en négligeant les termes d'ordre deux :

$$0 = \frac{-c^2 l^2}{1 - \frac{2GM}{rc^2}} + \left(1 + \frac{2GM}{rc^2} \right) \left[\left(\frac{dr}{d\varphi} \right)^2 + r^2 \right] \dot{\varphi}^2$$

¹⁸ Elle conserve cependant les vertus de l'espace-temps de Minkowski, dont elle constitue une version enrichie, et permet de prendre en compte les corrections de relativité restreinte si la particule est dotée d'une vitesse importante dans le référentiel considéré.

$$\left[\left(\frac{dr}{d\varphi} \right)^2 + r^2 \right] \dot{\varphi}^2 = c^2 l^2$$

$$\left(\frac{dr}{d\varphi} \right)^2 = \frac{c^2 l^2}{h^2} r^4 \left(1 + \frac{4GM}{rc^2} \right) - r^2. \quad (7.59)$$

Supposons qu'asymptotiquement, avant la déviation, la lumière se propage parallèlement à l'axe des x selon la droite d'équation $y = p$. La quantité p est le paramètre d'impact de la trajectoire, c'est-à-dire la coordonnée radiale minimale qui serait atteinte si la lumière n'était pas déviée (figure 7.2).

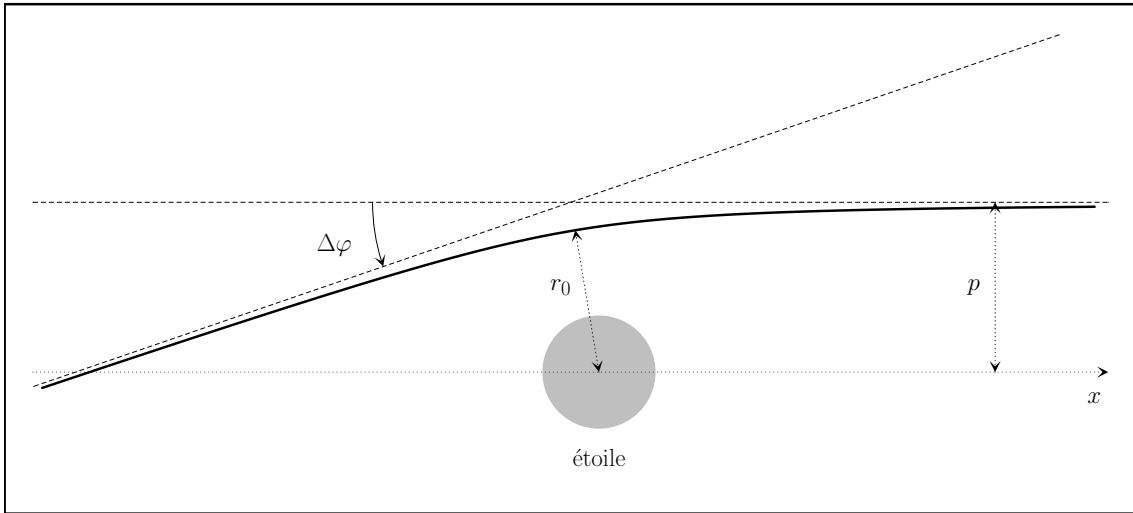


FIGURE 7.2: Déviation $\Delta\varphi$ d'un rayon lumineux passant à proximité d'une étoile sphérique.

En introduisant dans (7.59) cette condition limite, qui peut s'écrire

$$r(\varphi) = p / \sin \varphi,$$

et en passant à la limite pour r tendant vers l'infini, donc φ tendant vers zéro, on obtient

$$\frac{c^2 l^2}{h^2} = \frac{1}{p^2}.$$

L'équation (7.59) prend ainsi la forme

$$\left(\frac{dr}{d\varphi} \right)^2 = \frac{r^4}{p^2} \left(1 + \frac{4GM}{rc^2} \right) - r^2. \quad (7.60)$$

La valeur minimale de r , notée r_0 , s'obtient en annulant le membre de gauche de (7.60) :

$$r_0 = p \left(1 - \frac{2GM}{pc^2} \right).$$

Posant $r = 1/u$, l'équation (7.60) se réduit à

$$\frac{d\varphi}{du} = \pm \frac{p}{\sqrt{-p^2 u^2 + 4GMu/c^2 + 1}}.$$

Choisissant le signe positif, on obtient, toujours au premier ordre :

$$\frac{d\varphi}{du} = \frac{p}{\sqrt{1 - \left(pu - \frac{2GM}{pc^2}\right)^2}}.$$

Comme la valeur maximale de u , notée u_0 , est donnée par

$$u_0 = \frac{1}{p} \left(1 + \frac{2GM}{pc^2}\right),$$

le terme entre parenthèses dans le radicand est inférieur ou égal à l'unité et on aboutit au résultat

$$\sin[\varphi(u) + K] = pu - \frac{2GM}{pc^2}, \quad (7.61)$$

où K est une constante d'intégration.

Avant la déviation, loin du déflecteur, on a $u \rightarrow 0$ et $\varphi \rightarrow 0$, donc

$$K = -2GM/pc^2 \simeq -2GM/r_0c^2.$$

Après la collision, on a $\sin(\varphi - 2GM/r_0c^2) = -2GM/r_0c^2$, donc $\varphi = \pi + 4GM/r_0c^2$. On obtient ainsi la valeur cherchée pour l'angle de déviation :

$$\Delta\varphi = \frac{4GM}{r_0c^2}. \quad (7.62)$$

Comparons ce résultat à celui de la théorie newtonienne. Soit un corpuscule lumineux de vitesse asymptotique égale à c . Si le point désigne la dérivée par rapport au temps absolu, l'équation de conservation de l'énergie mécanique totale s'écrit

$$\frac{1}{2} (\dot{r}^2 + r^2 \dot{\varphi}^2) - \frac{GM}{r} = \frac{1}{2} c^2. \quad (7.63)$$

On dispose également de l'équation de conservation du moment cinétique

$$r^2 \dot{\varphi} = pc, \quad (7.64)$$

ce qui permet d'écrire (7.63) sous la forme

$$\left(\frac{dr}{d\varphi}\right)^2 = \frac{r^4}{p^2} \left(1 + \frac{2GM}{rc^2}\right) - r^2. \quad (7.65)$$

De la comparaison de (7.60) et (7.65), il résulte qu'en relativité générale, tout se passe comme si l'astre déflecteur était deux fois plus massif qu'en mécanique newtonienne.

Pour un rayon lumineux frôlant le disque solaire on a $r_0 = 6.95 \times 10^8$ m, ce qui donne $\Delta\varphi = 1.75$ secondes d'arc. Ce résultat, annoncé par Einstein le 18 novembre 1915, est donc le double de la valeur prédite par la théorie de Newton. Il fut vérifié lors de l'éclipse totale de soleil du 29 mai 1919, ce qui contribua largement à populariser la relativité générale et son auteur. Les mesures ont été réalisées sur plusieurs étoiles de magnitude 5 ou 6 dans le groupe des Hyades, depuis Sobral (nord du Brésil) et l'île de Principe (au large du Gabon), sous la direction générale de A. Eddington : on a trouvé respectivement $\Delta\varphi = 1.98 \pm 0.16''$ et $\Delta\varphi = 1.61 \pm 0.40''$. Depuis lors, maintes observations ont étayé ces premiers tests. Actuellement, il n'est plus nécessaire d'attendre une éclipse totale car il suffit que le soleil passe à proximité angulaire d'une radiosource.

7.7 Champ fort loin des sources

Le champ de gravitation engendré par des sources fortement relativistes ne peut être connu qu'en résolvant les équations d'Einstein exactes. Cependant, si les sources sont localisées, l'espace-temps est asymptotiquement celui de Minkowski et il est possible de connaître la solution à grande distance, c'est-à-dire dans les régions telles que l'on puisse utiliser l'approximation des champs faibles (7.25). Utilisons des coordonnées (ct, x, y, z) centrées à l'intérieur des sources et asymptotiquement cartésiennes. Raisonnons enfin dans le cadre de l'approximation quasi-statique (pas d'émission de rayonnement).

Les équations linéarisées (7.34) se réduisent à

$$\nabla^2 \bar{h}^{\mu\nu} = 0,$$

dont la solution au premier ordre est

$$\bar{h}^{\mu\nu} = \frac{A^{\mu\nu}}{r} + O(1/r^2)$$

où les $A^{\mu\nu}$ sont des constantes et $r = \sqrt{x^2 + y^2 + z^2}$. La condition de jauge de Lorenz s'écrit

$$0 = \bar{h}^{\mu\nu}{}_{,\nu} = \bar{h}^{\mu j}{}_{,j} = -\frac{A^{\mu 1}x + A^{\mu 2}y + A^{\mu 3}z}{r^3} + O(1/r^3)$$

et implique

$$A^{\mu j} = 0,$$

donc

$$|\bar{h}^{00}| \gg |\bar{h}^{0j}|, \quad |\bar{h}^{00}| \gg |\bar{h}^{ij}|.$$

Ainsi, comme dans la section 7.6.3, le champ à grande distance se présente comme un champ newtonien de potentiel

$$\Phi = -\frac{c^2 \bar{h}^{00}}{4}$$

et se trouve décrit par la métrique (7.38) :

$$ds^2 \approx -\left(1 + \frac{2\Phi}{c^2}\right) c^2 dt^2 + \left(1 - \frac{2\Phi}{c^2}\right) (dx^2 + dy^2 + dz^2). \quad (7.66)$$

Par analogie avec l'expression asymptotique du potentiel newtonien, on écrira donc

$$\Phi = -\frac{GM}{r}$$

où la constante M , déductible, par l'observation, de la troisième loi de Kepler, *définit* en fait ce que l'on appelle *la masse totale des sources relativistes*.

En résumé, la solution engendrée par des sources relativistes quasi-statiques est donnée, à grande distance, par

$$ds^2 \approx -\left(1 - \frac{2GM}{rc^2}\right) c^2 dt^2 + \left(1 + \frac{2GM}{rc^2}\right) (dx^2 + dy^2 + dz^2). \quad (7.67)$$

Il s'agit de l'approximation au premier ordre (monopolaire) de la métrique de Schwarzschild en coordonnées isotropes [équation (9.10)].

7.8 L'approximation post-newtonienne

La gravitation gravite : cette formule exprime que l'énergie d'interaction gravitationnelle contribue elle-même aux sources du champ. L'approximation post-newtonienne est une méthode permettant d'explicitier cette caractéristique, tout en autorisant le calcul des premières corrections exigées par la relativité générale quant au mouvement des particules matérielles ou des fluides parfaits.

La procédure est la suivante. On sait que l'on peut récupérer la gravitation classique en évaluant le tenseur d'énergie-impulsion des fluides parfaits dans l'espace-temps de Minkowski rapporté à des coordonnées « cartésiennes », et en ne conservant que la source dominante, $T^{00} \simeq \mu c^2$. Ce terme engendre la métrique (7.38), qui permet de retrouver l'équation de Newton pour une particule matérielle (section 7.6.4) ou un fluide parfait (section 7.6.6). Mais elle permet également de calculer l'intervalle de temps propre à l'approximation suivante (post-minkowskienne), et donc la quadrivitesse et le tenseur d'énergie-impulsion qui y sont associés. Dans les sources des équations du champ (« membre de droite ») à l'ordre suivant interviennent ainsi, non seulement les termes minkowskiens T^{00} et T^{0i} , mais également des termes spécifiques générés dans T^{00} par le champ gravitationnel newtonien. Quant au « membre de gauche », il est quasi-linéaire en la dérivée seconde de la métrique post-newtonienne. Or, à côté de ce terme se manifestent également, au même degré d'approximation, et notamment du fait de la non-linéarité, des contributions provenant de l'ordre newtonien.

Des équations ainsi obtenues on déduit le tenseur métrique au-delà de son approximation newtonienne, ce qui en retour permet d'insérer, dans l'équation de mouvement des particules matérielles ou des fluides parfaits, les premières corrections dues à la relativité générale.

Il est évidemment possible de poursuivre la logique de ce développement en forme de « bootstrap » afin de prendre en compte les corrections d'ordre supérieur (approximation *post-post-newtonienne* et suivantes).

Cet étripage des équations d'Einstein mutile aussi bien la covariance générale que l'interprétation géométrique de la gravitation. Toutefois, il possède l'avantage concret de conduire à des équations approchées gérables, au besoin avec l'aide de l'ordinateur. Bien entendu, les équations obtenues à chaque niveau d'itération sont profondément inesthétiques, entendons par là, d'apparence arbitraire. Il est difficile de justifier la présence de tel ou tel terme par des arguments un tant soit peu coercitifs. Mais ce tohu-bohu n'est que le miroir brisé d'une harmonie supérieure, perceptible seulement à hauteur du tenseur d'Einstein. Il constitue le prix à payer si l'on ne peut affronter globalement les équations du champ dans leur terrible complexité mathématique.

Passons aux détails techniques¹⁹. Notons tout d'abord qu'à chaque étape de calcul, il importera de maîtriser l'ordre de grandeur des termes négligés, lesquels seront exprimés en termes de puissances de c . Afin notamment de pouvoir comparer correctement l'énergie volumique ρ et la pression P , nous adopterons l'hypothèse qu'il est possible de séparer l'énergie volumique en deux termes :

$$\rho = \mu c^2 + \mu \Pi,$$

où μ désigne comme d'habitude la masse volumique propre et où $\mu \Pi$ est l'énergie volumique

19. S. Chandrasekhar, *Astrophysical Journal* **142**, pp. 1488 à 1512 (1965) ; C.W. Misner, K.S. Thorne, J.A. Wheeler, *op.cit.*, pp. 1089-1090.

interne, au sens de la thermodynamique ; autrement dit, Π représente l'énergie interne par unité de masse propre. Cette séparation est valide dans un grand nombre de situations.

Récrivons maintenant les équations du champ à la limite newtonienne, afin de préciser les termes négligés dans la section 7.6.3. Pour rappel, on raisonne dans l'espace-temps de Minkowski rapporté à des coordonnées cartésiennes, et l'on pose $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, où les indices de $h_{\mu\nu}$ sont élevés ou abaissés à l'aide du tenseur de Minkowski. Dans le contexte des calculs post-newtoniens, il est d'usage d'imposer un système de coordonnées tel que

$$\frac{\partial h^{0k}}{\partial x^k} + \frac{1}{2} \frac{\partial h^k_k}{\partial x^0} = O(c^{-5}), \quad \frac{\partial h^{jk}}{\partial x^k} - \frac{1}{2} \frac{\partial h^\lambda_\lambda}{\partial x^j} = O(c^{-4}). \quad (7.68)$$

Cette jauge diffère de celle de Lorenz par sa composante temporelle.²⁰ Dans l'approximation linéaire, on calcule

$$R_{00} = -\frac{1}{2} \nabla^2 h_{00}, \quad R_{0i} = -\frac{1}{2} \square h_{0i} - \frac{1}{4} \frac{\partial^2 h_{00}}{\partial x^0 \partial x^i}, \quad R_{ij} = -\frac{1}{2} \square h_{ij}. \quad (7.69)$$

À l'ordre le plus bas, le tenseur d'impulsion-énergie est donné par sa valeur dans l'espace-temps de Minkowski [équations (6.14) à (6.16)] :

$$T^{00} = \mu c^2 + O(c^0), \quad T^{0i} = c \mu V^i + O(c^{-1}), \quad T^{ij} = \mu V^i V^j + P \delta^{ij} + O(c^{-2}). \quad (7.70)$$

Cependant, au niveau des sources du champ newtonien, on peut se limiter aux informations suivantes :

$$T^{00} = \mu c^2 + O(c^0), \quad T^{0i} = O(c), \quad T^{ij} = O(c^0). \quad (7.71)$$

Ainsi, à l'ordre newtonien les équations du champ, écrites sous la forme

$$R^{\mu\nu} = \frac{8\pi G}{c^4} \left(T^{\mu\nu} - \frac{T}{2} g^{\mu\nu} \right),$$

sont

$$\begin{aligned} \nabla^2 h_{00} &= -\frac{8\pi G \mu}{c^2} + O(c^{-4}), \\ \nabla^2 h_{0i} - h_{0i,00} + \frac{1}{2} h_{00,0i} &= O(c^{-3}), \\ \nabla^2 h_{ij} - h_{ij,00} &= -\frac{8\pi G \mu}{c^2} \delta_{ij} + O(c^{-4}). \end{aligned}$$

Posons

$$h_{00} = -\frac{2\Phi}{c^2} + O(c^{-4}), \quad (7.72)$$

où, d'après la première équation, Φ est le potentiel newtonien. En n'oubliant pas que chaque dérivée par rapport à x^0 génère un facteur $1/c$, on obtient

$$h_{0i} = O(c^{-3}), \quad h_{ij} = -\frac{2\Phi}{c^2} \delta_{ij} + O(c^{-4}). \quad (7.73)$$

20. Définissant les $\tilde{h}_{\mu\nu}$ comme en (7.31), elle revient en fait à exiger

$$\tilde{h}^{0\nu}{}_{,\nu} = \frac{1}{2} h_{00,0} + O(c^{-5}) \quad \text{et} \quad \tilde{h}^{j\nu}{}_{,\nu} = O(c^{-4}).$$

Cette approche permet de préciser les termes tronqués dans la métrique (7.38) :

$$g_{00} = - \left(1 + \frac{2\Phi}{c^2} \right) + O(c^{-4}), \quad g_{0i} = O(c^{-3}), \quad g_{ij} = \left(1 - \frac{2\Phi}{c^2} \right) \delta_{ij} + O(c^{-4}). \quad (7.74)$$

Et c'est ici qu'on en arrive au niveau post-newtonien. L'équation (7.74) permet de calculer l'intervalle de temps propre $d\tau$ [c'est le résultat (7.52)], et donc la quadrivitesse U^α et le tenseur d'énergie-impulsion à l'approximation suivant (7.70) :

$$T^{00} = \mu c^2 + \mu \left(-2\Phi + V^2 + \Pi \right) + O(c^{-2})$$

$$T^{0i} = c \mu V^i + \frac{1}{c} \mu V^i \left(-2\Phi + V^2 + \Pi + P/\mu \right) + O(c^{-3})$$

$$T^{ij} = \mu V^i V^j + P \delta^{ij} + \frac{1}{c^2} \left[\mu V^i V^j \left(-2\Phi + V^2 + \Pi + P/\mu \right) + 2P\Phi \delta^{ij} \right] + O(c^{-4}).$$

A l'approximation post-newtonienne il faut développer les sources jusqu'à l'ordre en $1/c^2$ au-delà de (7.71) :

$$T^{00} = \mu c^2 + \mu \left(-2\Phi + V^2 + \Pi \right) + O(c^{-2}) \quad (7.75)$$

$$T^{0i} = c \mu V^i + O(c^{-1}) \quad (7.76)$$

$$T^{ij} = \mu V^i V^j + P \delta^{ij} + O(c^{-2}). \quad (7.77)$$

Comme la constante de couplage est en $1/c^4$, les composantes ci-dessus permettent d'écrire l'équation pour R_{00} jusqu'aux termes en c^{-4} inclus et celle pour R_{0i} jusqu'aux termes en c^{-3} inclus. On en déduit g_{00} jusqu'à l'ordre en c^{-4} inclus et g_{0i} jusqu'à l'ordre en c^{-3} inclus. Ces informations s'avèrent suffisantes pour calculer les symboles de Christoffel avec la précision requise pour les équations post-newtoniennes du mouvement. En fait, ces dernières n'exigent pas la connaissance des g_{ij} au-delà des termes en c^{-2} . (On remarque un phénomène semblable au stade de l'approximation newtonienne, où pour étudier le mouvement il suffit de connaître les g_{ij} à l'ordre minkowskien en c^0 .) Comme ces termes sont déjà connus, il est inutile de raffiner le calcul de R_{ij} . Cependant, les T^{ij} interviennent dans la trace T , et donc dans les sources pour R_{00} , au même titre que d'autres termes en c^0 : c'est pourquoi les composantes spatiales doivent malgré tout être connues jusqu'à l'ordre en c^0 inclus.

Concrètement on obtient, toujours dans la jauge (7.68) :

$$R_{00} = \nabla^2 \left(-\frac{1}{2} h_{00} - \frac{\Phi^2}{c^4} \right) + \frac{4}{c^4} \Phi \nabla^2 \Phi + O(c^{-6}). \quad (7.78)$$

Quant à R_{0i} , il résulte de (7.69), (7.72) et (7.73) :

$$R_{0i} = -\frac{1}{2} \nabla^2 h_{0i} + \frac{1}{2c^3} \frac{\partial^2 \Phi}{\partial t \partial x^i} + O(c^{-5}). \quad (7.79)$$

Utilisant (7.74) et (7.75) à (7.77), on calcule

$$T_{00} - \frac{T}{2} g_{00} = \frac{\mu c^2}{2} + \mu \left(\Phi + V^2 + \frac{\Pi}{2} + \frac{3P}{2\mu} \right) + O(c^{-2}).$$

En tenant compte de l'équation de Poisson $\nabla^2\Phi = 4\pi G\mu$, la connaissance (7.78) de R_{00} fournit ainsi l'équation pour h_{00} à l'ordre post-newtonien :

$$\nabla^2 \left(-\frac{1}{2} h_{00} - \frac{\Phi}{c^2} - \frac{\Phi^2}{c^4} \right) = \frac{8\pi G\mu}{c^4} \left(-\Phi + V^2 + \frac{\Pi}{2} + \frac{3P}{2\mu} \right) + O(c^{-6}).$$

Posant

$$\nabla^2\Psi = -4\pi G\mu \left(-\Phi + V^2 + \frac{\Pi}{2} + \frac{3P}{2\mu} \right)$$

et supposant que h_{00} , Φ , μ et P s'annulent à l'infini, on en déduit

$$h_{00} = -\frac{2\Phi}{c^2} + \frac{2}{c^4} (2\Psi - \Phi^2) + O(c^{-6}). \quad (7.80)$$

La théorie des distributions permet d'exprimer Φ et Ψ sous la forme d'une intégrale sur les sources :

$$\begin{aligned} \Phi(t, \vec{x}) &= -G \int \frac{\mu(t, \vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}' \\ \Psi(t, \vec{x}) &= G \int \frac{(-\mu\Phi + \mu V^2 + \mu\Pi/2 + 3P/2)(t, \vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}'. \end{aligned}$$

On calcule de même

$$T_{0i} - \frac{T}{2} g_{0i} = -c\mu V^i + O(c^{-1}).$$

Ce résultat, associé à la connaissance (7.79) de R_{0i} , fournit l'équation pour h_{0i} à l'ordre post-newtonien :

$$-\frac{1}{2} \nabla^2 h_{0i} + \frac{1}{2c^3} \frac{\partial^2 \Phi}{\partial t \partial x^i} = \frac{-8\pi G\mu V^i}{c^3} + O(c^{-5}).$$

Posant

$$\nabla^2 \chi = 2\Phi \quad \text{et} \quad \nabla^2 X_i = -4\pi G\mu V^i,$$

et supposant que h_{0i} , Φ et μ s'annulent à l'infini, on obtient

$$h_{0i} = \frac{1}{c^3} \left(\frac{1}{2} \frac{\partial^2 \chi}{\partial t \partial x^i} - 4 X_i \right) + O(c^{-5}).$$

Or, on peut montrer que

$$X_i(t, \vec{x}) = G \int \frac{\mu(t, \vec{x}') V^i(t, \vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}'$$

$$\chi(t, \vec{x}) = -G \int \mu(t, \vec{x}') |\vec{x} - \vec{x}'| d\vec{x}', \quad \frac{\partial^2 \chi}{\partial t \partial x^i} = X_i - W_i + O(c^{-2}),$$

où

$$W_i(t, \vec{x}) = G \int \mu(t, \vec{x}') \left[(\vec{x} - \vec{x}') \cdot \vec{V}(t, \vec{x}') \right] \frac{x^i - x'^i}{|\vec{x} - \vec{x}'|^3} d\vec{x}'.$$

Par conséquent,

$$h_{0i} = -\frac{1}{2c^3} (W_i + 7X_i) + O(c^{-5}). \quad (7.81)$$

L'expression de la métrique à l'ordre post-newtonien résulte de (7.80), (7.81) et (7.73) :

$$g_{00} = -1 - \frac{2\Phi}{c^2} + \frac{2}{c^4} (2\Psi - \Phi^2) + O(c^{-6}), \quad (7.82)$$

$$g_{0i} = -\frac{1}{2c^3} (W_i + 7X_i) + O(c^{-5}), \quad (7.83)$$

$$g_{ij} = \left(1 - \frac{2\Phi}{c^2}\right) \delta_{ij} + O(c^{-4}). \quad (7.84)$$

Elle permet de calculer les symboles de Christoffel intervenant dans les équations du mouvement pour une particule matérielle, ou dans celles $T^{\mu\nu}_{;\nu} = 0$ d'un fluide parfait, équations que l'on peut ainsi développer jusqu'aux termes en $1/c^2$ inclus.

7.9 La tentation de l'espace absolu

La cascade d'approximations et de réinterprétations permettant le passage de la relativité générale à la théorie newtonienne de la gravitation illustre clairement les immenses différences conceptuelles entre les deux approches. Or, toute la mécanique newtonienne repose sur l'existence supposée d'un espace absolu. A la lumière de la relativité générale, il est intéressant de récapituler en quelques mots comment cette illusion a pu s'imposer pendant des siècles.

L'invariance des lois constitue un élément de réponse à l'épineuse question du *réalisme* en physique. En effet, elle exprime une convergence des points de vue sur une situation donnée, un peu comme la convergence des impressions fournies par nos cinq sens contribue à nous suggérer la « réalité » de tel ou tel phénomène. La relativité restreinte admet l'invariance des lois par rapport au groupe de Lorentz mais il ne s'agit là, en définitive, que d'un *principe de symétrie* reflétant les isométries de l'espace-temps absolu de Minkowski. La relativité générale va plus loin. Est qualifié de « fait objectif » toute coïncidence spatio-temporelle entre deux événements (une aiguille en face d'une graduation sur un appareil de mesure, par exemple). Pour Einstein, qui à plusieurs reprises s'est exprimé sur ces questions fondamentales, ces coïncidences constituaient d'ailleurs les seuls « faits » auxquels la science pût prétendre. Or, la « réalité » d'une telle coïncidence ne peut être créditée que si elle est maintenue lors d'une transformation *quelconque* de coordonnées. C'est pourquoi le principe de covariance générale a joué un rôle essentiel de fil conducteur dans l'édification de la relativité générale.

Dans un premier temps, tâchons de comprendre dans quelle mesure les lois de la mécanique newtonienne peuvent être qualifiées de covariantes.

7.9.1 L'équation du mouvement en mécanique newtonienne

Le concept d'espace-temps n'est évidemment pas l'apanage de la physique relativiste. En associant le temps absolu t aux trois coordonnées spatiales cartésiennes \vec{x} , on crée un ensemble de coordonnées spatio-temporelles $(x^\alpha) = (t, \vec{x})$. Leurs différentielles constituent un quadrivecteur, que nous noterons $d\vec{r}$, pour les transformations de coordonnées

$$\underline{t} = t, \quad \underline{\vec{x}} = \vec{f}(t, \vec{x}) \quad (7.85)$$

(le groupe plus général où l'on changerait la coordonnée temporelle n'a pas d'intérêt physique dans le cadre de la mécanique classique). On définit une quadrivitesse $\vec{U} = d\vec{r}/dt$

de composantes $(1, \vec{V})$ où \vec{V} est la vitesse spatiale $d\vec{x}/dt$, et une quadriaccélération $d\vec{U}/dt$ de composantes $(0, \vec{a})$ où \vec{a} est l'accélération $d^2\vec{x}/dt^2$. Enfin, à partir de la force \vec{F} , elle-même vectorielle pour les transformations $\underline{x} = \vec{f}(t, \vec{x})$ sur les sections spatiales à temps t constant, on crée un quadrivecteur \vec{K} de composantes $(0, \vec{F})$.

En coordonnées cartésiennes, l'équation fondamentale de la dynamique $\vec{F} = m\vec{a}$ prend ainsi la forme spatio-temporelle ($\alpha = 0, 1, 2, 3$)

$$K^\alpha = m \frac{d^2 x^\alpha}{dt^2}. \quad (7.86)$$

Elle peut également s'écrire

$$\vec{K} = m \nabla_{\vec{U}} \vec{U}, \quad (7.87)$$

qui est clairement tensorielle dans l'espace-temps pour le groupe des transformations du type (7.85). La coordonnée temporelle t joue le rôle de paramètre affine pour le membre de droite, qui n'est autre que le terme géodésique. Dans un système de coordonnées générales \underline{x}^α , l'équation (7.87) s'écrit

$$\underline{K}^\alpha = m \left(\frac{d^2 \underline{x}^\alpha}{dt^2} + \Gamma_{\mu\nu}^\alpha \frac{d\underline{x}^\mu}{dt} \frac{d\underline{x}^\nu}{dt} \right). \quad (7.88)$$

Dans cette formulation, les accélérations fictives d'inertie sont cachées dans les symboles de connexion affine. Rappelons leur loi de transformation (5.46) :

$$\Gamma_{\xi\nu}^\rho = \frac{\partial^2 x^\pi}{\partial \underline{x}^\xi \partial \underline{x}^\nu} \frac{\partial \underline{x}^\rho}{\partial x^\pi} + \frac{\partial x^\sigma}{\partial \underline{x}^\xi} \frac{\partial x^\varphi}{\partial \underline{x}^\nu} \frac{\partial \underline{x}^\rho}{\partial x^\pi} \Gamma_{\sigma\varphi}^\pi. \quad (7.89)$$

Cette relation montre que les $\Gamma_{\mu\nu}^0$ sont tous nuls ; donc, comme annoncé ci-dessus, dans tous les référentiels la relation $\alpha = 0$ est bien l'identité $0 = 0$.

A titre d'exemple, considérons, dans le plan rapporté à des coordonnées cartésiennes (x, y) , un système de coordonnées cartésiennes $(\underline{x}, \underline{y})$ d'origine C et de vitesse angulaire constante ω , et donc reliées aux coordonnées précédentes par le changement de variables

$$\begin{pmatrix} t \\ \underline{x} \\ \underline{y} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\omega t) & \sin(\omega t) \\ 0 & -\sin(\omega t) & \cos(\omega t) \end{pmatrix} \begin{pmatrix} t \\ x - x_C(t) \\ y - y_C(t) \end{pmatrix}.$$

On calcule les matrices jacobiniennes (α est l'indice de ligne et β l'indice de colonne) :

$$\left(\frac{\partial x^\alpha}{\partial \underline{x}^\beta} \right) = \begin{pmatrix} 1 & 0 & 0 \\ \dot{x}_C - \omega \underline{x} \sin(\omega t) - \omega \underline{y} \cos(\omega t) & \cos(\omega t) & -\sin(\omega t) \\ \dot{y}_C + \omega \underline{x} \cos(\omega t) - \omega \underline{y} \sin(\omega t) & \sin(\omega t) & \cos(\omega t) \end{pmatrix},$$

$$\left(\frac{\partial \underline{x}^\alpha}{\partial x^\beta} \right)_{\vec{x}(\vec{x})} = \begin{pmatrix} 1 & 0 & 0 \\ -\dot{x}_C \cos(\omega t) - \dot{y}_C \sin(\omega t) + \omega \underline{y} & \cos(\omega t) & \sin(\omega t) \\ \dot{x}_C \sin(\omega t) - \dot{y}_C \cos(\omega t) - \omega \underline{x} & -\sin(\omega t) & \cos(\omega t) \end{pmatrix}.$$

A partir de ces informations, on vérifie que

$$\begin{aligned} \vec{U} &= \vec{e}_t + \dot{x} \vec{e}_x + \dot{y} \vec{e}_y \\ &= \underline{\vec{e}}_t + \underline{\dot{x}} \underline{\vec{e}}_x + \underline{\dot{y}} \underline{\vec{e}}_y. \end{aligned}$$

D'autre part,

$$\dot{x} \vec{e}_x + \dot{y} \vec{e}_y = \dot{x}_C \vec{e}_x + \dot{y}_C \vec{e}_y + \underline{\dot{x}} \vec{e}_x + \underline{\dot{y}} \vec{e}_y - \omega \underline{y} \vec{e}_x + \omega \underline{x} \vec{e}_y.$$

Cette relation constitue la loi de Poisson pour les vitesses : posant

$$\vec{\omega} = \omega \vec{e}_x \wedge \vec{e}_y,$$

elle peut en effet s'écrire

$$\vec{V} = \vec{V}_C + \vec{V} + \vec{\omega} \wedge \underline{\vec{x}}.$$

La relation (7.89) permet de calculer les symboles de connexion affine des coordonnées \underline{x}^α : les seules quantités non nulles sont

$$\begin{aligned} \underline{\Gamma}_{00}^1 &= \ddot{x}_C \cos(\omega t) + \ddot{y}_C \sin(\omega t) - \omega^2 \underline{x}, \\ \underline{\Gamma}_{00}^2 &= -\ddot{x}_C \sin(\omega t) + \ddot{y}_C \cos(\omega t) - \omega^2 \underline{y}, \\ \underline{\Gamma}_{01}^2 &= \underline{\Gamma}_{10}^2 = \omega, \\ \underline{\Gamma}_{02}^1 &= \underline{\Gamma}_{20}^1 = -\omega. \end{aligned}$$

Les équations (7.88) prennent ainsi la forme

$$\begin{aligned} \underline{\ddot{x}} + \ddot{x}_C \cos(\omega t) + \ddot{y}_C \sin(\omega t) - \omega^2 \underline{x} - 2\omega \underline{\dot{y}} &= \underline{F}^1/m \\ \underline{\ddot{y}} - \ddot{x}_C \sin(\omega t) + \ddot{y}_C \cos(\omega t) - \omega^2 \underline{y} + 2\omega \underline{\dot{x}} &= \underline{F}^2/m. \end{aligned}$$

A côté de l'accélération relative au système en rotation, apparaissent les accélérations fictives d'entraînement, centrifuge et de Coriolis.

7.9.2 Sur la covariance de la mécanique newtonienne

Les symboles de connexion affine $\Gamma_{\mu\nu}^\alpha$ sont des quantités non tensorielles, ectoplasmes susceptibles d'apparaître et de disparaître au gré des transformations de coordonnées auxquelles ils sont attachés. Cependant, *ils se transforment comme des tenseurs sous l'effet de transformations linéaires de coordonnées*, et seulement sous cette condition [relation (7.89)]. Ils vérifient donc le critère de « réalité » pour les transformations linéaires. Plus particulièrement, s'ils sont nuls dans un système de coordonnées, ils le restent dans tout autre système obtenu à partir du premier par une transformation linéaire : l'équation du mouvement y conserve la forme $K^\alpha = m\ddot{x}^\alpha$. Newton légitime cette propriété en inventant le concept de *référentiel inertiel* : la nature distinguerait une classe de référentiels liés les uns aux autres par des transformations *linéaires* de coordonnées et dont le statut privilégié ne peut se justifier qu'en appelant à la rescousse la notion métaphysique d'espace absolu. Est en effet réputé « inertiel » tout référentiel en translation rectiligne et uniforme par rapport à cet espace absolu.²¹

Ainsi, Newton a mis l'accent sur la covariance ou non d'un morceau seulement de l'équation (7.88), à savoir le terme en \ddot{x}^α . Historiquement donc, sa mécanique fut myope.

21. Ce relent d'invariance des lois de la mécanique (covariance par rapport aux transformations de Galilée) rend en lui-même ces idées assez suspectes : on s'attendrait plutôt à ce que seuls les référentiels *immobiles* par rapport à un espace absolu homogène et isotrope présentent entre eux des similitudes. A cet égard, l'absence de covariance des équations de Maxwell pour le groupe de Galilée constitue un argument historiquement important en faveur de l'espace absolu, et l'on comprend mieux les enjeux liés à la recherche de l'éther au XIX^e siècle.

Elle n'a pas perçu les symboles de connexion, lesquels se transforment comme le tenseur nul dans les référentiels inertiels. Pourtant, d'un point de vue conceptuel, c'est-à-dire en prenant suffisamment de recul mathématique, les équations du mouvement ne sont pas plus simples dans un système de coordonnées où s'annulent les symboles de connexion, que dans n'importe quel autre : elles ne font jamais que traduire la même propriété géométrique $\vec{K} = m \nabla_{\vec{U}} \vec{U}$.

Cependant, Newton eût-il connu la géométrie différentielle qu'il n'aurait pas modifié son interprétation. En effet, bien que les équations (7.88) aient la même forme dans *tous* les systèmes de coordonnées, on ne peut les qualifier de *généralement covariantes*. La raison en est la suivante. Les symboles de connexion affine ne peuvent être évalués dans un système de coordonnées que par référence à ce que l'on connaît déjà dans un autre système de coordonnées : au travers de son premier terme, la loi de transformation (7.89) fait intervenir explicitement le changement de coordonnées. L'interdépendance des référentiels est donc patente et il est naturel de s'attacher à un référentiel premier à partir duquel les calculs pourront être effectués dans n'importe quel autre. La notion d'espace absolu était incontournable.

La restriction de la covariance aux référentiels inertiels apparaît clairement dans le formalisme lagrangien : la mécanique newtonienne est incapable de fournir un lagrangien scalaire pour des transformations quelconques de coordonnées.²²

7.9.3 Sur la covariance de la relativité restreinte

Quelle nouveauté la théorie de la relativité restreinte introduit-elle ? La réponse tient en une phrase : *elle fait usage d'un espace-temps métrique, l'espace-temps de Minkowski*. A première vue, l'équation du mouvement ne s'y distingue pas trop de son homologue newtonien, si ce n'est que le temps absolu t a été remplacé par le temps propre τ de la particule :

$$K^\alpha = m_0 \left(\frac{d^2 x^\alpha}{d\tau^2} + \Gamma_{\mu\nu}^\alpha \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \right). \quad (7.90)$$

Pourtant, cette équation est maintenant généralement covariante. Certes, les symboles de Christoffel ne sont toujours pas des tenseurs ; mais désormais ils peuvent s'exprimer en fonction du tenseur métrique, par la relation (5.77) des connexions riemanniennes :

$$\Gamma_{\beta\mu}^\delta = \frac{1}{2} g^{\delta\alpha} (g_{\alpha\beta,\mu} + g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha}). \quad (7.91)$$

Ainsi, le terme géodésique *complet*, donc le membre de droite de (7.90), peut être calculé « de l'intérieur », c'est-à-dire intrinsèquement à partir de la valeur des composantes du tenseur métrique dans chaque système de coordonnées. Et le lagrangien du mouvement est maintenant un invariant scalaire, c'est-à-dire « réel » (pour les particules libres, il s'agit essentiellement du temps propre).

Malgré ces différences, Einstein et Newton sont d'accord pour reconnaître l'existence de certains systèmes de coordonnées où les équations du mouvement prennent une forme techniquement moins lourde. Or, les mathématiques connues à l'époque d'Einstein ont

22. Le lagrangien d'une particule libre n'est autre que son énergie cinétique $mV^2/2$. A strictement parler, il n'est même pas invariant pour le groupe de Galilée. Cependant, le passage d'un système inertiel à un autre ne fait qu'ajouter au lagrangien initial une dérivée totale par rapport au temps : ce terme supplémentaire est donc invisible lorsque l'on calcule les variations de l'action. On réalise que l'invariance des lois de la mécanique par rapport au groupe de Galilée est, en quelque sorte, accidentelle.

permis de hisser le débat au niveau des propriétés géométriques de l'espace-temps lui-même. En effet, elles enseignent que, dans un espace euclidien, et dans un espace euclidien seulement, il est toujours possible de trouver des coordonnées annulant les symboles de connexion affine (théorème 5.9). Si la variété est riemannienne et la métrique réduite à la forme canonique, ces coordonnées sont interprétées comme référentiel inertiel et qualifiées de *cartésiennes*. On voit que les notions de référentiel inertiel et d'espace euclidien sont intimement corrélées.

Ainsi, bien que la relativité restreinte fasse un pas essentiel dans la direction d'une physique généralement covariante, la notion de référentiel inertiel l'imprègne tout autant que la mécanique newtonienne. Les deux théories se cramponnent finalement à la même idée : un espace absolu pour Newton, et son avatar, l'espace-temps absolu de Minkowski pour Einstein.²³ Et comme le parallélisme demeure absolu, on peut, sans ambiguïté, comparer les composantes d'un tenseur défini en un point avec celles d'un tenseur défini en n'importe quel autre point ; ce qui préserve l'idée d'une *physique globale* à laquelle la mécanique newtonienne nous a accoutumés avec ses référentiels rigides déployés partout.

En résumé, la révolution démocratique opérée par la relativité restreinte ne doit pas voiler le maintien au pouvoir d'un monarque absolu de droit divin : l'espace-temps de Minkowski, référence ultime incarnant la loi gravée dans le marbre euclidien, despote inflexible servi par une caste de pâles courtisans, les référentiels inertiels, lesquels entendent bien conserver leurs prérogatives. Et si l'accélération spatiale est maintenant une notion relative (même au sein de la famille des référentiels inertiels), le quadrivecteur accélération d'une particule est quant à lui toujours bel et bien défini par rapport à un substrat universel et immuable.

7.9.4 Sur la covariance de la relativité générale

C'est ici que la relativité générale effectue un bond magistral et radicalise l'entreprise de chasse aux prérogatives amorcée par la relativité restreinte. Comme on le sait, son fer de lance est le *principe d'équivalence* : en condamnant la notion de référentiel inertiel, le principe d'équivalence retire à l'espace-temps de Minkowski son statut ontologiquement mystérieux et le réduit à *un cas particulier parmi une infinité d'espaces-temps possibles soumis aux équations d'Einstein*. L'espace-temps général n'est plus absolu, il est modelé par la matière. Il n'est plus euclidien que *localement*. Identifié à l'espace-temps, le champ gravitationnel impose une physique essentiellement locale, car la courbure de la variété riemannienne interdit la comparaison de tenseurs en des points distants.

Les « forces fictives d'inertie » ne décrivent plus qu'un champ de gravitation particulier, celui de l'espace-temps euclidien considéré dans un système de coordonnées quelconques. Et l'équation (7.46) boucle la boucle : la relativité générale interprète le « véritable » champ de gravitation newtonien $\vec{g} = -\vec{\nabla}\Phi$ comme un triplet de symboles de Christoffel ! Remarquons une fois encore que, chez Newton, la « réalité » du champ gravitationnel, tout comme le caractère « fictif » des forces d'inertie, sont justifiés par une chimère, l'espace absolu. Chez Einstein, comme nous l'avons dit, le critère de réalité est un principe d'invariance géométrique, c'est-à-dire de tensorialité. La gravitation (au sens newtonien) et l'inertie sont de même nature et tout aussi irréels l'un que l'autre : seul le tenseur métrique, synthèse

23. En somme, c'est le fantasme d'un temps absolu qui conduit à l'illusion d'un espace-temps objectivement découpé en un espace absolu et un temps absolu. En physique relativiste, le caractère absolu du temps-coordonnée s'efface au profit de l'invariance du temps propre, et le couple « espace et temps absolus » cède la place à une seule entité, l'espace-temps absolu.

des deux, possède un sens géométrique.

En guise de conclusion, on appréciera cette réflexion d'Einstein, exprimée dans une lettre à Arnold Sommerfeld datée du 29 octobre 1912 :²⁴

« Je travaille ces temps-ci exclusivement sur le problème de la gravitation et je pense que, avec l'aide d'un ami mathématicien ici, je vais pouvoir en surmonter toutes les difficultés. Mais, une chose est sûre - jamais auparavant je n'avais été autant tourmenté par un problème. J'ai acquis un grand respect pour les mathématiques, alors qu'auparavant j'avais tendance à considérer les techniques mathématiques sophistiquées comme un luxe inutile ! En comparaison avec le problème actuel, la théorie de la relativité restreinte était un jeu d'enfants. »

24. *The Collected Papers of Albert Einstein*, Volume 5, pages 505-506, edited by Martin J. Klein, A.J. Kox and Robert Schulmann, Princeton University Press, 1993. Cité en français par Thibault Damour, *Si Einstein m'était conté*, Editions du Cherche-midi, Paris, 2012, pp. 102-103.

Chapitre 8

Les univers de Friedmann-Lemaître

8.1 Les difficultés de la cosmologie newtonienne

On peut affirmer que la cosmologie, cette discipline ayant pour ambition d'étudier l'univers dans son ensemble, n'a acquis un statut scientifique qu'avec l'avènement de la relativité générale.

Ainsi que Newton l'avait déjà remarqué, un univers fini verrait toute sa matière se concentrer en un point. Rejeter ses limites à l'infini présentait, outre l'avantage d'éviter l'épineux problème de la frontière, celui d'empêcher cet effondrement : Newton pensait qu'un univers infini et homogène pouvait se maintenir en équilibre instable en admettant seulement des condensations locales.

Cette hypothèse d'une infinité d'étoiles uniformément réparties est cependant lourde de conséquences. Depuis la Terre on devrait, en effet, rencontrer une étoile dans chaque direction de visée (une seulement si l'on suppose qu'une étoile donnée éclipe toutes celles qui sont derrière elle). L'éclairement est inversement proportionnel au carré de la distance mais il en va de même pour l'angle solide sous-tendu par l'étoile. Si l'on suppose que toutes les étoiles ont la même taille et la même brillance que le soleil, alors chaque portion du ciel devrait être aussi lumineuse que si l'astre du jour s'y trouvait ! Or, la brillance du ciel nocturne est de l'ordre de 1% de celle de la Voie lactée...

Ce paradoxe a été popularisé sous le nom de H. Olbers, qui l'a formulé en 1823, mais on le trouve déjà mentionné chez J. Kepler (1610) et J.-Ph. Loys de Chéseaux (1744). Sa résolution n'est pas aisée : de nombreuses tentatives d'explication ont été fournies au fil des siècles et il n'est même pas certain que l'on puisse prétendre actuellement à une réponse définitive¹. Il paraît cependant incontournable de devoir conclure à la *finitude*, sinon de l'univers, en tout cas de la partie qui nous en est accessible.

D'autre part, la section 2.7 suggère que les effets relativistes dans une sphère homogène de masse m et rayon R deviennent importants si $Gm/(Rc^2) \approx 1$, c'est-à-dire si

$$R \approx \sqrt{\frac{3c^2}{4\pi G\mu}}$$

où μ est la masse volumique. (Ce critère sera conforté par l'étude de la solution de Schwarzschild.) Autrement dit, la théorie newtonienne de la gravitation s'avère tout à fait inadéquate pour traiter le cas d'une sphère homogène, même de très faible densité, si son rayon

1. E. Harrison, *Le noir de la nuit, une énigme du cosmos*, Collection Points Sciences S123, Seuil, 1990.

est suffisamment grand. C'est la raison pour laquelle il faut s'attendre à ce que la discussion des questions cosmologiques relève plutôt de la relativité générale.

Or, justement, cette théorie permet non seulement d'envisager des modèles d'univers *finis* et *illimités* mais, plus généralement, *évolutifs* et donne à la notion d'*âge de l'univers* une signification positive. En effet, dans la plupart des modèles il apparaît que l'on ne peut remonter l'histoire cosmique au-delà d'une singularité - « le Big Bang » - survenue il y a un temps fini. Compte tenu de la finitude de la vitesse de la lumière, nous n'aurions effectivement ainsi accès qu'à une partie finie de l'univers, limitée par l'*horizon cosmologique*.

8.2 Quelques citations

Le thème de la cosmologie ravive souvent les convictions intimes de tout un chacun, suscitant des commentaires allant des plus modérés aux plus fantaisistes, les uns sages et les autres, parfois, arbitraires. A titre d'exemple, on trouvera ci-dessous quelques prises de position extraites, un peu au hasard, de textes dont certains sont destinés au grand public.

« L'expansion constitue la première preuve du big bang.

[...]

Mais la preuve ultime de l'évolution cosmique était connue dès 1965, l'année où on a détecté un rayonnement diffus venant de toutes les directions du ciel, le rayonnement de fond cosmologique.

[...]

A la limite, on peut imaginer que tous les univers logiquement possibles existent, dont le nôtre, en équilibre quelque part entre le néant et l'absolu. »

M. Séguin, *La cosmologie en 10 questions*, dans *Ciel et Espace* 339, août 1998, pp. 46-52.

« En tant que construction scientifique, l'univers de G. Lemaître échappe précisément à cette antinomie; la science a su concevoir une structure spatiale sur laquelle l'antinomie kantienne n'a plus de prise. »

F. Gonseth, préface du livre de G. Lemaître, *L'hypothèse de l'atome primitif*, Editions Culture et Civilisation, Bruxelles, 1972, p. 21.

« ... notre imagination est irrémédiablement euclidienne... »

G. Lemaître, *Ibidem*, pp. 51-52.

« L'Univers s'étend-il à l'infini? Est-il fini et clos? Heine a répondu à cette question dans un de ses poèmes en disant que seuls les imbéciles se la posent. »

A. Einstein, lettre à W. de Sitter, 12 mars 1917. Cité par Fang Lizhi, Li Shuxian, *La naissance de l'Univers*, InterEditions, Paris, 1990, p. 80.

« Jamais, les problèmes métaphysiques n'ont été aussi clairement posés à l'intelligence humaine. [...] Le renouveau de la métaphysique viendra de la part des scientifiques parce que eux au moins ont le point de départ correct : la réalité elle-même. Ils savent que le point de départ de la pensée rationnelle, c'est l'expérience. [...] Aujourd'hui, en cette fin du XX^e siècle, tout est à reprendre, tout est à repenser, en métaphysique, sur la base expérimentale nouvelle, à nouveaux frais [...] »

C. Tresmontant, cité par J. Demaret, *Univers. Les théories de la cosmologie contemporaine*, Le Mail, 1991, pp. 310-311.

« Considérés métaphysiquement, les objets des sens sont des phénomènes, mais pour la physique ce sont les choses en elles-mêmes qui affectent les sens. »

E. Kant, cité par P. Clavier, *Kant, Les idées cosmologiques*, Presses Universitaires de France, Paris, 1997, p. 113.

« Que veut dire « mesurer » en cosmologie? [...] Les termes employés couramment en cosmologie sont ceux d'une physique locale étendus abusivement à tout l'Univers. »

Ch. Vilain, *Genèse d'une idée : l'expansion de l'Univers*, dans *La Recherche* 284, février 1996, pp. 104-107.

« Même s'il n'existe qu'un ensemble unique de lois possibles, ce n'est qu'un ensemble d'équations. Qu'est-ce qui insuffle le feu aux équations et leur donne un univers à gouverner? L'ultime théorie unifiée est-elle si irrésistible qu'elle cause sa propre existence? Bien que la science puisse résoudre le problème de la manière dont a commencé l'Univers, elle ne pourra pas répondre à la question : pourquoi l'Univers s'est-il donné la peine d'exister? »

S. Hawking, *Trous noirs et bébés univers*, Odile Jacob, 2000, pp. 96-97.

Dans l'annexe E, on trouvera quelques pistes de réflexion relatives aux liens qu'entretiennent la physique et la philosophie, au travers de l'œuvre d'un philosophe dont la pensée nous semble toujours pertinente aujourd'hui, Emmanuel Kant.

8.3 Les modèles cosmologiques spatialement homogènes et isotropes

8.3.1 Motivation

Les modèles cosmologiques les plus simples du point de vue mathématique décrivent la géométrie d'un univers spatialement homogène et isotrope (ces deux notions feront l'objet d'une définition mathématique précise ci-dessous). C'est incontestablement à Alexandre Friedmann que revient le mérite d'en avoir, le premier, développé l'étude entre 1922 et 1924, suivi indépendamment par Georges Lemaître dès 1927.

Du point de vue observationnel, l'homogénéité spatiale de l'univers est évidemment difficile à vérifier. On ne peut, de toute façon, la postuler qu'à des échelles de distances supérieures aux dimensions caractéristiques des superamas de galaxies, de l'ordre de 100 millions d'années-lumière. On a même découvert des structures plus grandes encore, le *grand attracteur* et le *grand mur*, d'une étendue de 500 millions d'années-lumière chacun². Il ne paraît cependant pas absurde de maintenir l'hypothèse selon laquelle, à partir d'une échelle de distances de l'ordre de 1 gigaparsec, la matière de l'univers pourrait être décrite en termes de fluide homogène dont les entités de base joueraient le rôle de molécules.

Il s'avère également que les observations effectuées à de telles distances sont indépendantes de la direction et, sauf retour penaud à un géocentrisme scabreux, on peut imaginer l'univers isotrope en tout point. Or, un univers isotrope en tout point doit être homogène (l'inverse n'étant pas correct, penser aux cristaux). On peut donc étayer l'hypothèse d'homogénéité à l'aide d'un argument de type « copernicien » : l'univers est homogène s'il nous apparaît isotrope et que nous n'occupons pas de position privilégiée.

2. J. Demaret, *Univers. Les théories de la cosmologie contemporaine*, Le Mail, 1991, p. 63.

8.3.2 Notion de groupe d'isométries

Soit un groupe de déplacements agissant sur une variété riemannienne et appliquant chaque point P sur un point \underline{P} . On réserve l'appellation d'*isométries* aux transformations telles que

$$g_{\alpha\beta}(P) dx^\alpha dx^\beta = g_{\mu\nu}(\underline{P}) d\underline{x}^\mu d\underline{x}^\nu. \quad (8.1)$$

Cette définition revient à imposer

$$\begin{aligned} g_{\mu\nu}(\underline{P}) &= g_{\alpha\beta}(P) \frac{\partial x^\alpha}{\partial \underline{x}^\mu} \frac{\partial x^\beta}{\partial \underline{x}^\nu} \\ &= \underline{g}_{\mu\nu}(P) \end{aligned}$$

(on peut simuler l'effet d'une transformation de coordonnées *active* à l'aide de la même transformation envisagée du point de vue *passif*). Plus précisément :

$$g_{\mu\nu}(\underline{\vec{x}}) = \frac{\partial x^\alpha}{\partial \underline{x}^\mu} \frac{\partial x^\beta}{\partial \underline{x}^\nu} g_{\alpha\beta}[\vec{x}(\underline{\vec{x}})]. \quad (8.2)$$

Pour les transformations infinitésimales

$$\underline{x}^\alpha = x^\alpha + \epsilon \xi^\alpha(\vec{x}) \quad (\epsilon \ll 1), \quad (8.3)$$

l'équation (8.2) est équivalente à

$$\xi_{\alpha;\beta} + \xi_{\beta;\alpha} = 0. \quad (8.4)$$

Celle-ci est appelée *équation de W. Killing* (1892). Elle permet de calculer les vecteurs $\vec{\xi}$ engendrant les isométries locales d'une métrique en un point donné, eux-mêmes nommés *vecteurs de Killing*. Munis du commutateur, ils déterminent l'*algèbre de Lie* associée au groupe d'isométries considéré.

A titre d'exemple, considérons, dans l'espace euclidien à trois dimensions rapporté à des coordonnées cartésiennes, une rotation de $\pi/2$ autour de l'axe des x :

$$\underline{y} = -z, \quad \underline{z} = y.$$

Intuitivement, il est évident que cette application constitue une isométrie sur la sphère centrée à l'origine. Montrons-le par calcul. Les coordonnées sphériques se transforment comme suit :

$$\sin \underline{\theta} \sin \underline{\varphi} = -\cos \theta, \quad \cos \underline{\theta} = \sin \theta \sin \varphi,$$

qui s'inverse aisément en

$$\begin{aligned} \underline{\theta} &= \arccos(\sin \theta \sin \varphi) \\ \underline{\varphi} &= \arcsin \frac{-\cos \theta}{\sin \underline{\theta}} = \arcsin \frac{-\cos \theta}{\sqrt{1 - \sin^2 \theta \sin^2 \varphi}} \end{aligned}$$

si $\underline{\varphi}$ est situé dans le premier quadrant. Dans les deuxième et troisième quadrants on a $\underline{\varphi} = \pi - \arcsin(\dots)$ et dans le quatrième quadrant $\underline{\varphi} = 2\pi + \arcsin(\dots)$. On obtient ainsi

$$\begin{aligned} d\underline{\theta}^2 &= \frac{(\cos \theta \sin \varphi d\theta + \sin \theta \cos \varphi d\varphi)^2}{1 - \sin^2 \theta \sin^2 \varphi}, \\ \sin^2 \underline{\theta} d\underline{\varphi}^2 &= \frac{(\cos \varphi d\theta - \sin \theta \cos \theta \sin \varphi d\varphi)^2}{1 - \sin^2 \theta \sin^2 \varphi}, \end{aligned}$$

et, par conséquent,

$$d\underline{\theta}^2 + \sin^2 \underline{\theta} d\underline{\varphi}^2 = d\theta^2 + \sin^2 \theta d\varphi^2,$$

conformément à l'exigence (8.1).

8.3.3 Notion d'espace à symétrie maximale

Brossons quelques définitions importantes relatives aux symétries des variétés riemanniennes (sans entrer dans le formalisme technique de la théorie des vecteurs de Killing).

Une variété est *homogène* si elle est l'orbite d'un groupe d'isométries. Cela signifie que, partant d'un point arbitraire, on peut atteindre n'importe quel autre point au moyen d'une isométrie au moins.

La variété est *isotrope* en un point si, dans toutes les directions, il existe un sous-groupe du groupe d'isométries laissant ce point invariant. Plus précisément, mais toujours intuitivement, dans chacune des $C_2^n = n(n - 1)/2$ surfaces associées à chaque couple de coordonnées, il existe une isométrie (de type *rotation ordinaire* ou « *boost* », selon les valeurs propres du tenseur métrique) laissant ce point invariant.

Un espace isotrope en tout point et homogène est appelé *espace à symétrie maximale* (il suffit pour cela qu'il soit isotrope en un point, et homogène). Son groupe d'isométries admet au total $n(n + 1)/2$ paramètres : n degrés de liberté pour rendre compte de l'homogénéité³, plus $n(n - 1)/2$ paramètres relatifs au sous-groupe d'isotropie.

Exercice 8.1 Trouver les vecteurs de Killing de l'espace-temps de Minkowski.

Solution. Cette variété est suffisamment familière pour que l'on puisse deviner la réponse : en coordonnées cartésiennes $\{t, x, y, z\}$ (on a posé $c = 1$), les isométries sont les translations le long des quatre axes, les rotations spatiales autour des trois axes spatiaux et les « boosts » selon les trois axes spatiaux ; soit dix vecteurs en tout. Il ne peut y en avoir davantage : nous avons affaire à un espace à symétrie maximale. Cherchons à les écrire explicitement.

Une translation infinitésimale de ϵ mètres-lumière dans le temps s'exprime par la loi $\underline{t} = t + \epsilon$, $\underline{x}^i = x^i$. On en déduit un vecteur de Killing de composantes $\xi^\alpha = \delta_0^\alpha$, c'est-à-dire le vecteur $\underline{\xi} = \partial_t$. Les trois autres translations fournissent respectivement $\partial_x, \partial_y, \partial_z$.

Une rotation spatiale d'angle ϵ autour de l'axe des x se traduit par

$$\begin{pmatrix} \underline{y} \\ \underline{z} \end{pmatrix} = \begin{pmatrix} \cos \epsilon & -\sin \epsilon \\ \sin \epsilon & \cos \epsilon \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} \simeq \begin{pmatrix} 1 & -\epsilon \\ \epsilon & 1 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}$$

(les coordonnées t et x restant inchangées). Il en résulte le vecteur $y \partial_z - z \partial_y$. Les deux autres rotations donnent respectivement $z \partial_x - x \partial_z$ et $x \partial_y - y \partial_x$.

Enfin, un « boost » de paramètre ϵ le long de l'axe des x se traduit par

$$\begin{pmatrix} \underline{t} \\ \underline{x} \end{pmatrix} = \begin{pmatrix} \cosh \epsilon & \sinh \epsilon \\ \sinh \epsilon & \cosh \epsilon \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} \simeq \begin{pmatrix} 1 & \epsilon \\ \epsilon & 1 \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix}$$

(les coordonnées y et z restant inchangées). Il en résulte le vecteur $x \partial_t + t \partial_x$. Les deux autres « boosts » donnent respectivement $y \partial_t + t \partial_y$ et $z \partial_t + t \partial_z$.

Ces dix vecteurs de Killing engendrent l'algèbre de Lie associée au *groupe de Poincaré*. Son sous-groupe d'isotropie, constitué des rotations spatiales et des « boosts » (et, d'un point de vue physique, lié au principe de relativité restreinte), est le groupe de Lorentz.

□

3. Cela ne signifie nullement que le groupe d'isométries contient un « sous-groupe d'homogénéité » à n paramètres. Par exemple, la sphère commune est l'orbite du groupe des rotations $SO(3)$, à trois paramètres, mais ce dernier n'admet aucun sous-groupe à deux paramètres.

8.3.4 Espaces avec sous-espaces à symétrie maximale

Nous admettrons le théorème suivant⁴ :

Théorème 8.1 (Weinberg) *Soit une variété riemannienne de dimension n décomposable en sous-espaces à symétrie maximale de dimension m . Appelons $\{v^a\}$ ($a=1 \dots n-m$) les coordonnées désignant chacun de ces sous-espaces, et $\{u^i\}$ ($i=1 \dots m$) les coordonnées à l'intérieur de ces sous-espaces. Il est alors possible de choisir les $\{u^i\}$ pour que l'élément de longueur à n dimensions se décompose comme suit :*

$$ds^2 = g_{ab}(v) dv^a dv^b + f(v) \tilde{g}_{ij}(u) du^i du^j, \quad (8.5)$$

où $\tilde{g}_{ij}(u)$ est la métrique d'un espace à symétrie maximale de dimension m .

□

Considérons le cas particulier d'un espace-temps à quatre dimensions ($n = 4$) décomposable en hypersurfaces spatiales homogènes et isotropes ($m = 3$) : on peut alors choisir les coordonnées pour que

$$ds^2 = g(v) dv^2 + f(v) \tilde{g}_{ij}(u) du^i du^j,$$

où $g(v) < 0$. Posant $\sqrt{-g(v)} dv = c dt$ et $f(v) = R^2(t)$ on a ainsi

$$ds^2 = -c^2 dt^2 + R^2(t) \tilde{g}_{ij}(u) du^i du^j. \quad (8.6)$$

Connaître la métrique de l'espace-temps revient à établir la valeur des \tilde{g}_{ij} , c'est-à-dire la métrique sur une variété spatiale homogène et isotrope à trois dimensions.

8.3.5 Les variétés spatiales homogènes et isotropes à trois dimensions

Comme ces espaces sont partout isotropes, ils le sont notamment autour d'un point donné arbitraire. Cela permet d'envisager l'espace comme un empilement de sphères, puisque chacun de ses points est associé à une 2-sphère de points équivalents. Or, la sphère constitue un espace à symétrie maximale. Donc, d'après le théorème 8.1, il existe des coordonnées θ, φ telles que l'élément de longueur prenne la forme

$$dl^2 = dr'^2 + f(r') (d\theta^2 + \sin^2 \theta d\varphi^2) \quad (8.7)$$

où $f(r') > 0$. Les lignes de coordonnée r' (correspondant à des valeurs données θ_0 et φ_0) sont orthogonales aux sphères [intuitivement, on peut faire rouler chaque sphère sur elle-même de façon à aligner les points (θ_0, φ_0)]. Il est alors commode d'utiliser la variable radiale $r = \sqrt{f(r')}$ afin que la mesure des sphères prenne la valeur habituelle $4\pi r^2$. On a ainsi

$$dl^2 = e^{2\lambda(r)} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2).$$

Cela étant, si l'espace est homogène sa courbure scalaire est forcément constante :

$${}^{(3)}R = \frac{2}{r^2} \left[1 - \frac{d}{dr} (r e^{-2\lambda(r)}) \right] = K$$

4. S. Weinberg, *Gravitation and Cosmology*, John Wiley & Sons, 1972, chapitre 13, section 5, p. 395.

où K est une constante réelle. Cette condition détermine la fonction $\lambda(r)$:

$$e^{2\lambda(r)} = \frac{1}{1 - \frac{K r^2}{6} + \frac{A}{r}} \quad ,$$

A étant une constante de primitivation que l'on doit annuler ⁵. Enfin, il est loisible de renormaliser la variable r pour que K prenne les valeurs canoniques $k = 1, 0, -1$. Ainsi,

$$dl^2 = R^2 \left[\frac{dr^2}{1 - k r^2} + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2) \right] \quad (8.8)$$

où R est une constante que l'on peut absorber dans la fonction $R(t)$ de (8.6).

Si $k = 1$, alors r doit être restreint à l'intervalle $[0, 1]$ (nous allons voir ci-dessous que l'on peut lever la singularité apparente de la métrique en $r = 1$ par un changement de coordonnées). Si $k = 0$ ou $k = -1$, alors $r \in [0, +\infty[$.

Jusqu'ici nous n'avons traduit que des conditions nécessaires. Montrons à présent qu'elles sont suffisantes, c'est-à-dire que les trois espaces décrits par (8.8) sont effectivement homogènes et isotropes en tout point.

Si $k = 0$, il s'agit de l'espace euclidien \mathbb{R}^3 rapporté aux coordonnées sphériques. Posant, pour uniformiser les notations, $r = \chi \in [0, +\infty[$, on a en effet

$$dl^2 = R^2 \left[d\chi^2 + \chi^2 (d\theta^2 + \sin^2 \theta d\varphi^2) \right]. \quad (8.9)$$

Manifestement homogène et isotrope, cette variété est l'orbite du groupe euclidien $E(3)$.

Si $k = 1$, le changement de variable $r = \sin \chi$ permet d'écrire l'élément de longueur sous la forme

$$dl^2 = R^2 \left[d\chi^2 + \sin^2 \chi (d\theta^2 + \sin^2 \theta d\varphi^2) \right]. \quad (8.10)$$

On reconnaît la métrique intrinsèque sur la sphère ⁶ S^3 d'équation $u^2 + x^2 + y^2 + z^2 = R^2$, plongée dans l'espace euclidien de métrique $ds^2 = du^2 + dx^2 + dy^2 + dz^2$. Pour s'en convaincre, il suffit de poser

$$u = R \cos \chi, \quad x = R \sin \chi \sin \theta \cos \varphi, \quad y = R \sin \chi \sin \theta \sin \varphi, \quad z = R \sin \chi \cos \theta.$$

Homogène et isotrope comme toutes les sphères, celle-ci est l'orbite du groupe $O(4)$.

5. La longueur physique d'une petite circonférence centrée à l'origine et de coordonnée radiale constante $\mathcal{R} \ll 1$ vaut $2\pi\mathcal{R}$, tandis que la mesure physique de son rayon donne $\mathcal{R} \sqrt{g_{rr}(r=0)}$ (si les coordonnées sont régulières en $r = 0$). L'exigence d'un espace localement euclidien en $r = 0$ impose donc

$$g_{rr}(0) = \lim_{r \rightarrow 0} e^{2\lambda(r)} = 1.$$

6. Dans toute cette section nous ne raisonnons que localement en négligeant les aspects topologiques, dont les enjeux pourraient cependant s'avérer importants en cosmologie. Ainsi, l'expression du dl^2 dans (8.10) est compatible, notamment, avec les deux topologies suivantes : l'espace sphérique, c'est-à-dire la sphère S^3 , simplement connexe, où χ varie entre 0 et π , et l'espace elliptique, multiplement connexe, obtenu par identification des points antipodes sur la sphère, de sorte que χ varie entre 0 et $\pi/2$.

Pour la classification des espaces euclidiens à trois dimensions, on pourra consulter P.-G. Sprimont, *Aspects topologiques en cosmologie*, mémoire de licence en sciences physiques, Université de Liège, 2001 ;

N. Matagne, *Mirages topologiques dans les modèles cosmologiques d'Einstein-de Sitter non compacts*, mémoire de licence en sciences physiques, Université de Liège, 2002.

Si $k = -1$, le changement de variable $r = \sinh \chi$, $\chi \in [0, +\infty[$, permet d'écrire l'élément de longueur sous la forme

$$dl^2 = R^2 \left[d\chi^2 + \sinh^2 \chi (d\theta^2 + \sin^2 \theta d\varphi^2) \right]. \quad (8.11)$$

Dans l'espace de Minkowski de métrique $ds^2 = -du^2 + dx^2 + dy^2 + dz^2$, on peut l'interpréter comme la métrique sur l'hyperboloïde à deux nappes H^3 constitué des points situés à égale distance $-u^2 + x^2 + y^2 + z^2 = -R^2$ de l'origine. On s'en rend compte en posant

$$u = \pm R \cosh \chi, \quad x = R \sinh \chi \sin \theta \cos \varphi, \quad y = R \sinh \chi \sin \theta \sin \varphi, \quad z = R \sinh \chi \cos \theta.$$

Considérons alors, par exemple sur la nappe supérieure, un point P quelconque de coordonnées (u_P, x_P, y_P, z_P) . Il peut être passivement appliqué sur l'origine ($\underline{u}_P = R, \underline{x}_P = \underline{y}_P = \underline{z}_P = 0$) par un changement de coordonnées (une transformation de Lorentz) qui préserve

- l'équation de l'hyperboloïde :

$$-u^2 + x^2 + y^2 + z^2 = -R^2 = -\underline{u}^2 + \underline{x}^2 + \underline{y}^2 + \underline{z}^2;$$

- le tenseur métrique :

$$ds^2 = -du^2 + dx^2 + dy^2 + dz^2 = -d\underline{u}^2 + d\underline{x}^2 + d\underline{y}^2 + d\underline{z}^2.$$

Puisque l'élément de longueur, calculé le long de l'hyperboloïde, n'est autre que la quantité $ds^2 = -du^2 + dx^2 + dy^2 + dz^2$ avec la contrainte $-u^2 + x^2 + y^2 + z^2 = -R^2$, on conclut que, du point de vue de ses propriétés métriques, chaque point P de l'hyperboloïde est équivalent à l'origine O ($u_O = R, x_O = y_O = z_O = 0$) : *cet espace est donc homogène* (le groupe de Lorentz constitue un groupe d'isométries). Manifestement isotrope à l'origine, il est par conséquent *isotrope en tout point*. Ainsi, chacune des deux nappes de l'hyperboloïde peut être vue comme l'orbite du groupe de Lorentz orthochrone (préservant l'orientation temporelle) $O^+(1, 3)$. Elle modélise l'espace de Lobatchevski à trois dimensions. La figure 8.1 illustre tous ces propos à deux dimensions spatiales.

En résumé, nous avons montré que les contraintes mathématiques d'homogénéité et d'isotropie fixent complètement la métrique (à un facteur multiplicatif constant près), par exemple sous la forme (8.8).

8.3.6 Les modèles cosmologiques spatialement homogènes et isotropes

Les hypothèses à la base de ces modèles d'univers sont les suivantes.

- L'espace-temps peut être découpé en sections spatiales (hypersurfaces) homogènes et isotropes. Compte tenu de (8.6) et de (8.8), la métrique de ces modèles, appelés *modèles de Friedmann-Lemaître*, peut s'écrire sous la forme suivante, dite de *Friedmann-Lemaître-Robertson-Walker*⁷ :

$$\boxed{ds^2 = -c^2 dt^2 + R^2(t) \left[\frac{dr^2}{1 - k r^2} + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2) \right]} \quad (8.12)$$

où k peut prendre les valeurs 1, 0, -1 et où le *facteur d'échelle* $R(t)$ est solution des équations d'Einstein. Si $k = 1$ on parle d'un univers *fermé*, sinon on dit qu'il est *ouvert*.

7. J.-P. Luminet, *L'invention du Big Bang*, in A. Friedmann et G. Lemaître, *Essais de cosmologie*, Seuil, 1997, pp. 91-96.

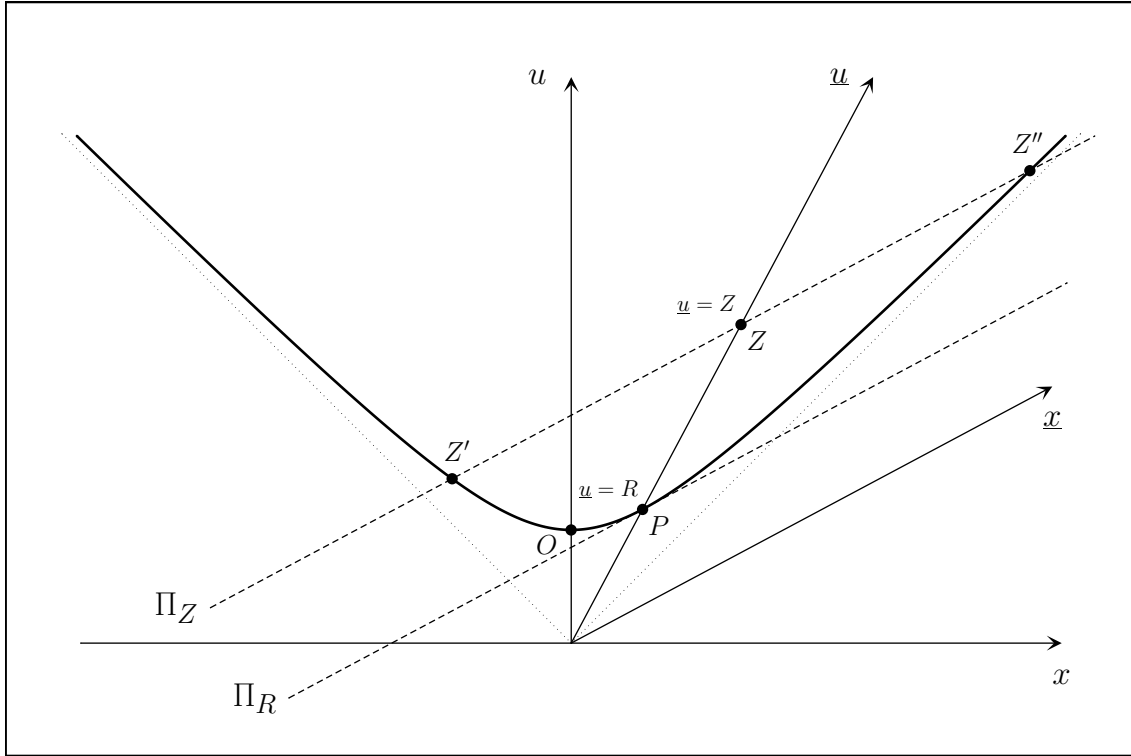


FIGURE 8.1: Coupe dans le plan $y = 0$ de la nappe supérieure de l'hyperboloïde H^2 d'équation $-u^2 + x^2 + y^2 = -R^2$. Tout point P de coordonnées $(u_P, x_P, y_P = 0)$ appartenant à cette surface peut passivement être appliqué sur l'origine $(\underline{u}_P = R, \underline{x}_P = 0, \underline{y}_P = 0)$ par le « boost » $\underline{u} = u \cosh \theta - x \sinh \theta$, $\underline{x} = -u \sinh \theta + x \cosh \theta$, $\underline{y} = y$, où $\theta = \operatorname{arcsinh}(x_P/R)$. Dans les coordonnées barrées, le plan Π_Z d'équation $\underline{u} = Z$ ($Z \geq R$) intersecte l'hyperboloïde selon un cercle \mathcal{C}_Z de centre $(\underline{u} = Z, \underline{x} = 0, \underline{y} = 0)$ et de rayon $\sqrt{Z^2 - R^2}$. En effet, étant donné que $\underline{u} = Z \Leftrightarrow u \cosh \theta - x \sinh \theta = Z$, dans Π_Z l'équation de l'hyperboloïde se ramène à $(-u \sinh \theta + x \cosh \theta)^2 + y^2 = Z^2 - R^2$, c'est-à-dire $\underline{x}^2 + \underline{y}^2 = Z^2 - R^2$. L'existence des cercles \mathcal{C}_Z pour tout $Z \geq R$ met en lumière l'isotropie de la variété au point P . Leur trace dans le plan $y = 0$ est représentée par les segments $Z'Z''$. En particulier, \mathcal{C}_R se réduit à un point : le plan Π_R est tangent à H^2 . Evidemment, le caractère « circulaire » de ces courbes d'intersection ne transparaît pas dans les coordonnées (u, x, y) puisque la métrique de l'espace d'immersion n'est pas définie positive. Cependant, la congruence des segments ZZ' et ZZ'' est patente (c'est dû au fait qu'ils sont tous deux parallèles à l'axe des \underline{x}).

Comme les lignes de temps sont orthogonales aux hypersurfaces et que g_{00} ne dépend pas des coordonnées spatiales, la distance (de nature temporelle) entre deux hypersurfaces données est constante, c'est-à-dire indépendante des coordonnées spatiales. D'autre part, il est facile de montrer que les lignes de temps constituent des *géodésiques* de l'espace-temps. On résume la situation en disant que les hypersurfaces spatiales sont *géodésiquement parallèles*.

- La matière est localement au repos par rapport à ces hypersurfaces. En d'autres termes, le quadrivecteur vitesse de chaque particule est orthogonal aux hypersurfaces et donc tangent aux lignes géodésiques de temps. (Si ce n'était pas le cas, le fluide cosmique déterminerait des directions privilégiées sur les hypersurfaces spatiales, incompatibles avec l'hypothèse

d'isotropie.) On adopte pour la matière les coordonnées spatiales *comobiles* (r, θ, φ) ou (χ, θ, φ) . Remarquons l'existence d'un *temps physique universel*.

Il est remarquable que des modèles aussi simples permettent de rendre compte d'un grand nombre de résultats observationnels.

8.4 Cosmologie et inertie

L'hypothèse d'un univers spatialement homogène et isotrope, régi par la métrique (8.12) et apte à modéliser les observations effectuées à l'échelle cosmologique, permet de débrouiller une difficulté conceptuelle que nous avons laissée en suspens jusqu'ici et que l'on peut résumer comme suit. Au travers de la relativité restreinte, la physique fait constamment référence à l'espace-temps de Minkowski et à des repères inertiels déterminés expérimentalement. Mais comment la théorie assure-t-elle le raccord avec l'expérience en établissant ce qui est inertiel et ce qui ne l'est pas? Bien entendu, c'est la métrique qui fournit le critère, mais comment justifier les conditions aux limites permettant de la calculer? Par exemple, quand on affirme que la Terre tourne autour de l'axe des pôles, par rapport à quel absolu camouflé définit-on cette rotation? Pour répondre à ces questions, il faut hisser le débat jusqu'au niveau de la cosmologie.

Dans son traité « Die Mechanik in ihrer Entwicklung » (1883), Ernst Mach récusait la conception newtonienne de l'inertie, où l'accélération d'un corps est définie par rapport à un espace absolu qu'il jugeait métaphysique. Mach préféra supposer que ce cadre absolu possède une origine matérielle et résulte d'une interaction du corps avec le fond des étoiles lointaines. Ce point de vue a fortement influencé Einstein dans son cheminement vers la relativité générale : cette théorie requiert en effet que le mouvement inertiel soit réalisé par les géodésiques d'un espace-temps qui, loin d'être absolu, se trouve courbé par le tenseur d'énergie-impulsion de la matière ou des champs.

Or, le théorème 5.4 énonce qu'autour d'une géodésique se déploie un « tube » d'espace-temps localement euclidien. Il est possible de préciser cette propriété dans le cas particulier des modèles cosmologiques spatialement homogènes et isotropes. Intuitivement, en effet, on peut approximer la variété riemannienne par l'espace-temps de Minkowski le long d'une géodésique temporelle de vecteur tangent $\partial/\partial t$, durant des intervalles de temps

$$\Delta t \ll R/\dot{R}$$

(où le point désigne la dérivée par rapport à t) et dans un domaine d'étendue spatiale d petite par rapport au facteur d'échelle :

$$d \ll R(t).$$

Ainsi, en omettant bien sûr la présence de courbure au voisinage d'étoiles sporadiques, on peut, en pratique, raisonner dans le cadre illusoire d'un espace-temps absolu pour des durées dépassant de beaucoup celle d'une vie humaine et jusqu'à des distances spatiales qui excèdent largement la dimension de notre galaxie.

Soit alors la tétrade orthonormée constituée des vecteurs de base

$$\left\{ \vec{e}_0 = \frac{1}{c} \frac{\partial}{\partial t}, \vec{e}_1 = \frac{\sqrt{1 - kr^2}}{R(t)} \frac{\partial}{\partial r}, \vec{e}_2 = \frac{1}{R(t)r} \frac{\partial}{\partial \theta}, \vec{e}_3 = \frac{1}{R(t)r \sin \theta} \frac{\partial}{\partial \varphi} \right\}.$$

On montre facilement que

$$\nabla_0 \vec{e}_\alpha = 0 \quad \forall \alpha. \quad (8.13)$$

Par exemple, en recourant aux symboles de Christoffel pour la métrique (8.12), on a

$$\begin{aligned}\nabla_0 \vec{e}_1 &= -\frac{\sqrt{1-kr^2} \dot{R}}{cR^2} \frac{\partial}{\partial r} + \frac{\sqrt{1-kr^2}}{R} \Gamma_{ctr}^r \frac{\partial}{\partial r} \\ &= -\frac{\sqrt{1-kr^2} \dot{R}}{cR^2} \frac{\partial}{\partial r} + \frac{\sqrt{1-kr^2}}{R} \frac{\dot{R}}{cR} \frac{\partial}{\partial r} \\ &= 0.\end{aligned}$$

Compte tenu du résultat (8.13), le théorème 5.4 garantit que les \vec{e}_α sont tangents aux lignes de coordonnées d'un référentiel localement inertiel à tout instant t . Dans l'esprit (sinon la lettre) de l'intuition de Mach, il reste donc à spécifier le lien entre ce référentiel et la matière.

Considérons un quadrivecteur spatial $\vec{S} = S^1 \vec{e}_1 + S^2 \vec{e}_2 + S^3 \vec{e}_3$. Ce vecteur est attaché au système de coordonnées comobiles (r, θ, φ) (ce qui signifie, par exemple, que la géodésique nulle de direction spatiale initiale \vec{S} passe constamment par les mêmes particules du fluide cosmique) si et seulement si les quantités S^i n'évoluent pas temporellement. Mais d'après (8.13), on a

$$\dot{S}^\alpha = 0 \iff \nabla_0 \vec{S} = 0. \quad (8.14)$$

Au fil du temps le vecteur \vec{S} est donc transporté parallèlement, c'est-à-dire, en l'occurrence, sans rotation spatiale. Notant $\vec{U} = c \vec{e}_0$ la quadrivitesse du fluide cosmique, on peut écrire la seconde relation de (8.14) sous la forme

$$\nabla_{\vec{U}} \vec{S} = 0. \quad (8.15)$$

Envisageons maintenant un gyroscope en translation rectiligne et uniforme de quadrivitesse \vec{U} dans l'espace-temps de Minkowski. Sans entrer dans les détails de la théorie relativiste du moment cinétique, nous admettrons que son spin peut être décrit par un quadrivecteur spatial \vec{S} . Dans le référentiel comobile, la composante temporelle S^0 est nulle (ce qui, incidemment, implique $\vec{U} \cdot \vec{S} = 0$); et pour les composantes spatiales S^i , on s'en tiendra à l'image mentale d'un gyroscope newtonien intégré dans une suspension à cardans, comme dans le montage historique de Foucault (1852) : les S^i sont alors constants et alignés sur le vecteur de rotation. Au total on a donc

$$\frac{dS^\alpha}{d\tau} = 0 \quad \forall \alpha, \quad (8.16)$$

où τ est le temps propre du centre de masse.

En présence d'un champ de gravitation, un tel gyroscope parcourt une géodésique de l'espace-temps courbe. Son équation résulte directement du principe d'équivalence appliqué à (8.16).⁸ Elle s'avère identique à (8.15). Dans le cadre spécifique de la métrique (8.12), un gyroscope en mouvement libre de quadrivitesse orthogonale aux hypersurfaces spatialement homogènes et isotropes pointe donc en permanence vers une particule donnée du fluide cosmique. Ainsi, les repères inertiels sont dépourvus de rotation vis-à-vis de la matière.

8. Le gyroscope doit être suffisamment petit pour qu'il ne soit pas nécessaire de prendre en compte les forces de marée. Du point de vue de la relativité générale, cela revient à négliger des termes faisant intervenir la courbure de l'espace-temps. Cette hypothèse serait donc trop restrictive si on voulait, par exemple, traiter la précession des équinoxes.

8.5 Evolution de l'univers et décalage des spectres

Les observations montrent que le spectre des objets lointains présente un décalage vers le rouge donné par

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e},$$

où λ_e est la longueur d'onde propre d'une raie donnée mesurée lors de l'émission en $t = t_e$ et λ_o sa valeur observée dans nos détecteurs à l'époque actuelle $t = t_o$.

Il est possible d'établir une relation simple entre le décalage vers le rouge z et le facteur d'échelle R . Choisissons les coordonnées comobiles pour que la position de l'observateur soit donnée par $\chi = 0$ et celle de l'objet par $\chi = \chi_e$. Supposons que les signaux lumineux se déplacent sur des géodésiques nulles radiales ($d\theta = 0$, $d\varphi = 0$), conformément à

$$0 = ds^2 = -c^2 dt^2 + R^2(t) d\chi^2.$$

Cette hypothèse est raisonnable tant que l'interaction entre la radiation et la matière est négligeable, donc, de toute façon, après le découplage électromagnétique. On a ainsi

$$\frac{c dt}{R(t)} = -d\chi,$$

le signe négatif provenant du fait que la lumière *s'approche* de l'origine lorsque le temps t augmente. Par conséquent,

$$c \int_{t_e}^{t_o} \frac{1}{R(t)} dt = - \int_{\chi_e}^0 d\chi = \chi_e. \quad (8.17)$$

Or, le terme de droite ne change pas si le signal lumineux est émis en $t_e + dt_e$ et recueilli en $t_o + dt_o$. Dès lors,

$$\int_{t_e}^{t_o} \frac{1}{R(t)} dt = \int_{t_e+dt_e}^{t_o+dt_o} \frac{1}{R(t)} dt.$$

On en déduit, quel que soit le signe de k :

$$\frac{dt_o}{R(t_o)} = \frac{dt_e}{R(t_e)}$$

$$\boxed{\frac{\lambda_o}{\lambda_e} = \frac{R(t_o)}{R(t_e)}} \quad (8.18)$$

et

$$z = \frac{R(t_o)}{R(t_e)} - 1 \quad (8.19)$$

$$R(t_o) = (z + 1) R(t_e). \quad (8.20)$$

Un *décalage vers le rouge* ($z > 0$) trouve ainsi une interprétation naturelle dans le cadre d'un *univers en expansion* [$R(t_o) > R(t_e)$].

8.6 La loi de Hubble-Lemaître

La distance physique entre l'origine $\chi = 0$ et un objet de coordonnée radiale χ , évaluée le long d'une géodésique ($d\theta = 0$, $d\varphi = 0$), est donnée par

$$l(t) = R(t) \chi. \quad (8.21)$$

Si le point désigne la dérivée par rapport au temps propre, on a donc

$$\dot{l}(t) = \frac{\dot{R}(t)}{R(t)} l(t)$$

En notant $V(t)$ la « vitesse » $\dot{l}(t)$ et en posant

$$H(t) = \frac{\dot{R}(t)}{R(t)}, \quad (8.22)$$

on peut écrire cette relation sous la forme

$$V(t) = H(t)l(t). \quad (8.23)$$

La « vitesse » d'un corps à un instant donné est proportionnelle à sa distance : c'est la *loi de Hubble-Lemaître*. Le coefficient de proportionnalité H ne dépend que du temps. On l'appelle *paramètre de Hubble*.

Il faut lire la relation (8.23) avec précaution. La quantité $V(t)$ exprime la variation de la distance d'un corps, résultant de l'évolution du facteur d'échelle de l'univers, et ne doit pas être comprise en termes traditionnels de vitesse du corps par rapport à un référentiel donné. Elle peut, d'ailleurs, être supérieure à la vitesse de la lumière.

Historiquement Hubble, contrairement à Lemaître, a proposé et interprété cette relation dans un contexte étranger aux questions théoriques de cosmologie. Pour des objets proches (mais suffisamment lointains pour participer au mouvement de récession cosmologique), elle se prête en effet à une formulation intuitive accessible à l'observation, que nous allons maintenant préciser.⁹

Définissons le *paramètre de décélération*

$$q(t) = -\frac{\ddot{R}(t) R(t)}{\dot{R}^2(t)}. \quad (8.24)$$

Il s'agit d'une quantité sans dimensions accessible à l'observation, comme nous le verrons plus loin. Notant t_0 l'âge actuel de l'univers (c'est-à-dire au moment de la réception du photon, lui-même émis par l'objet de coordonnée χ au temps t), $R_0 = R(t_0)$, $H_0 = H(t_0)$ et $q_0 = q(t_0)$, on peut développer le facteur d'échelle en série de Taylor comme suit :

$$\begin{aligned} R(t) &= R[t_0 - (t_0 - t)] \\ &= R_0 - (t_0 - t) \dot{R}(t_0) + \frac{1}{2} (t_0 - t)^2 \ddot{R}(t_0) + O(t_0 - t)^3 \\ &= R_0 \left[1 - H_0 (t_0 - t) - \frac{1}{2} q_0 H_0^2 (t_0 - t)^2 \right] + O(3). \end{aligned}$$

9. M.P. Hobson, G. Efstathiou, A.N. Lasenby, *General Relativity, An Introduction for Physicists*, Cambridge University Press, 2006, pp. 368-371.

Le décalage vers le rouge de cet objet est donné par (8.19) :

$$\begin{aligned} z &= \frac{R_0}{R} - 1 \\ &= \frac{1}{\left[1 - H_0(t_0 - t) - \frac{1}{2}q_0 H_0^2(t_0 - t)^2\right]} - 1 + O(3) \\ &= H_0(t_0 - t) + H_0^2(t_0 - t)^2 \left(1 + \frac{1}{2}q_0\right) + O(3) \end{aligned}$$

On peut inverser cette relation. Posant $t_0 - t = Az + Bz^2 + O(z^3)$, on calcule facilement les valeurs de A et B :

$$t_0 - t = \frac{1}{H_0}z - \frac{1 + q_0/2}{H_0}z^2 + O(z^3). \quad (8.25)$$

Utilisant (8.17), on en déduit les premiers termes du développement en série de la fonction $\chi(z)$:

$$\begin{aligned} \chi &= \int_t^{t_0} \frac{c}{R(t')} dt' \\ &= \frac{c}{R_0} \int_t^{t_0} [1 + H_0(t_0 - t') + O(2)] dt' \\ &= \frac{c}{R_0} \left[(t_0 - t) + \frac{H_0}{2}(t_0 - t)^2 + O(3) \right] \\ &= \frac{c}{R_0 H_0} \left(z - \frac{1 + q_0}{2} z^2 \right) + O(3). \end{aligned} \quad (8.26)$$

La distance propre de l'objet, mesurée à l'époque actuelle, vaut $l_0 = R_0 \chi$, en vertu de (8.21), donc

$$l_0 = \frac{c}{H_0} \left(z - \frac{1 + q_0}{2} z^2 \right) + O(3). \quad (8.27)$$

Au premier ordre en z , c'est-à-dire pour des décalages vers le rouge suffisamment petits, on obtient ainsi la loi

$$z \cong \frac{H_0 l_0}{c}, \quad (8.28)$$

déduite de l'observation par Hubble en 1929. En fait, celui-ci disposait de données pour des objets éloignés tout au plus d'une distance $l_0 \cong 30$ millions d'années-lumière¹⁰. Si l'on prend $H_0 \approx 71 \text{ km}/(\text{s Mpc}) \approx 2.30 \times 10^{-18} \text{ /s}$, la relation (8.28) donne $z \approx 2.2 \times 10^{-3}$: l'approximation consistant à négliger les termes d'ordre 2 est alors cohérente. Elle restitue, du reste, la relation intuitive $l_0 \cong c(t_0 - t)$, ainsi que le montrent les relations (8.25) et (8.27).

Hubble interpréta le décalage vers le rouge (8.28) du spectre d'un corps en termes d'effet Doppler-Fizeau résultant du mouvement non relativiste ($V \ll c$) de ce corps dans un espace euclidien :

$$z \cong \frac{V}{c} \ll 1. \quad (8.29)$$

Ce point de vue est cependant erroné conceptuellement. Combinant les résultats (8.28) et (8.29), il déduisit la loi

$$V \cong H_0 l_0 \quad (8.30)$$

selon laquelle la vitesse de récession d'un objet est proportionnelle à sa distance.

10. D.W. Sciama, *Modern Cosmology*, Cambridge University Press, 1971, pp. 44-45.

8.7 Ecriture des équations d'Einstein

Nous considérerons que la matière peut raisonnablement être décrite par un fluide parfait de tenseur d'énergie-impulsion

$$T^{\mu\nu} = \frac{1}{c^2} (\rho + P) U^\mu U^\nu + P g^{\mu\nu}$$

où la pression P et l'énergie volumique ρ ne dépendent que du temps, et où le fluide est comobile ($U^0 = c$, $U^i = 0$). Les équations d'Einstein avec constante cosmologique Λ

$$R^{\mu\nu} - \frac{R}{2} g^{\mu\nu} + \Lambda g^{\mu\nu} = \frac{8\pi G}{c^4} T^{\mu\nu}$$

s'écrivent

$$\boxed{\frac{3}{R^2} (\dot{R}^2 + k c^2) = \frac{8\pi G \rho}{c^2} + \Lambda c^2} \quad (8.31)$$

$$\boxed{2 \frac{\ddot{R}}{R} + \frac{\dot{R}^2}{R^2} + \frac{k c^2}{R^2} = -\frac{8\pi G P}{c^2} + \Lambda c^2.} \quad (8.32)$$

De (8.31) et (8.32), on déduit

$$\boxed{\frac{d}{dt} (\rho R^3) = -P \frac{d}{dt} R^3} \quad (8.33)$$

qui peut être lue comme une conséquence des équations de conservation $T^{\mu\nu}{}_{;\nu} = 0$. Cette équation est susceptible d'une interprétation physique intéressante rappelant le premier principe de la thermodynamique : *la variation de l'énergie contenue dans un système donné est égale au travail de la pression.*

On n'a que deux équations pour, apparemment, trois inconnues R , P et ρ . Cependant, P et ρ sont liés par une équation d'état, en général l'équation barotrope (6.13). Dans un grand nombre de cas de figure, on peut d'ailleurs considérer que le fluide est constitué d'un mélange de matière incohérente sans pression (« poussières », $\gamma = 1$) et de rayonnement ($m_0 = 0$, $\gamma = 4/3$), et l'on décompose l'énergie volumique totale en deux termes notés respectivement ρ_m et ρ_r :

$$\rho = \rho_m + \rho_r.$$

La pression est alors celle du rayonnement seul :

$$P = \rho_r/3.$$

D'autre part, on peut définir une énergie volumique associée à la constante cosmologique par la relation

$$\rho_\Lambda = \frac{c^4 \Lambda}{8\pi G}.$$

Cette interprétation est en vogue depuis que les observations semblent indiquer que l'univers se trouve en expansion accélérée. Cependant, la nature de l'énergie associée à Λ demeure problématique. Faute de mieux, on l'a qualifiée d'« énergie sombre ».

L'équation (8.31) prend ainsi la forme

$$\frac{\dot{R}^2}{R^2} + \frac{k c^2}{R^2} = \frac{8\pi G}{3c^2} (\rho_m + \rho_r + \rho_\Lambda), \quad (8.34)$$

c'est-à-dire

$$k = \frac{R^2 H^2}{c^2} \left(\frac{\rho_m + \rho_r + \rho_\Lambda}{\rho_c} - 1 \right), \quad (8.35)$$

où la quantité

$$\rho_c(t) = \frac{3c^2 H^2(t)}{8\pi G} \quad (8.36)$$

porte pour des raisons évidentes le nom d'« énergie volumique critique ».

Définissons les trois *paramètres de densité* associés respectivement à la matière, à la radiation et à la constante cosmologique, par

$$\Omega_m(t) = \frac{\rho_m(t)}{\rho_c(t)}, \quad \Omega_r(t) = \frac{\rho_r(t)}{\rho_c(t)} \quad \text{et} \quad \Omega_\Lambda(t) = \frac{\rho_\Lambda}{\rho_c(t)}. \quad (8.37)$$

Il s'avère que la simple connaissance des quatre paramètres H_0 , Ω_{m0} , Ω_{r0} et $\Omega_{\Lambda0}$ (où l'indice 0 caractérise les quantités à l'époque actuelle t_0) fournit des informations capitales sur l'évolution du modèle spatialement homogène et isotrope, en permettant de calculer le signe de k , la valeur du paramètre de décélération et l'âge de l'univers t_0 .

En effet, si l'on pose $R_0 = R(t_0)$, l'équation (8.35) s'écrit

$$k = \frac{R_0^2 H_0^2}{c^2} (\Omega_{m0} + \Omega_{r0} + \Omega_{\Lambda0} - 1). \quad (8.38)$$

L'univers est donc ouvert si $\Omega_{m0} + \Omega_{r0} + \Omega_{\Lambda0} < 1$ et fermé si $\Omega_{m0} + \Omega_{r0} + \Omega_{\Lambda0} > 1$. Or, les valeurs admises pour l'époque actuelle sont ¹¹

$$\Omega_{m0} \approx 0.266, \quad \Omega_{r0} \approx 8 \times 10^{-5}, \quad \Omega_{\Lambda0} \approx 0.732 \quad (8.39)$$

(remarquons en passant que la constante cosmologique serait donc positive). Dès lors $k = -1$, mais l'univers aurait une énergie volumique (toutes origines confondues) proche de l'énergie volumique critique, dont la valeur actuelle serait

$$\rho_{c0} \approx 8.5 \times 10^{-10} \text{ J/m}^3$$

si l'on admet que

$$\begin{aligned} H_0 &\approx 71 \text{ km/(s Mpc)} \\ &\approx 2.30 \times 10^{-18} \text{ /s.} \end{aligned} \quad (8.40)$$

Couplée à (8.38), cette information permet de calculer le facteur d'échelle actuel :

$$R_0 \approx 3.0 \times 10^{27} \text{ m.}$$

Venons-en au paramètre de décélération. La combinaison linéaire (8.31)–3×(8.32) donne

$$-\frac{\ddot{R}}{R} = \frac{4\pi G}{3c^2} (\rho_m + 2\rho_r - 2\rho_\Lambda)$$

11. Ces chiffres sont déduits de la relation entre le décalage vers le rouge et la magnitude des supernovae de type Ia, ainsi que de l'étude détaillée de l'anisotropie du rayonnement fossile. Voir par exemple J.B. Hartle, *Gravity, An Introduction to Einstein's General Relativity*, Addison Wesley, 2003, p. 402. Or, la masse volumique actuelle des baryons, déduite de la théorie de la nucléosynthèse primordiale du deutérium et de son abondance observée à notre époque, vaut tout au plus $4 \times 10^{-31} \text{ g/cm}^3$ (J. Demaret, *op. cit.*, p. 80), correspondant à un paramètre de densité égal à 0.04. La valeur nettement plus élevée obtenue pour Ω_{m0} suggère que l'univers contient une grande quantité de matière « noire » non baryonique.

donc

$$\begin{aligned}
 q &= -\frac{\ddot{R}R}{\dot{R}^2} \\
 &= -\frac{\ddot{R}}{R H^2} \\
 &= \frac{4\pi G}{3c^2 H^2} (\rho_m + 2\rho_r - 2\rho_\Lambda) \\
 &= \frac{1}{2} \Omega_m + \Omega_r - \Omega_\Lambda.
 \end{aligned} \tag{8.41}$$

Remarquons que la pression tend à décélérer l'expansion (ce qui est logique puisqu'elle correspond à de l'énergie volumique), tandis qu'une constante cosmologique positive tend à l'accélérer. Les valeurs observées (8.39) impliquent $q_0 < 0$: l'univers serait donc actuellement dans une phase d'expansion accélérée.

8.8 Evolution qualitative du facteur d'échelle

Comme on le voit, la seule écriture des équations d'Einstein permet déjà de tirer certaines conclusions intéressantes. Cela étant, on peut établir à peu de frais la relation liant $\rho(t)$ à $R(t)$ dans le cas où le fluide parfait est du type « poussière » ou « rayonnement ». De (8.33) et (6.13) on déduit en effet :

$$\frac{\dot{\rho}}{\rho} = -3\gamma \frac{\dot{R}}{R}$$

donc

$$\rho = \frac{K}{R^{3\gamma}}$$

avec K constant. Pour $\gamma = 1$ et $\gamma = 4/3$ on trouve respectivement¹²

$$\rho_m = \frac{K_1}{R^3} \tag{8.42}$$

$$\rho_r = \frac{K_2}{R^4}. \tag{8.43}$$

La relation (8.34) prend la forme

$$\frac{\dot{R}^2}{R^2} - \frac{8\pi G}{3c^2} \left(\frac{\rho_{m0} R_0^3}{R^3} + \frac{\rho_{r0} R_0^4}{R^4} + \rho_\Lambda \right) = -\frac{kc^2}{R^2}.$$

Introduisant le facteur d'échelle normalisé

$$x(t) = \frac{R(t)}{R_0}, \tag{8.44}$$

et tenant compte de (8.37), on obtient successivement

$$\frac{\dot{x}^2}{x^2} - \frac{8\pi G \rho_{c0}}{3c^2} \left(\frac{\Omega_{m0}}{x^3} + \frac{\Omega_{r0}}{x^4} + \Omega_{\Lambda 0} \right) = -\frac{kc^2}{R_0^2 x^2}$$

12. Intuitivement, la relation (8.43) peut être comprise comme suit. Le nombre de photons par unité de volume décroît comme $1/R^3$ et l'énergie par photon, proportionnelle à la fréquence, décroît comme $1/R$ [voir (8.18)]. Au total, la densité d'énergie du gaz de photons décroît bien comme $1/R^4$.

$$\dot{x}^2 - H_0^2 \left(\frac{\Omega_{r0}}{x^2} + \frac{\Omega_{m0}}{x} + \Omega_{\Lambda 0} x^2 \right) = -\frac{kc^2}{R_0^2},$$

où l'on a exploité la définition de la densité critique (8.36) à l'époque actuelle. Introduisons enfin le temps normalisé

$$t' = H_0 t. \quad (8.45)$$

Si le « prime » désigne la dérivée par rapport à t' , on trouve, en utilisant (8.38) pour simplifier le membre de droite, l'équation dite *de Friedmann* :

$$\boxed{x'^2 - \left(\frac{\Omega_{r0}}{x^2} + \frac{\Omega_{m0}}{x} + \Omega_{\Lambda 0} x^2 \right) = 1 - (\Omega_{r0} + \Omega_{m0} + \Omega_{\Lambda 0})}. \quad (8.46)$$

Cette équation est très utile car elle permet de ramener l'étude qualitative de l'évolution du facteur d'échelle à celle du mouvement $x(t')$ d'une particule plongée dans un potentiel conservatif d'équation

$$V_{eff}(x) = -\frac{1}{2} \left(\frac{\Omega_{r0}}{x^2} + \frac{\Omega_{m0}}{x} + \Omega_{\Lambda 0} x^2 \right). \quad (8.47)$$

Remarquons cependant que l'« énergie totale »

$$\frac{1}{2} [1 - (\Omega_{r0} + \Omega_{m0} + \Omega_{\Lambda 0})] \quad (8.48)$$

n'est pas arbitraire puisqu'elle se trouve fixée par les trois paramètres de densité. Les figures 8.2 à 8.4 illustrent quelques comportements possibles.

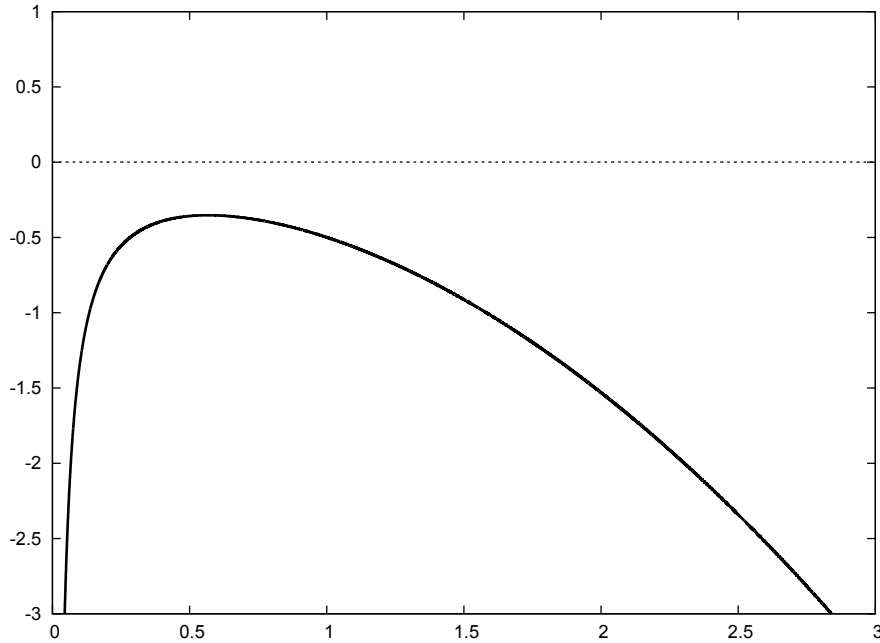


FIGURE 8.2: En continu, on a représenté la courbe du potentiel effectif (8.47) pour les valeurs actuellement admises : $\Omega_{m0} = 0.266$, $\Omega_{r0} = 8 \times 10^{-5}$ et $\Omega_{\Lambda 0} = 0.732$. En pointillés, l'« énergie totale » (8.48). Ce modèle d'univers, du type $k = -1$, admet une singularité en $x = 0$ et se trouve en expansion illimitée.

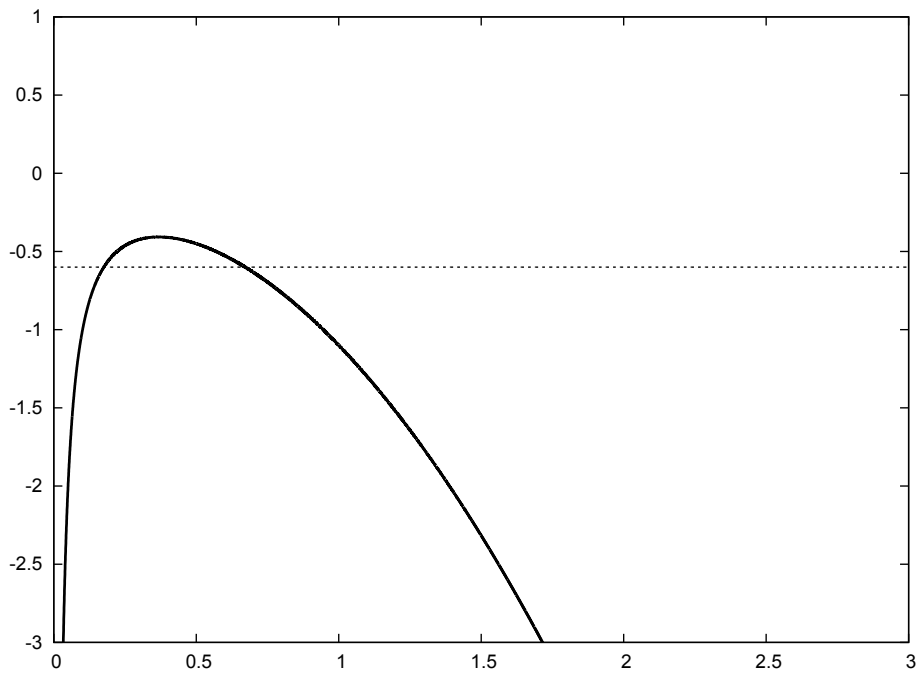


FIGURE 8.3: Le modèle d'univers $\Omega_{m0} = 0.2$, $\Omega_{r0} = 0$ et $\Omega_{\Lambda 0} = 2$. Du type $k = 1$, il n'admet pas de singularité et se trouve en expansion illimitée.

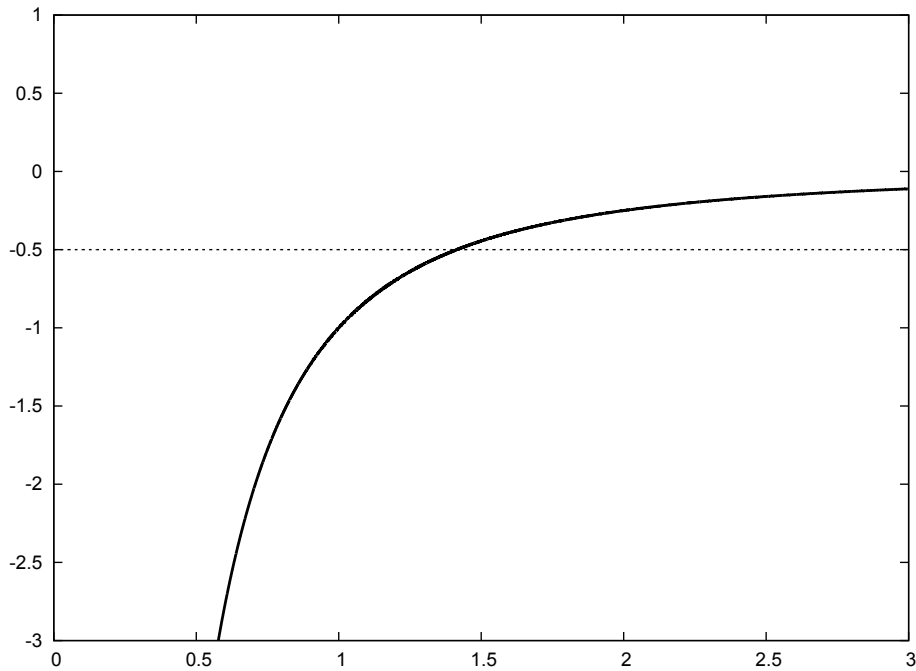


FIGURE 8.4: Le modèle d'univers $\Omega_{m0} = 0$, $\Omega_{r0} = 2$ et $\Omega_{\Lambda 0} = 0$. Du type $k = 1$, il admet une singularité initiale, passe par un facteur d'échelle maximal et se recontracte en une singularité finale.

Le cas $\Omega_{r0} = 0$ (univers sans pression) et $\Lambda > 0$ a joué un rôle historique important. Sous ces hypothèses en effet, la fonction $V_{eff}(x)$ possède un maximum facile à calculer, situé en

$$\left(x_M = \sqrt[3]{\frac{\Omega_{m0}}{2\Omega_{\Lambda 0}}}, V_M = -\frac{3}{4} \sqrt[3]{2\Omega_{m0}^2 \Omega_{\Lambda 0}} \right). \quad (8.49)$$

Si l'« énergie totale » est égale à V_M , c'est-à-dire si

$$\frac{1}{2} (1 - \Omega_{m0} - \Omega_{\Lambda 0}) = -\frac{3}{4} \sqrt[3]{2\Omega_{m0}^2 \Omega_{\Lambda 0}}, \quad (8.50)$$

on rencontre alors une situation critique [appelée *modèle d'Eddington-Lemaître* (1927) si $x > x_M$]. La condition (8.50) exige que le membre de gauche soit négatif, ce qui, compte tenu de (8.38), impose $k = 1$. D'autre part, associée à (8.38), elle revient à supposer que

$$\frac{4\pi G}{c^4} \rho_{m0} R_0^3 = \frac{4\pi G}{c^4} \rho_m(t) R^3(t) = \frac{1}{\sqrt{\Lambda}}.$$

Ce modèle critique tend (par la gauche ou par la droite) vers une solution asymptotique correspondant au point

$$\left(x_M = \frac{1}{R_0 \sqrt{\Lambda}}, V_M = \frac{-3\Omega_{\Lambda 0}}{2\Lambda R_0^2} \right).$$

C'est le *modèle statique d'Einstein* (1917), lequel s'avère ainsi *instable*. Son facteur d'échelle est une constante

$$R_E = 1/\sqrt{\Lambda}.$$

En introduisant cette valeur dans l'équation du champ (8.31), on trouve $\rho_m = 2\rho_\Lambda$.

Les équations (8.44), (8.45) et (8.46) permettent de calculer l'âge de l'univers, c'est-à-dire le temps cosmologique écoulé entre $x = 0$ et $x = 1$:

$$t_0(H_0, \Omega_{r0}, \Omega_{m0}, \Omega_{\Lambda 0}) = \frac{1}{H_0} \int_0^1 \frac{dx}{\sqrt{1 - \Omega_{r0} - \Omega_{m0} - \Omega_{\Lambda 0} + \frac{\Omega_{r0}}{x^2} + \frac{\Omega_{m0}}{x} + \Omega_{\Lambda 0} x^2}}, \quad (8.51)$$

pour autant bien sûr que l'intégrale existe (si le radicand est négatif dans un intervalle $[x_1 \geq 0, x_2 < 1]$, alors on a affaire à un univers sans singularité initiale). En reprenant les valeurs numériques $H_0 = 2.30 \times 10^{-18}$ /s, $\Omega_{r0} = 0$, $\Omega_{m0} = 0.266$ et $\Omega_{\Lambda 0} = 0.732$, on obtient $t_0 = 13.7$ milliards d'années.

8.9 Résolution des équations d'Einstein ($\Lambda = 0$)

Dans les phases reculées de l'histoire de l'univers, la constante cosmologique joue un rôle négligeable car l'énergie volumique ρ_Λ qui lui est associée s'efface progressivement devant ρ_m et ρ_r . Il est alors possible d'intégrer explicitement les équations d'Einstein.

8.9.1 Cas $P = 0$ (matière incohérente)

D'après (8.42), on a $\rho = K_1/R^3$. Injectant ce résultat dans l'équation (8.31) avec $\Lambda = 0$, on déduit

$$\frac{dR}{dt} = \pm c \sqrt{\frac{8\pi G K_1}{3Rc^4} - k}$$

donc

$$\int \frac{dR}{\sqrt{\frac{8\pi GK_1}{3Rc^4} - k}} = \pm c \int dt.$$

On peut calculer les primitives mais pas inverser la relation ainsi obtenue afin d'expliciter R en fonction de t . C'est la raison pour laquelle on utilise habituellement une autre coordonnée temporelle η (sans dimension), définie par

$$c dt = R(\eta) d\eta. \quad (8.52)$$

Posant

$$\mathcal{R}_0 = \frac{4\pi GK_1}{3c^4},$$

on a alors

$$\int \frac{dR}{\sqrt{2\mathcal{R}_0 R - kR^2}} = \pm \int d\eta. \quad (8.53)$$

Choisissant les constantes arbitraires d'intégration, additives, de telle sorte que la singularité cosmologique $R = 0$ se produise en $\eta = 0$ (ainsi qu'en $\eta = 2\pi$ si $k = 1$), cette valeur correspondant elle-même à $t = 0$, on obtient le tableau suivant.¹³

$k = 1$	$k = 0$	$k = -1$
$\eta \in]0, 2\pi[$	$\eta \in]0, +\infty[$	$\eta \in]0, +\infty[$
$R = \mathcal{R}_0 (1 - \cos \eta)$	$R = \mathcal{R}_0 \eta^2 / 2$	$R = \mathcal{R}_0 (\cosh \eta - 1)$
$ct = \mathcal{R}_0 (\eta - \sin \eta)$	$ct = \mathcal{R}_0 \eta^3 / 6$	$ct = \mathcal{R}_0 (\sinh \eta - \eta)$

Pour $k = 1$, les relations $R(\eta)$ et $ct(\eta)$ constituent les équations paramétriques d'une cycloïde $R(t)$. Tous ces résultats sont illustrés en figure 8.5.

On calcule facilement les paramètres de Hubble et de décélération :

$$H = \frac{c}{R^2} \frac{dR}{d\eta}, \quad q = 1 - R \frac{d^2 R / d\eta^2}{(dR/d\eta)^2}.$$

$k = 1$	$k = 0$	$k = -1$
$H = \frac{c \sin \eta}{\mathcal{R}_0 (1 - \cos \eta)^2}$	$H = \frac{4c}{\mathcal{R}_0 \eta^3}$	$H = \frac{c \sinh \eta}{\mathcal{R}_0 (\cosh \eta - 1)^2}$
$q = \frac{1}{1 + \cos \eta}$	$q = \frac{1}{2}$	$q = \frac{1}{1 + \cosh \eta}$

Remarquons que, si $k = 0$, la quantité \mathcal{R}_0 est dénuée de sens physique. En effet, dans l'équation (8.12) elle peut être absorbée dans la coordonnée radiale r , ce qui revient à fixer sa valeur à l'unité¹⁴.

13. Pour les modèles $k = 0$ et $k = -1$, le choix du domaine de η revient à imposer le signe + dans l'équation (8.53), c'est-à-dire un facteur d'échelle croissant avec le temps.

14. Cela impose

$$K_1 = \frac{3c^4}{4\pi G} \times 1 \text{ m}.$$

Une mesure de ρ_m fixe ainsi, par (8.42), le facteur d'échelle $R(t)$; ce qui permet de fixer l'âge de l'univers puisque

$$R(t) = \sqrt[3]{9c^2/2} t^{2/3} \times 1 \text{ m}^{1/3}.$$

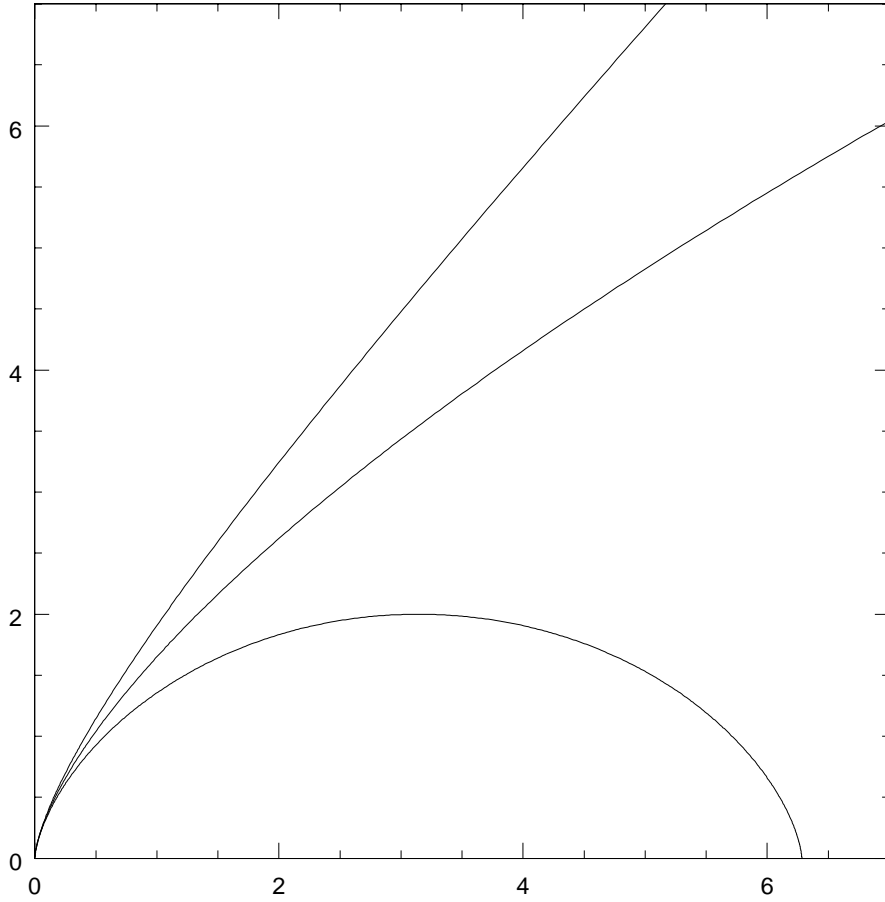


FIGURE 8.5: Evolution du facteur d'échelle dans les modèles cosmologiques spatialement homogènes et isotropes à pression nulle. En unités \mathcal{R}_0 on a porté en abscisse le temps ct et en ordonnée le facteur d'échelle $R(t)$, solution des équations de Friedmann-Lemaître sans constante cosmologique, dans les trois cas $k = 1, 0, -1$.

8.9.2 Cas $P = \rho/3$ (rayonnement)

D'après (8.43), on a $\rho = K_2/R^4$. Injectant ce résultat dans l'équation (8.31) avec $\Lambda = 0$, on déduit

$$\int \frac{dR}{\sqrt{\frac{8\pi GK_2}{3R^2 c^4} - k}} = \pm c \int dt.$$

On pourrait, cette fois, calculer explicitement R en fonction de t mais il s'avère malgré tout plus avantageux d'utiliser la coordonnée temporelle η , toujours définie par (8.52). Posant

$$\mathcal{R}_0 = \sqrt{\frac{8\pi GK_2}{3c^4}},$$

on obtient

$$\int \frac{dR}{\sqrt{\mathcal{R}_0^2 - kR^2}} = \pm \int d\eta. \quad (8.54)$$

Si, comme ci-dessus, on choisit les constantes arbitraires d'intégration de telle sorte que la singularité cosmologique $R = 0$ se produise en $\eta = 0$ (ainsi qu'en $\eta = \pi$ si $k = 1$), cette valeur correspondant elle-même à $t = 0$, on obtient le tableau suivant (figure 8.6).

$k = 1$	$k = 0$	$k = -1$
$\eta \in]0, \pi[$	$\eta \in]0, +\infty[$	$\eta \in]0, +\infty[$
$R = \mathcal{R}_0 \sin \eta$	$R = \mathcal{R}_0 \eta$	$R = \mathcal{R}_0 \sinh \eta$
$ct = \mathcal{R}_0 (1 - \cos \eta)$	$ct = \mathcal{R}_0 \eta^2/2$	$ct = \mathcal{R}_0 (\cosh \eta - 1)$
$H = \frac{c \cos \eta}{\mathcal{R}_0 \sin^2 \eta}$	$H = \frac{c}{\mathcal{R}_0 \eta^2}$	$H = \frac{c \cosh \eta}{\mathcal{R}_0 \sinh^2 \eta}$
$q = \frac{1}{\cos^2 \eta}$	$q = 1$	$q = \frac{1}{\cosh^2 \eta}$

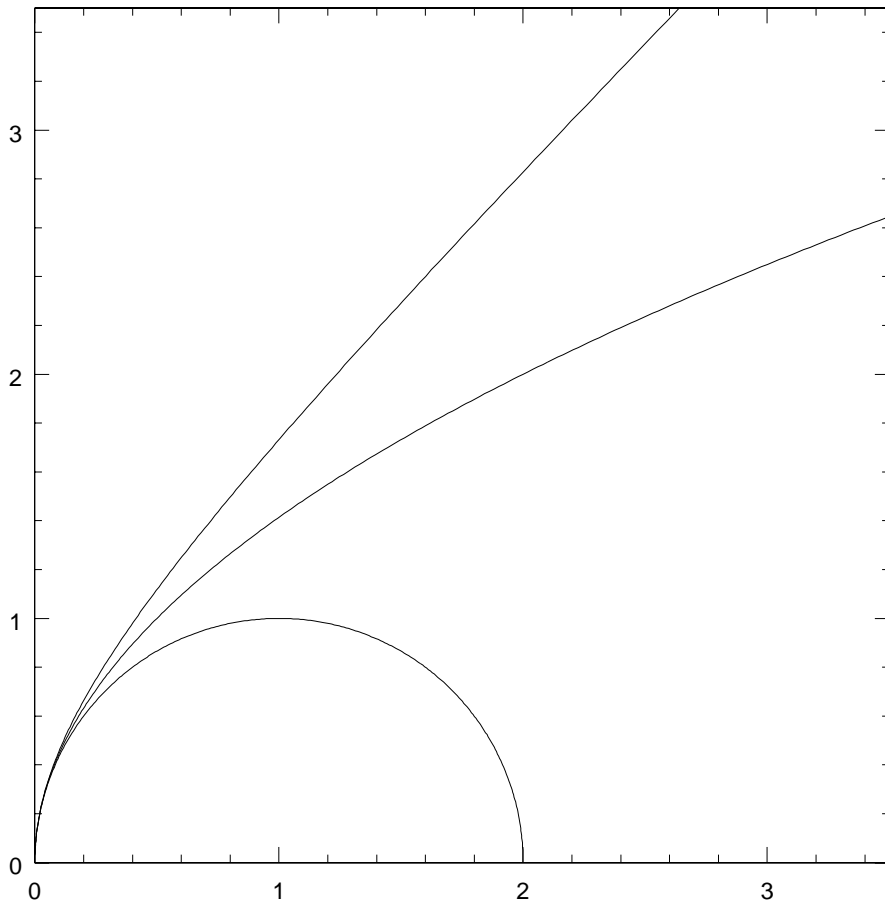


FIGURE 8.6: Même légende que pour la figure 8.5, mais pour un univers rempli de radiation vérifiant l'équation d'état $P = \rho/3$. Les courbes représentées sont, en fait, des fragments d'hyperbole ($k = -1$), de parabole ($k = 0$) et de cercle ($k = 1$).

8.9.3 Exercices

Exercice 8.2 *Supposons que l'observation permette de conclure au caractère ouvert de l'univers et admettons l'équation d'état $P = 0$ (donc $\Omega_r = \Omega_\Lambda = 0$). Etablir les relations donnant l'âge et le facteur d'échelle de l'univers en fonction des paramètres H et q .*

Solution. Comme $k = -1$ et $P = 0$, on a, d'après la section 8.9.1,

$$q = \frac{1}{1 + \cosh \eta}, \quad \text{donc} \quad \cosh \eta = \frac{1 - q}{q}.$$

On en déduit

$$\eta = \ln \left[\frac{1 - q + \sqrt{1 - 2q}}{q} \right] \quad (8.55)$$

$$\mathcal{R}_0 = \frac{c \sinh \eta}{H (\cosh \eta - 1)^2} = \frac{cq}{H \sqrt{(1 - 2q)^3}}. \quad (8.56)$$

Les deux relations demandées sont donc

$$t = \frac{\mathcal{R}_0}{c} (\sinh \eta - \eta) = \frac{\sqrt{1 - 2q} - q \ln \left[\frac{1 - q + \sqrt{1 - 2q}}{q} \right]}{H \sqrt{(1 - 2q)^3}} \quad (8.57)$$

$$R = \mathcal{R}_0 (\cosh \eta - 1) = \frac{c}{H \sqrt{1 - 2q}}. \quad (8.58)$$

D'après (8.51), on doit avoir

$$t = \frac{1}{H} \int_0^1 \frac{dx}{\sqrt{1 - \Omega_m + \Omega_m/x}}$$

où la relation (8.41) impose $\Omega_m = 2q$. Sachant que, à une constante additive près,

$$\int \frac{dx}{\sqrt{1 + a/x}} = \sqrt{x(x+a)} - a \ln(\sqrt{x} + \sqrt{x+a}),$$

on retrouve le résultat (8.57).

Exercice 8.3 *Dans le modèle d'univers décrit à l'exercice 8.2, on observe un quasar dont le spectre présente un décalage vers le rouge égal à z .*

a) *Exprimer les paramètres de Hubble et de décélération à l'époque de l'émission, respectivement notés H_e et q_e , en fonction de z et de ces mêmes paramètres à l'époque actuelle, notés H_o et q_o .*

b) *En plaçant l'observateur à l'origine, calculer la coordonnée radiale comobile χ_e du quasar et sa distance propre en fonction du temps.*

Solution.

a) En fonction du temps η , la relation (8.20) s'écrit

$$R(\eta_o) = (z + 1) R(\eta_e). \quad (8.59)$$

Dans le cas particulier de l'univers ouvert $k = -1$, on en déduit

$$\begin{aligned}\cosh \eta_o - 1 &= (z + 1)(\cosh \eta_e - 1) \\ \cosh \eta_e - 1 &= \frac{1 - 2q_o}{q_o(z + 1)} = \frac{1 - 2q_e}{q_e}.\end{aligned}$$

Dès lors,

$$q_e = \frac{q_o(z + 1)}{1 + 2q_o z}.$$

Vu (8.58) et (8.59), on calcule

$$H_e = H_o(z + 1)\sqrt{1 + 2q_o z}.$$

b) On a

$$\int_{\eta_e}^{\eta_o} d\eta = - \int_{\chi_e}^0 d\chi = \chi_e,$$

ce qui, compte tenu du résultat (8.55), permet de calculer la coordonnée radiale du quasar :

$$\chi_e = \eta_o - \ln \left[\frac{1 - q_e + \sqrt{1 - 2q_e}}{q_e} \right].$$

Au temps η , la distance spatiale à l'origine $\chi = 0$ d'un point de coordonnée radiale χ_e , évaluée le long d'une géodésique ($d\theta = 0$, $d\varphi = 0$), se calcule par la relation

$$l(\eta) = R(\eta) \int_0^{\chi_e} d\chi = R(\eta) \chi_e.$$

Exercice 8.4 (l'horizon cosmologique) *Dans le modèle d'univers hyperbolique dominé par le rayonnement, établir la loi exprimant le rayon physique de la sphère d'influence de l'origine $\chi = 0$ en fonction du temps propre cosmologique t .*

Solution. Compte tenu de (8.12), on peut écrire

$$ds^2 = R^2(\eta) \left[-d\eta^2 + d\chi^2 + \sinh^2 \chi (d\theta^2 + \sin^2 \theta d\varphi^2) \right].$$

La matière étant alors vraisemblablement ionisée (voir exercice 8.6), la lumière n'emprunte certainement pas une trajectoire du type « géodésique nulle ». Cependant, celle-ci permet de déterminer une borne supérieure pour la distance parcourue, puisque rien ne peut se propager plus vite que la lumière dans le vide. Posons ainsi $ds^2 = 0$, et supposons qu'un photon est émis radialement en $\eta = 0$ depuis l'origine $\chi = 0$. Cela implique $d\eta = d\chi$, donc

$$\begin{aligned}\int_0^\eta d\eta' &= \int_0^\chi d\chi' \\ \chi(\eta) &= \eta.\end{aligned}$$

Ce résultat représente ce que l'on appelle, en cosmologie, l'*horizon-particule* : il s'agit de la coordonnée radiale maximale des particules susceptibles d'avoir interagi avec l'origine depuis la singularité initiale en $\eta = 0$. En fonction du temps propre t , on a, en vertu des résultats obtenus dans la section 8.9.2,

$$\chi(t) = \operatorname{arcosh} \left(1 + \frac{ct}{\mathcal{R}_0} \right)$$

où, pour mémoire,

$$\operatorname{arcosh} x = \ln \left(x + \sqrt{x^2 - 1} \right).$$

La distance physique séparant le photon d'un objet de coordonnée comobile χ_0 , mesurée le long d'une géodésique radiale, vaut $R(t) [\chi_0 - \chi(t)]$, avec

$$R(t) = \mathcal{R}_0 \sqrt{\left(1 + \frac{ct}{\mathcal{R}_0}\right)^2 - 1}.$$

Comme le montre le graphique représenté sur la figure 8.7, cette expression présente un maximum : *l'expansion de l'univers est telle que, pendant un certain temps, la distance entre le photon parti de l'origine et l'objet situé en χ_0 croît avec le temps cosmologique, bien que le photon fasse route vers l'objet en question.*

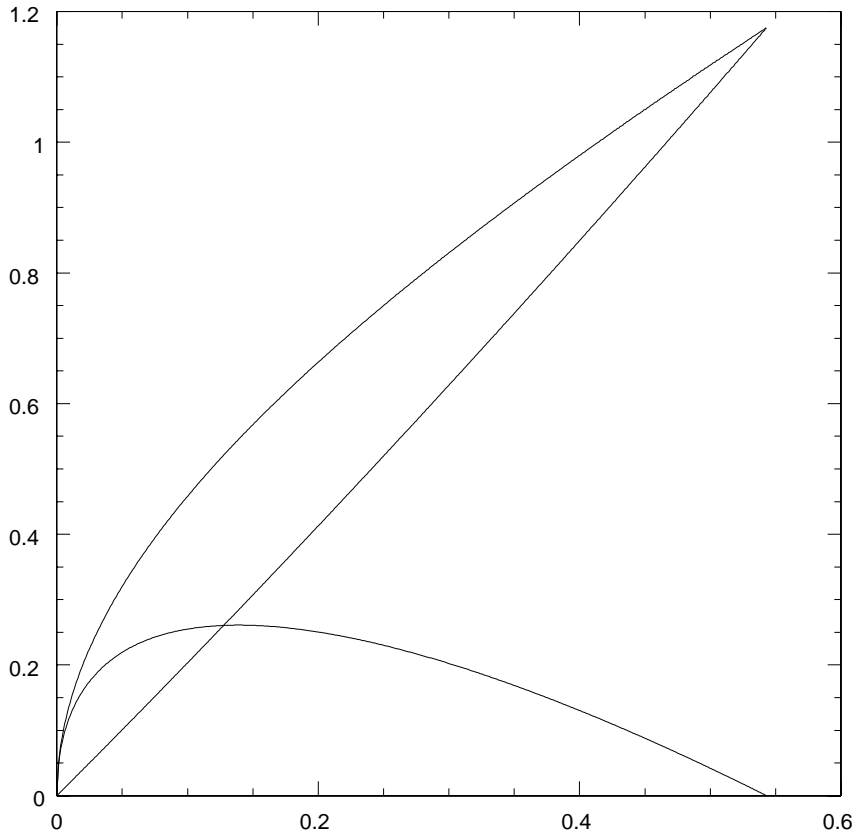


FIGURE 8.7: On a respectivement porté, en unités \mathcal{R}_0 et en fonction du temps cosmologique ct , la distance physique à l'origine d'une particule de coordonnée radiale $\chi_0 = 1$ (expansion de l'univers), la distance physique à l'origine d'un photon émis en $t = 0$ (horizon-particule) et la différence entre ces deux quantités.

La distance propre totale parcourue par le photon est donnée par

$$d(t) = \int_0^\eta R(\eta') d\chi = \int_0^t c dt' = ct,$$

résultat facilement accessible à l'intuition.

Exercice 8.5 Lorsque l'on remonte le temps vers la singularité cosmologique, le modèle d'« univers de poussière » n'est manifestement plus correct à l'époque du découplage de la radiation. Il est légitime de le raccorder à un « univers de rayonnement », en supposant que le facteur d'échelle et sa dérivée sont continus. Dans le cadre des modèles spatialement euclidiens, montrer que l'époque du découplage calculée selon l'équation d'état $P = 0$ est égale aux $4/3$ du résultat obtenu via l'équation d'état $P = \rho/3$.

Suggestion. D'après les sections 8.9.1 et 8.9.2, on a respectivement :

$$R(t) = \left(9 \mathcal{R}_0 c^2 / 2\right)^{1/3} t^{2/3} \quad \text{si } P = 0$$

et

$$R(t) = (2 \mathcal{R}_0 c)^{1/2} t^{1/2} \quad \text{si } P = \rho/3.$$

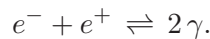
La conclusion en découle immédiatement.

8.10 Cosmologie physique

8.10.1 Histoire thermique de l'univers

Les notions de mécanique statistique nécessaires à la compréhension de cette section sont rappelées en annexe D.

Le modèle dit *standard* postule tout d'abord que l'univers peut être décrit dans le cadre des géométries spatialement homogènes et isotropes de Friedmann-Lemaître. On suppose ensuite que la matière est constituée de particules élémentaires en équilibre thermodynamique à tout instant. Cela revient à admettre que les temps caractéristiques des réactions entre les particules sont largement inférieurs au temps caractéristique d'expansion de l'univers, donné par $1/H$, de sorte que les équilibres entre les concentrations des diverses particules sont déplacés « en temps réel ». ¹⁵ Notamment, à température suffisamment élevée, chaque type de particule se trouve en équilibre par rapport à la formation de paires. Par exemple, pour les électrons et les positrons on a



Or, on peut montrer que, si plusieurs types de particules A_i sont liées par la réaction chimique (au sens large et présentée comme une équation algébrique)

$$\sum_i \nu_i A_i = 0,$$

alors le potentiel chimique est conservé additivement ¹⁶ :

$$\sum_i \nu_i \mu_i = 0.$$

¹⁵. Cette hypothèse n'est donc certainement pas valable au voisinage de la singularité cosmologique. La théorie de la nucléosynthèse des éléments légers montre qu'elle est également incorrecte pour $T < 10^9$ K. Voir par exemple S. Weinberg, *op. cit.*, pp. 552-553 ; S. Weinberg, *Les trois premières minutes de l'univers*, Seuil, 1978, figure 9.

¹⁶. Voir par exemple L. Landau et E. Lifchitz, *Physique théorique, Tome 5, Physique statistique*, Mir, 1967, § 103.

Le potentiel chimique des photons étant nul par (D.8), les particules et les antiparticules d'une espèce donnée ont ainsi des potentiels chimiques opposés.

Dans le cadre du modèle standard, on admet en première étude que le nombre $N^+(\mu_i^+)$ de particules d'un type i donné, est égal au nombre $N^-(\mu_i^-)$ d'antiparticules correspondantes. D'autre part, d'après la conclusion du paragraphe précédent, $\mu_i^+ = -\mu_i^-$. On en déduit, par (D.2), l'annulation du potentiel chimique de chaque catégorie de particules :

$$\mu_i = 0 \quad \forall i. \quad (8.60)$$

Les grandeurs caractéristiques associées à chaque type de particules ne dépendent alors plus que de la température et non pas des détails des réactions entre particules. La densité, notamment, se calcule via la relation (D.2) avec $\mu=0$, c'est-à-dire

$$\frac{N}{V} = \frac{g}{2\pi^2} \left(\frac{kT}{\hbar c} \right)^3 \int_{\frac{m_0 c^2}{kT}}^{+\infty} \frac{z \sqrt{z^2 - \left(\frac{m_0 c^2}{kT} \right)^2}}{e^z \pm 1} dz.$$

Pour $kT \ll m_0 c^2$ elle varie comme $\exp[-m_0 c^2/(kT)]$ et peut donc être négligée. Pour $kT \gg m_0 c^2$, les relations (D.2) et (D.3) particularisées à $\mu = 0$ s'écrivent

$$\frac{N}{V} = \frac{g}{2\pi^2} \left(\frac{kT}{\hbar c} \right)^3 \int_0^{+\infty} \frac{z^2}{e^z \pm 1} dz \quad (8.61)$$

$$E = \frac{gV}{2\pi^2} \frac{(kT)^4}{(\hbar c)^3} \int_0^{+\infty} \frac{z^3}{e^z \pm 1} dz \quad (8.62)$$

(la valeur de toutes les intégrales est donnée en annexe D.3). Le gaz est alors ultrarelativiste car $E/N \approx kT \gg m_0 c^2$ et sa pression se déduit de l'équation d'état (D.11) $P = \rho/3$. En pratique, en première étude, on considère que chaque type de particules n'est présent que pour $kT \geq m_0 c^2$ et se trouve alors décrit par les relations (8.61) et (8.62).

La loi de Stefan-Boltzmann (8.62) implique le résultat $\rho = KT^4$ où la « constante » K dépend en fait des types de particules ultrarelativistes en équilibre avec la radiation à la température T , par le biais de leurs poids statistiques et de leur qualité de fermion ou de boson. Si l'on rappelle la relation (8.43) selon laquelle $\rho = K_2/R^4$ (où il convient maintenant de noter que la « constante » K_2 change, il est vrai sans modifier son ordre de grandeur, chaque fois qu'une famille de particules s'annihile avec les antiparticules correspondantes), on en déduit la « constance » de la quantité RT . Ce résultat permet de préciser quantitativement la notion d'*histoire thermique de l'univers*.

Nous n'envisagerons l'histoire thermique de l'univers que pour des époques postérieures à celle où $T \approx 10^{13}$ K, vers la fin de l'ère *hadronique*, caractérisée par l'annihilation massive des baryons et des antibaryons (les plus légers des hadrons, les pions, s'annihilent pour $T \approx 1.5 \times 10^{12}$ K).

La température $T \approx 10^{12}$ K marque le début de l'ère *leptonique*. Cohabitent alors, en équilibre thermique, des photons, des muons, des antimuons, des électrons, des positrons, des neutrinos et des antineutrinos des trois types (électronique, muonique, tauonique), plus quelques protons et neutrons ainsi que des pions. Ces derniers achèvent leur annihilation et seront bientôt imités par les muons. A la température $T \approx 10^{10}$ K se produit le découplage de l'interaction faible, c'est-à-dire des neutrinos et des antineutrinos. La faiblesse de leur interaction mutuelle ne permet pas leur annihilation et ces particules mènent dès lors une existence indépendante.

La situation sera différente pour les électrons et les positrons qui, soumis en plus à l'interaction électromagnétique, vont s'annihiler en $T \approx 5 \times 10^9$ K, marquant l'avènement de l'ère *radiative*.¹⁷ Cette dernière cède le relais à l'ère *matérielle* au moment où la densité d'énergie lumineuse devient inférieure à la densité d'énergie matérielle.

Exercice 8.6 *La mesure de la température du rayonnement de fond cosmologique fournit la valeur $T = 2.73$ K.*

- Déduire la densité de photons dans l'univers à l'époque actuelle.*
- Calculer leur densité d'énergie.*
- On estime que la valeur actuelle de la masse volumique de la matière baryonique est égale à 4×10^{-31} g/cm³, correspondant à $\rho_m = 3.6 \times 10^{-11}$ J/m³. Déterminer l'époque critique de transition entre l'ère radiative et l'ère matérielle.*
- Quelle était alors la température du rayonnement ?*

Solution.

- De (D.9) on tire immédiatement $N/V = 416$ photons/cm³.
- Vu (D.10), $\rho_r = 4.2 \times 10^{-14}$ J/m³.
- Puisque l'énergie volumique de la matière ρ_m évolue comme $1/R^3$ et celle de la radiation ρ_r comme $1/R^4$ [relations (8.42) et (8.43)], le quotient ρ_r/ρ_m évolue comme $1/R$. Les deux densités d'énergie étant actuellement dans un rapport de 1 à 1000 environ, on en déduit qu'elles étaient du même ordre de grandeur lorsque le facteur d'échelle était environ 1000 fois plus petit, donc pour $z \approx 1000$.
- Rappelons la constance de la quantité RT . Par conséquent, à la fin de l'ère radiative la température était 1000 fois plus élevée qu'à l'époque actuelle, donc était de l'ordre de 3000 K. Typiquement, il s'agit de la température à laquelle la matière est passée de l'état ionisé à l'état atomique, donc la température de découplage de la radiation. Cette coïncidence doit être considérée comme accidentelle.

8.10.2 L'entropie de l'univers

L'utilisation, dans les équations d'Einstein, du tenseur d'énergie-impulsion d'un fluide parfait (donc sans dissipation d'énergie) impose la constance de l'entropie d'un élément du fluide¹⁸. En effet, l'équation de conservation (8.33)

$$\frac{d}{dt}(\rho R^3) + P \frac{d}{dt} R^3 = 0$$

peut s'écrire sous la forme

$$dE + P dV = 0. \quad (8.63)$$

Or, l'entropie S vérifie

$$T dS = dE + P dV - \sum_i \mu_i dN_i.$$

Compte tenu de (8.63) et de l'annulation (8.60) des potentiels chimiques μ_i , on a ainsi

$$dS = 0. \quad (8.64)$$

17. Cette ultime annihilation s'est produite aux environs de la première seconde. Depuis lors, les « constantes » K_1 et K_2 n'ont plus évolué. On commet donc une erreur négligeable, dans le calcul de l'âge de l'univers, si on les suppose inchangées depuis le Big Bang jusqu'à l'époque actuelle. Cela justifie la procédure utilisée dans l'équation (8.51).

18. Cette propriété est bien connue dans le cas où le nombre de particules du fluide parfait est conservé : voir p. 132.

Dans le cadre du modèle standard, l'entropie S d'un covolume V ne change pas au cours de l'expansion de l'univers : cette dernière est donc réversible¹⁹.

Remarquons que l'annulation de l'énergie libre de Gibbs $\Phi = E - TS + PV = \sum N_i \mu_i$ impose

$$S = \frac{1}{T} (E + PV).$$

Par conséquent, dans l'univers en évolution l'entropie $S(V, T, N_i)$ d'un système de particules en équilibre à la température T dans un covolume V est donnée par

$$S(V, T, N_i) = \frac{1}{T} [\rho(T)V + P(T)V]. \quad (8.65)$$

8.10.3 Les neutrinos cosmologiques

L'annihilation des électrons et des positrons se produit vers 5×10^9 K. La diminution d'entropie qui en résulte est compensée par l'augmentation d'entropie du gaz de photons (en effet, comme on l'a vu plus haut, dans un covolume donné l'entropie *totale* du gaz d'électrons, de positrons et de photons est conservée). Il s'ensuit une augmentation du nombre de photons par unité de covolume, c'est-à-dire une augmentation de la température du gaz de photons. Cela n'affecte cependant pas les neutrinos qui, à cette température, sont découplés du gaz d'électrons, de positrons et de photons. Leur entropie dans un covolume donné reste ainsi constante, et dès lors leur température continue à varier proportionnellement à l'inverse du facteur d'échelle R .

Il est assez facile de calculer le rapport entre la température des neutrinos et celle des photons à l'issue de l'annihilation électron-positron. Pour cela, remarquons que l'entropie partielle de chacun de ces constituants peut être calculée par la relation (8.65)

$$S(V, T) = \frac{1}{T} (\rho V + PV)$$

avec $P = \rho/3$ dans le cas ultrarelativiste. Donc

$$S = \frac{4\rho V}{3T}.$$

Pour les photons, on a

$$\rho_\gamma = \frac{1}{\pi^2} \frac{(kT)^4}{(\hbar c)^3} \int_0^{+\infty} \frac{z^3}{e^z - 1} dz.$$

Si les indices av et ap caractérisent les quantités *avant* et *après* l'annihilation, alors

$$\begin{aligned} (\rho_{e^-})_{av} &= (\rho_{e^+})_{av} = \frac{1}{\pi^2} \frac{(kT)_{av}^4}{(\hbar c)^3} \int_0^{+\infty} \frac{z^3}{e^z + 1} dz = \frac{7}{8} (\rho_\gamma)_{av} \\ (\rho_{e^-})_{ap} &= (\rho_{e^+})_{ap} = 0. \end{aligned}$$

La conservation de l'entropie totale implique

$$\left(\frac{7}{8} + \frac{7}{8} + 1 \right) \frac{4(\rho_\gamma)_{av} R_{av}^3}{3(T_\gamma)_{av}} = \frac{4(\rho_\gamma)_{ap} R_{ap}^3}{3(T_\gamma)_{ap}}.$$

19. Il n'en est pas moins vrai que les étoiles, quant à elles, évoluent de façon irréversible. Elles ont ainsi accru l'entropie cosmique d'environ une partie pour mille depuis l'émission du rayonnement fossile : pour 1000 photons qui parcourent le ciel aujourd'hui, 999 voyagent depuis l'époque du découplage électromagnétique. Voir par exemple H. Reeves, *Dernières nouvelles du cosmos, Tome 1, Vers la première seconde*, Seuil, 1994, pp. 109-110 et 128.

On en déduit

$$\frac{(RT_\gamma)_{ap}}{(RT_\gamma)_{av}} = \sqrt[3]{\frac{11}{4}}.$$

Comme, d'autre part, la température T_ν des neutrinos vérifie la relation

$$(RT_\gamma)_{av} = (RT_\nu)_{av} = (RT_\nu)_{ap},$$

on obtient

$$\frac{(T_\gamma)_{ap}}{(T_\nu)_{ap}} = \sqrt[3]{\frac{11}{4}}.$$

Ainsi, depuis l'annihilation des électrons et des positrons, la température des neutrinos et celle des photons évoluent selon la loi

$$RT_\nu = \sqrt[3]{\frac{4}{11}} RT_\gamma = \text{constante}, \quad \text{donc} \quad T_\nu = \sqrt[3]{\frac{4}{11}} T_\gamma,$$

de sorte qu'à l'époque actuelle

$$T_\nu = \sqrt[3]{\frac{4}{11}} \times 2.73 = 1.95 \text{ K}.$$

La densité de chaque espèce de neutrinos est égale à celle des photons multipliée par 1/2 (un seul état d'hélicité) $\times 3/4$ [fermions au lieu de bosons, comparer les intégrales (D.5) et (D.4)] $\times 4/11$ (cube du quotient des températures), soit, pour l'époque actuelle :

$$416 \times \frac{3}{22} \approx 57 \text{ neutrinos/cm}^3.$$

Tous ces résultats sont valides pour autant que les neutrinos demeurent ultrarelativistes. C'est évidemment toujours le cas si leur masse propre est nulle, comme on l'a longtemps admis. Mais on dispose maintenant de solides arguments théoriques et observationnels à l'encontre de cette hypothèse (oscillations des neutrinos). Si leur masse est suffisamment élevée, les neutrinos auraient pu voir leur température actuelle descendre sous la température de seuil et donc, devenir non relativistes.

8.10.4 Matière et antimatière

Désignons respectivement par n_γ , n et \bar{n} le nombre de photons, de baryons (protons et neutrons) et d'antibaryons (antiprotons et antineutrons) par unité de volume. On définit le nombre baryonique volumique par la quantité

$$n_B = n - \bar{n}.$$

Calculons le rapport n_γ/n_B pour l'époque actuelle. Etant donné que l'antimatière semble largement minoritaire par rapport à la matière, on peut négliger \bar{n} devant n , ce qui revient à écrire $n_B \approx n$. Or, la masse du proton, approximativement égale à celle du neutron, vaut 1.67×10^{-24} g. En admettant la valeur 4×10^{-31} g/cm³ pour la masse volumique de la matière baryonique, on obtient

$$n_B \approx n \approx \frac{4 \times 10^{-31}}{1.67 \times 10^{-24}} \approx 2.4 \times 10^{-7}.$$

Ainsi, vu les résultats de l'exercice 8.6,

$$\frac{n_\gamma}{n_B} \approx \frac{420}{2.4 \times 10^{-7}} \approx 10^9.$$

Comme, d'après les lois rappelées dans la section D.4, le nombre n_γ est proportionnel à l'entropie volumique du gaz de photons, on dit que l'*entropie par baryon* de l'univers actuel est de l'ordre du milliard.

Discutons l'évolution du rapport n_γ/n_B au cours de l'histoire thermique de l'univers. D'une part, on postule habituellement la conservation du nombre baryonique dans un covolume donné. La quantité $n_B R^3$ est donc constante. D'autre part, le nombre de photons dans un covolume donné, proportionnel à $n_\gamma R^3$, est également constant depuis que la température est tombée en-dessous de la valeur $T \approx 5 \times 10^9$ K, c'est-à-dire après l'annihilation des électrons et des positrons (remarquons que cette condition de conservation de l'entropie du gaz de photons est plus exigeante que la conservation de l'entropie *totale* de l'univers ; elle implique que la température T du rayonnement varie alors comme l'inverse du facteur d'échelle R). Avant cette annihilation, ce nombre était 4/11 fois aussi élevé, ce qui ne change pas son ordre de grandeur (un phénomène semblable se produit lors de l'annihilation particule-antiparticule de chaque espèce). Dans ces conditions, la quantité $n_\gamma R^3$ peut *grosso modo* être considérée comme constante dans le cadre du modèle standard. Il s'ensuit la constance du rapport n_γ/n_B , qui, compte tenu de sa valeur actuelle estimée ci-dessus, vaut à peu près un milliard.

De tout cela on peut tirer des conclusions intéressantes concernant la dissymétrie entre la matière et l'antimatière. En effet, au-delà de la température d'annihilation des nucléons et des antinucléons, c'est-à-dire pour $T > 10^{13}$ K, on a

$$n \approx \bar{n} \approx n_\gamma,$$

donc

$$\frac{\bar{n}}{n - \bar{n}} \approx 10^9$$

ou encore

$$\frac{n}{\bar{n}} \approx 1 + 10^{-9}.$$

A cette époque, il y avait donc un excès de un nucléon par milliard de paires nucléon-antinuécléon.

Il est également possible de jauger la pertinence de l'hypothèse d'annulation du nombre baryonique adoptée dans cette section pour décrire l'époque actuelle. On peut en effet calculer que, s'il y avait *exactement* autant de matière que d'antimatière, alors, par 10^{18} photons, une seule paire nucléon-antinuécléon aurait échappé à la destruction survenue vers la fin de l'ère hadronique ; et ce rapport n'a pas évolué depuis la fin de l'annihilation.²⁰ Appelons x le nombre actuel de nucléons par couple nucléon-antinuécléon rescapé de la fournaise. Il y a donc $x + 1$ nucléons pour 10^{18} photons. Or, comme on l'a vu, on trouve actuellement 10^9 photons par nucléon. On en déduit que $x = 10^9$. Ainsi, la quantité résiduelle d'antimatière dans l'univers est de l'ordre de un antinucléon par milliard de nucléons.

²⁰. Voir par exemple

J. Demaret, *op. cit.*, p. 113 ;

J. Barrow et J. Silk, *La main gauche de la création*, Londreys, 1985, pp. 96-97.

Chapitre 9

La solution de Schwarzschild

9.1 L'espace-temps à symétrie sphérique spatiale

En 1916, K. Schwarzschild obtint la première solution exacte des équations d'Einstein. Historiquement, cette métrique spécifiait la géométrie engendrée par une étoile sphérique en équilibre hydrostatique et de masse volumique uniforme. Plus généralement, nous nous proposons ici d'écrire les équations d'Einstein lorsque les sources possèdent la symétrie sphérique, quelle que soit leur équation d'état et quel que soit leur mouvement, pourvu qu'il demeure central et symétrique.

Etablissons tout d'abord la métrique générale d'un espace-temps invariant sous le groupe d'isométries $SO(3)$. Cela veut dire que chaque point P appartient à une 2-sphère s de points équivalents obtenus par l'action du groupe des rotations :

$$s = \{M(P)\}, \quad M \in SO(3).$$

Les sphères sont des espaces à symétrie maximale. Le théorème 8.1 (avec $n = 4$ et $m = 2$) garantit alors l'existence de coordonnées $\{v^a\} = \{T, R\}$ et $\{u^i\} = \{\theta, \varphi\}$ adaptées à cette symétrie, et permettant d'exprimer l'élément de longueur spatio-temporelle sous la forme

$$ds^2 = g_{TT}(T, R) dT^2 + 2g_{TR}(T, R) dT dR + g_{RR}(T, R) dR^2 + f(T, R) d\Omega^2 \quad (9.1)$$

où T et R sont des coordonnées respectivement temporelle et spatiale ; $f(T, R) > 0$; et

$$d\Omega^2 = d\theta^2 + \sin^2 \theta d\varphi^2$$

est la métrique sur la sphère unité.

[Tâchons de décrire intuitivement le résultat (9.1). On a harmonisé les coordonnées angulaires θ et φ pour que, sur une section spatiale (c'est-à-dire à coordonnée temporelle T constante), chaque ligne orthogonale aux sphères rencontre des coordonnées angulaires constantes. Ces lignes sont les lignes de coordonnée R puisqu'obtenues en gelant T , θ et φ . Ce choix implique l'annulation des termes en $dR d\theta$ et $dR d\varphi$. D'autre part, il est naturel de s'attendre à ce que la métrique demeure inchangée sous l'isométrie

$$\theta \rightarrow \pi - \theta, \quad \varphi \rightarrow 2\pi - \varphi,$$

correspondant à une rotation d'angle π autour de l'axe x des coordonnées sphériques et revenant à permuter les pôles nord et sud des sphères. Mais suite à cette opération, les termes croisés en $dT d\theta$ et $dT d\varphi$ changent de signe. Ils sont donc nuls, ce qui signifie

que les lignes de coordonnée temporelle T , obtenues en fixant R , θ et φ , sont elles aussi orthogonales aux sphères. Au demeurant, s'il était impossible de choisir les lignes de temps orthogonales aux sphères, leur projection sur lesdites sphères indiquerait une direction privilégiée dans l'espace. Cela contredirait l'hypothèse de l'invariance par rotation (on rencontre cette situation si les sources sont animées d'un mouvement de rotation).]

Pour tout T , il est avantageux de poser $r = \sqrt{f(T, R)}$, afin que l'aire d'une sphère associée à la coordonnée r vaille $4\pi r^2$. Il en résulte

$$ds^2 = g_{TT}(T, r) dT^2 + 2g_{Tr}(T, r) dT dr + g_{rr}(T, r) dr^2 + r^2 d\Omega^2.$$

Enfin, un changement de variable temporelle $T = g(t, r)$ permet d'éliminer le terme croisé. En effet, la relation $dT = \dot{g} dt + g' dr$ (où le point désigne la dérivée partielle par rapport à t et le « prime » la dérivée partielle par rapport à r) implique

$$ds^2 = g_{TT} \dot{g}^2 dt^2 + 2\dot{g} (g_{TT} g' + g_{Tr}) dt dr + (g_{TT} g'^2 + 2g_{Tr} g' + g_{rr}) dr^2 + r^2 d\Omega^2,$$

et si $g(t, r)$ satisfait l'équation différentielle $g_{TT}(g, r) g' + g_{Tr}(g, r) = 0$ on en déduit

$$ds^2 = g_{TT} \dot{g}^2 dt^2 + (g_{rr} - g_{TT} g'^2) dr^2 + r^2 d\Omega^2.$$

On obtient ainsi l'expression canonique de l'élément de longueur pour un espace-temps à symétrie sphérique spatiale dans les coordonnées (ct, r, θ, φ) dites « de Schwarzschild » :

$$ds^2 = -e^{2\nu(t,r)} c^2 dt^2 + e^{2\lambda(t,r)} dr^2 + r^2 d\Omega^2. \quad (9.2)$$

Cette forme est préservée lors d'un changement de paramètre sur les lignes de temps. On est donc autorisé à ajouter à $\nu(t, r)$ une fonction arbitraire dépendant du temps sans modifier la situation physique. Cette latitude sera exploitée dans la suite. Remarquons aussi que l'hypothèse de symétrie sphérique spatiale n'implique pas toujours l'existence d'un « centre » : la borne inférieure de la coordonnée radiale r n'est pas nécessairement nulle.

On calcule facilement les composantes mixtes $G^\alpha_\beta = G_\beta^\alpha$ du tenseur d'Einstein. En voici les composantes non identiquement nulles, où la dérivée temporelle est notée par un point et la dérivée radiale par un « prime » :

$$G_0^0 = \frac{1 - 2r\lambda' - e^{2\lambda}}{r^2 e^{2\lambda}} \quad (9.3)$$

$$G_1^1 = \frac{1 + 2r\nu' - e^{2\lambda}}{r^2 e^{2\lambda}} \quad (9.4)$$

$$G_1^0 = \frac{-2\dot{\lambda}}{r e^{2\nu}} \quad G_0^1 = \frac{2\dot{\lambda}}{r e^{2\lambda}} \quad (9.5)$$

$$G_2^2 = G_3^3 = \frac{\nu'' + \nu'^2 - \lambda'\nu'}{e^{2\lambda}} - \frac{\ddot{\lambda} + \dot{\lambda}^2 - \dot{\lambda}\dot{\nu}}{e^{2\nu}} + \frac{\nu' - \lambda'}{r e^{2\lambda}}. \quad (9.6)$$

Le tenseur d'Einstein est nul en-dehors des sources, c'est-à-dire dans les régions de l'espace-temps où le tenseur d'énergie-impulsion est égal à zéro. Attention, cette condition n'est pas réalisée aux endroits dépourvus de « matière » mais où règne un champ électromagnétique, par exemple en-dehors d'une sphère métallique chargée. Pour connaître le champ de gravitation généré dans ce cas de figure, il faut mener une étude spécifique qui sort du cadre de ce cours (métrique de *Reissner-Nordström*, 1918). Dans les sections qui suivent, nous allons résoudre les équations d'Einstein dans le vide et à l'intérieur d'une sphère constituée d'un fluide parfait incompressible.

9.2 Solution dans le vide

Compte tenu des relations (9.3) à (9.5), les équations d'Einstein dans le vide se réduisent respectivement à

$$1 - 2r\lambda' - e^{2\lambda} = 0 \quad (9.7)$$

$$\nu' = -\lambda' \quad (9.8)$$

$$\dot{\lambda} = 0 \quad (9.9)$$

[l'équation (9.6) est alors automatiquement vérifiée]. Vu (9.9), λ ne dépend que de la variable radiale. La relation (9.8) fournit $\nu(t, r) = -\lambda(r) + f(t)$ où $f(t)$ est une fonction arbitraire du temps qui peut être résorbée en calibrant judicieusement la coordonnée temporelle, ainsi que nous l'avons expliqué dans la section 9.1. Il vient ainsi $\nu(r) = -\lambda(r)$. On trouve leur expression en intégrant (9.7) :

$$e^{-2\lambda(r)} = e^{2\nu(r)} = 1 + \frac{K}{r}$$

où K est une constante. On a donc la solution

$$ds^2 = -(1 + K/r)c^2 dt^2 + \frac{dr^2}{1 + K/r} + r^2 d\Omega^2.$$

Si l'origine $r = 0$ se trouve dans le vide, alors la variété riemannienne est singulière en ce point, ce qui est inadmissible, sauf si $K = 0$. L'espace-temps est alors minkowskien. Ainsi, le champ de gravitation est nul à l'intérieur d'une cavité sphérique centrée sur l'origine, comme chez Newton.

Si le domaine du vide s'étend jusqu'à l'infini, le changement de variable radiale

$$r = \underline{r} \left(1 - \frac{K}{4\underline{r}}\right)^2$$

fait apparaître la solution extérieure sous la forme spatialement isotrope

$$ds^2 = - \left[\frac{1 + K/4\underline{r}}{1 - K/4\underline{r}} \right]^2 c^2 dt^2 + \left(1 - \frac{K}{4\underline{r}}\right)^4 (d\underline{r}^2 + \underline{r}^2 d\Omega^2) \quad (9.10)$$

que l'on peut développer en série de puissances de $1/\underline{r}$: en considérant que $\underline{r}, \theta, \varphi$ sont les coordonnées sphériques associées à des coordonnées cartésiennes x, y, z , on trouve asymptotiquement une métrique

$$ds^2 = - \left[1 + \frac{K}{\underline{r}} + O(2) \right] c^2 dt^2 + \left[1 - \frac{K}{\underline{r}} + O(2) \right] (dx^2 + dy^2 + dz^2)$$

du type « champ loin des sources » (7.66)

$$ds^2 = - \left(1 + \frac{2\Phi}{c^2} \right) c^2 dt^2 + \left(1 - \frac{2\Phi}{c^2} \right) (dx^2 + dy^2 + dz^2).$$

Le potentiel gravitationnel Φ engendré par une masse newtonienne M à symétrie sphérique est donné par $\Phi = -GM/\underline{r}$. Par identification, il est donc naturel de poser $K = -2GM/c^2$. Le paramètre M est la quantité qui intervient dans la troisième loi de Kepler appliquée au

mouvement d'une particule test loin de la source, même si celle-ci est un champ fort, et on continuera à l'appeler la *masse* de la source. La solution extérieure s'écrit ainsi

$$\boxed{ds^2 = - \left(1 - \frac{2GM}{rc^2}\right) c^2 dt^2 + \frac{dr^2}{1 - \frac{2GM}{rc^2}} + r^2 d\Omega^2.} \quad (9.11)$$

La quantité $r_S = 2GM/c^2$ a reçu le nom de *rayon de Schwarzschild* ou encore, *rayon gravitationnel*.

En relativité générale, il n'existe pas d'autre solution extérieure à symétrie sphérique spatiale : c'est le *théorème de G.D. Birkhoff* (1923). La géométrie de l'espace-temps est statique en-dehors du rayon d'une étoile en mouvement radial, en effondrement par exemple ou bien en pulsation (notamment, il ne peut y avoir émission d'ondes gravitationnelles monopolaires). A nouveau, on reconnaîtra facilement l'homologue de ce résultat en mécanique newtonienne.

Comme t est une coordonnée temporelle et r une coordonnée spatiale, la solution (9.11) n'est *a priori* pas cohérente si le domaine du vide pénètre à l'intérieur du rayon de Schwarzschild. Aussi supposons-nous, dans un premier temps, que la matière n'est pas entièrement contenue dans une sphère de paramètre r inférieur à $2GM/c^2$. D'un point de vue astrophysique, cette hypothèse semble à peine restrictive (par exemple, le rayon de Schwarzschild du soleil vaut 3 km). Nous la relâcherons dans la section 9.11.

9.3 Solution intérieure

La relation (9.3) fournit l'équation d'Einstein

$$\frac{1 - 2r\lambda' - e^{2\lambda}}{r^2 e^{2\lambda}} = \frac{8\pi G}{c^4} T_0^0,$$

qui peut prendre la forme

$$\left[\frac{r(e^{-2\lambda} - 1)}{r^2} \right]' = \frac{8\pi G}{c^4} T_0^0.$$

On en connaît la solution générale pour des sources statiques. Lorsque $r \rightarrow 0$, $\lambda(r)$ doit s'annuler (au moins comme r^2), sinon T_0^0 serait singulier en $r = 0$.¹ Dès lors,

$$\lambda(r) = -\frac{1}{2} \ln \left[1 + \frac{8\pi G}{rc^4} \int_0^r T_0^0(r') r'^2 dr' \right]. \quad (9.12)$$

Notons

$$m(r) = -\frac{4\pi}{c^2} \int_0^r T_0^0(r') r'^2 dr', \quad (9.13)$$

ce qui revient à poser

$$e^{2\lambda(r)} = \frac{1}{1 - \frac{2Gm(r)}{rc^2}}. \quad (9.14)$$

La fonction λ doit être continue lorsque l'on raccorde la solution intérieure (9.14) avec la solution extérieure (9.11). Notons R la coordonnée radiale de la sphère délimitant la

1. On peut également reprendre l'argument de la note 5 p. 179.

matière. La relation (9.13) détermine l'expression de la masse totale des sources par le tenseur d'énergie-impulsion :

$$M = -\frac{4\pi}{c^2} \int_0^R T_0^0(r') r'^2 dr'. \quad (9.15)$$

Particularisons encore en supposant que les sources constituent un fluide parfait. On calcule facilement

$$\begin{aligned} U^0 &= c e^{-\nu(r)}, & U^i &= 0 \\ U_0 &= -c e^{\nu(r)}, & U_i &= 0 \\ T_0^0 &= -\rho(r), & T_i^i &= T_0^i = 0, & T_1^1 &= T_2^2 = T_3^3 = P(r). \end{aligned}$$

Les quantités ρ et P sont respectivement l'énergie volumique et la pression du fluide (mesurées, par définition, dans une tétrade locale orthonormée immobile par rapport à ce dernier, c'est-à-dire, dont le vecteur temporel est égal à la quadrivitesse du fluide divisée par c).

La relation (9.13) donne alors

$$m(r) = \frac{4\pi}{c^2} \int_0^r \rho(r') r'^2 dr'. \quad (9.16)$$

Compte tenu de (9.4) et de (9.14), on obtient l'équation pour ν :

$$\nu'(r) = \frac{G [4\pi r^3 P(r) + m(r)c^2]}{rc^2 [rc^2 - 2Gm(r)]}. \quad (9.17)$$

Enfin, à l'aide de (9.6), (9.14), (9.16) et (9.17) on trouve l'équation d'équilibre hydrostatique de *J.R. Oppenheimer - G.M. Volkoff* :

$$P'(r) = \frac{-G [4\pi r^3 P(r) + m(r)c^2] [\rho(r) + P(r)]}{rc^2 [rc^2 - 2Gm(r)]}. \quad (9.18)$$

Les équations (9.14), (9.16), (9.17) et (9.18), couplées à une équation d'état liant la pression et l'énergie volumique, déterminent complètement la structure interne d'une étoile sphérique, statique et constituée d'un fluide parfait. En combinant (9.17) et (9.18), on obtient

$$P'(r) = -[\rho(r) + P(r)] \nu'(r),$$

qui est en fait identique à $T^{1\mu}_{;\mu} = 0$, la seule des équations de mouvement qui ne se réduise pas à une identité.

On peut intégrer les équations pour un fluide homogène et incompressible de masse volumique constante $\mu = \rho/c^2$. L'équation (9.16) donne immédiatement

$$m(r) = \frac{4}{3} \pi r^3 \mu. \quad (9.19)$$

L'équation (9.18) prend alors la forme

$$P'(r) = \frac{-4\pi G r [\mu c^2 + P(r)] [\mu c^2 + 3P(r)]}{3c^4 - 8\pi G \mu r^2 c^2}.$$

On l'intègre facilement : en notant $P(0) = P_c$ la valeur de la pression au centre de la sphère, on trouve

$$\frac{\mu c^2 + 3P(r)}{\mu c^2 + P(r)} = \frac{\mu c^2 + 3P_c}{\mu c^2 + P_c} \sqrt{1 - \frac{8\pi G \mu r^2}{3c^2}}. \quad (9.20)$$

Comme $P(R) = 0$ (la pression s'annule à la surface), on a

$$P_c = -\mu c^2 \frac{1 - \sqrt{1 - \frac{2GM}{Rc^2}}}{1 - 3\sqrt{1 - \frac{2GM}{Rc^2}}}. \quad (9.21)$$

Remarquons que la pression au centre augmente si le rayon diminue. Elle tend vers l'infini si le dénominateur tend vers zéro, c'est-à-dire si le rayon tend vers les 9/8 du rayon de Schwarzschild. Physiquement, cela signifie que, du point de vue de la relativité générale, il ne peut exister d'étoile sphérique, statique, homogène² et stable de rayon

$$R \leq \frac{9}{8} r_s.$$

Remplaçant dans (9.20) la pression centrale par sa valeur (9.21) en fonction du rayon, on obtient

$$P(r) = -\mu c^2 \frac{\sqrt{1 - \frac{2GM r^2}{R^3 c^2}} - \sqrt{1 - \frac{2GM}{Rc^2}}}{\sqrt{1 - \frac{2GM r^2}{R^3 c^2}} - 3\sqrt{1 - \frac{2GM}{Rc^2}}}. \quad (9.22)$$

L'injection de ce résultat dans l'équation (9.17) donne

$$\nu'(r) = \frac{2GM r}{R^3 c^2 \left(3\sqrt{1 - \frac{2GM}{Rc^2}} - \sqrt{1 - \frac{2GM r^2}{R^3 c^2}} \right) \sqrt{1 - \frac{2GM r^2}{R^3 c^2}}},$$

donc

$$\nu(r) = \ln \left(3\sqrt{1 - \frac{2GM}{Rc^2}} - \sqrt{1 - \frac{2GM r^2}{R^3 c^2}} \right) + \ln K.$$

Pour raccorder continûment $\nu(r)$ à la solution extérieure (9.11) en $r = R$, il faut que

$$\nu(R) = \ln \sqrt{1 - \frac{2GM}{Rc^2}},$$

donc $K = 1/2$ et

$$e^{\nu(r)} = \frac{3}{2} \sqrt{1 - \frac{2GM}{Rc^2}} - \frac{1}{2} \sqrt{1 - \frac{2GM r^2}{R^3 c^2}}.$$

On obtient ainsi la solution intérieure pour une étoile statique, sphérique et constituée d'un fluide parfait incompressible :

$$ds^2 = -c^2 dt^2 \left(\frac{3}{2} \sqrt{1 - \frac{2GM}{Rc^2}} - \frac{1}{2} \sqrt{1 - \frac{2GM r^2}{R^3 c^2}} \right)^2 + \frac{dr^2}{1 - \frac{8\pi G \mu r^2}{3c^2}} + r^2 d\Omega^2. \quad (9.23)$$

2. Un théorème dû à Buchdahl stipule que cette conclusion est, en fait, indépendante de l'équation d'état du fluide. Voir H.A. Buchdahl, *Physical Review* **116**, pp. 1027 à 1034 (1959).

9.4 Visualisation des sections spatiales

Les sections spatiales de l'espace-temps de Schwarzschild en $\theta = \pi/2$ ont pour élément de longueur

$$ds^2 = \frac{dr^2}{1 - \frac{2Gm(r)}{rc^2}} + r^2 d\varphi^2 \quad (9.24)$$

où

$$m(r) = \frac{4}{3}\pi r^3 \mu \quad (r \leq R)$$

$$m(r) = \frac{4}{3}\pi R^3 \mu = M \quad (r \geq R).$$

La relation (4.15) permet de calculer leur courbure gaussienne :

$$K = \frac{2GM}{R^3 c^2} \quad (r < R)$$

$$K = -\frac{GM}{r^3 c^2} \quad (r > R).$$

Il est possible de réaliser la métrique (9.24) de façon intuitive à l'aide d'une surface de révolution $z = z(r)$ immergée dans un espace euclidien de métrique

$$ds^2 = dr^2 + r^2 d\varphi^2 + dz^2$$

$$= dr^2 \left\{ 1 + \left[\frac{dz(r)}{dr} \right]^2 \right\} + r^2 d\varphi^2. \quad (9.25)$$

Comparant (9.24) et (9.25), on obtient

$$\frac{dz}{dr} = \pm \sqrt{\frac{1}{1 - \frac{2Gm(r)}{rc^2}} - 1}.$$

Choisissant le signe positif et imposant $z(0) = 0$, on trouve

$$z(r) = \int_0^r \frac{d\xi}{\sqrt{\frac{\xi c^2}{2Gm(\xi)} - 1}},$$

c'est-à-dire

$$z(r) = R \left[\sqrt{\frac{Rc^2}{2GM}} - \sqrt{\frac{Rc^2}{2GM} - \left(\frac{r}{R}\right)^2} \right] \quad (r \leq R)$$

$$z(r) = \sqrt{\frac{8GMr}{c^2} \left(1 - \frac{2GM}{rc^2}\right)} + L \quad (r \geq R)$$

avec

$$L = R \sqrt{\frac{Rc^2}{2GM}} \left(1 - \sqrt{1 - \frac{2GM}{Rc^2}}\right) - \sqrt{\frac{8GMR}{c^2} \left(1 - \frac{2GM}{Rc^2}\right)}.$$

Du point de vue géométrique, on raccorde une sphère de rayon

$$\mathcal{R} = \sqrt{\frac{R^3 c^2}{2GM}}$$

(solution intérieure) à la surface obtenue par rotation d'une parabole d'axe horizontal autour de l'axe des z (solution extérieure) : voir la figure 9.1.

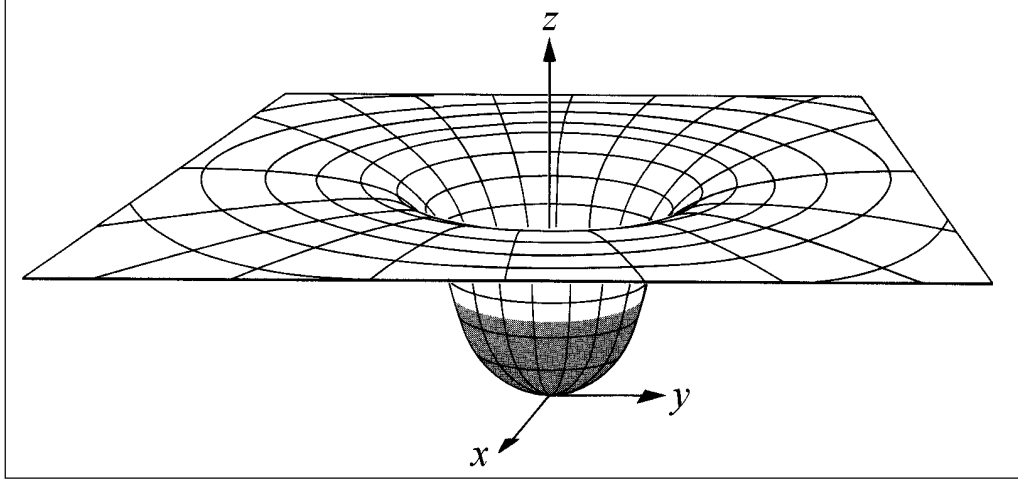


FIGURE 9.1: *Visualisation de la métrique de Schwarzschild engendrée par une étoile sphérique de rayon $R > 2GM/c^2$: sections spatiales (c'est-à-dire à temps t constant) en $\theta = \pi/2$. La partie correspondant à l'intérieur de l'astre est en gris. D'après C.W. Misner, K.S. Thorne, J.A. Wheeler, Gravitation, Freeman, 1973, figure 23.1.*

Dans le cadre de la solution intérieure, on peut établir la relation exacte $s(r)$ exprimant la distance physique entre le centre et un point de coordonnée radiale r . En effet, pour $r \leq R$ la distance physique mesurée le long d'un diamètre est donnée par

$$s(r) = \int_0^r \frac{d\xi}{\sqrt{1 - \frac{8\pi G\mu}{3c^2} \xi^2}}.$$

Notons que

$$\sqrt{\frac{8\pi G\mu}{3c^2}} \xi < 1 \quad \text{si} \quad 0 \leq \xi \leq R > 2GM/c^2.$$

Par conséquent,

$$s(r) = \sqrt{\frac{3c^2}{8\pi G\mu}} \arcsin \left(\sqrt{\frac{8\pi G\mu}{3c^2}} r \right).$$

On en déduit la quantité $\pi_{eff}(r)$, définie par le rapport effectif $C(r)/D(r)$ où $D(r) = 2s(r)$ est la longueur physique du diamètre d'un cercle de circonférence $C(r) = 2\pi r$.

Dans le cas limite où

$$\sqrt{\frac{8\pi G\mu}{3c^2}} r \ll 1, \quad \text{c. - à - d.} \quad \sqrt{\frac{2GM}{Rc^2}} \frac{r}{R} \ll 1,$$

certainement réalisé si $R \gg GM/c^2$, on peut approximer l'arc sinus par les deux premiers termes non nuls de son développement en série et écrire

$$\begin{aligned} s(r) &= r + \frac{4\pi G\mu}{9c^2} r^3 + O(5) \\ &= r \left(1 + \frac{GM}{3Rc^2} \frac{r^2}{R^2} \right) + O(5). \end{aligned}$$

A la surface de l'étoile, par définition $r = R$ et

$$\pi_{eff}(R) \cong \pi \left(1 - \frac{GM}{3Rc^2} \right).$$

Dans le cas du soleil,³ on a $GM/c^2 \cong 1.5$ km et $R \cong 7 \times 10^5$ km, ce qui donne

$$\pi_{eff}(R) \cong 3.141590\dots$$

si l'on suppose notre étoile homogène.

Au Palais de la Découverte à Paris se trouve une salle circulaire de cinq mètres de rayon, sur les murs de laquelle sont affichées les sept cents premières décimales du nombre π . Si, en son centre, on plaçait une sphère homogène de masse égale à cent kilogs et de circonférence égale à $2\pi \times 5$ m, la vingt-sixième décimale serait, en fait, incorrecte.⁴

9.5 Equations du mouvement et intégrales premières dans le vide

Soit une particule test dans une région de l'espace-temps décrite par la métrique extérieure (9.11). Son mouvement est piloté par les équations des géodésiques, écrites ci-dessous en fonction du paramètre affine τ — que l'on choisira naturellement égal au temps propre s'il s'agit d'une particule matérielle :

$$\begin{aligned} \frac{d^2t}{d\tau^2} + \frac{2GM}{r^2c^2 \left(1 - \frac{2GM}{rc^2} \right)} \frac{dt}{d\tau} \frac{dr}{d\tau} &= 0 \\ \frac{d^2r}{d\tau^2} + \left(1 - \frac{2GM}{rc^2} \right) \left[\frac{GM}{r^2} \left(\frac{dt}{d\tau} \right)^2 - \frac{GM}{r^2c^2 \left(1 - \frac{2GM}{rc^2} \right)^2} \left(\frac{dr}{d\tau} \right)^2 - r \frac{d\Omega^2}{d\tau^2} \right] &= 0 \\ \frac{d^2\theta}{d\tau^2} + \frac{2}{r} \frac{dr}{d\tau} \frac{d\theta}{d\tau} - \sin\theta \cos\theta \left(\frac{d\varphi}{d\tau} \right)^2 &= 0 \\ \frac{d^2\varphi}{d\tau^2} + \frac{2}{r} \frac{dr}{d\tau} \frac{d\varphi}{d\tau} + \frac{2}{\tan\theta} \frac{d\theta}{d\tau} \frac{d\varphi}{d\tau} &= 0. \end{aligned}$$

3. Pour rappel, la valeur du coefficient M utilisé dans la métrique de Schwarzschild est obtenue par la théorie newtonienne appliquée au mouvement des planètes suffisamment éloignées du soleil. D'autre part, on peut prendre pour R , donc pour μ , les valeurs observationnelles déduites de la mécanique classique car une correction relativiste supplémentaire à ce niveau induirait des termes d'ordre supérieur en r .

4. P. Couderc, *La relativité*, Que Sais-je ? 37, Presses Universitaires de France, 1977, p. 110.

La première et la quatrième équation peuvent respectivement s'écrire sous la forme

$$\frac{d}{d\tau} \left[\left(1 - \frac{2GM}{rc^2} \right) \frac{dt}{d\tau} \right] = 0$$

$$\frac{d}{d\tau} \left(r^2 \sin^2 \theta \frac{d\varphi}{d\tau} \right) = 0.$$

Deux intégrales premières s'en déduisent immédiatement :

$$\left(1 - \frac{2GM}{rc^2} \right) \frac{dt}{d\tau} = l \quad (9.26)$$

$$r^2 \sin^2 \theta \frac{d\varphi}{d\tau} = h, \quad (9.27)$$

où l et h sont des constantes. On arrive directement à cette conclusion en inspectant le lagrangien (3.20) pour une particule libre

$$L = -m_0 c \sqrt{c^2 \left(1 - \frac{2GM}{rc^2} \right) \left(\frac{dt}{d\lambda} \right)^2 - \frac{1}{1 - \frac{2GM}{rc^2}} \left(\frac{dr}{d\lambda} \right)^2 - r^2 \frac{d\Omega^2}{d\lambda^2}},$$

qui ne dépend explicitement ni de t , ni de φ . Les équations de Lagrange imposent alors

$$\frac{\partial L}{\partial (dt/d\lambda)} = \text{constante}, \quad \frac{\partial L}{\partial (d\varphi/d\lambda)} = \text{constante},$$

respectivement équivalentes à (9.26) et (9.27). Le problème étant à symétrie sphérique, on peut sans perte de généralité supposer que le mouvement se fait dans le plan $\theta = \pi/2$. C'est évidemment compatible avec l'équation d'évolution pour θ .

Par analogie avec la mécanique classique, nous interpréterons h comme *moment cinétique par unité de masse* de la particule. Pour connaître la signification physique de l , rappelons d'abord que l'énergie E d'une particule de quadrivecteur énergie-impulsion \vec{p} , mesurée dans le référentiel d'un observateur de quadrivitesse \vec{U}_{obs} , est donnée par

$$E = -\vec{p} \cdot \vec{U}_{obs} \quad (9.28)$$

[(équation (2.16)]. Or, d'une part, pour une particule matérielle on a

$$p^0 = m_0 \frac{cdt}{d\tau} = \frac{m_0 c l}{1 - \frac{2GM}{rc^2}}.$$

D'autre part, pour un observateur immobile de coordonnée radiale r dans la métrique extérieure de Schwarzschild, on a

$$U_{obs}^0 = \frac{c}{\sqrt{1 - \frac{2GM}{rc^2}}}, \quad U_{obs}^i = 0. \quad (9.29)$$

Ainsi,

$$E(r) = m_0 \left(1 - \frac{2GM}{rc^2} \right) \frac{c^2 l}{\left(1 - \frac{2GM}{rc^2} \right) \sqrt{1 - \frac{2GM}{rc^2}}} = \frac{m_0 c^2 l}{\sqrt{1 - \frac{2GM}{rc^2}}},$$

d'où

$$l = \frac{E(r)}{m_0 c^2} \sqrt{1 - \frac{2GM}{rc^2}}. \quad (9.30)$$

Si la particule est lâchée depuis la coordonnée r_0 avec une vitesse initiale nulle, alors $E(r_0) = m_0 c^2$ et

$$l = \sqrt{1 - \frac{2GM}{r_0 c^2}} < 1. \quad (9.31)$$

Dans ce cas, l'énergie mesurée par un observateur immobile en r vaut

$$E(r) = m_0 c^2 \frac{\sqrt{1 - \frac{2GM}{r_0 c^2}}}{\sqrt{1 - \frac{2GM}{rc^2}}}. \quad (9.32)$$

Si la particule vient de l'infini et qu'un observateur immobile y mesure sa vitesse égale à V_∞ , on a

$$E(\infty) = \frac{m_0 c^2}{\sqrt{1 - V_\infty^2/c^2}}$$

donc

$$l = \frac{1}{\sqrt{1 - V_\infty^2/c^2}} > 1. \quad (9.33)$$

9.6 Equation des trajectoires

En utilisant l'expression de l'élément de longueur et les intégrales premières, on peut trouver une équation pour $dr/d\tau$. Comme $ds^2 = -c^2 d\tau^2$ pour une particule matérielle et $ds^2 = 0$ pour un photon, on obtient

$$k = \frac{l^2}{1 - \frac{2GM}{rc^2}} - \frac{1}{c^2 \left(1 - \frac{2GM}{rc^2}\right)} \left(\frac{dr}{d\tau}\right)^2 - \frac{h^2}{r^2 c^2}$$

(où k est égal à 1 pour une particule matérielle et 0 pour un photon), c'est-à-dire

$$\left(\frac{dr}{d\tau}\right)^2 = c^2 l^2 - c^2 \left(1 - \frac{2GM}{rc^2}\right) \left(k + \frac{h^2}{r^2 c^2}\right). \quad (9.34)$$

Or,

$$\frac{dr}{d\tau} = \frac{dr}{d\varphi} \frac{d\varphi}{d\tau} = \frac{h}{r^2} \frac{dr}{d\varphi},$$

compte tenu de (9.27). Donc

$$\left(\frac{dr}{d\varphi}\right)^2 = \frac{r^4 c^2}{h^2} \left[l^2 - \left(1 - \frac{2GM}{rc^2}\right) \left(k + \frac{h^2}{r^2 c^2}\right) \right]. \quad (9.35)$$

Tout comme en mécanique classique, il s'avère utile de poser $u = 1/r$:

$$\left(\frac{du}{d\varphi}\right)^2 = \frac{c^2(l^2 - k)}{h^2} + \frac{2kGM}{h^2} u - u^2 + \frac{2GM}{c^2} u^3. \quad (9.36)$$

Cette équation est résoluble directement en termes d'intégrales elliptiques mais il est plus simple de considérer sa dérivée par rapport à φ qui, outre les orbites circulaires $du/d\varphi = 0$, fournit l'équation des trajectoires :

$$\frac{d^2u}{d\varphi^2} + u = \frac{kGM}{h^2} + \frac{3GM}{c^2} u^2. \quad (9.37)$$

Nous en établirons des solutions approchées dans les sections 9.9.2 ($k = 1$) et 9.9.3 ($k = 0$). En faisant tendre formellement la vitesse de la lumière vers l'infini, on retrouve l'équation linéaire de J. Binet :

$$\frac{d^2u}{d\varphi^2} + u = \frac{GM}{h^2},$$

valable, selon Newton, aussi bien pour une particule matérielle que pour un corpuscule de lumière.

9.7 Le potentiel effectif

On peut écrire l'équation (9.34) sous la forme

$$\frac{dr}{d\tau} = \pm \sqrt{c^2(l^2 - k) - V_{eff}(k, r, h)} \quad (9.38)$$

en définissant le *potentiel effectif* par

$$V_{eff}(k, r, h) = \frac{-2GMk}{r} + \frac{h^2}{r^2} - \frac{2GMh^2}{r^3c^2}. \quad (9.39)$$

Cette équation rappelle son homologue newtonien : une particule de masse m et d'énergie E vérifie

$$\frac{dr}{d\tau} = \sqrt{\frac{2E}{m} - V_{eff,N}(r, h)}$$

où

$$V_{eff,N}(r, h) = \frac{-2GM}{r} + \frac{h^2}{r^2}$$

est le potentiel effectif intervenant dans la théorie newtonienne de la gravitation.⁵ Remarquons que

$$\lim_{c \rightarrow +\infty} V_{eff}(1, r, h) = V_{eff,N}(r, h).$$

La courbe $V_{eff}(0, r, h)$ (donc relative aux photons) présente un maximum en

$$V_{max} = V_{eff}\left(0, \frac{3GM}{c^2}, h\right) = \frac{h^2c^4}{27G^2M^2} \quad (9.40)$$

(voir figure 9.2). Pour $c^2l^2 = V_{max}$ la trajectoire est une orbite circulaire instable située en $r = 3GM/c^2$. Pour $c^2l^2 < V_{max}$, la trajectoire est libre. Si c^2l^2 est légèrement inférieur à V_{max} , le photon s'enroule un grand nombre de fois autour du cercle de rayon $r = 3GM/c^2$ avant de repartir vers l'infini (figure 9.3). Si c^2l^2 est légèrement supérieur à V_{max} , le photon s'enroule également un grand nombre de fois autour du cercle de rayon $r = 3GM/c^2$ mais finit par le franchir (figure 9.4).

5. Voir par exemple L. Landau, E. Lifchitz, *Physique théorique, Tome 1, Mécanique*, Mir, 1982, §§ 14-15.

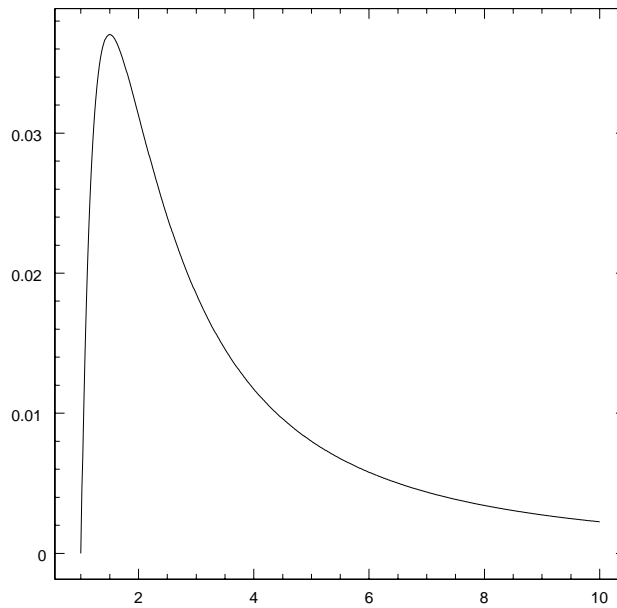


FIGURE 9.2: Courbe de potentiel effectif $V_{eff}(0, r, h)/h^2$ pour un photon dans la métrique extérieure de Schwarzschild, en unités $GM = 1$, $c = 1$; axe des abscisses gradué en unités r_s .

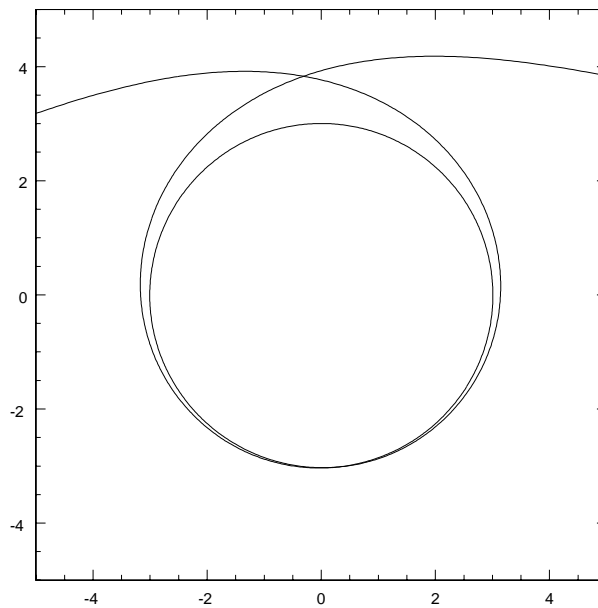


FIGURE 9.3: Intégration numérique de l'équation (9.37) avec $k = 0$, lorsque $c^2 l^2$ est légèrement inférieur à V_{max} , c'est-à-dire $c^2 l^2 / h^2$ un peu plus petit que $c^4 / 27 G^2 M^2$. On vérifie à l'aide de (9.36) que cette condition est bien réalisée pour les conditions initiales $[u(\varphi)]_0 = c^2 / 20 GM$, $[du/d\varphi]_0 = 0.1865126 c^2 / GM$. En abscisse on a porté $x = r \cos \varphi$ et en ordonnée $y = r \sin \varphi$. Unités $GM = 1$, $c = 1$.

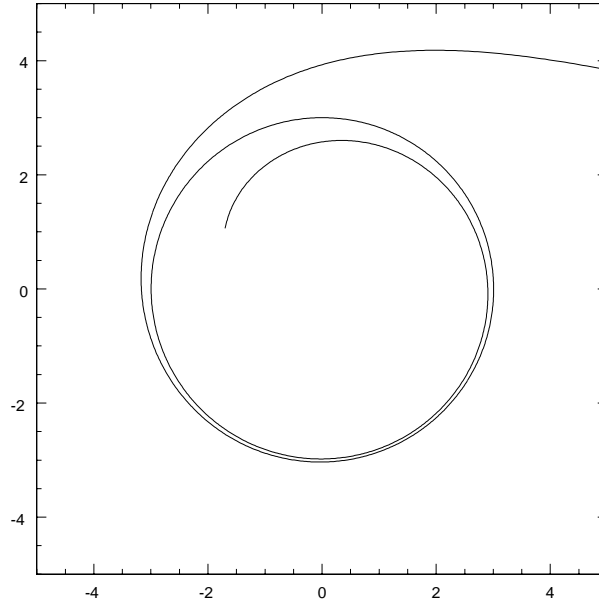


FIGURE 9.4: *Intégration numérique de l'équation (9.37) avec $k = 0$, lorsque $c^2 l^2$ est légèrement supérieur à V_{max} . Conditions initiales : $[u(\varphi)]_0 = c^2/20GM$, $[du/d\varphi]_0 = 0.186513 c^2/GM$. On a stoppé l'intégration numérique au moment où le photon franchit le rayon de Schwarzschild.*

Etudions maintenant la fonction $V_{eff}(1, r, h)$. D'une part,

$$\frac{dV_{eff}(1, r, h)}{dr} = \frac{2GM}{r^2} - \frac{2h^2}{r^3} + \frac{6GMh^2}{r^4 c^2}.$$

- Pour $h < 2\sqrt{3} GM/c$, la dérivée ne s'annule en aucun r réel.
- Pour $h = 2\sqrt{3} GM/c$, la dérivée s'annule en $r = 6GM/c^2$.
- Pour $h > 2\sqrt{3} GM/c$, $V_{eff}(1, r, h)$ admet deux extrema distincts : un maximum en

$$r_- = \frac{h^2}{2GM} \left(1 - \sqrt{1 - \frac{12G^2 M^2}{h^2 c^2}} \right) \quad (9.41)$$

et un minimum en

$$r_+ = \frac{h^2}{2GM} \left(1 + \sqrt{1 - \frac{12G^2 M^2}{h^2 c^2}} \right). \quad (9.42)$$

On a

$$\frac{3GM}{c^2} < r_- < \frac{6GM}{c^2} < r_+. \quad (9.43)$$

D'autre part,

$$\frac{d^2 V_{eff}(1, r, h)}{dr^2} = -\frac{4GM}{r^3} + \frac{6h^2}{r^4} - \frac{24GMh^2}{r^5 c^2}.$$

- Pour $h < \sqrt{32/3} GM/c$, cette expression ne peut s'annuler.
- Pour $h = \sqrt{32/3} GM/c$, la dérivée seconde et la dérivée troisième de $V_{eff}(1, r, h)$ s'annulent en $r = 8GM/c^2$.

– Pour $h > \sqrt{32/3} GM/c$, la dérivée seconde s’annule en

$$r = \frac{3h^2}{4GM} \left(1 \pm \sqrt{1 - \frac{32G^2M^2}{3c^2h^2}} \right),$$

qui constituent des points d’inflexion.

Remarquons que

$$\frac{2\sqrt{3}GM}{c} > \sqrt{\frac{32}{3}} \frac{GM}{c}.$$

La présence d’un maximum et d’un minimum implique donc celle des deux points d’inflexion. Pour $h = 2\sqrt{3} GM/c$ la fonction admet un point d’inflexion à tangente horizontale en $r = 6GM/c^2$ et un autre point d’inflexion en $r = 12GM/c^2$. Quelques-unes de ces courbes sont présentées en figure 9.5.

Cette analyse révèle plusieurs propriétés inconnues de la physique classique concernant le comportement des particules matérielles dans un champ de gravitation à symétrie sphérique spatiale :

– Pas de périastre si $h \leq 2\sqrt{3} GM/c$ (en théorie newtonienne, une telle situation ne se rencontre que dans le cas $h = 0$, correspondant à la chute radiale) et, pour $|l|$ suffisamment grand, même si $h > 2\sqrt{3} GM/c$.

– La valeur limite minimale du périastre d’une trajectoire bornée (c’est-à-dire, pour laquelle il existe un apoastre) est obtenue pour $h \rightarrow 4GM/c$ et vaut $r_- = 4GM/c^2$: on a

$$V_{eff}(1, r = \frac{4GM}{c^2}, h = \frac{4GM}{c}) = 0 = \lim_{r \rightarrow +\infty} V_{eff}(1, r, h) \quad \forall h.$$

– Pas d’orbite circulaire si $r \leq 3GM/c^2$ et pas d’orbite circulaire stable si $r \leq 6GM/c^2$. En effet, les orbites circulaires instables et stables correspondent respectivement aux valeurs de r_- et r_+ , dont le domaine est donné par la relation (9.43).

– Pour $h \gg 2\sqrt{3} GM/c$ et $|l|$ pas trop grand on retrouve les orbites classiques, à la précession du périastre près (voir section 9.9.2).

Exercice 9.1 *Montrer que les orbites circulaires vérifient la troisième loi de Kepler :*

$$\frac{d\varphi}{dt} = \sqrt{\frac{GM}{r^3}}.$$

Solution. Utiliser les relations (9.26), (9.27), (9.34) et (9.42) pour $\theta = \pi/2$ et r constant.

9.8 La capture gravitationnelle

9.8.1 Section efficace de capture

Soit une particule libre s’approchant d’une masse déflectrice ponctuelle⁶ M située à l’origine, et en provenance de l’infini où son angle polaire $\varphi \rightarrow 0^+$ (ce qui n’est pas une restriction : il suffit pour cela d’adapter l’origine des φ). Son *paramètre d’impact* p est défini par l’ordonnée de la droite asymptotiquement tangente à la trajectoire :

$$p = \lim_{r \rightarrow +\infty} [r \sin \varphi(r)] \geq 0. \tag{9.44}$$

6. Il s’agit là d’une hypothèse de travail permettant de traiter en une fois tous les cas de figure. A strictement parler, elle revient à supposer que la masse engendre un trou noir (voir section 9.11).

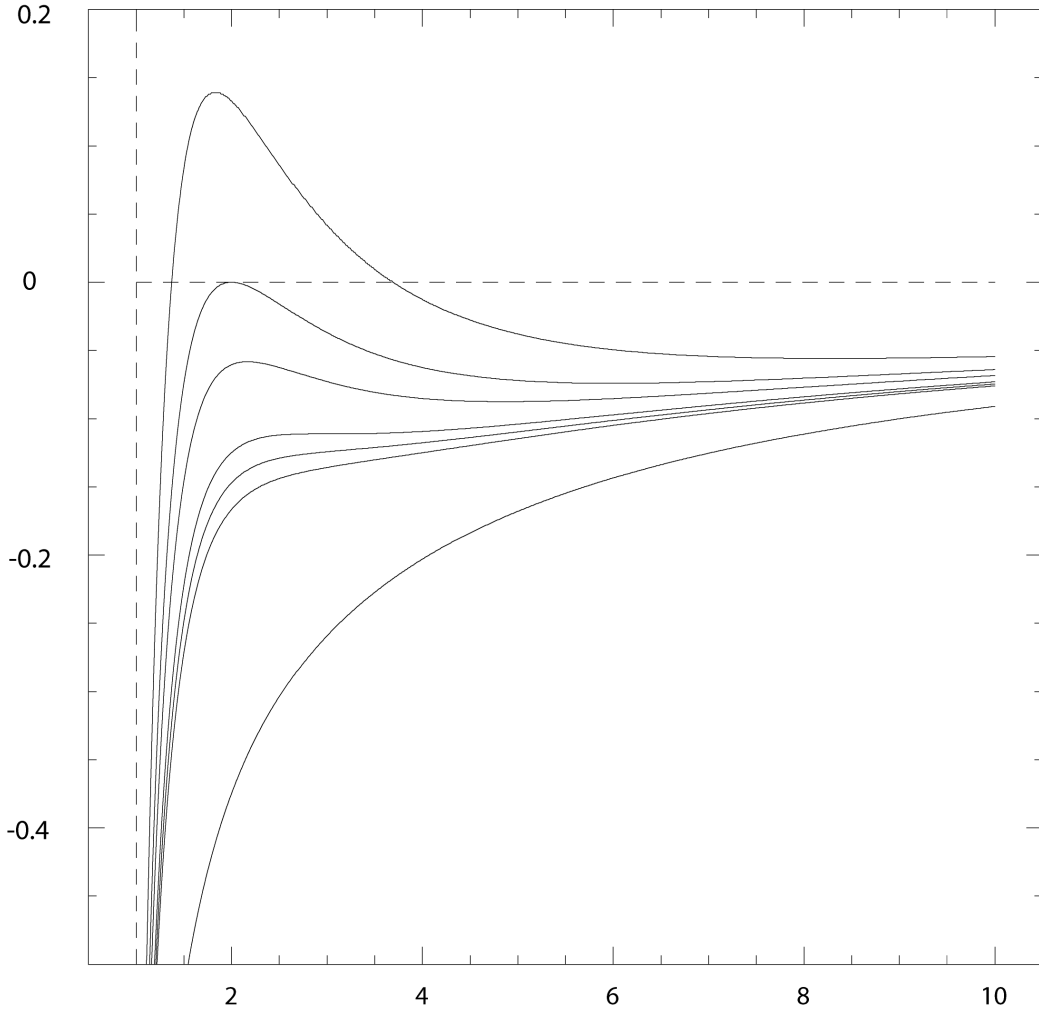


FIGURE 9.5: Courbes de potentiel effectif pour une particule matérielle dans la métrique extérieure de Schwarzschild. En abscisse on a porté r , en unités r_s . En ordonnée on a $V_{eff}(1, r, h)$, en unités c^2 . De bas en haut on a représenté les courbes correspondant respectivement, en unités GM/c , à $h_1 = 2$, $h_2 = \sqrt{32/3}$, $h_3 = 3.36$, $h_4 = 2\sqrt{3}$, $h_5 = 3.75$, $h_6 = 4$, $h_7 = 4.5$.

En mécanique newtonienne, la particule ne peut s'effondrer sur M que si sa chute est radiale, c'est-à-dire si $p = 0$. Ce résultat n'est plus correct en relativité générale : il existe un paramètre d'impact critique p_c en-dessous duquel la collision se produit, donc une section efficace de capture gravitationnelle non nulle

$$\sigma = \pi p_c^2. \quad (9.45)$$

Nous allons la calculer pour une particule matérielle et pour un photon.

On a

$$\lim_{r \rightarrow +\infty} \frac{dr}{d\tau} = \lim_{r \rightarrow +\infty} \left(\frac{dr}{d\varphi} \frac{d\varphi}{d\tau} \right) = \lim_{r \rightarrow +\infty} \left(\frac{-p \cos \varphi}{\sin^2 \varphi} \frac{h}{r^2} \right) = \frac{-h}{p}.$$

Dès lors, de la relation (9.34) :

$$\left(\frac{dr}{d\tau}\right)^2 = c^2 l^2 - c^2 \left(1 - \frac{2GM}{rc^2}\right) \left(k + \frac{h^2}{r^2 c^2}\right), \quad (9.46)$$

on tire

$$\frac{h^2}{p^2} = c^2(l^2 - k),$$

ce qui permet de relier le paramètre d'impact p aux constantes h et l :

$$p^2 = \frac{h^2}{c^2(l^2 - k)}. \quad (9.47)$$

Dans le cas des photons, k est nul et le second terme du membre de droite de (9.46), c'est-à-dire le potentiel effectif (9.39), présente un maximum d'ordonnée

$$V_{max} = \frac{h^2 c^4}{27G^2 M^2}$$

[équation (9.40)]. Il y a collision si

$$l^2 > l_c^2(h) = \frac{h^2 c^2}{27G^2 M^2}. \quad (9.48)$$

Compte tenu de (9.47), le paramètre d'impact critique est donc donné par

$$p_c = \frac{3\sqrt{3}GM}{c^2}$$

et la section efficace de capture vaut

$$\sigma = \frac{27\pi G^2 M^2}{c^4}. \quad (9.49)$$

Pour les particules matérielles, k est égal à 1 et $l^2 > 1$, en vertu de (9.33). De l'étude de $V_{eff}(1, r, h)$ (section 9.7), il résulte alors que

– Si $h < 4GM/c$ alors $\forall r$ on a $V_{eff}(1, r, h) < 0 < c^2(l^2 - 1)$: la particule sera capturée quel que soit l .

– Si $h \geq 4GM/c$ il y aura capture si $V_{eff}(1, r_-, h) < c^2(l^2 - 1)$, où

$$r_- = \frac{h^2}{2GM} \left(1 - \sqrt{1 - \frac{12G^2 M^2}{h^2 c^2}}\right)$$

est la valeur de r maximisant $V_{eff}(1, r, h)$ [(cf. (9.41)]. On calcule

$$V_{eff}(1, r_-, h) = c^2 \left(\frac{2}{3} - \frac{1}{3} \sqrt{1 - \frac{12G^2 M^2}{h^2 c^2}}\right) \left[\frac{2}{3} + \frac{h^2 c^2}{18G^2 M^2} \left(1 + \sqrt{1 - \frac{12G^2 M^2}{h^2 c^2}}\right)\right] - c^2.$$

La condition s'écrit ainsi

$$l^2 > l_c^2(h) = \frac{2}{3} + \frac{h^2 c^2}{54G^2 M^2} + \left(\frac{h^2 c^2}{54G^2 M^2} - \frac{2}{9}\right) \sqrt{1 - \frac{12G^2 M^2}{h^2 c^2}}. \quad (9.50)$$

En inversant cette équation on peut calculer le moment cinétique critique h_c , et par conséquent la section efficace de capture, en fonction de l , donc de la vitesse de la particule à l'infini. Le calcul est assez simple dans les deux cas extrêmes que nous allons maintenant envisager : lorsque la particule est asymptotiquement non relativiste ou, au contraire, ultra-relativiste.

D'après (9.33), pour une particule non relativiste à l'infini on a $l^2 \simeq 1 + V_\infty^2/c^2 \simeq 1$. Il y a donc capture pour des valeurs de h telles que $V_{eff}(1, r_-, h) \simeq 0$, c'est-à-dire

$$h = \frac{4GM}{c} (1 + \epsilon)$$

avec $0 < \epsilon \ll 1$. Injectant cette expression dans (9.50), on trouve par un calcul au premier ordre

$$l_c^2(h) \simeq 1 + \epsilon.$$

Donc $\epsilon = V_\infty^2/c^2$ et

$$h_c(l) \simeq \frac{4GM}{c} l^2.$$

En utilisant (9.45) et (9.47), il vient finalement

$$\sigma = \frac{\pi h_c^2(l)}{c^2(l^2 - 1)} \simeq \frac{16\pi G^2 M^2}{c^2 V_\infty^2}. \quad (9.51)$$

Pour une particule ultra-relativiste on a $l^2 \gg 1$, toujours en vertu de (9.33). Or, selon (9.50), l_c^2 est une fonction croissante de $h^2 c^2 / G^2 M^2$. Par conséquent, on a également $h^2 c^2 / G^2 M^2 \gg 1$. En se limitant au premier ordre du développement de la racine carrée dans (9.50) en série de puissances de $G^2 M^2 / h^2 c^2 \ll 1$, on trouve

$$l_c^2(h) \simeq \frac{h^2 c^2}{27G^2 M^2} + \frac{1}{3}.$$

On en déduit

$$h_c^2(l) \simeq \frac{27G^2 M^2}{c^2} \left[l^2 - \frac{1}{3} + O\left(\frac{1}{l^2}\right) \right],$$

$$\sigma = \frac{\pi h_c^2(l)}{c^2(l^2 - 1)} \simeq \frac{27\pi G^2 M^2}{c^4} \left[1 + \frac{2}{3l^2} + O\left(\frac{1}{l^4}\right) \right] \simeq \frac{9\pi G^2 M^2}{c^4} \left(5 - 2 \frac{V_\infty^2}{c^2} \right). \quad (9.52)$$

On retrouve la section efficace de capture des photons (9.49) en prenant dans cette expression la limite pour $V_\infty \rightarrow c$.

9.8.2 Angle de capture

L'angle de capture est l'angle d'ouverture, en un point donné, du cône des trajectoires qui se concluent par une collision. Supposons que depuis un point A de coordonnées (r, φ) un photon soit lâché dans une direction faisant un angle α avec la direction radiale. Attention, quoique la géométrie euclidienne soit localement valable, l'élément de longueur n'est pas celui des coordonnées polaires, mais bien

$$ds^2 = \frac{dr^2}{1 - \frac{2GM}{rc^2}} + r^2 d\varphi^2.$$

Il faut en tenir compte pour le calcul des distances spatiales et des angles.

De la figure 9.6, représentant une coupe de l'espace-temps de Schwarzschild à t et θ constants, il résulte

$$\frac{dr}{d\varphi} = \frac{r \sqrt{1 - 2GM/rc^2}}{\tan \alpha}.$$

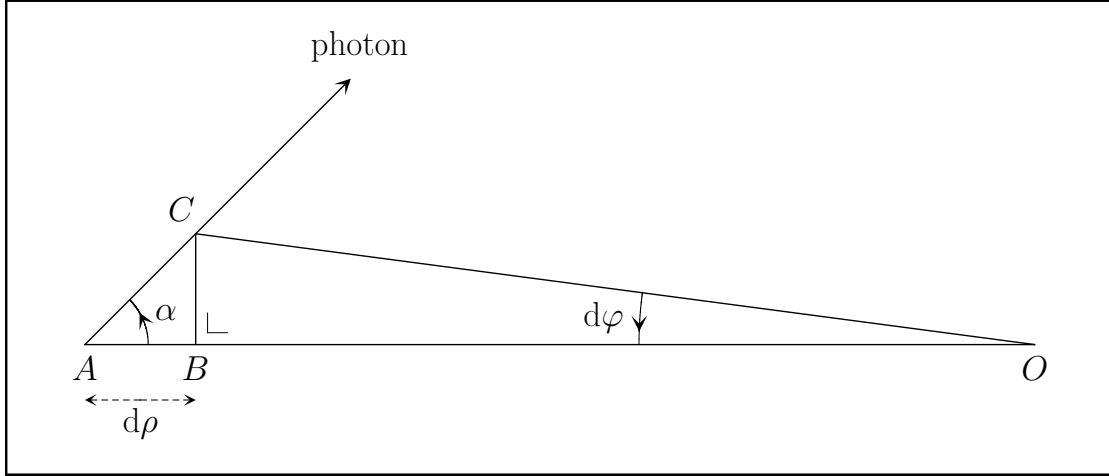


FIGURE 9.6: La masse déflectrice M est à l'origine des coordonnées polaires, en O . La distance physique spatiale entre $A(r, \varphi)$ et $B(r - dr, \varphi)$ vaut $d\rho = dr/\sqrt{1 - 2GM/rc^2}$. Au premier ordre, la distance physique spatiale entre $B(r - dr, \varphi)$ et $C(r - dr, \varphi - d\varphi)$ vaut $rd\varphi = d\rho \tan \alpha$.

En introduisant cette relation dans l'équation (9.35), on obtient

$$\sin^2 \alpha = \frac{h^2 \left(1 - \frac{2GM}{rc^2}\right)}{r^2 c^2 \left[l^2 - k \left(1 - \frac{2GM}{rc^2}\right)\right]}.$$

Il y a capture si $\alpha < \alpha_c(k, r, h)$ où

$$\sin^2 \alpha_c(k, r, h) = \frac{h^2 \left(1 - \frac{2GM}{rc^2}\right)}{r^2 c^2 \left[l_c^2(h) - k \left(1 - \frac{2GM}{rc^2}\right)\right]}. \quad (9.53)$$

La quantité $l_c^2(h)$ est donnée par (9.48) si $k = 0$ et par (9.50) si $k = 1$. Pour les photons, l'angle critique α_c ne dépend, en fait, que de la coordonnée radiale initiale r :

$$\sin^2 \alpha_c(0, r, h) = \frac{27G^2 M^2}{r^2 c^4} \left(1 - \frac{2GM}{rc^2}\right). \quad (9.54)$$

On obtient naturellement $\alpha_c \rightarrow 0$ pour $r \rightarrow +\infty$; si $r \in]3GM/c^2, +\infty[$, alors $\alpha_c \in]0, \pi/2[$; en $r = 3GM/c^2$, $\alpha = \pi/2$; et lorsque $r \in]2GM/c^2, 3GM/c^2[$, on a $\alpha_c \in]\pi/2, \pi[$. Le statut exact du cas $r = 2GM/c^2$ ne peut être appréhendé correctement dans les coordonnées de Schwarzschild (voir section 9.11).

9.9 Les tests de la relativité générale

9.9.1 L'effet Einstein

Soit une distribution sphérique de matière de rayon $R > 2GM/c^2$, et un observateur immobile en la coordonnée radiale $r > R$. Son temps propre T est donné par

$$dT = dt \sqrt{1 - \frac{2GM}{rc^2}}. \quad (9.55)$$

Si $r = R + \xi$ avec $R \gg \xi$ et si $R \gg 2GM/c^2$, on peut écrire

$$dT \simeq \left(1 + \frac{GM\xi}{R^2c^2}\right) \left(1 - \frac{GM}{Rc^2}\right) dt.$$

En posant $g = GM/R^2$ et $\tau = t(1 - GM/Rc^2)$, on trouve la relation

$$dT = \left(1 + \frac{g\xi}{c^2}\right) d\tau,$$

obtenue dans l'approche heuristique du chapitre 2 consacré au champ gravitationnel constant. Du reste, le raisonnement de la section 2.4.3 est entièrement applicable à la relation (9.55). Si un photon se déplace dans la métrique extérieure de Schwarzschild (radialement ou non), un observateur immobile de coordonnée r en mesurera la fréquence égale à $\nu(r)$, avec

$$\nu(r) \sqrt{1 - \frac{2GM}{rc^2}} = \text{constante} = \nu(\infty). \quad (9.56)$$

Cette équation constitue la version rigoureuse de (2.18).

Exercice 9.2 *Soit, en r_0 , une source émettant des photons radialement vers le bas. Un observateur se laisse tomber sans vitesse initiale depuis ce point. Trouver l'expression de la fréquence mesurée tout au long de sa chute.*

Solution. Raisonnons dans le plan $\theta = \pi/2$ et appelons ν_0 la fréquence initiale. Notons ici $\vec{U}_{stat}(r)$ la quadrivitesse de l'observateur statique donnée dans la relation (9.29). Par (9.26) et (9.34), le quadrivecteur $\vec{p}_{photon}(r)$ doit être du type

$$K \left(\frac{1}{1 - \frac{2GM}{rc^2}}, -1, 0, 0 \right),$$

avec K constant, et par (9.28), on doit avoir

$$-\vec{p}_{photon}(r_0) \cdot \vec{U}_{stat}(r_0) = 2\pi\hbar\nu_0.$$

On en déduit

$$\left(p_{photon}^\alpha\right)(r) = \frac{2\pi\hbar\nu_0}{c} \sqrt{1 - \frac{2GM}{r_0c^2}} \left(\frac{1}{1 - \frac{2GM}{rc^2}}, -1, 0, 0 \right). \quad (9.57)$$

D'autre part, les relations (9.26), (9.31) et (9.34) imposent, pour l'observateur en chute libre,

$$(U_{obs}^\alpha)(r) = c \left(\frac{\sqrt{1 - \frac{2GM}{r_0 c^2}}}{1 - \frac{2GM}{rc^2}}, -\sqrt{\frac{2GM}{rc^2} - \frac{2GM}{r_0 c^2}}, 0, 0 \right).$$

On a

$$\vec{U}_{obs} \cdot \vec{U}_{stat} = -c^2 \frac{\sqrt{1 - \frac{2GM}{r_0 c^2}}}{\sqrt{1 - \frac{2GM}{rc^2}}}.$$

Ce produit scalaire est indépendant du repère choisi. Or, la théorie de la relativité restreinte enseigne que, dans la base orthonormée statique, il peut s'écrire sous la forme

$$\frac{-c^2}{\sqrt{1 - V^2/c^2}}$$

où V est la vitesse spatiale radiale de l'observateur, mesurée dans cette base. Ainsi,

$$(U_{obs}^\alpha)(r) = \frac{1}{\sqrt{1 - V^2/c^2}} \left(\frac{c}{\sqrt{1 - \frac{2GM}{rc^2}}}, -V \sqrt{1 - \frac{2GM}{rc^2}}, 0, 0 \right) \quad (V > 0). \quad (9.58)$$

L'énergie du photon mesurée par l'observateur, donc la fréquence demandée, est calculable par (9.28). Combinant (9.57) et (9.58), on obtient

$$\nu = \nu_0 \sqrt{\frac{1 - \frac{2GM}{r_0 c^2}}{1 - \frac{2GM}{rc^2}}} \sqrt{\frac{1 - V/c}{1 + V/c}}.$$

Les deux premiers termes du membre de droite décrivent l'effet Einstein. Le troisième est dû à l'effet Doppler longitudinal (page 19 avec $\varphi = 0$).

9.9.2 Précession du périastre des planètes

Etudions la trajectoire d'une planète du système solaire assimilée à une particule-test ponctuelle décrivant une géodésique de l'espace-temps (on négligera donc notamment le rayonnement émis sous forme d'ondes gravitationnelles).

Il s'agit de résoudre l'équation (9.37) avec $k = 1$:

$$\frac{d^2 u}{d\varphi^2} + u = \frac{1}{p} + \frac{3GM}{c^2} u^2, \quad (9.59)$$

en posant $p = h^2/GM$. Remarquons d'emblée que la variable φ n'y apparaît pas explicitement : la solution générale de (9.59) est donc une fonction de $\varphi - \varphi_0$, où φ_0 est une constante arbitraire d'intégration que l'on peut toujours annuler en redéfinissant les axes auxquels sont rapportées les coordonnées polaires (r, φ) .

En théorie newtonienne, le second terme du membre de droite est absent. La solution générale prend alors la forme

$$u(\varphi) = \frac{1 + e \cos \varphi}{p}$$

où e est l'unique constante arbitraire d'intégration subsistante. Cette solution représente une conique de paramètre p admettant l'origine pour foyer.⁷ Nous pouvons sans restriction supposer $e \geq 0$, ce qui revient à choisir les axes pour que le périhélie de la trajectoire se situe en $\varphi = 0$. Si $e < 1$ il s'agit d'une ellipse (voir figure 9.7), si $e = 1$ c'est une parabole et si $e > 1$ on a une branche d'hyperbole.

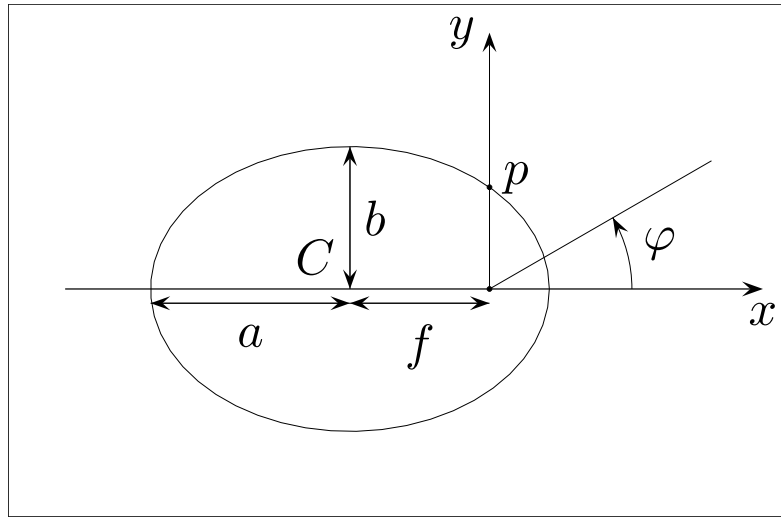


FIGURE 9.7: *Ellipse. Centre C; demi-grand axe a; demi-petit axe b; distance focale $f = \sqrt{a^2 - b^2}$; excentricité $e = f/a$; paramètre $p = a(1 - e^2)$.*

En pratique, dans le système solaire, le second terme de l'équation (9.59) est négligeable par rapport au premier. En effet,

$$\frac{3GMu^2/c^2}{1/p} = \frac{3GMu^2/c^2}{GM/h^2} = \frac{3h^2u^2}{c^2} = 3 \left(\frac{V_\varphi}{c} \right)^2 \ll 1,$$

où $V_\varphi = r d\varphi/d\tau$ est la vitesse orthoradiale de la planète. On peut donc résoudre l'équation (9.59) par une méthode de perturbations. Comme seules les trajectoires planétaires nous intéressent ici, nous considérerons dans la suite que $e < 1$. Introduisons le paramètre sans dimension

$$\epsilon = \frac{3GM/c^2}{p}.$$

On a $\epsilon \ll 1$ car, pour les planètes, $p \simeq r \gg 3GM/c^2$.

Cherchons la solution sous la forme

$$u(\varphi) = u_0(\varphi) + \epsilon u_1(\varphi) + \epsilon^2 u_2(\varphi) + \dots$$

7. Le paramètre d'une conique est la longueur du segment de droite abaissé perpendiculairement à l'axe (le grand axe dans le cas d'une ellipse) depuis un foyer jusqu'à la courbe.

Nous nous limiterons aux termes $u_0(\varphi)$ et $u_1(\varphi)$. On obtient

$$\frac{d^2 u_0}{d\varphi^2} + u_0 = \frac{1}{p} \tag{9.60}$$

$$\frac{d^2 u_1}{d\varphi^2} + u_1 = p(u_0)^2. \tag{9.61}$$

La solution générale de (9.60) est la trajectoire newtonienne

$$u_0(\varphi) = \frac{1 + e \cos \varphi}{p}.$$

La solution générale de l'équation non homogène (9.61) est la solution générale de l'équation homogène correspondante [on peut considérer qu'elle est déjà incluse dans $u_0(\varphi)$] plus une solution particulière. Recherchons donc une solution de

$$\frac{d^2 u_1}{d\varphi^2} + u_1 = \frac{(1 + e \cos \varphi)^2}{p}$$

par la méthode de variation des constantes.⁸ On a

$$u_1(\varphi) = \cos \varphi \int C_1(\varphi) d\varphi + \sin \varphi \int C_2(\varphi) d\varphi$$

où $C_1(\varphi)$ et $C_2(\varphi)$ sont solutions du système

$$C_1(\varphi) \cos \varphi + C_2(\varphi) \sin \varphi = 0$$

8. Rappelons ici cette méthode. Soit

$$u(x) = \sum_i C_i u_i(x)$$

la solution générale de l'équation différentielle linéaire à coefficients constants

$$L(D) u(x) = 0$$

avec

$$L(D) = A_0 + A_1 D + \dots + A_p D^p,$$

les A_i étant des constantes. Alors, une solution particulière $u(x)$ de l'équation

$$L(D) u(x) = f(x)$$

est donnée par

$$u(x) = \sum_{i=1}^p \left[\int C_i(x) dx \right] u_i(x)$$

où les fonctions $C_i(x)$ constituent la solution (unique) du système

$$\sum_i C_i(x) u_i(x) = 0$$

$$\sum_i C_i(x) D u_i(x) = 0$$

...

$$\sum_i C_i(x) D^{p-1} u_i(x) = \frac{f(x)}{A_p}.$$

$$-C_1(\varphi) \sin \varphi + C_2(\varphi) \cos \varphi = \frac{(1 + e \cos \varphi)^2}{p}.$$

On calcule

$$u_1(\varphi) = \frac{1}{p} \left(\frac{e \cos \varphi}{2} + 1 + \frac{e^2}{3} + \frac{e^2 \sin^2 \varphi}{3} + e\varphi \sin \varphi \right).$$

Remarquons que le premier terme du membre de droite est solution de l'équation homogène pour $u_1(\varphi)$. On obtient donc une autre solution particulière tout aussi valable en le supprimant. Ainsi,

$$\begin{aligned} u(\varphi) &= \frac{1}{p} \left[1 + e \cos \varphi + \epsilon \left(1 + \frac{e^2}{3} + \frac{e^2 \sin^2 \varphi}{3} + e\varphi \sin \varphi \right) \right] \\ &\simeq \frac{1}{p} \left\{ 1 + \epsilon \left(1 + \frac{e^2}{3} + \frac{e^2 \sin^2 \varphi}{3} \right) + e \cos[\varphi(1 - \epsilon)] \right\} + O(\epsilon^2). \end{aligned}$$

La trajectoire n'est plus une courbe fermée : l'ellipse newtonienne est décalée à chaque révolution d'une quantité $\Delta\varphi$ donnée par

$$u(\varphi + 2\pi + \Delta\varphi) = u(\varphi),$$

c'est-à-dire

$$\frac{\epsilon e^2}{3} \sin^2(\varphi + 2\pi + \Delta\varphi) + e \cos[(\varphi + 2\pi + \Delta\varphi)(1 - \epsilon)] = \frac{\epsilon e^2}{3} \sin^2 \varphi + e \cos[\varphi(1 - \epsilon)].$$

Cette quantité $\Delta\varphi$ est du premier ordre au moins en ϵ . Développant l'équation précédente jusqu'à cet ordre, on obtient

$$\Delta\varphi = 2\pi\epsilon = \frac{6\pi GM}{a(1 - e^2)c^2}. \quad (9.62)$$

On dispose des données suivantes :⁹

astre	a (u. a.)	e	période (années)	$\Delta\varphi$ calculé (sec. d'arc/siècle)	$\Delta\varphi$ observé (sec. d'arc/siècle)
Mercure	0.3870	0.2056	0.241	43.03	43.11 ± 0.45
Vénus	0.7233	0.0068	0.615	8.63	8.4 ± 4.8
Terre	1.0000	0.0167	1.000	3.84	5.0 ± 1.2
Mars	1.5237	0.0933	1.881	1.35	
astéroïde Icare	1.0688	0.8265	1.119	10.3	9.8 ± 0.8

La précession du périastre est bien visible en figure 9.8. Cette correction d'origine relativiste s'ajoute aux perturbations dues aux autres planètes et à l'aplatissement du soleil, pour lesquelles un traitement newtonien suffit. L'anomalie du périhélie de Mercure, étudiée par U. Le Verrier dès 1842 mais demeurée inexpliquée, constitua historiquement le premier test de la relativité générale. Le calcul, communiqué par Einstein¹⁰ le 18 novembre 1915, repose sur une approximation de la solution des équations du champ dans le vide jusqu'à l'ordre 2 inclus : la métrique de Schwarzschild ne fut en effet découverte que quelques semaines plus tard.

9. G. Pascoli, *La gravitation*, Que Sais-je ? 2489, Presses universitaires de France, 1995, p. 42.

10. *Albert Einstein, œuvres choisies 2*, textes choisis et présentés par F. Balibar, O. Darrigol, B. Jech et J. Stachel, Seuil, 1993, pp. 169-177.

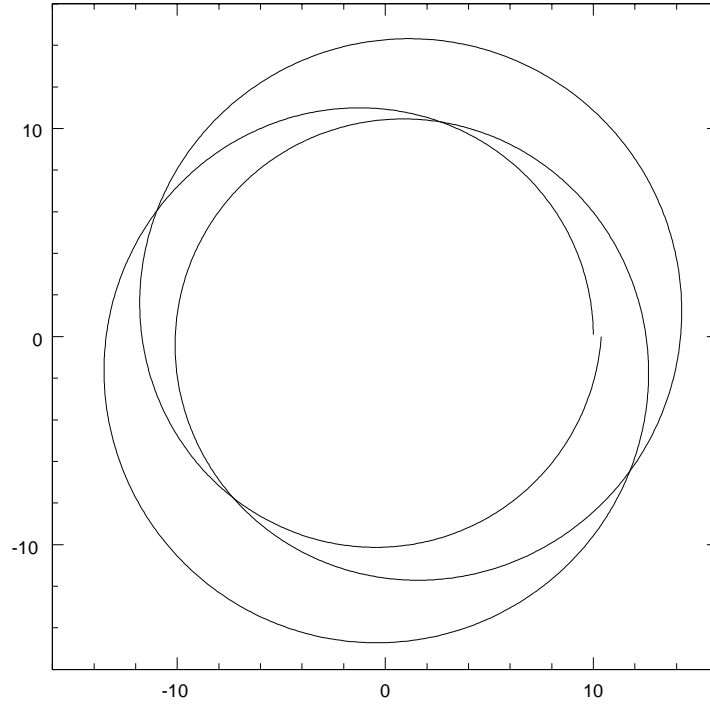


FIGURE 9.8: *Intégration numérique de l'équation (9.59), pour $h = 4GM/c$ et moyennant les conditions initiales $[u(\varphi)]_0 = c^2/10GM$, $[du/d\varphi]_0 = 0$. En abscisse on a porté $x = r \cos \varphi$ et en ordonnée $y = r \sin \varphi$. Unités $GM = 1$, $c = 1$.*

9.9.3 Déviation des rayons lumineux par le soleil

Rappelons l'équation (9.37), particularisée au mouvement de la lumière :

$$\frac{d^2u}{d\varphi^2} + u = \frac{3GM}{c^2} u^2.$$

Appelons p le paramètre d'impact, supposé tel que $p \gg GM/c^2$, et posons

$$\frac{3GM}{pc^2} = \epsilon \ll 1.$$

On obtient ainsi l'équation

$$\frac{d^2u}{d\varphi^2} + u = \epsilon pu^2$$

dont on peut chercher une solution approchée par la méthode des perturbations. Nous ne détaillerons pas les calculs de la déviation $\Delta\varphi$ car ils feraient double emploi avec la section 7.6.8. Au premier ordre en ϵ , on retrouve le résultat (7.62) :

$$\Delta\varphi = \frac{4GM}{r_0 c^2} \quad (9.63)$$

où r_0 est la coordonnée radiale minimale atteinte.

9.9.4 Retard des échos RADAR

La section précédente s'occupait des *trajectoires* des rayons lumineux dans le système solaire. On peut également avancer des prédictions concernant la *durée* d'un trajet. Effectuons le calcul pour un signal RADAR émis depuis la Terre et passant près du soleil avant de se réfléchir sur une planète en conjonction supérieure (figure 9.9).

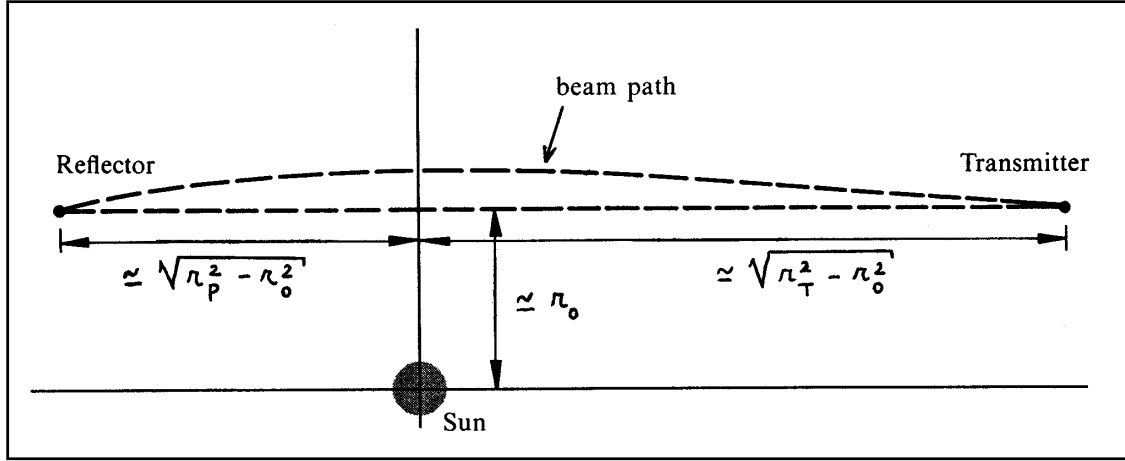


FIGURE 9.9: Trajectoire d'un rayon lumineux émis par la Terre (« Transmitter ») et réfléchi sur une planète en conjonction supérieure (« Reflector »). D'après C.W. Misner, K.S. Thorne, J.A. Wheeler, op. cit., figure 40.3.

Supposons que le mouvement de la lumière s'effectue dans le plan $\theta = \pi/2$. Les deux intégrales premières (9.26) et (9.27) s'écrivent

$$\left(1 - \frac{2GM}{rc^2}\right) \frac{dt}{d\tau} = l \quad \text{et} \quad r^2 \frac{d\varphi}{d\tau} = h,$$

et on a

$$r^2 \left(\frac{d\varphi}{dt}\right)^2 = \frac{h^2}{l^2 r^2} \left(1 - \frac{2GM}{rc^2}\right)^2. \quad (9.64)$$

Si l'on injecte cette égalité dans la relation $0 = ds^2$ écrite sous la forme

$$0 = -c^2 \left(1 - \frac{2GM}{rc^2}\right) + \frac{(dr/dt)^2}{1 - \frac{2GM}{rc^2}} + r^2 (d\varphi/dt)^2,$$

et si l'on appelle r_0 la coordonnée radiale minimale, annulant donc dr/dt , on obtient

$$\frac{h^2}{l^2} = \frac{r_0^2 c^2}{1 - \frac{2GM}{r_0 c^2}},$$

ce qui permet d'écrire (9.64) sous la forme

$$r^2 \left(\frac{d\varphi}{dt}\right)^2 = c^2 \frac{r_0^2}{r^2} \frac{\left(1 - \frac{2GM}{rc^2}\right)^2}{1 - \frac{2GM}{r_0 c^2}}.$$

En injectant cette relation dans $0 = ds^2$ écrite sous la forme

$$c^2 dt^2 = \frac{dr^2}{\left(1 - \frac{2GM}{rc^2}\right)^2} + r^2 \frac{(d\varphi/dt)^2}{1 - \frac{2GM}{rc^2}} dt^2,$$

on trouve une relation entre r et t où n'intervient que le seul paramètre r_0 :

$$c^2 dt^2 \left(1 - \frac{r_0^2}{r^2} \frac{1 - \frac{2GM}{rc^2}}{1 - \frac{2GM}{r_0 c^2}}\right) = \frac{dr^2}{\left(1 - \frac{2GM}{rc^2}\right)^2}.$$

Développons cette équation au premier ordre en $GM/r_0 c^2$ et GM/rc^2 : comme

$$\begin{aligned} 1 - \frac{r_0^2}{r^2} \frac{1 - \frac{2GM}{rc^2}}{1 - \frac{2GM}{r_0 c^2}} &\simeq 1 - \frac{r_0^2}{r^2} \left(1 - \frac{2GM}{rc^2}\right) \left(1 + \frac{2GM}{r_0 c^2}\right) \\ &\simeq 1 - \frac{r_0^2}{r^2} - \frac{2GM r_0}{r^2 c^2} + \frac{2GM r_0^2}{r^3 c^2} \\ &= \left(1 - \frac{r_0^2}{r^2}\right) \left[1 - \frac{2GM r_0}{r(r+r_0)c^2}\right], \end{aligned}$$

on a

$$\begin{aligned} c^2 dt^2 &\simeq dr^2 \left(1 + \frac{4GM}{rc^2}\right) \frac{1 + \frac{2GM r_0}{r(r+r_0)c^2}}{1 - r_0^2/r^2} \\ &\simeq \frac{dr^2}{1 - r_0^2/r^2} \left[1 + \frac{4GM}{rc^2} + \frac{2GM r_0}{r(r+r_0)c^2}\right] \\ c dt &\simeq \frac{\pm 1}{\sqrt{1 - r_0^2/r^2}} \left[1 + \frac{2GM}{rc^2} + \frac{GM r_0}{r(r+r_0)c^2}\right] dr. \end{aligned}$$

Posons

$$f(r) = \frac{1}{\sqrt{1 - r_0^2/r^2}} \left[1 + \frac{2GM}{rc^2} + \frac{GM r_0}{r(r+r_0)c^2}\right].$$

Si r_T désigne la coordonnée radiale de Schwarzschild de la Terre, r_P celle d'une planète (Mercure ou Vénus) sur laquelle se réfléchit le signal RADAR et t le temps coordonnée nécessaire à la lumière pour effectuer un aller-retour, on obtient donc la relation

$$t = \frac{2}{c} \left[\int_{r_0}^{r_T} f(r) dr \pm \int_{r_0}^{r_P} f(r) dr \right],$$

dans laquelle il faut prendre le signe négatif si la Terre et la planète sont du même côté du soleil, et le signe positif sinon. Les effets relativistes n'étant importants que dans ce dernier cas, c'est celui-là que nous considérerons. On vérifie facilement la fonction primitive

$$\int f(r) dr = \sqrt{r^2 - r_0^2} + \frac{2GM}{c^2} \ln \frac{r + \sqrt{r^2 - r_0^2}}{r_0} + \frac{GM}{c^2} \sqrt{\frac{r - r_0}{r + r_0}}.$$

Ainsi,

$$t = \frac{2}{c} \left(\sqrt{r_T^2 - r_0^2} + \sqrt{r_P^2 - r_0^2} \right) + \Delta t$$

où le premier terme, dominant, correspond au temps coordonnée que mettrait la lumière pour parcourir l'aller-retour Terre-planète en ligne droite à vitesse constante si $M = 0$, c'est-à-dire dans l'espace de Minkowski (figure 9.9) ; et le second est la correction due à la gravitation :

$$\Delta t = \frac{4GM}{c^3} \ln \frac{(r_T + \sqrt{r_T^2 - r_0^2})(r_P + \sqrt{r_P^2 - r_0^2})}{r_0^2} + \frac{2GM}{c^3} \left(\sqrt{\frac{r_T - r_0}{r_T + r_0}} + \sqrt{\frac{r_P - r_0}{r_P + r_0}} \right). \quad (9.65)$$

Si $r_T \gg r_0$ et $r_P \gg r_0$, on peut écrire

$$\Delta t \simeq \frac{4GM}{c^3} \left[\ln \left(\frac{4r_T r_P}{r_0^2} \right) + 1 \right]. \quad (9.66)$$

Le rayon lumineux est d'autant plus retardé qu'il s'aventure à proximité du soleil. La raison principale en est l'effet Einstein, que l'on peut grossièrement résumer en disant que « plus la coordonnée radiale est petite, plus lentement s'écoule le temps ». Remarquons que, selon Newton, une particule devrait plutôt *accélérer* en se rapprochant du soleil.¹¹ Dans le cas d'une réflexion sur la planète Mercure, Δt est de l'ordre de 240 microsecondes (correspondant à une distance de 72 km) pour un temps de parcours total de l'ordre de 20 minutes. On trouvera en figure 9.10 une confrontation entre la théorie et l'expérience.

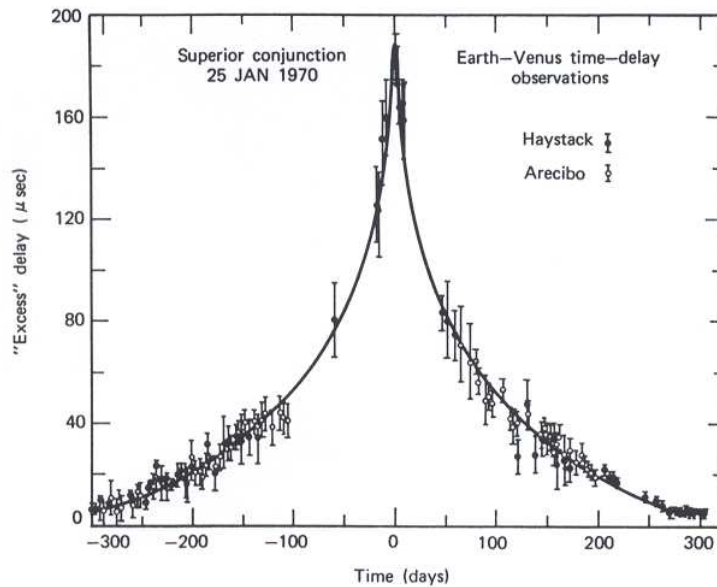


FIGURE 9.10: Retard d'un écho RADAR réfléchi sur la planète Vénus : expérience effectuée par I. Shapiro en 1970. D'après S. Weinberg, *Gravitation and Cosmology*, John Wiley & Sons, 1972, figure 8.3.

11. La courbure de la trajectoire influence peu le résultat : on retrouve l'essentiel du retard si l'on considère que le rayon se propage selon la ligne droite $r \sin \varphi = r_0$. Voir par exemple R. D'Inverno, *op. cit.*, p. 204.

Si l'on néglige les effets du champ gravitationnel terrestre, le temps propre τ s'écoulant sur Terre est relié au temps coordonnée par la relation

$$\tau = t \sqrt{1 - \frac{2GM}{r_T c^2}} \simeq t \left(1 - \frac{GM}{r_T c^2}\right).$$

Cette correction n'affecte que le terme dominant

$$\frac{2}{c} \left(\sqrt{r_T^2 - r_0^2} + \sqrt{r_P^2 - r_0^2} \right).$$

9.10 Particules au voisinage du rayon gravitationnel

Soit un photon émis un peu au-delà du rayon de Schwarzschild, en supposant que l'espace-temps y est toujours décrit par la solution extérieure (9.11). La relation (9.54) exprime qu'un tel photon ne peut échapper à l'astre central qu'en empruntant une trajectoire quasiment radiale et dirigée vers l'extérieur. Ainsi, la lumière elle-même peine à s'extraire d'une sphère de paramètre r légèrement supérieur à r_S . Cette propriété suscite quelques interrogations quant à la voracité de l'espace-temps en cet endroit. Par exemple, dans quelle mesure pourrait-on arrimer une particule matérielle au bout d'une corde suffisamment solide, et la maintenir immobile au voisinage du rayon de Schwarzschild? Pour fournir une réponse convaincante, il est utile de raisonner dans le langage de la question, en termes newtoniens de *force*.

D'après le principe d'équivalence et la relation fondamentale (1.36) de la dynamique relativiste, une particule de masse propre m_0 en mouvement quelconque vérifie l'équation

$$K^\alpha = m_0 \left(\frac{dU^\alpha}{d\tau} + \Gamma_{\beta\gamma}^\alpha U^\beta U^\gamma \right) \quad (9.67)$$

où K^α peut s'interpréter comme la quadriforce accélérant la particule, c'est-à-dire l'écartant de la trajectoire géodésique. Rappelons à cet égard que la relativité générale bannit le concept de force gravitationnelle : une particule « soumise seulement à son inertie et à la gravitation » n'accélère pas, elle suit « une droite » (une géodésique) de l'espace-temps et le membre de gauche de (9.67) est nul.

Pour une particule immobile en la coordonnée r , (9.29) impose

$$K^\alpha = m_0 \left(0 + \Gamma_{00}^\alpha \frac{c^2}{1 - \frac{2GM}{rc^2}} \right) = \frac{GMm_0}{r^2} \delta_1^\alpha.$$

Sa vitesse spatiale étant nulle, en vertu de (1.39) on a $K^i = F^i$ où les F^i sont les composantes de la force intuitive exercée par la corde, dans les coordonnées de Schwarzschild. Ainsi,

$$\vec{F} = \frac{GMm_0}{r^2} \vec{e}_1.$$

Ce résultat ressemble à la réponse newtonienne, à ceci près que \vec{e}_1 n'est pas normé. La force est radiale, tournée vers les r croissants et sa norme

$$\frac{GMm_0}{r^2 \sqrt{1 - \frac{2GM}{rc^2}}}$$

tend vers l'infini si $r \rightarrow r_S^+$. Aucun câble ne peut donc résister à la gravitation, suffisamment près du rayon gravitationnel.

Il n'empêche qu'une particule en chute libre ne ressent rien de particulier, même dans la région $r \leq 2GM/c^2$. Par contre, un corps étendu se trouvera tirailé par les variations du champ gravitationnel selon la position des diverses particules qui le constituent ; en d'autres mots, par les dérivées des symboles de Christoffel ; donc par le tenseur de Riemann. Dans la terminologie newtonienne on parlerait de *forces de marée*. Ce raisonnement est confirmé par les mathématiques. En effet, si les particules n'interagissaient pas, elles suivraient chacune une géodésique. Or, l'écart entre des géodésiques proches est donné par l'équation de déviation géodésique (5.121), elle-même régie par le tenseur de Riemann. Selon le cas, elles ont tendance à se rapprocher ou à s'écartier les unes des autres. Le rôle des forces de cohésion dans le corps est alors de lutter contre ces tendances à la compression ou à la dilatation dues à la gravitation.

Evaluons les tensions subies par un corps étendu en chute libre radiale vers un trou noir. Les quantités physiques doivent être mesurées dans une base orthonormée, ce qui n'est pas le cas des formes et vecteurs de base naturels associés aux coordonnées de Schwarzschild :

$$\begin{aligned}\tilde{\omega}^0 &= c dt, \quad \tilde{\omega}^1 = dr, \quad \tilde{\omega}^2 = d\theta, \quad \tilde{\omega}^3 = d\varphi \\ \vec{e}_0 &= \frac{\partial}{\partial t}, \quad \vec{e}_1 = \frac{\partial}{\partial r}, \quad \vec{e}_2 = \frac{\partial}{\partial \theta}, \quad \vec{e}_3 = \frac{\partial}{\partial \varphi}.\end{aligned}$$

Les formes et les vecteurs déterminant le repère orthonormé statique (c'est-à-dire le repère minkowskien local) associé à la métrique extérieure de Schwarzschild, ont pour expression

$$\begin{aligned}\tilde{\omega}^{0'} &= \sqrt{1 - \frac{2GM}{rc^2}} \tilde{\omega}^0, \quad \tilde{\omega}^{1'} = \frac{1}{\sqrt{1 - \frac{2GM}{rc^2}}} \tilde{\omega}^1, \quad \tilde{\omega}^{2'} = r \tilde{\omega}^2, \quad \tilde{\omega}^{3'} = r \sin \theta \tilde{\omega}^3 \\ \vec{e}_0' &= \frac{1}{\sqrt{1 - \frac{2GM}{rc^2}}} \vec{e}_0, \quad \vec{e}_1' = \sqrt{1 - \frac{2GM}{rc^2}} \vec{e}_1, \quad \vec{e}_2' = \frac{1}{r} \vec{e}_2, \quad \vec{e}_3' = \frac{1}{r \sin \theta} \vec{e}_3.\end{aligned}$$

La quadrivitesse d'une particule en chute libre a pour composantes

$$(U^\alpha) = \left(\frac{dx^\alpha}{d\tau} \right) = c \left(\frac{l}{1 - \frac{2GM}{rc^2}}, -\sqrt{l^2 - 1 + \frac{2GM}{rc^2}}, 0, 0 \right).$$

Puisque $\vec{U} = U^\alpha \vec{e}_\alpha = U^{\beta'} \vec{e}_{\beta}'$, on en déduit ses composantes dans la base orthonormée statique :

$$(U^{\beta'}) = \frac{c}{\sqrt{1 - \frac{2GM}{rc^2}}} \left(l, -\sqrt{l^2 - 1 + \frac{2GM}{rc^2}}, 0, 0 \right).$$

Or,

$$(U^{\beta'}) = \frac{1}{\sqrt{1 - V^2/c^2}} (c, \vec{V})$$

où V est la norme de la vitesse spatiale, radiale par hypothèse. On en déduit

$$V = \frac{c}{l} \sqrt{l^2 - 1 + \frac{2GM}{rc^2}}.$$

Passons au repère orthonormé concomitant $\{\tilde{\omega}^\alpha\}, \{\tilde{e}_\alpha\}$, tel que $\vec{U} = c \tilde{e}_0$, via la transformation de Lorentz

$$\begin{pmatrix} \tilde{\omega}^0 \\ \tilde{\omega}^1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1-V^2/c^2}} & \frac{V/c}{\sqrt{1-V^2/c^2}} \\ \frac{V/c}{\sqrt{1-V^2/c^2}} & \frac{1}{\sqrt{1-V^2/c^2}} \end{pmatrix} \begin{pmatrix} \tilde{\omega}^{0'} \\ \tilde{\omega}^{1'} \end{pmatrix}$$

$$\begin{pmatrix} \tilde{e}_0 \\ \tilde{e}_1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1-V^2/c^2}} & \frac{-V/c}{\sqrt{1-V^2/c^2}} \\ \frac{-V/c}{\sqrt{1-V^2/c^2}} & \frac{1}{\sqrt{1-V^2/c^2}} \end{pmatrix} \begin{pmatrix} \tilde{e}_0' \\ \tilde{e}_1' \end{pmatrix}$$

(les autres quantités demeurant inchangées). Il n'est pas difficile de prouver que

$$\nabla_{\vec{U}} \tilde{e}_\alpha = 0 \quad \forall \alpha.$$

Si $\vec{S} = \underline{S}^\alpha \tilde{e}_\alpha$ désigne un champ vectoriel quelconque, on a donc

$$\nabla_{\vec{U}} \nabla_{\vec{U}} \vec{S} = \frac{d^2 \underline{S}^\alpha}{d\tau^2} \tilde{e}_\alpha.$$

Choisissons pour \vec{S} le vecteur connectant deux événements de temps propre identique situés sur deux géodésiques proches. L'équation de déviation géodésique (5.121) prend ainsi la forme

$$\frac{d^2 \underline{S}^\alpha}{d\tau^2} = c^2 \underline{R}^\alpha{}_{00\beta} \underline{S}^\beta. \quad (9.68)$$

A partir de (9.11), on calcule sans difficulté les composantes du tenseur de Riemann dans les coordonnées naturelles (ct, r, θ, φ) :

$$R_{0101} = -\frac{2GM}{r^3 c^2}, \quad R_{0202} = \left(1 - \frac{2GM}{rc^2}\right) \frac{GM}{rc^2}, \quad R_{0303} = \left(1 - \frac{2GM}{rc^2}\right) \frac{GM}{rc^2} \sin^2 \theta,$$

$$R_{1212} = \frac{-GM}{rc^2 \left(1 - \frac{2GM}{rc^2}\right)}, \quad R_{1313} = \frac{-GM}{rc^2 \left(1 - \frac{2GM}{rc^2}\right)} \sin^2 \theta, \quad R_{2323} = \frac{2GM}{c^2} r \sin^2 \theta. \quad (9.69)$$

On peut en déduire ses composantes dans les divers repères définis ci-dessus, à l'aide des relations

$$\begin{aligned} \mathbf{R} = R_{\alpha\beta\gamma\delta} \tilde{\omega}^\alpha \otimes \tilde{\omega}^\beta \otimes \tilde{\omega}^\gamma \otimes \tilde{\omega}^\delta &= R'_{\lambda\mu\nu\xi} \tilde{\omega}^{\lambda'} \otimes \tilde{\omega}^{\mu'} \otimes \tilde{\omega}^{\nu'} \otimes \tilde{\omega}^{\xi'} \\ &= \underline{R}_{\lambda\mu\nu\xi} \tilde{\omega}^\lambda \otimes \tilde{\omega}^\mu \otimes \tilde{\omega}^\nu \otimes \tilde{\omega}^\xi. \end{aligned}$$

On obtient

$$R'_{0101} = -\frac{2GM}{r^3 c^2}, \quad R'_{0202} = R'_{0303} = \frac{GM}{r^3 c^2}, \quad R'_{1212} = R'_{1313} = -\frac{GM}{r^3 c^2}, \quad R'_{2323} = \frac{2GM}{r^3 c^2}$$

$$\underline{R}_{\lambda\mu\nu\xi} = R'_{\lambda\mu\nu\xi}.$$

Le développement de (9.68) donne ainsi

$$\frac{d^2 \underline{S}^0}{d\tau^2} = 0 \quad (9.70)$$

$$\frac{d^2 \underline{S}^1}{d\tau^2} = \frac{2GM}{r^3} \underline{S}^1 \quad (9.71)$$

$$\frac{d^2 \underline{S}^2}{d\tau^2} = -\frac{GM}{r^3} \underline{S}^2 \quad (9.72)$$

$$\frac{d^2 \underline{S}^3}{d\tau^2} = -\frac{GM}{r^3} \underline{S}^3. \quad (9.73)$$

La deuxième équation exprime une tendance à l'écartement radial. Les deux suivantes traduisent un rapprochement « latéral ». Cela s'accorde bien avec la symétrie sphérique de la situation.

Modélisons un voyageur humain par un parallélépipède rectangle homogène de dimensions propres $b \times b \times h$ et de masse propre m_0 . A une hauteur x au-dessus du centre de masse, une tranche de matière d'épaisseur dx subit, par rapport à celui-ci, une force

$$\frac{m_0}{b^2 h} b^2 dx \frac{2GM}{r^3} x$$

(analogue à une force centrifuge¹²). La force totale vers le haut vaut donc

$$F_{\parallel} = \int_0^{h/2} \frac{2GMm_0 x}{r^3 h} dx = \frac{GMm_0 h}{4r^3}.$$

On en déduit la force radiale par unité de surface, c'est-à-dire la *tension radiale*

$$T_{\parallel} = \frac{GMm_0 h}{4b^2 r^3}. \quad (9.74)$$

On calcule de même

$$F_{\perp} = \int_0^{b/2} \frac{GMm_0 x}{r^3 b} dx = \frac{GMm_0 b}{8r^3},$$

d'où la *pression latérale*

$$T_{\perp} = \frac{GMm_0}{8hr^3}. \quad (9.75)$$

Remplaçant par les valeurs $b = 0.2$ m, $h = 1.9$ m et $m_0 = 80$ kg, on trouve

$$T_{\parallel} = 1.3 \times 10^9 \frac{M/M_{\odot}}{(r/1 \text{ km})^3} \text{ atmosphères}$$

$$T_{\perp} = 6.9 \times 10^6 \frac{M/M_{\odot}}{(r/1 \text{ km})^3} \text{ atmosphères.}$$

Le corps humain peut résister à des tensions de l'ordre de 100 atmosphères. Or, le rayon de Schwarzschild d'un trou noir de x masses solaires est égal à $3x$ km. Un être humain pourra donc franchir le rayon de Schwarzschild d'un trou noir de x masses solaires sans être déchiré ni broyé, si

$$1.3 \times 10^9 \frac{x}{(3x)^3} \leq 10^2 \quad \text{donc} \quad x \geq 7 \times 10^2.$$

Plus importante est la masse du trou noir, moins dangereuse est la traversée du rayon gravitationnel.

12. L'équation (9.71) rappelle celle de Newton pour une particule solidaire d'un parallélépipède rectangle en rotation uniforme de vitesse angulaire $\omega = \sqrt{2GM/r^3}$ autour d'un axe principal d'inertie passant par son centre de masse, lui-même supposé inertiel. Dans ce référentiel, la particule ne pourra rester immobile que si la force centrifuge est compensée par la tension au sein du barreau.

9.11 Le trou noir : solution extérieure

9.11.1 Pathologies des coordonnées de Schwarzschild

On appelle *trou noir statique* une région de l'espace-temps décrite par la solution extérieure de Schwarzschild, avec la condition $r < r_S = 2GM/c^2$. La possibilité de trouver cette condition réalisée dans la nature fut jugée absurde en 1916. En effet, il apparaît immédiatement que la métrique (9.11) diverge en $r = r_S$; et d'autres difficultés apparaissent, que nous allons passer en revue.

L'équation de mouvement d'un photon radial s'écrit

$$c \frac{dt}{dr} = \pm \frac{1}{1 - \frac{2GM}{rc^2}}.$$

Dès lors, les cônes de lumière radiaux sont de plus en plus écrasés lorsque $r \rightarrow r_S$. Les lignes d'univers des particules matérielles étant à l'intérieur de ces cônes, on peut déjà affirmer qu'une particule mettra un temps t infini pour atteindre le rayon de Schwarzschild en provenant de l'extérieur.

Effectuons le calcul explicite. Soit une particule matérielle lâchée sans vitesse initiale depuis la coordonnée $r_0 > 2GM/c^2$. En combinant (9.34), (9.26) et (9.31), on trouve l'équation

$$c dt = - \frac{\sqrt{1 - \frac{2GM}{r_0 c^2}}}{\left(1 - \frac{2GM}{rc^2}\right) \sqrt{\frac{2GM}{rc^2} - \frac{2GM}{r_0 c^2}}} dr. \quad (9.76)$$

A l'aide du changement de variables $r = r_0(1 + \cos \eta)/2$, $\eta \in [0, \pi[$, on obtient

$$c dt = r_0 \sqrt{\frac{r_0 c^2}{2GM} - 1} \frac{\cos^2 \eta/2}{1 - \frac{2GM}{r_0 c^2 \cos^2 \eta/2}} d\eta$$

$$ct + K = \sqrt{\frac{r_0 c^2}{2GM} - 1} \left[\frac{r_0}{2} (\eta + \sin \eta) + \frac{2GM}{c^2} \eta \right] + \frac{2GM}{c^2} \ln \left| \frac{\sqrt{\frac{r_0 c^2}{2GM} - 1} + \tan \frac{\eta}{2}}{\sqrt{\frac{r_0 c^2}{2GM} - 1} - \tan \frac{\eta}{2}} \right| \quad (9.77)$$

où K est une constante arbitraire de primitivation (si $K = 0$, alors $t = 0$ en $\eta = 0$, correspondant à $r = r_0$). La quantité sous le logarithme diverge pour $r \rightarrow 2GM/c^2$. En effet, on a alors

$$\cos^2 \frac{\eta}{2} \rightarrow \frac{2GM}{r_0 c^2},$$

donc

$$\tan \frac{\eta}{2} \rightarrow \sqrt{\frac{r_0 c^2}{2GM} - 1}.$$

La courbe $ct(r)$ est représentée en figure 9.11, où l'on a posé $K = 0$ partout.

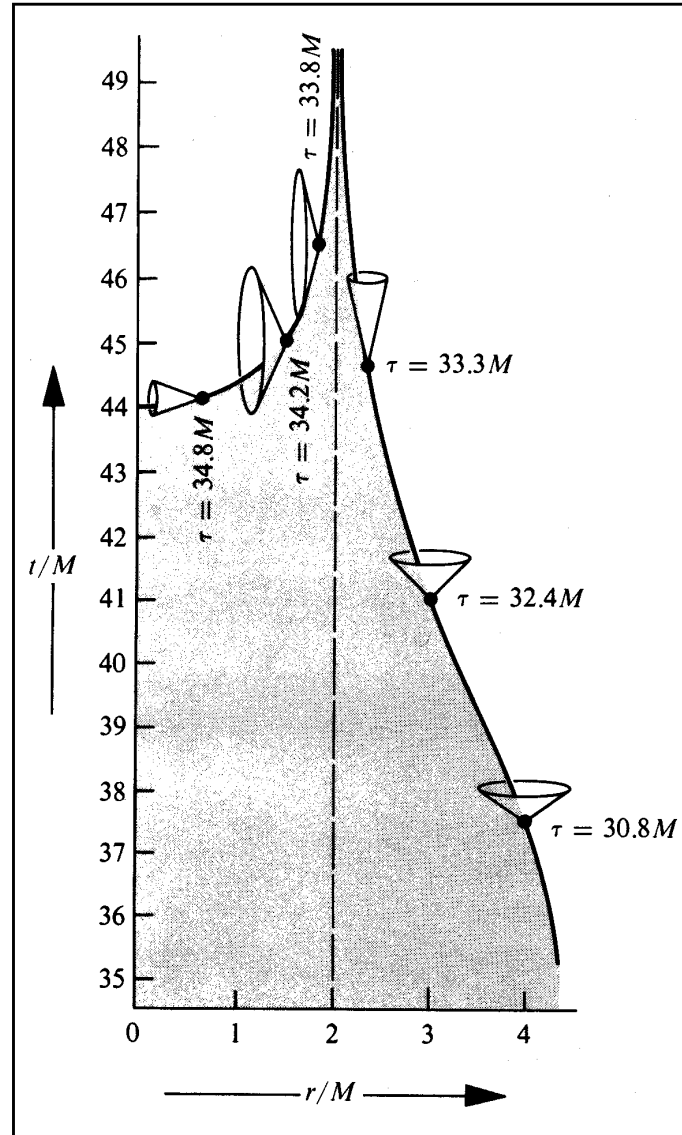


FIGURE 9.11: Ligne d'univers d'une particule à la surface d'une étoile sphérique, en coordonnées de Schwarzschild ($rc^2/GM, tc^3/GM$). L'étoile, de rayon initial $r_0 = 10 GM/c^2$, s'effondre sous l'effet de la gravitation et l'on suppose que la particule suit une géodésique radiale (c'est le cas si l'étoile est modélisée par une sphère de poussière, donc sans pression). On a adopté les unités géométriques $c = 1, G = 1$. En $r = 0, t/M = 14\pi$. Les événements survenant à l'intérieur de l'astre sont dans la région grise. D'après C.W. Misner, K.S. Thorne, J.A. Wheeler, op. cit., figure 32.1 (a).

Montrons que la particule atteint pourtant le rayon gravitationnel en un temps propre fini. De (9.34) et (9.31), il résulte

$$c d\tau = \frac{-dr}{\sqrt{\frac{2GM}{rc^2} - \frac{2GM}{r_0c^2}}}. \quad (9.78)$$

La primitive est connue¹³ mais on simplifie les calculs en utilisant le même changement de variable qu'au paragraphe précédent : en posant

$$r = \frac{r_0}{2} (1 + \cos \eta), \quad \eta \in [0, \pi[, \quad (9.79)$$

on trouve successivement

$$\begin{aligned} d\tau &= \frac{r_0}{2} \sqrt{\frac{r_0}{2GM}} (1 + \cos \eta) d\eta \\ \tau &= \frac{r_0}{2} \sqrt{\frac{r_0}{2GM}} (\eta + \sin \eta), \end{aligned} \quad (9.80)$$

si l'on fixe l'origine des τ pour que $\eta = 0$ corresponde à $\tau = 0$. Il est évident que le temps propre reste toujours borné. Les relations (9.79) et (9.80) constituent les équations paramétriques de la *cycloïde* générée par une particule d'ordonnée initiale $r = r_0$ et solidaire d'un cercle de rayon $r_0/2$ roulant sur l'axe des abscisses $\tau\sqrt{2GM/r_0}$.

D'après (9.79) et (9.80), une particule lancée radialement sans vitesse initiale depuis $r_0 > r_S$ tend vers la singularité centrale $r = 0$ ($\eta \rightarrow \pi^-$) en le temps propre

$$\tau = \frac{\pi r_0}{2} \sqrt{\frac{r_0}{2GM}}.$$

Si $r_0 \rightarrow r_S^+$ on a

$$\tau \rightarrow \frac{\pi GM}{c^3}. \quad (9.81)$$

Dans ces conditions, $l \rightarrow 0^+$ et $1 - 2GM/rc^2$ est négatif durant tout le voyage : d'après (9.34), la trajectoire radiale ($h = 0$) réalise alors le temps propre maximal. Le résultat (9.81) constitue l'espérance de vie maximale à laquelle peut prétendre une particule au moment où elle franchit le rayon de Schwarzschild. Pour un astre de trois masses solaires, elle vaut 4.5×10^{-5} s.

Dans le cas particulier où $r_0 \rightarrow +\infty$, il est possible d'exprimer analytiquement les relations $r(t)$ et $r(\tau)$ au voisinage du rayon de Schwarzschild.

En fonction du temps coordonnée, on trouve à partir de (9.76) :

$$c dt = \frac{-dr}{\left(1 - \frac{2GM}{rc^2}\right) \sqrt{\frac{2GM}{rc^2}}}.$$

13. Si $0 < r < r_0$, alors à l'aide du changement de variable $\tan \omega = \sqrt{r_0/r - 1}$ on trouve

$$\int \frac{dr}{\sqrt{\frac{r_0}{r} - 1}} = -r_0 \arctan \sqrt{\frac{r_0}{r} - 1} - r \sqrt{\frac{r_0}{r} - 1}.$$

Si $r_0 < r$, en posant $\tanh \omega = \sqrt{1 - r_0/r}$ on trouve

$$\int \frac{dr}{\sqrt{1 - \frac{r_0}{r}}} = r_0 \operatorname{arctanh} \sqrt{1 - \frac{r_0}{r}} + r \sqrt{1 - \frac{r_0}{r}}.$$

En effectuant le changement de variables $\sqrt{r} = u$, on se ramène au calcul de la primitive d'une fraction rationnelle en u . Notant K la constante arbitraire d'intégration, on obtient

$$\frac{c^3(t+K)}{2GM} = -\frac{2}{3}\sqrt{\left(\frac{rc^2}{2GM}\right)^3} - 2\sqrt{\frac{rc^2}{2GM}} + \ln \left| \frac{\sqrt{\frac{rc^2}{2GM} + 1}}{\sqrt{\frac{rc^2}{2GM} - 1}} \right|.$$

Posons

$$r = \frac{2GM}{c^2} (1 + \epsilon), \quad |\epsilon| \ll 1, \quad (9.82)$$

et développons au premier ordre en ϵ . Il vient

$$\frac{c^3(t+K)}{2GM} \simeq -\ln|\epsilon|$$

$$r(t) = \frac{2GM}{c^2} \left[1 + K' \exp\left(\frac{-c^3 t}{2GM}\right) \right]$$

où K' est une constante reliée à K et possédant le même signe que ϵ . Cette relation précise comment la coordonnée radiale tend vers le rayon de Schwarzschild, asymptotiquement en fonction du temps coordonnée.

En fonction du temps propre, en vertu de (9.78) on peut écrire

$$d\tau = -\sqrt{\frac{r}{2GM}} dr.$$

Choisissant la constante d'intégration pour que $r = 0$ corresponde à $\tau = 0$, on obtient

$$\tau = -\frac{2}{3}\sqrt{\frac{r^3}{2GM}}.$$

En développant cette relation au premier ordre en ϵ à partir de (9.82), on trouve

$$r(\tau) = \frac{2GM}{c^2} - c(\tau - \tau_S)$$

où la quantité

$$\tau_S = \frac{-4GM}{3c^3}$$

est l'âge de la particule au moment où elle atteint le rayon de Schwarzschild. L'approche se fait linéairement en fonction du temps propre.

Pour compléter l'étude des géodésiques radiales, il reste à envisager les géodésiques spatiales. Comme paramètre affine, on choisira naturellement la longueur d'arc s . Vu l'intégrale première (9.26) et l'expression du ds^2 , on dispose des équations

$$c \frac{dt}{ds} = \frac{l}{1 - \frac{2GM}{rc^2}}$$

$$\frac{dr}{ds} = \pm \sqrt{l^2 + 1 - \frac{2GM}{rc^2}}.$$

Si $l \neq 0$, l'analyse ci-dessus reste valable et ces géodésiques franchissent le rayon gravitationnel en un temps t infini. Seul le cas $l = 0$ n'a pas encore été traité : il s'agit des géodésiques radiales d'équation $t = \text{constante}$. Elles ne sont en fait spatiales que si $r > r_S$. Il s'agit là d'une difficulté supplémentaire des coordonnées de Schwarzschild, car ces courbes s'interrompent brutalement lorsqu'elles atteignent le rayon gravitationnel.

En conclusion, hormis un ensemble de mesure nulle ($l = 0$), toutes les géodésiques radiales passent par le « point à l'infini » ($r = r_S, t = \infty$).

Signalons encore quelques bizarreries :

- Pour $r < r_S$, la coordonnée r est temporelle et t est spatiale. L'élément de longueur ds^2 est strictement positif sur l'hypersurface $r = \text{constante} < r_S$. Celle-ci possède donc le caractère *spatial* et aucune particule ne peut s'y maintenir. Tout cela n'est guère intuitif.
- La sous-variété $r = r_S$ possède un volume nul : pour cette valeur de r on a

$$\int \sqrt{|g_{00} g_{22} g_{33}|} dt d\theta d\varphi = 0.$$

9.11.2 Les coordonnées de Kruskal-Szekeres

Une étude mathématique plus fouillée suggère que l'espace-temps n'est pas davantage singulier en $r = r_S$ que la Terre ne l'est au pôle nord : en fait, ce sont les coordonnées temporelle et spatiale de Schwarzschild qui sont inadéquates pour décrire les phénomènes au voisinage du rayon gravitationnel.

Le calcul de quantités géométriques scalaires, donc indépendantes du choix des coordonnées, fournit un premier argument en faveur de cette affirmation. Il est naturel de choisir une expression liée au tenseur de Riemann (car les quantités formées à partir du tenseur de Ricci sont nulles en vertu des équations d'Einstein), par exemple

$$I = R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta}.$$

A partir de (9.69), on calcule sans peine

$$I = \frac{48 G^2 M^2}{r^6 c^4}.$$

Cet invariant ne diverge pas lorsque $r \rightarrow r_S$. Par contre, $r = 0$ s'annonce comme une singularité irréductible.

Un meilleur système de coordonnées devrait démêler cet imbroglio. Au fil des décennies plusieurs changements de variables ont été proposés, parmi lesquels nous retiendrons celui découvert en 1960 par M. Kruskal et G. Szekeres :

Région I ($r > r_S$)

$$\begin{aligned} u &= \sqrt{\frac{rc^2}{2GM} - 1} \exp\left(\frac{rc^2}{4GM}\right) \cosh\left(\frac{c^3 t}{4GM}\right) \\ v &= \sqrt{\frac{rc^2}{2GM} - 1} \exp\left(\frac{rc^2}{4GM}\right) \sinh\left(\frac{c^3 t}{4GM}\right) \end{aligned}$$

Région II ($r < r_S$)

$$\begin{aligned} u &= \sqrt{1 - \frac{rc^2}{2GM}} \exp\left(\frac{rc^2}{4GM}\right) \sinh\left(\frac{c^3 t}{4GM}\right) \\ v &= \sqrt{1 - \frac{rc^2}{2GM}} \exp\left(\frac{rc^2}{4GM}\right) \cosh\left(\frac{c^3 t}{4GM}\right) \end{aligned}$$

Région III ($r > r_S$)

$$\begin{aligned} u &= -\sqrt{\frac{rc^2}{2GM} - 1} \exp\left(\frac{rc^2}{4GM}\right) \cosh\left(\frac{c^3 t}{4GM}\right) \\ v &= -\sqrt{\frac{rc^2}{2GM} - 1} \exp\left(\frac{rc^2}{4GM}\right) \sinh\left(\frac{c^3 t}{4GM}\right) \end{aligned}$$

Région IV ($r < r_S$)

$$\begin{aligned} u &= -\sqrt{1 - \frac{rc^2}{2GM}} \exp\left(\frac{rc^2}{4GM}\right) \sinh\left(\frac{c^3 t}{4GM}\right) \\ v &= -\sqrt{1 - \frac{rc^2}{2GM}} \exp\left(\frac{rc^2}{4GM}\right) \cosh\left(\frac{c^3 t}{4GM}\right). \end{aligned}$$

Ces relations peuvent aisément être inversées. Dans les régions I et III, on trouve

$$\begin{aligned} \left(\frac{rc^2}{2GM} - 1\right) \exp\left(\frac{rc^2}{2GM}\right) &= u^2 - v^2 \\ ct &= \frac{4GM}{c^2} \operatorname{arctanh}\left(\frac{v}{u}\right), \end{aligned}$$

tandis que dans les régions II et IV on obtient

$$\begin{aligned} \left(\frac{rc^2}{2GM} - 1\right) \exp\left(\frac{rc^2}{2GM}\right) &= u^2 - v^2 \\ ct &= \frac{4GM}{c^2} \operatorname{arctanh}\left(\frac{u}{v}\right). \end{aligned}$$

Dans les nouvelles coordonnées, l'élément de longueur s'écrit

$$ds^2 = \frac{32}{r} \left(\frac{GM}{c^2}\right)^3 \exp\left(\frac{-rc^2}{2GM}\right) (-dv^2 + du^2) + r^2 d\Omega^2. \quad (9.83)$$

Il apparaît que la seule véritable singularité se trouve en $r = 0$.

On pourrait s'inquiéter de la pertinence d'un résultat provenant de changements de variables singuliers effectués sur des métriques elles-mêmes singulières. Mais il est possible de vérifier directement que la métrique (9.83), qui est bien à symétrie sphérique spatiale, et où la coordonnée r doit être exprimée (ne fût-ce que numériquement) en fonction de u et de v , satisfait les équations du champ dans le vide. On dit que *la variété de Kruskal-Szekeres constitue l'extension analytique maximale de la variété de Schwarzschild*.

La figure 9.12 représente l'espace-temps dans les coordonnées de Schwarzschild et de Kruskal. Les coordonnées θ et φ ont été supprimées : chaque point représente, en réalité, une 2-sphère. On voit que la singularité centrale correspond à la branche d'hyperbole $v^2 - u^2 = 1$ (plus exactement, il s'agit d'une hypersurface spatiale si l'on tient compte des coordonnées angulaires). D'autre part, la « ligne » ($r = r_S, -\infty < t < +\infty$) possède la topologie d'un *point* (en fait d'une surface à deux dimensions), tout comme, sur une sphère, la « ligne » ($\theta = 0, 0 \leq \varphi < 2\pi$) représente un point, le pôle nord. Quant au « point à

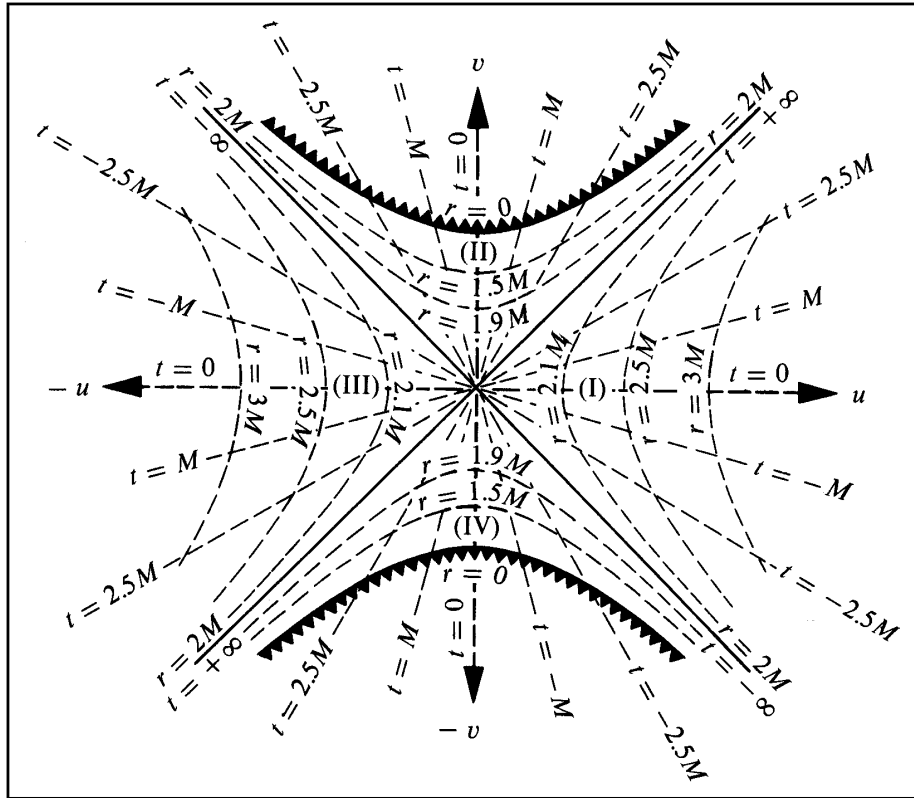


FIGURE 9.12: Lien entre les coordonnées de Schwarzschild et celles de Kruskal. D'après C.W. Misner, K.S. Thorne, J.A. Wheeler, op. cit., figure 31.3.

l'infini » ($r = r_S, t = \infty$), il possède la topologie d'une ligne (*i.e.*, une hypersurface) nulle.

Les coordonnées de Kruskal offrent un grand nombre d'avantages : comparer, par exemple, les figures 9.11 et 9.13 représentant une même situation physique, respectivement dans chacun des deux types de coordonnées. Notamment, en n'importe quel point les cônes de lumière radiaux ont une équation du type

$$v - K_2 = \pm(u - K_1).$$

Cela permet de comprendre le caractère irréversible de la traversée du rayon gravitationnel. Une particule qui s'aventurerait à le franchir dans le sens des r décroissants n'aurait d'autre destin que de rencontrer la singularité centrale. On dit que l'hypersurface $v = \pm u$, c'est-à-dire $r = 2GM/c^2$, constitue un *horizon-événement* : rien de ce qui se passe au-delà de cette frontière (en venant de l'extérieur) ne peut être communiqué en deçà. Cette notion est totalement inconnue de la mécanique classique (le *trou noir laplacien* est simplement un astre newtonien dont la vitesse de libération est supérieure ou égale à celle de la lumière : il ne ressemble donc que très superficiellement à son homologue relativiste).

L'hypothèse de l'existence des trous noirs n'a cessé de gagner du crédit suite aux progrès de l'astrophysique théorique et observationnelle (découverte des naines blanches et des étoiles à neutrons). Depuis la fin des années 1960, elle a reçu l'adhésion de la majorité des spécialistes et depuis lors, les arguments observationnels en sa faveur ont continué à s'accumuler.

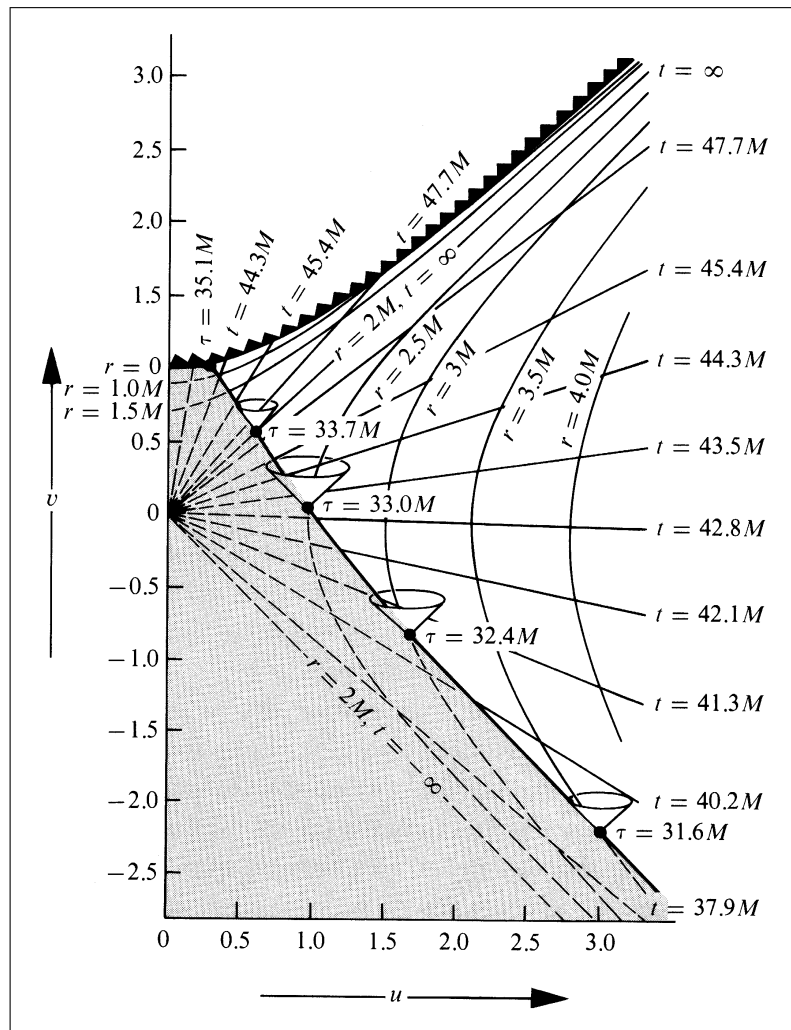


FIGURE 9.13: Môme situation qu'à la figure 9.11, dans les coordonnées (u, v) de Kruskal (sauf que, dans le changement de coordonnées, on a remplacé t par $t - 42.8 \text{ GM}/c^3$). D'après C.W. Misner, K.S. Thorne, J.A. Wheeler, op. cit., figure 32.1 (b).

9.12 Le trou noir : solution intérieure

Le trou noir décrit par la solution extérieure de Schwarzschild (celui de la figure 9.12) est posé comme solution des équations d'Einstein existant, en quelque sorte, de toute éternité. Aussi instructives soient-elles, ses propriétés ne permettent donc pas de se forger une idée générale de la formation d'un trou noir *réaliste*, par exemple en astrophysique. Il faut pour cela s'atteler à l'étude d'une solution intérieure, dont les détails dépendent du modèle adopté pour la description de la matière. Or, il se trouve que l'on connaît une solution analytique si l'équation d'état est celle d'un fluide parfait homogène à pression nulle (sphère de poussière). Nous allons d'abord en faire brièvement l'étude newtonienne. Cela nous permettra d'acquérir une première intuition qu'il faudra confronter, ensuite, aux exigences de la relativité générale.

9.12.1 Effondrement radial d'une sphère de poussière selon Newton

Soit une sphère de poussière initialement au repos et de masse volumique homogène μ_0 . Il est facile d'écrire l'équation de conservation de l'énergie pour une particule de coordonnée initiale $r = r_0$:

$$\frac{1}{2} V^2 - \frac{Gm_0}{r} = E,$$

en posant

$$m_0 = \frac{4}{3} \pi r_0^3 \mu_0, \quad E = -\frac{Gm_0}{r_0}$$

(seule la sphère intérieure à la particule contribue au champ de gravitation). On en déduit

$$dr = -\sqrt{\frac{r_0}{r} - 1} \sqrt{-2E} dt.$$

Cette équation est identique à (9.78) où l'on a remplacé M par m_0 . En résolvant de la même manière, on trouve

$$r(\eta) = \frac{r_0}{2} (1 + \cos \eta) \tag{9.84}$$

$$\sqrt{-2E} t(\eta) = \frac{r_0}{2} (\eta + \sin \eta)$$

avec $\eta \in [0, \pi[$. La deuxième relation peut également s'écrire

$$t = \sqrt{\frac{3}{32\pi G\mu_0}} (\eta + \sin \eta). \tag{9.85}$$

On en tire deux conclusions majeures :

– Toutes les particules atteignent le centre en $\eta_c = \pi$. Cela correspond à un temps de chute *indépendant de r_0* :

$$t_c = \sqrt{\frac{3\pi}{32G\mu_0}}. \tag{9.86}$$

– La sphère reste homogène à tout instant. En effet,

$$\begin{aligned} \mu(\eta) &= \frac{m_0}{4/3 \pi r(\eta)^3} \\ &= \frac{8\mu_0}{(1 + \cos \eta)^3}. \end{aligned} \tag{9.87}$$

9.12.2 Effondrement radial d'une sphère de poussière selon Einstein

La solution à ce problème fut obtenue en 1934 par R. Tolman et remaniée en 1939 par J. Oppenheimer et H. Snyder.¹⁴ Il s'avère que la géométrie intérieure engendrée par une sphère de poussière en effondrement gravitationnel est décrite par la métrique de Friedmann-Lemaître-Robertson-Walker avec $k = 1$ (chapitre 8), c'est-à-dire, en coordonnées comobiles :

$$ds^2 = -c^2 d\tau^2 + R^2(\tau) \left[d\chi^2 + \sin^2 \chi (d\theta^2 + \sin^2 \theta d\varphi^2) \right], \quad \chi \in [0, \chi_{surf}]. \tag{9.88}$$

14. J.R. Oppenheimer et H. Snyder, *On Continued Gravitational Contraction*, Phys. Rev. **56**, pp. 455-459 (1939).

L'origine $\chi = 0$ est associée au centre de la sphère matérielle, la quantité χ_{surf} correspondant à la surface (nous calculerons sa valeur plus loin). Quant au facteur d'échelle R , il est donné dans la section 8.9 avec $P = 0$, en fonction du temps η défini par $c d\tau = R(\eta) d\eta$. Nous allons quelque peu adapter les notations en posant $\eta' = \eta - \pi$, pour que le facteur d'échelle soit maximal en $\eta' = 0$ (début de l'effondrement) et que la singularité finale corresponde à $\eta' \rightarrow \pi^-$. Dans la suite, on omettra les « prime ». Notons également $R_0 = 2\mathcal{R}_0$ et $\rho = \mu c^2$.

En résumé, on a

$$R(\eta) = \frac{R_0}{2} (1 + \cos \eta) \quad (9.89)$$

$$c\tau(\eta) = \frac{R_0}{2} (\eta + \sin \eta) \quad (9.90)$$

$$\mu(\eta) = \frac{3R_0 c^2}{8\pi G R^3(\eta)}.$$

On voit que $R(0) = R_0$. Posant $\mu(0) = \mu_0$, on obtient les relations

$$R_0 = \sqrt{\frac{3c^2}{8\pi G \mu_0}} \quad (9.91)$$

$$\mu(\eta) = \frac{8\mu_0}{(1 + \cos \eta)^3}. \quad (9.92)$$

Remarquer la similitude des relations (9.89), (9.90) et (9.92) avec la solution newtonienne (9.84), (9.85) et (9.87).¹⁵ Les conclusions sont donc identiques. Notamment, le temps propre nécessaire pour atteindre la singularité « centrale » est donné par (9.86), quelle que soit la particule. Dans le cadre du modèle simple que nous avons adopté, la singularité du trou noir s'apparente donc à une singularité cosmologique de type « Big Crunch », située dans le futur : on comprend qu'elle soit décrite par une hypersurface de type spatial, comme nous l'avons vu dans la section 9.11.2. Il ne faut pas concevoir le *collapse* comme la coalescence de toute la matière en un point spatial donné (ce qui serait d'ailleurs peu conforme à l'esprit de la théorie), mais comme une sorte de mayonnaise qui prend simultanément en tout point et devient de plus en plus dense.

Pour achever la construction de l'espace-temps, il reste à raccorder la solution (9.88) sur la surface $\chi = \chi_{surf}$, avec la solution extérieure dans le vide. Or, d'après le théorème de Birkhoff, celle-ci ne peut être que la métrique de Schwarzschild (9.11). Notons r_{surf0} la coordonnée radiale de Schwarzschild de la surface au temps initial $\eta = 0$, strictement supérieure au rayon de Schwarzschild $r_S = 2GM/c^2$. Les particules de la surface au temps η quelconque vérifient les équations (9.79) et (9.80) :

$$r_{surf}(\eta) = \frac{r_{surf0}}{2} (1 + \cos \eta) \quad (9.93)$$

$$\tau_{surf}(\eta) = \frac{r_{surf0}}{2} \sqrt{\frac{r_{surf0}}{2GM}} (\eta + \sin \eta). \quad (9.94)$$

15. En mécanique newtonienne les paramètres r_0 et μ_0 sont indépendants, tandis qu'en relativité générale R_0 et μ_0 sont reliés par (9.91) : le deuxième paramètre est en fait χ_{surf} .

Comparant (9.94) et (9.90), on trouve¹⁶

$$R_0 = \sqrt{\frac{r_{surf0}^3 c^2}{2GM}}. \quad (9.95)$$

A l'intérieur de la matière, introduisons la coordonnée

$$r = R(\eta) \sin \chi. \quad (9.96)$$

Cette relation revient à exiger que les sphères de coordonnée χ ont une aire égale à $4\pi r^2$, comme dans la solution extérieure. Par conséquent, à tout instant la coordonnée intérieure r de la surface est identique à la coordonnée radiale extérieure de Schwarzschild. La relation (9.96) prend donc la forme

$$r_{surf0} = R_0 \sin \chi_{surf}. \quad (9.97)$$

Compte tenu de la valeur (9.95) de R_0 , il vient

$$\sin \chi_{surf} = \sqrt{\frac{r_S}{r_{surf0}}} < 1. \quad (9.98)$$

9.12.3 La formation de l'horizon

Considérons des photons quittant le centre de la sphère de poussière en différents instants. Il existe un temps critique η_{crit} pour lequel les photons correspondants parviendront à la surface au moment précis où celle-ci atteint le rayon gravitationnel r_S . D'après les résultats de la section 9.11, ils y resteront alors indéfiniment. Cela permet de partager l'espace-temps en deux régions : tous les photons émis depuis le centre en $\eta < \eta_{crit}$ pourront se propager jusqu'à l'infini ; et ceux émis en $\eta > \eta_{crit}$ retomberont vers la singularité. Par définition, les lignes d'univers des photons critiques déterminent l'*horizon-événement*.

Il est facile d'établir l'équation intérieure d'un photon émis depuis le centre $\chi = 0$ en $\eta = \eta_0$, en posant $ds^2 = 0$ dans (9.88). On obtient $d\chi = d\eta$, donc

$$\chi(\eta) = \eta - \eta_0.$$

Vu la définition (9.96) et les équations (9.89) et (9.90), les points de la trajectoire ont pour coordonnées intérieures

$$r(\eta) = \frac{R_0}{2} (1 + \cos \eta) \sin(\eta - \eta_0) \quad (9.99)$$

$$c\tau(\eta) = \frac{R_0}{2} (\eta + \sin \eta). \quad (9.100)$$

Le photon atteint la surface au temps

$$\eta^* = \eta_0 + \chi_{surf}.$$

16. En confrontant (9.95) et (9.91), on obtient

$$M = \frac{4}{3} \pi \mu_0 r_{surf0}^3.$$

Ce résultat n'est pas aussi anodin qu'il en a l'air. D'abord, parce que le volume propre de la sphère n'est pas donné par la formule euclidienne ; ensuite, parce que la masse n'est pas additive (les poussières sont en interaction gravitationnelle).

Compte tenu du résultat (9.97), il s'agit de l'événement

$$r^* = \frac{r_{surf0}}{2} [1 + \cos(\eta_0 + \chi_{surf})] \quad (9.101)$$

$$c\tau^* = \frac{R_0}{2} [\eta_0 + \chi_{surf} + \sin(\eta_0 + \chi_{surf})]. \quad (9.102)$$

Par définition, les photons critiques sont ceux pour lesquels $r^* = r_S$. Donc, vu la valeur de χ_{surf} donnée en (9.98) :

$$\eta_{crit} = \arccos\left(\frac{2r_S}{r_{surf0}} - 1\right) - \arcsin\sqrt{\frac{r_S}{r_{surf0}}}. \quad (9.103)$$

La figure 9.14 donne quelques exemples de trajectoires intérieures.

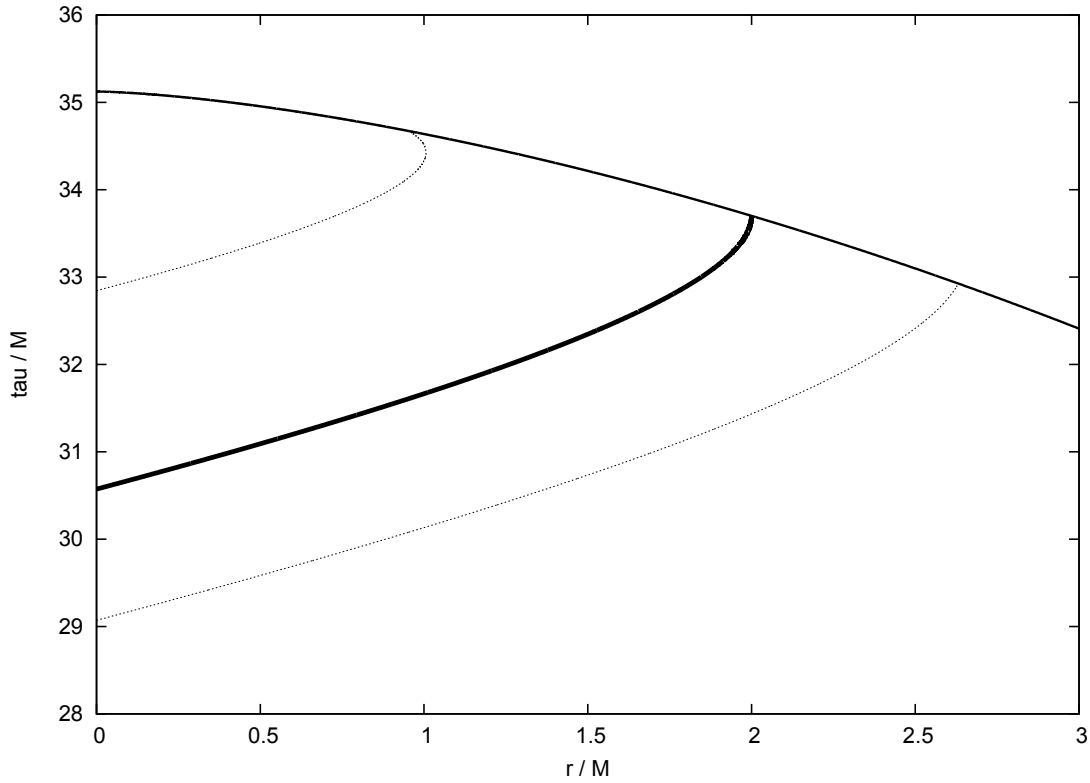


FIGURE 9.14: Formation de l'horizon à l'intérieur d'une sphère de poussière en effondrement gravitationnel. Unités $c = 1$, $G = 1$. En abscisse on a porté r/M , où r est donné par (9.96). L'ordonnée est τ/M , où τ exprime le temps propre « cosmologique » (c'est-à-dire celui d'une particule comobile). Le rayon initial est donné par $r_{surf0} = 10M$, comme dans les figures (9.11) et (9.13). La ligne supérieure représente la surface. Toutes les poussières atteignent simultanément la singularité en $\tau_c = 35.12 M$. L'horizon est tracé en ligne continue grasse. Par (9.103) on calcule $\eta_{crit} = 1.75$, correspondant à $\tau_{crit} = 30.57 M$ en vertu de (9.100). Pour les photons critiques on a $r^* = r_S = 2M$. A la surface la tangente à leur trajectoire $\tau(r)$ est verticale. De part et d'autre, on a également représenté un photon piégé sous l'horizon ($\eta_0 = 2.05$, $\tau_0 = 32.84 M$) et un photon libre ($\eta_0 = 1.60$, $\tau_0 = 29.07 M$). Toutes les trajectoires ont été calculées à partir des équations (9.99) et (9.100).

Dans les coordonnées extérieures de Schwarzschild, l'événement $(c\tau^*, r^*)$ s'exprime sous la forme

$$[t^* = t(\eta^*), r^*]$$

où la fonction $t(\eta)$ est donnée par (9.77). Le mouvement extérieur du photon est ainsi régi par l'équation

$$dr = \pm \left(1 - \frac{2GM}{rc^2}\right) cdt$$

avec la condition initiale $r(t^*) = r^*$. Le signe $-$ distingue un photon qui se déplace vers la gauche dans le diagramme de Kruskal (voir figure 9.13); et le signe $+$ indique un photon progressant vers la droite (c'est le cas qui nous intéresse ici, car la lumière *sort* de l'étoile). En utilisant le changement de variables $cdt = r d\eta$, on peut écrire la solution explicitement sous forme paramétrique :

$$r(\eta) = (r^* - r_S) \exp(\eta - \eta_0 - \chi_{surf}) + r_S \tag{9.104}$$

$$ct(\eta) = ct^* + (r^* - r_S) [\exp(\eta - \eta_0 - \chi_{surf}) - 1] + r_S (\eta - \eta_0 - \chi_{surf}). \tag{9.105}$$

Les trajectoires extérieures sont illustrées en figure 9.15.

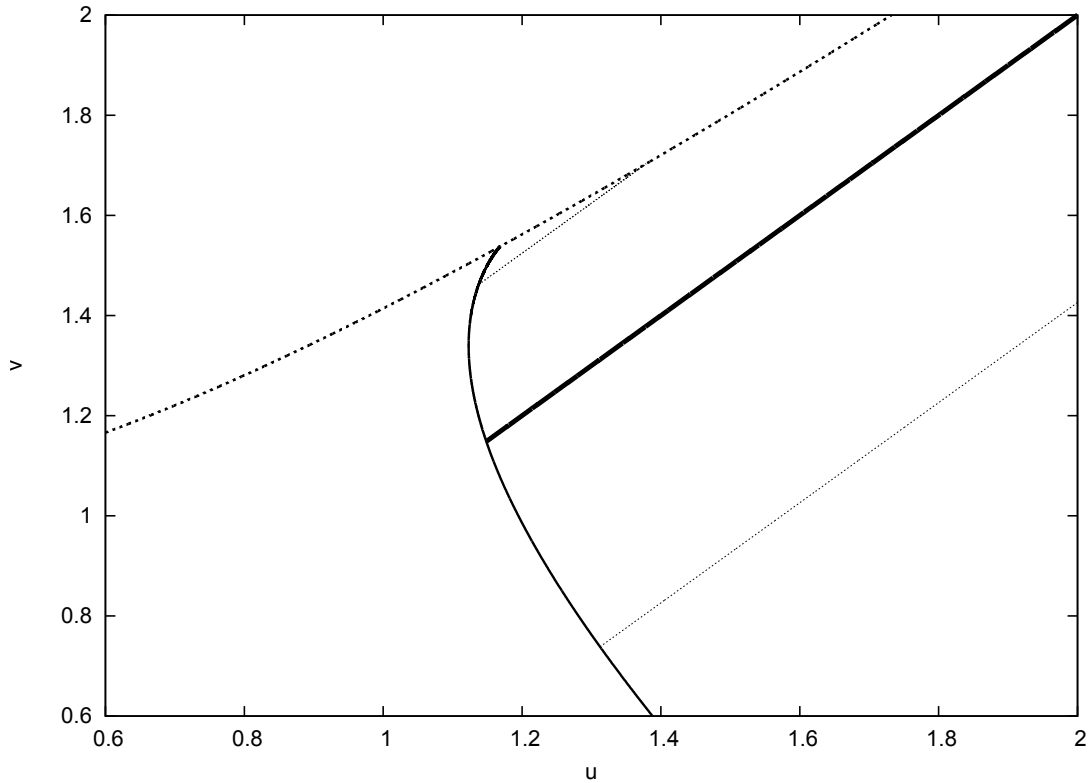


FIGURE 9.15: « Suite » de la figure 9.14, à l'extérieur de l'étoile. La ligne du dessus représente la singularité. Les trajectoires des photons ont été dessinées à l'aide des équations (9.104) et (9.105) converties dans les coordonnées (u, v) de Kruskal. Bien entendu, les photons critiques se maintiennent sur l'horizon $v = u$ (en gras). Afin d'assurer une meilleure lisibilité, lors du calcul des temps t^* on a choisi $K = 40.0$ dans l'équation (9.77).

Annexe A

L'électromagnétisme en Système International

A.1 Notations et équations fondamentales

Nous utiliserons le système d'unités internationales et des coordonnées cartésiennes $(x^\alpha) = (ct, x, y, z)$ dans l'espace-temps de Minkowski. Soit ϵ_0 la permittivité diélectrique du vide et μ_0 sa perméabilité magnétique. La vitesse de la lumière s'écrit

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}.$$

Soit Φ le potentiel scalaire et \vec{A} le potentiel vecteur. Ces quantités permettent de calculer le champ électrique \vec{E} et le champ magnétique \vec{B} via les relations

$$\vec{E} = -\vec{\nabla}\Phi - \frac{\partial \vec{A}}{\partial t}, \quad \vec{B} = \vec{\nabla} \wedge \vec{A}.$$

Notons $U^\mu = dx^\mu/d\tau$ la quadrivitesse d'un élément de fluide chargé, et ρ_0 sa densité de charges dans le référentiel propre. On définit un quadrivecteur « densité de courant » par la relation

$$j^\mu = \rho_0 U^\mu.$$

Dans un référentiel inertiel quelconque, soit ρ la densité de charges, \vec{V} la vitesse du courant de charges et $\vec{j} = \rho \vec{V}$ la densité de courant habituelle dans l'espace à trois dimensions. Ces quantités s'associent pour former le quadrivecteur densité de courant :

$$(j^\mu) = (c\rho, \vec{j}).$$

Dans la jauge [dite « de Lorenz », voir plus loin la définition (A.13)] telle que

$$\frac{1}{c^2} \frac{\partial \Phi}{\partial t} + \vec{\nabla} \cdot \vec{A} = 0,$$

les équations de Maxwell s'écrivent

$$\square \frac{\Phi}{c} = -\mu_0 c\rho, \quad \square \vec{A} = -\mu_0 \vec{j}.$$

Comme le d'alembertien est invariant sous la transformation de Lorentz et que les membres de droite constituent le quadrivecteur $-\mu_0 j^\mu$, on en déduit que Φ/c et \vec{A} s'associent pour constituer un quadrivecteur « potentiel vecteur » de composantes

$$(A^\mu) = \left(\frac{\Phi}{c}, \vec{A} \right). \quad (\text{A.1})$$

Le tenseur de Faraday a pour expression

$$(F_{\alpha\beta}) = (\partial_\alpha A_\beta - \partial_\beta A_\alpha) = \begin{pmatrix} 0 & -E_1/c & -E_2/c & -E_3/c \\ E_1/c & 0 & B_3 & -B_2 \\ E_2/c & -B_3 & 0 & B_1 \\ E_3/c & B_2 & -B_1 & 0 \end{pmatrix} \quad (\text{A.2})$$

et satisfait aux quatre relations (première paire des équations de Maxwell)

$$\partial_\nu F_{\rho\xi} + \partial_\rho F_{\xi\nu} + \partial_\xi F_{\nu\rho} = 0. \quad (\text{A.3})$$

On a la seconde paire des équations de Maxwell

$$\partial_\nu F^{\mu\nu} = \mu_0 j^\mu. \quad (\text{A.4})$$

L'équation de conservation de la charge en découle automatiquement, compte tenu de l'antisymétrie du tenseur de Faraday et de la symétrie des dérivées partielles :

$$\partial_\mu j^\mu = 0. \quad (\text{A.5})$$

Enfin, les équations de mouvement d'une particule de charge q s'écrivent

$$m_0 \frac{dU^\mu}{d\tau} = K^\mu \quad (\text{A.6})$$

avec

$$K^\mu = q F^\mu{}_\beta U^\beta. \quad (\text{A.7})$$

Leurs composantes temporelle et spatiales prennent respectivement la forme

$$\frac{d}{dt}(mc^2) = q \vec{V} \cdot \vec{E} \quad (\text{A.8})$$

$$\frac{d}{dt}(m\vec{V}) = q(\vec{E} + \vec{V} \wedge \vec{B}). \quad (\text{A.9})$$

A.2 L'invariance de jauge

Les champs \vec{E} et \vec{B} sont invariants pour la transformation

$$\Phi^{new} = \Phi^{old} - \frac{\partial f}{\partial t} \quad (\text{A.10})$$

$$\vec{A}^{new} = \vec{A}^{old} + \vec{\nabla} f. \quad (\text{A.11})$$

De la définition (A.1), il résulte

$$(A^{new\ \mu}) = \left(\frac{\Phi^{old}}{c} - \frac{1}{c} \frac{\partial f}{\partial t}, \vec{A}^{old} + \vec{\nabla} f \right),$$

c'est-à-dire

$$A^{new}_{\mu} = A^{old}_{\mu} + \partial_{\mu} f. \quad (\text{A.12})$$

La fonction arbitraire f constitue un degré de liberté permettant d'imposer une contrainte mathématique aux équations de l'électromagnétisme, sans pour autant modifier la situation physique qu'elles décrivent : c'est l'*invariance de jauge*. On choisit souvent la *jauge de Lorenz*, compatible avec l'invariance relativiste :

$$\partial_{\mu} A^{new \mu} = 0. \quad (\text{A.13})$$

Il est facile de trouver la fonction f permettant de travailler dans cette jauge :

$$\begin{aligned} 0 &= \partial_{\mu} A^{new \mu} \\ &= \partial^{\mu} A^{new}_{\mu} \\ &= \partial^{\mu} (A^{old}_{\mu} + \partial_{\mu} f) \\ &= \partial_{\mu} A^{old \mu} + \square f; \end{aligned}$$

elle doit donc vérifier

$$\square f = -\partial_{\mu} A^{old \mu}. \quad (\text{A.14})$$

L'invariance de jauge est liée à une caractéristique mathématique fondamentale des équations de Maxwell (A.4). Pour comprendre cela, écrivons ces dernières en fonction du potentiel vecteur, via la définition (A.2) :

$$\begin{aligned} \mu_0 j^{\mu} &= \eta^{\mu\alpha} \eta^{\nu\beta} \partial_{\nu} (\partial_{\alpha} A_{\beta} - \partial_{\beta} A_{\alpha}) \\ &= \partial^{\mu} \partial_{\nu} A^{\nu} - \partial^{\nu} \partial_{\nu} A^{\mu}. \end{aligned} \quad (\text{A.15})$$

Imaginons que l'on cherche à déterminer le potentiel vecteur en tout point de l'espace-temps, connaissant les inconnues et leurs dérivées temporelles premières sur une hypersurface spatiale correspondant à un temps initial $t = 0$. Sans entrer dans les arcanes de l'analyse (problème de Cauchy pour les équations aux dérivées partielles), tâchons de discuter sommairement dans quelle mesure les conditions initiales garantissent l'unicité de la solution du système (A.15).

Les équations spatiales sont des équations d'évolution pour les quantités A^i , du second ordre en leurs dérivées temporelles. Mais il s'avère que l'équation temporelle

$$\partial^0 \partial_i A^i - \partial^i \partial_i A^0 = \mu_0 j^0 \quad (\text{A.16})$$

ne fait pas intervenir de dérivées temporelles pour A^0 . Il s'agit, en fait, de la première équation de Maxwell, que l'on peut écrire sous la forme

$$\vec{\nabla} \cdot \left(-\vec{\nabla} \Phi - \frac{\partial \vec{A}}{\partial t} \right) = \frac{\rho}{\epsilon_0}.$$

Elle jouit d'une propriété remarquable : si elle est vérifiée à l'instant initial $t = 0$, alors elle est vérifiée à tout instant, comme conséquence des équations spatiales. En effet, l'identité

$$\partial_{\mu} \partial_{\nu} (\partial^{\mu} A^{\nu} - \partial^{\nu} A^{\mu}) = 0$$

implique

$$\partial_0 (\partial_{\nu} \partial^0 A^{\nu} - \partial_{\nu} \partial^{\nu} A^0) + \partial_i (\partial_{\nu} \partial^i A^{\nu} - \partial_{\nu} \partial^{\nu} A^i) = 0.$$

Or, d'après (A.15), le second terme du membre de gauche est égal à $\mu_0 \partial_i j^i$, lui-même égal à $-\mu_0 \partial_0 j^0$ vu l'équation (A.5) de conservation de la charge. Par conséquent,

$$\partial_\nu \partial^0 A^\nu - \partial_\nu \partial^\nu A^0 - \mu_0 j^0 = \text{constante}, \quad (\text{A.17})$$

c'est-à-dire

$$\partial^0 \partial_i A^i - \partial^i \partial_i A^0 - \mu_0 j^0 = \text{constante}.$$

L'équation (A.16) se réduit donc à une *contrainte sur les conditions initiales*. Mais si elle est satisfaite, le système (A.15) admet alors une infinité de solutions pour un ensemble de conditions initiales données, puisque rien ne permet de calculer A^0 pour $t > 0$! Cette conclusion paraît inadmissible. Heureusement, un théorème permet de redresser la situation.

Théorème A.1 *Les différentes solutions mathématiques de l'équation (A.15)*

$$\partial^\mu \partial_\nu A^{old \nu} - \partial^\nu \partial_\nu A^{old \mu} = \mu_0 j^\mu \quad (\text{A.18})$$

avec les conditions initiales de Cauchy

$$\left(A^{old \mu} \right)_{t=0}, \quad \left(\frac{\partial A^{old \mu}}{\partial t} \right)_{t=0} \quad (\text{A.19})$$

soumises à la contrainte (A.16), correspondent en fait à la même réalité physique, c'est-à-dire ne diffèrent entre elles que par un changement de jauge du type (A.12). Plus précisément, toute configuration électromagnétique peut être décrite par les équations

$$\partial^\nu \partial_\nu A^{new \mu} = -\mu_0 j^\mu \quad (\text{A.20})$$

avec une contrainte qui s'écrit

$$\partial_\mu A^{new \mu} = 0 \quad (\text{A.21})$$

et des conditions initiales de Cauchy pour A^{new} déduites de (A.19).

Démonstration. Remarquons d'abord que la condition initiale

$$(\partial_\mu A^{new \mu})_{t=0} = 0 \quad (\text{A.22})$$

peut toujours être satisfaite car elle exprime le choix de la jauge de Lorenz sur l'hypersurface initiale.¹ D'autre part, la contrainte (A.16)

$$\left(\partial^0 \partial_\nu A^{old \nu} - \partial^\nu \partial_\nu A^{old 0} \right)_{t=0} = \mu_0 \left(j^0 \right)_{t=0}$$

1. On trouve une procédure particulièrement judicieuse dans R. Wald, *General Relativity*, The University of Chicago Press, 1984, p. 254. Si \vec{E} et \vec{B} désignent respectivement le champ électrique et le champ magnétique en $t = 0$, alors sur cette hypersurface on peut imposer

$$\Phi^{new} = 0$$

et choisir pour \vec{A}^{new} n'importe quelle solution de

$$\vec{\nabla} \wedge \vec{A}^{new} = \vec{B}.$$

Il s'ensuit les conditions initiales sur les dérivées premières : en $t = 0$ on a

$$\partial \vec{A}^{new} / \partial t = -\vec{E} \quad \text{et} \quad \partial \Phi^{new} / \partial t = -c^2 \vec{\nabla} \cdot \vec{A}^{new}.$$

Dans le langage de la relation (A.12), la condition sur Φ^{new} impose $(\partial_0 f)(0) = -A^{old}_0(0)$. Par ailleurs, on peut toujours supposer que f est constante en $t = 0$ [c'est-à-dire que $\vec{A}^{new}(0) = \vec{A}^{old}(0)$]. Ces conditions initiales déterminent univoquement la fonction $f(\vec{x})$ solution de (A.14).

est équivalente à

$$\left(\partial^0 \partial_\nu A^{new\nu} - \partial^\nu \partial_\nu A^{new0}\right)_{t=0} = \mu_0 (j^0)_{t=0}.$$

Or, l'équation (A.20) impose

$$\left(\partial^\nu \partial_\nu A^{new0}\right)_{t=0} = -\mu_0 (j^0)_{t=0}.$$

Il en résulte la condition initiale

$$(\partial_0 \partial_\mu A^{new\mu})_{t=0} = 0. \quad (\text{A.23})$$

Cela étant, d'après l'équation (A.20) on a

$$\square (\partial_\mu A^{new\mu}) = \partial_\mu \square A^{new\mu} = -\mu_0 \partial_\mu j^\mu = 0. \quad (\text{A.24})$$

La solution unique de cette équation en la variable $\partial_\mu A^{new\mu}$, pour les conditions initiales (A.22) et (A.23), est $\partial_\mu A^{new\mu} = 0$. Cela signifie que, si la condition de jauge est vérifiée en $t = 0$, les équations d'évolution (A.20) assurent qu'elle est vérifiée $\forall t$: elle a valeur de contrainte. A ce titre, (A.20) est équivalente à (A.18).

En résumé, les équations (A.18) pour $A^{old\mu}$, avec les conditions initiales (A.19) soumises à la contrainte (A.16), sont physiquement équivalentes aux équations (A.20) pour $A^{new\mu}$, avec de nouvelles conditions initiales soumises à la contrainte de jauge de Lorenz (A.21). Le problème mathématique se trouve alors convenablement posé car les équations (A.20) admettent une solution *unique* pour un ensemble de conditions initiales données. Deux solutions différentes $A^{old\mu}$, associées à des conditions initiales identiques, sont ainsi physiquement équivalentes à la solution unique $A^{new\mu}$ déduite des nouvelles conditions initiales. L'électromagnétisme est bien une théorie causale, régie par l'équation

$$\square A^\mu = -\mu_0 j^\mu \quad (\text{A.25})$$

soumise à la condition de jauge (A.21) ; ce qu'il fallait démontrer. \square

Une solution particulière de (A.25) est donnée par la théorie des potentiels retardés² :

$$A^\mu(t, \vec{x}) = \frac{\mu_0}{4\pi} \int \frac{j^\mu \left(t' = t - \frac{|\vec{x} - \vec{x}'|}{c}, \vec{x}' \right)}{|\vec{x} - \vec{x}'|} d\vec{x}'. \quad (\text{A.26})$$

On peut vérifier qu'elle satisfait la condition de jauge, comme il se doit. Si les sources sont immobiles, alors

$$\begin{aligned} j^\mu(t', \vec{x}') &= j^\mu(\vec{x}') \\ A^\mu(\vec{x}) &= \frac{\mu_0}{4\pi} \int \frac{j^\mu(\vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}'. \end{aligned} \quad (\text{A.27})$$

Comme les dérivées temporelles s'annulent, l'opérateur de d'Alembert dans l'équation (A.25) se réduit au laplacien :

$$\nabla^2 A^\mu = -\mu_0 j^\mu, \quad (\text{A.28})$$

et on retrouve les équations de base de l'électrostatique et de la magnétostatique

$$\nabla^2 \Phi = -\frac{\rho}{\varepsilon_0}, \quad \nabla^2 \vec{A} = -\mu_0 \vec{j}.$$

2. On obtient la solution complète en y ajoutant la solution générale de l'équation homogène et en imposant les conditions aux limites correspondant à la situation physique envisagée. Concrètement, la solution homogène décrit le champ de rayonnement extérieur dans lequel baigne le système, et peut être omise en l'absence d'un tel champ. Voir par exemple L. Landau et E. Lifchitz, *Physique théorique, Tome 2, Théorie des champs*, Mir, 1989, § 62.

A.3 Solutions quasi-statiques des équations de Maxwell

Il existe des situations physiques réalistes où des approximations raisonnables permettent de simplifier le calcul³ de l'intégrale retardée (A.26), que nous présenterons sous la forme

$$A^\mu(t, \vec{x}) = \frac{\mu_0}{4\pi} \int d\vec{x}' \int \frac{j^\mu(t', \vec{x}')}{|\vec{x} - \vec{x}'|} \delta\left(t' - t + \frac{|\vec{x} - \vec{x}'|}{c}\right) dt',$$

où δ représente la distribution de Dirac.

Supposons, pour fixer les idées, que les sources $j^\mu(t', \vec{x}')$ soient animées d'un *mouvement périodique* de période $T = 2\pi/\omega_0$. On peut alors décomposer ce mouvement en série de Fourier :

$$j^\mu(t', \vec{x}') = \sum_{n=-\infty}^{+\infty} J_n^\mu(\vec{x}') e^{-in\omega_0 t'} \quad (\text{A.29})$$

avec

$$J_n^\mu(\vec{x}') = \frac{1}{T} \int_0^T j^\mu(t', \vec{x}') e^{in\omega_0 t'} dt'. \quad (\text{A.30})$$

On obtient

$$\begin{aligned} A^\mu(t, \vec{x}) &= \frac{\mu_0}{4\pi} \int d\vec{x}' \int \frac{\sum J_n^\mu(\vec{x}') e^{-in\omega_0 t'}}{|\vec{x} - \vec{x}'|} \delta\left(t' - t + \frac{|\vec{x} - \vec{x}'|}{c}\right) dt' \\ &= \frac{\mu_0}{4\pi} \int \frac{\sum J_n^\mu(\vec{x}') e^{ik_n |\vec{x} - \vec{x}'|} e^{-in\omega_0 t}}{|\vec{x} - \vec{x}'|} d\vec{x}', \end{aligned}$$

où l'on a posé

$$k_n = \frac{n\omega_0}{c}.$$

En permutant l'intégrale et la série, on trouve ainsi le développement en série de Fourier⁴ du potentiel vecteur :

$$A^\mu(t, \vec{x}) = \sum_{n=-\infty}^{+\infty} A_n^\mu(\vec{x}) e^{-in\omega_0 t} \quad (\text{A.31})$$

avec

$$A_n^\mu(\vec{x}) = \frac{\mu_0}{4\pi} \int \frac{J_n^\mu(\vec{x}')}{|\vec{x} - \vec{x}'|} e^{ik_n |\vec{x} - \vec{x}'|} d\vec{x}'. \quad (\text{A.32})$$

3. J.D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, 1975, section 9-1 ;

L. Landau et E. Lifchitz, *op. cit.*, §§ 64, 66 ;

C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation*, Freeman, 1973, p. 997.

4. Si le mouvement des sources n'est pas périodique, il faut développer $j^\mu(t', \vec{x}')$ en intégrale de Fourier sur le temps :

$$j^\mu(t', \vec{x}') = \int J^\mu(\omega, \vec{x}') e^{-i\omega t'} d\omega.$$

On obtient alors

$$A^\mu(t, \vec{x}) = \int A^\mu(\omega, \vec{x}) e^{-i\omega t} d\omega$$

où l'on a dégagé le terme générique relatif à la fréquence angulaire $\omega = kc$ en posant

$$A^\mu(\omega, \vec{x}) = \frac{\mu_0}{4\pi} \int \frac{J^\mu(\omega, \vec{x}') \exp(ik |\vec{x} - \vec{x}'|)}{|\vec{x} - \vec{x}'|} d\vec{x}'.$$

Compte tenu de (A.30), on peut récrire (A.32) sous la forme

$$A_n^\mu(\vec{x}) = \frac{\mu_0}{4\pi T} \int \frac{e^{ik_n|\vec{x}-\vec{x}'|}}{|\vec{x}-\vec{x}'|} d\vec{x}' \int_0^T j^\mu(t', \vec{x}') e^{in\omega_0 t'} dt'. \quad (\text{A.33})$$

Si la source est constituée d'une seule particule de charge q , trajectoire $\vec{x}'(t')$ et vitesse $\vec{V}(t')$, alors

$$j^\mu(t', \vec{x}') = \mathcal{J}^\mu(t') \delta[\vec{x}' - \vec{x}'(t')], \quad (\text{A.34})$$

où l'on a posé

$$\mathcal{J}^0 = qc, \quad \mathcal{J}^i = qV^i.$$

En permutant les intégrales dans (A.33), on en déduit l'expression du potentiel vecteur généré en tout point de l'espace-temps par une charge ponctuelle animée d'un mouvement périodique de période T :

$$A^\mu(t, \vec{x}) = \sum_{n=-\infty}^{+\infty} A_n^\mu(\vec{x}) e^{-in\omega_0 t} = \sum_{n=-\infty}^{+\infty} \frac{\mu_0 e^{-in\omega_0 t}}{4\pi T} \int_0^T \frac{\mathcal{J}^\mu(t') e^{in\omega_0 t'}}{|\vec{x} - \vec{x}'(t')|} e^{ik_n|\vec{x}-\vec{x}'(t')|} dt'. \quad (\text{A.35})$$

Le champ quasi-statique.

Revenons à l'intégrale (A.32) et plaçons-nous en un point \vec{x} donné. Si, pour chaque élément de source situé en \vec{x}' on a

$$|\vec{x} - \vec{x}'| \ll \frac{cT}{n}, \quad (\text{A.36})$$

c'est-à-dire si la distance $|\vec{x} - \vec{x}'|$ entre l'observateur et chaque élément de source est beaucoup plus petite que la longueur d'onde $\lambda_n = 2\pi/k_n = cT/n$ du mode n , alors on peut négliger le terme $\exp(ik_n|\vec{x} - \vec{x}'|)$.

Or, il faut un temps $|\vec{x} - \vec{x}'|/c$ pour que parvienne jusqu'à l'observateur l'information relative à la position d'un élément de source. La relation (A.36) exige que ce délai soit beaucoup plus court que la période T/n du mouvement harmonique d'ordre n : cela revient donc à supposer que la source s'est déplacée d'une fraction négligeable de cycle pendant le transfert de l'information.

Dans ces conditions, les effets dus au retard sont négligeables et l'on peut écrire

$$A_n^\mu(\vec{x}) \approx \frac{\mu_0}{4\pi} \int \frac{J_n^\mu(\vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}'.$$

Cette expression rappelle la solution statique (A.27). Supposons dès lors que l'inégalité (A.36) se trouve satisfaite en tout point \vec{x} d'un domaine contenant les sources.⁵ Compte tenu de (A.28), on a ainsi

$$\nabla^2 A_n^\mu(\vec{x}) \approx -\mu_0 J_n^\mu(\vec{x}). \quad (\text{A.37})$$

5. Cette hypothèse a pour objectif de justifier les relations (A.37) et (A.38) dans toute la région occupée par les sources. Si l'on note d la dimension caractéristique de cette région, elle entraîne $d \ll cT/n$. Il s'ensuit que le mouvement des sources est non relativiste. En effet, la vitesse d'une particule animée d'un mouvement harmonique de période T et d'amplitude a est donnée par $V \sim a/T$. Or, $a \ll cT$ puisque $a \leq d \ll cT/n$, donc $V \ll c$.

Or, le terme n du développement en série de Fourier du potentiel vecteur est égal à $A_n^\mu(\vec{x}) \exp(-in\omega_0 t)$. En multipliant (A.37) par $\exp(-in\omega_0 t)$ et en sommant sur n , on trouve⁶

$$\nabla^2 A^\mu(t, \vec{x}) \approx -\mu_0 j^\mu(t, \vec{x}). \quad (\text{A.38})$$

Ainsi, l'opérateur de d'Alembert (A.25) se réduit au laplacien : les sources peuvent se déplacer mais l'observateur subit leur influence « en direct », comme si la vitesse de propagation de l'information était infinie. On dit que le champ est *quasi-statique* dans le domaine (A.36) considéré.

Calcul du champ à grande distance des sources.

A grande distance des sources, il est possible de préciser l'impact du terme retardé $e^{ik_n|\vec{x}-\vec{x}'|}$ dans l'intégrale (A.32).

Fixons l'origine O des coordonnées en un point quelconque *situé à l'intérieur des sources*, appelons d la dimension caractéristique de ces dernières et posons $r = |\vec{x}|$, $\vec{n} = \vec{x}/r$. On a (figure A.1)

$$\begin{aligned} |\vec{x} - \vec{x}'| &= \sqrt{|\vec{x}|^2 - 2\vec{x} \cdot \vec{x}' + |\vec{x}'|^2} \\ &= r \sqrt{1 - \frac{2}{r} \vec{n} \cdot \vec{x}' + \frac{|\vec{x}'|^2}{r^2}}. \end{aligned} \quad (\text{A.39})$$

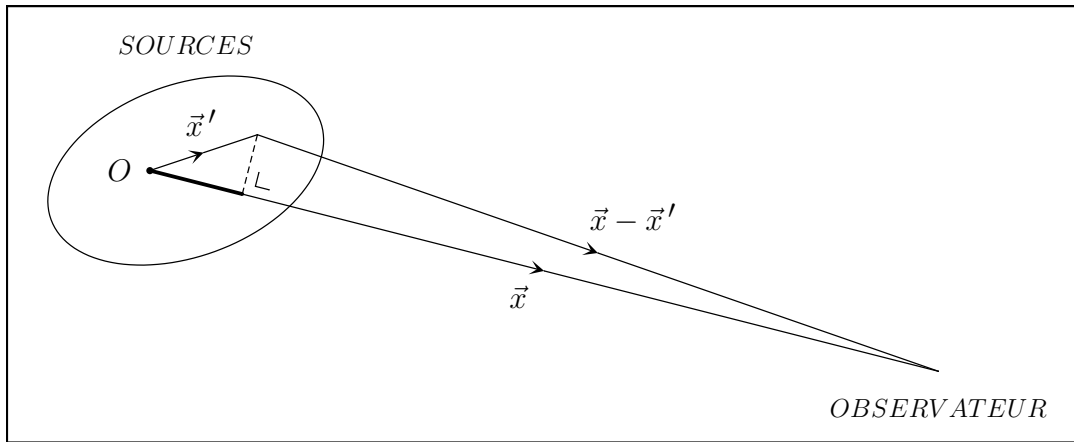


FIGURE A.1: Calcul approché de l'intégrale (A.32). En gras, la projection orthogonale de \vec{x}' sur \vec{x} , correspondant au vecteur $(\vec{n} \cdot \vec{x}') \vec{n}$.

Si l'observateur se trouve loin des sources, *i.e.* si

$$d \ll r, \quad (\text{A.40})$$

alors

$$\begin{aligned} |\vec{x} - \vec{x}'| &= r \left[1 - \frac{\vec{n} \cdot \vec{x}'}{r} + O\left(\frac{d^2}{r^2}\right) \right] \\ &\approx r - \vec{n} \cdot \vec{x}'. \end{aligned} \quad (\text{A.41})$$

6. Bien entendu, il est impossible de garantir la relation (A.36) pour des n arbitrairement élevés. Mais, en pratique, le développement en série de Fourier peut être limité à un nombre fini de termes.

Par conséquent,

$$e^{ik_n|\vec{x}-\vec{x}'|} \approx e^{ik_nr} e^{-ik_n\vec{n}\cdot\vec{x}'}. \quad (\text{A.42})$$

Le premier terme du membre de droite engendre le rayonnement généralement dominant, appelé *dipolaire électrique* ; le second détermine les rayonnements *quadrupolaire électrique* et *dipolaire magnétique*.

En injectant (A.41) et (A.42) dans (A.32), on obtient

$$A_n^\mu(\vec{x}) \approx \frac{\mu_0}{4\pi r} \left[e^{ik_nr} \int J_n^\mu(\vec{x}') e^{-ik_n\vec{n}\cdot\vec{x}'} d\vec{x}' + O\left(\frac{d}{r}\right) \right]. \quad (\text{A.43})$$

On y voit clairement apparaître le terme oscillant $\exp(ik_nr)$, de période $\lambda_n = cT/n$.

Si $k_nr \ll 1$, c'est-à-dire si

$$r \ll \lambda_n, \quad (\text{A.44})$$

alors $\exp(ik_nr) \approx 1$. Vu (A.40), cette hypothèse implique

$$d \ll \lambda_n, \quad (\text{A.45})$$

équivalente à $|k_n\vec{n}\cdot\vec{x}'| \ll 1$. La relation (A.45) exprime le fait que le mouvement des sources n'est pas relativiste, comme on l'a signalé dans la note 5.

En résumé, si $d \ll r \ll \lambda_n$, tous les termes retardés peuvent être négligés dans le calcul de $A_n^\mu(\vec{x})$: on dit que l'on se trouve dans la zone *quasi-statique*⁷ et on a

$$A_n^\mu(\vec{x}) \approx \frac{\mu_0}{4\pi r} \int J_n^\mu(\vec{x}') d\vec{x}', \quad (\text{A.46})$$

impliquant $\nabla^2 A_n^\mu(\vec{x}) \approx 0$, donc $\nabla^2 A^\mu(t, \vec{x}) \approx 0$.

On constate que

$$\left| \frac{\partial [A_n^\mu(\vec{x}) e^{-in\omega_0 t}]}{\partial t} \right| \sim n\omega_0 |A_n^\mu(\vec{x}) e^{-in\omega_0 t}|$$

$$\left| \frac{\partial [A_n^\mu(\vec{x}) e^{-in\omega_0 t}]}{\partial x^i} \right| \sim \frac{1}{r} |A_n^\mu(\vec{x}) e^{-in\omega_0 t}|$$

et, par conséquent,

$$\left| \frac{1}{c} \frac{\partial [A_n^\mu(\vec{x}) e^{-in\omega_0 t}]}{\partial t} \right| \ll \left| \frac{\partial [A_n^\mu(\vec{x}) e^{-in\omega_0 t}]}{\partial x^i} \right| \quad (\text{A.47})$$

car $r \ll \lambda_n$ est équivalent à $n\omega_0/c \ll 1/r$. On peut ainsi négliger les dérivées temporelles devant les dérivées spatiales pourvu que la longueur d'onde soit grande par rapport à la distance d'observation : cette approximation n'est donc pas correcte aux trop hautes fréquences.

7. Plus généralement, on distingue trois régions spatiales :

- la zone *quasi-statique* : $d \ll r \ll \lambda_n$;
- la zone *d'induction* : $d \ll r \sim \lambda_n$;
- la zone *radiative* : $d \ll \lambda_n \ll r$.

Annexe B

Les coordonnées normales

Théorème B.1 (Coordonnées normales de Riemann) *En un point donné P_0 d'une variété différentielle affine sans torsion, il est toujours possible de choisir un système de coordonnées dans lequel s'annulent tous les symboles de connexion affine.*

La démonstration est constructive.

Soit en P_0 une base de n vecteurs linéairement indépendants $\{\vec{e}_\alpha\}$, $\alpha = 1, \dots, n$. Considérons les géodésiques émanant du point P_0 . Appelons λ un paramètre affine sur chacune d'elles, d'origine $\lambda = 0$ située en P_0 , et \vec{v} leur vecteur tangent en P_0 :

$$\begin{aligned}\vec{v} &= \frac{d}{d\lambda} \\ &= v^\alpha \vec{e}_\alpha.\end{aligned}\tag{B.1}$$

Un point P donné suffisamment proche de P_0 se trouve sur une de ces géodésiques, et une seule car celles-ci « n'ont pas encore eu le temps de se croiser ». Cette géodésique est associée à un certain vecteur \vec{v} , et la connaissance de λ y précise l'emplacement de P . On définit les *coordonnées normales de Riemann* de P par les n nombres

$$x^\alpha = \lambda v^\alpha,\tag{B.2}$$

indépendants du paramétrage affine arbitraire choisi sur la géodésique.¹

Ecrivons alors l'équation d'une quelconque de ces géodésiques :

$$\frac{d^2 x^\alpha}{d\lambda^2} + \Gamma_{\beta\gamma}^\alpha(P) \frac{dx^\beta}{d\lambda} \frac{dx^\gamma}{d\lambda} = 0.$$

1. En effet, si $\lambda' = r\lambda$ ($r \in \mathbb{R}_0$), on a

$$\lambda \frac{d}{d\lambda} = \lambda' \frac{d}{d\lambda'}.$$

Remarquons incidemment que

$$\begin{aligned}\vec{v} &= \frac{dx^\alpha}{d\lambda} \frac{\partial}{\partial x^\alpha} \\ &= v^\alpha \frac{\partial}{\partial x^\alpha},\end{aligned}\tag{B.3}$$

la deuxième égalité résultant de (B.2). Comme les égalités (B.1) et (B.3) sont vérifiées pour un vecteur \vec{v} arbitraire, on a donc, à l'origine,

$$\vec{e}_\alpha = \frac{\partial}{\partial x^\alpha}.$$

Il se trouve que le premier terme est nul parce que x^α est linéaire en λ . Ainsi,

$$\Gamma_{\beta\gamma}^\alpha(P) v^\beta v^\gamma = 0.$$

Au point P_0 où toutes les géodésiques se croisent, on a donc

$$\Gamma_{\beta\gamma}^\alpha(P_0) v^\beta v^\gamma = 0 \quad \forall \vec{v}.$$

Comme les symboles de connexion affine sont symétriques, on en déduit

$$\Gamma_{\beta\gamma}^\alpha(P_0) = 0 \quad \forall \alpha, \beta, \gamma. \quad (\text{B.4})$$

□

Le théorème 5.4 généralise le théorème B.1.

Théorème B.2 *Dans les coordonnées normales de Riemann on peut expliciter les dérivées premières des symboles de connexion affine en fonction du tenseur de courbure :*

$$\Gamma_{\beta\gamma,\alpha}^\mu(P_0) = -\frac{1}{3} (R^\mu_{\beta\gamma\alpha} + R^\mu_{\gamma\beta\alpha})(P_0).$$

Preuve. Ecrivons l'équation de déviation géodésique (5.121)

$$\begin{aligned} (\nabla_{\vec{U}} \nabla_{\vec{U}} \vec{W})^\mu &= \mathbf{R}(\tilde{\omega}^\mu, \vec{U}, \vec{U}, \vec{W}) \\ &= R^\mu_{\beta\gamma\rho} U^\beta U^\gamma W^\rho. \end{aligned} \quad (\text{B.5})$$

Désignons par la lettre λ le paramètre affine le long de ces géodésiques et posons

$$\begin{aligned} (\vec{Z})^\mu &= (\nabla_{\vec{U}} \vec{W})^\mu \\ &= \frac{dW^\mu}{d\lambda} + \Gamma_{\beta\rho}^\mu U^\beta W^\rho. \end{aligned}$$

On a

$$\begin{aligned} (\nabla_{\vec{U}} \nabla_{\vec{U}} \vec{W})^\mu &= (U^\beta \nabla_{\beta} \vec{Z})^\mu \\ &= \frac{dZ^\mu}{d\lambda} + \Gamma_{\beta\rho}^\mu U^\beta Z^\rho \\ &= \frac{d^2 W^\mu}{d\lambda^2} + \frac{d}{d\lambda} (\Gamma_{\beta\rho}^\mu U^\beta W^\rho) + \Gamma_{\beta\rho}^\mu U^\beta \frac{dW^\rho}{d\lambda} + \Gamma_{\beta\rho}^\mu \Gamma_{\gamma\sigma}^\rho U^\beta U^\gamma W^\sigma. \end{aligned} \quad (\text{B.6})$$

Cette équation est entièrement générale.

Dans les coordonnées normales de Riemann, considérons toutes les géodésiques émanant de l'origine P_0 . Soit alors \vec{W} le champ vectoriel connectant les points de paramètre λ identique sur deux géodésiques infiniment proches dont les vecteurs tangents diffèrent seulement par la valeur de v^α pour un α donné. Ainsi,

$$\begin{aligned} \vec{W} &= \left(\frac{\partial}{\partial v^\alpha} \right)_\lambda = \left(\frac{\partial x^\beta}{\partial v^\alpha} \right)_\lambda \frac{\partial}{\partial x^\beta} = \lambda \delta_\alpha^\beta \frac{\partial}{\partial x^\beta} \\ \vec{U} &= \left(\frac{\partial}{\partial \lambda} \right)_{v^\alpha} = \left(\frac{\partial x^\beta}{\partial \lambda} \right)_{v^\alpha} \frac{\partial}{\partial x^\beta} = v^\beta \frac{\partial}{\partial x^\beta} \end{aligned}$$

et par conséquent

$$W^\beta = \lambda \delta_\alpha^\beta \quad (\text{B.7})$$

$$U^\beta = v^\beta. \quad (\text{B.8})$$

Comparant (B.5) et (B.6), on obtient, vu (B.7) et (B.8) :

$$0 + v^\beta \delta_\alpha^\rho \frac{d}{d\lambda} (\lambda \Gamma_{\beta\rho}^\mu) + \Gamma_{\beta\rho}^\mu v^\beta \delta_\alpha^\rho + \Gamma_{\beta\rho}^\mu \Gamma_{\gamma\sigma}^\rho v^\beta v^\gamma \lambda \delta_\alpha^\sigma = R_{\beta\gamma\rho}^\mu v^\beta v^\gamma \lambda \delta_\alpha^\rho. \quad (\text{B.9})$$

Or, le long d'une géodésique $\vec{v} = \text{constante}$, on a

$$\Gamma_{\beta\alpha}^\mu = \lambda v^\gamma \Gamma_{\beta\alpha,\gamma}^\mu + O(\lambda^2) \quad (\text{B.10})$$

où les dérivées des symboles de connexion affine sont implicitement calculées à l'origine P_0 .
Donc (B.9) s'écrit successivement

$$\begin{aligned} v^\beta v^\gamma \left[\frac{d}{d\lambda} (\lambda^2 \Gamma_{\beta\alpha,\gamma}^\mu) + \lambda \Gamma_{\beta\alpha,\gamma}^\mu \right] &= R_{\beta\gamma\alpha}^\mu v^\beta v^\gamma \lambda + O(\lambda^2) \\ 3 \Gamma_{\alpha\beta,\gamma}^\mu v^\beta v^\gamma &= R_{\beta\gamma\alpha}^\mu v^\beta v^\gamma + O(\lambda) \\ 3 (\Gamma_{\alpha\beta,\gamma}^\mu + \Gamma_{\alpha\gamma,\beta}^\mu) &= R_{\beta\gamma\alpha}^\mu + R_{\gamma\beta\alpha}^\mu + O(\lambda). \end{aligned} \quad (\text{B.11})$$

Les deux premiers termes du membre de droite de (B.11) valent

$$\Gamma_{\beta\alpha,\gamma}^\mu - \Gamma_{\beta\gamma,\alpha}^\mu + \Gamma_{\gamma\alpha,\beta}^\mu - \Gamma_{\gamma\beta,\alpha}^\mu + O(\lambda^2)$$

ce qui fait que (B.11) se ramène à

$$\Gamma_{\alpha\beta,\gamma}^\mu + \Gamma_{\gamma\alpha,\beta}^\mu + \Gamma_{\beta\gamma,\alpha}^\mu = O(\lambda). \quad (\text{B.12})$$

En utilisant (B.12), le membre de gauche de (B.11) se réduit à $-3\Gamma_{\beta\gamma,\alpha}^\mu + O(\lambda)$, donc (B.11) peut s'écrire

$$-3\Gamma_{\beta\gamma,\alpha}^\mu = R_{\beta\gamma\alpha}^\mu + R_{\gamma\beta\alpha}^\mu + O(\lambda).$$

Considérons ce résultat à l'origine $\lambda = 0$: la thèse est démontrée.

□

Annexe C

Les équations d'Einstein par la méthode d'Euler-Lagrange

Rappelons l'expression (7.24) de la densité lagrangienne effective :

$$L = g^{\rho\sigma} (\Gamma_{\rho\alpha}^{\beta} \Gamma_{\sigma\beta}^{\alpha} - \Gamma_{\rho\sigma}^{\alpha} \Gamma_{\alpha\beta}^{\beta}) \sqrt{-g}.$$

Posant

$$L_1 = g^{\rho\sigma} \Gamma_{\rho\alpha}^{\beta} \Gamma_{\sigma\beta}^{\alpha}, \quad L_2 = -g^{\rho\sigma} \Gamma_{\rho\sigma}^{\alpha} \Gamma_{\alpha\beta}^{\beta},$$

on a

$$dL = (dL_1 + dL_2) \sqrt{-g} + (L_1 + L_2) d\sqrt{-g}. \quad (\text{C.1})$$

• Calcul de dL_1 ¹.

$$dL_1 = dg^{\rho\sigma} \Gamma_{\rho\alpha}^{\beta} \Gamma_{\sigma\beta}^{\alpha} + g^{\rho\sigma} d\Gamma_{\rho\alpha}^{\beta} \Gamma_{\sigma\beta}^{\alpha} + g^{\rho\sigma} \Gamma_{\rho\alpha}^{\beta} d\Gamma_{\sigma\beta}^{\alpha}.$$

Les deuxième et troisième termes du membre de droite sont identiques, donc

$$\begin{aligned} dL_1 &= dg^{\rho\sigma} \Gamma_{\rho\alpha}^{\beta} \Gamma_{\sigma\beta}^{\alpha} + 2g^{\rho\sigma} \Gamma_{\rho\alpha}^{\beta} d\Gamma_{\sigma\beta}^{\alpha} \\ &= -dg^{\rho\sigma} \Gamma_{\rho\alpha}^{\beta} \Gamma_{\sigma\beta}^{\alpha} + 2\Gamma_{\rho\alpha}^{\beta} d(g^{\rho\sigma} \Gamma_{\sigma\beta}^{\alpha}). \end{aligned} \quad (\text{C.2})$$

Développons le deuxième terme :

$$\Gamma_{\rho\alpha}^{\beta} d(g^{\rho\sigma} \Gamma_{\sigma\beta}^{\alpha}) = \frac{1}{2} \Gamma_{\rho\alpha}^{\beta} d[g^{\rho\sigma} g^{\alpha\lambda} (g_{\lambda\sigma,\beta} + g_{\lambda\beta,\sigma} - g_{\sigma\beta,\lambda})].$$

La somme des deux derniers termes étant antisymétrique en λ et σ , l'expression

$$g^{\rho\sigma} g^{\alpha\lambda} (g_{\lambda\beta,\sigma} - g_{\sigma\beta,\lambda})$$

est antisymétrique en ρ et α . Compte tenu de la symétrie de $\Gamma_{\rho\alpha}^{\beta}$ sur ces deux indices, la quantité

$$\Gamma_{\rho\alpha}^{\beta} d[g^{\rho\sigma} g^{\alpha\lambda} (g_{\lambda\beta,\sigma} - g_{\sigma\beta,\lambda})]$$

s'annule et

$$\Gamma_{\rho\alpha}^{\beta} d(g^{\rho\sigma} \Gamma_{\sigma\beta}^{\alpha}) = \frac{1}{2} \Gamma_{\rho\alpha}^{\beta} d(g^{\alpha\lambda} g^{\rho\sigma} g_{\lambda\sigma,\beta}).$$

1. Nous suivons ici A. Einstein, « Die Grundlage der allgemeinen Relativitätstheorie », Annalen der Physik **49** (1916), pp. 804-805.

D'autre part, on a successivement

$$\begin{aligned} g_{\lambda\sigma} g^{\rho\sigma} &= \delta_\lambda^\rho \\ g_{\lambda\sigma} g^{\rho\sigma}{}_{,\beta} &= -g^{\rho\sigma} g_{\lambda\sigma,\beta} \\ g^{\alpha\lambda} g_{\lambda\sigma} g^{\rho\sigma}{}_{,\beta} &= -g^{\alpha\lambda} g^{\rho\sigma} g_{\lambda\sigma,\beta} \\ g^{\alpha\rho}{}_{,\beta} &= -g^{\alpha\lambda} g^{\rho\sigma} g_{\lambda\sigma,\beta}. \end{aligned}$$

Par conséquent,

$$\Gamma_{\rho\alpha}^\beta d(g^{\rho\sigma} \Gamma_{\sigma\beta}^\alpha) = -\frac{1}{2} \Gamma_{\rho\alpha}^\beta dg^{\rho\alpha}{}_{,\beta}.$$

En remplaçant dans (C.2), on obtient

$$\begin{aligned} dL_1 &= -\Gamma_{\rho\alpha}^\beta \Gamma_{\sigma\beta}^\alpha dg^{\rho\sigma} - \Gamma_{\rho\alpha}^\beta dg^{\rho\alpha}{}_{,\beta} \\ &= -\Gamma_{\mu\alpha}^\beta \Gamma_{\nu\beta}^\alpha dg^{\mu\nu} - \Gamma_{\mu\nu}^\lambda dg^{\mu\nu}{}_{,\lambda}. \end{aligned} \quad (\text{C.3})$$

• Calcul de dL_2 .

$$dL_2 = -dg^{\rho\sigma} \Gamma_{\rho\sigma}^\alpha \Gamma_{\alpha\beta}^\beta - g^{\rho\sigma} d\Gamma_{\rho\sigma}^\alpha \Gamma_{\alpha\beta}^\beta - g^{\rho\sigma} \Gamma_{\rho\sigma}^\alpha d\Gamma_{\alpha\beta}^\beta.$$

Or, d'une part,

$$\begin{aligned} g^{\rho\sigma} d\Gamma_{\rho\sigma}^\alpha &= g^{\rho\sigma} d\left[\frac{1}{2} g^{\alpha\xi} (g_{\xi\rho,\sigma} + g_{\xi\sigma,\rho} - g_{\rho\sigma,\xi})\right] \\ &= -\frac{1}{2} g^{\rho\sigma} d(g^{\alpha\xi}{}_{,\sigma} g_{\xi\rho} + g^{\alpha\xi}{}_{,\rho} g_{\xi\sigma}) - \frac{1}{2} g^{\rho\sigma} dg^{\alpha\xi} g_{\rho\sigma,\xi} - \frac{1}{2} g^{\rho\sigma} g^{\alpha\xi} dg_{\rho\sigma,\xi} \\ &= -g^{\rho\sigma} d(g^{\alpha\xi}{}_{,\sigma} g_{\xi\rho}) - \frac{1}{2} g^{\rho\sigma} dg^{\alpha\xi} g_{\rho\sigma,\xi} - \frac{1}{2} g^{\rho\sigma} g^{\alpha\xi} dg_{\rho\sigma,\xi}. \end{aligned}$$

On peut transformer le dernier terme en tenant compte du fait que

$$\begin{aligned} -\frac{1}{2} g^{\rho\sigma} dg_{\rho\sigma,\xi} &= \frac{1}{2} g^{\rho\sigma} d(g_{\rho\theta} g_{\sigma\varphi} g^{\theta\varphi}{}_{,\xi}) \\ &= \frac{1}{2} g^{\rho\sigma} (dg_{\rho\theta} g_{\sigma\varphi} g^{\theta\varphi}{}_{,\xi} + g_{\rho\theta} dg_{\sigma\varphi} g^{\theta\varphi}{}_{,\xi} + g_{\rho\theta} g_{\sigma\varphi} dg^{\theta\varphi}{}_{,\xi}) \\ &= dg_{\rho\theta} g^{\rho\theta}{}_{,\xi} + \frac{1}{2} g_{\rho\theta} dg^{\rho\theta}{}_{,\xi} \\ &= -g_{\rho\varphi} g_{\theta\chi} dg^{\varphi\chi} g^{\rho\theta}{}_{,\xi} + \frac{1}{2} g_{\rho\theta} dg^{\rho\theta}{}_{,\xi}. \end{aligned}$$

Donc

$$-\frac{1}{2} g^{\rho\sigma} dg_{\rho\sigma,\xi} = g_{\varphi\chi,\xi} dg^{\varphi\chi} + \frac{1}{2} g_{\rho\theta} dg^{\rho\theta}{}_{,\xi} \quad (\text{C.4})$$

et

$$\begin{aligned} g^{\rho\sigma} d\Gamma_{\rho\sigma}^\alpha &= -g^{\rho\sigma} d(g^{\alpha\xi}{}_{,\sigma} g_{\xi\rho}) - \frac{1}{2} g^{\rho\sigma} g_{\rho\sigma,\xi} dg^{\alpha\xi} + g_{\varphi\chi,\xi} g^{\alpha\xi} dg^{\varphi\chi} + \frac{1}{2} g^{\alpha\xi} g_{\rho\theta} dg^{\rho\theta}{}_{,\xi} \\ &= -g^{\rho\sigma} dg^{\alpha\xi}{}_{,\sigma} g_{\xi\rho} - g^{\rho\sigma} g^{\alpha\xi}{}_{,\sigma} dg_{\xi\rho} - \frac{1}{2} g^{\rho\sigma} g_{\rho\sigma,\xi} dg^{\alpha\xi} + g_{\varphi\chi,\xi} g^{\alpha\xi} dg^{\varphi\chi} + \frac{1}{2} g^{\alpha\xi} g_{\rho\theta} dg^{\rho\theta}{}_{,\xi} \\ &= -\delta_\xi^\sigma dg^{\alpha\xi}{}_{,\sigma} + g_{\xi\rho} g^{\alpha\xi}{}_{,\sigma} dg^{\rho\sigma} - \frac{1}{2} g^{\rho\sigma} g_{\rho\sigma,\xi} dg^{\alpha\xi} + g_{\varphi\chi,\xi} g^{\alpha\xi} dg^{\varphi\chi} + \frac{1}{2} g^{\alpha\xi} g_{\rho\theta} dg^{\rho\theta}{}_{,\xi}. \end{aligned}$$

Ainsi,

$$\begin{aligned}
 -g^{\rho\sigma} d\Gamma_{\rho\sigma}^{\alpha} \Gamma_{\alpha\beta}^{\beta} &= \delta_{\xi}^{\sigma} \Gamma_{\alpha\beta}^{\beta} dg^{\alpha\xi},_{\sigma} - g_{\xi\rho} g^{\alpha\xi},_{\sigma} \Gamma_{\alpha\beta}^{\beta} dg^{\rho\sigma} + \frac{1}{2} g^{\rho\sigma} g_{\rho\sigma,\xi} \Gamma_{\alpha\beta}^{\beta} dg^{\alpha\xi} \\
 &\quad - g_{\varphi\chi,\xi} g^{\alpha\xi} \Gamma_{\alpha\beta}^{\beta} dg^{\varphi\chi} - \frac{1}{2} g^{\alpha\xi} g_{\rho\theta} \Gamma_{\alpha\beta}^{\beta} dg^{\rho\theta},_{\xi} \\
 &= (\delta_{\nu}^{\lambda} \Gamma_{\mu\beta}^{\beta} - \frac{1}{2} g^{\alpha\lambda} g_{\mu\nu} \Gamma_{\alpha\beta}^{\beta}) dg^{\mu\nu},_{\lambda} \\
 &\quad - (g_{\xi\mu} g^{\alpha\xi},_{\nu} \Gamma_{\alpha\beta}^{\beta} - \frac{1}{2} g^{\rho\sigma} g_{\rho\sigma,\nu} \Gamma_{\mu\beta}^{\beta} + g_{\mu\nu,\xi} g^{\alpha\xi} \Gamma_{\alpha\beta}^{\beta}) dg^{\mu\nu}.
 \end{aligned}$$

D'autre part,

$$\begin{aligned}
 d\Gamma_{\alpha\beta}^{\beta} &= d[\frac{1}{2} g^{\beta\xi} (g_{\xi\alpha,\beta} + g_{\xi\beta,\alpha} - g_{\alpha\beta,\xi})] \\
 &= \frac{1}{2} d(g^{\beta\xi} g_{\beta\xi,\alpha}) \\
 &= \frac{1}{2} dg^{\beta\xi} g_{\beta\xi,\alpha} + \frac{1}{2} g^{\beta\xi} dg_{\beta\xi,\alpha}.
 \end{aligned}$$

On peut transformer le second terme à l'aide de (C.4) pour obtenir

$$d\Gamma_{\alpha\beta}^{\beta} = -\frac{1}{2} g_{\varphi\chi,\alpha} dg^{\varphi\chi} - \frac{1}{2} g_{\beta\xi} dg^{\beta\xi},_{\alpha}$$

donc

$$-g^{\rho\sigma} \Gamma_{\rho\sigma}^{\alpha} d\Gamma_{\alpha\beta}^{\beta} = \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\alpha} g_{\mu\nu,\alpha} dg^{\mu\nu} + \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\lambda} g_{\mu\nu} dg^{\mu\nu},_{\lambda}.$$

Il s'ensuit

$$\begin{aligned}
 dL_2 &= -\Gamma_{\mu\nu}^{\alpha} \Gamma_{\alpha\beta}^{\beta} dg^{\mu\nu} + (\delta_{\nu}^{\lambda} \Gamma_{\mu\beta}^{\beta} - \frac{1}{2} g^{\alpha\lambda} g_{\mu\nu} \Gamma_{\alpha\beta}^{\beta}) dg^{\mu\nu},_{\lambda} \\
 &\quad - (g_{\xi\mu} g^{\alpha\xi},_{\nu} \Gamma_{\alpha\beta}^{\beta} - \frac{1}{2} g^{\rho\sigma} g_{\rho\sigma,\nu} \Gamma_{\mu\beta}^{\beta} + g_{\mu\nu,\xi} g^{\alpha\xi} \Gamma_{\alpha\beta}^{\beta}) dg^{\mu\nu} \\
 &\quad + \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\alpha} g_{\mu\nu,\alpha} dg^{\mu\nu} + \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\lambda} g_{\mu\nu} dg^{\mu\nu},_{\lambda},
 \end{aligned}$$

c'est-à-dire

$$\begin{aligned}
 dL_2 &= (\delta_{\nu}^{\lambda} \Gamma_{\mu\beta}^{\beta} - \frac{1}{2} g^{\alpha\lambda} g_{\mu\nu} \Gamma_{\alpha\beta}^{\beta} + \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\lambda} g_{\mu\nu}) dg^{\mu\nu},_{\lambda} \\
 &\quad - (\Gamma_{\mu\nu}^{\alpha} \Gamma_{\alpha\beta}^{\beta} - g^{\alpha\xi} g_{\xi\mu,\nu} \Gamma_{\alpha\beta}^{\beta} - \frac{1}{2} g^{\rho\sigma} g_{\rho\sigma,\nu} \Gamma_{\mu\beta}^{\beta} + g_{\mu\nu,\xi} g^{\alpha\xi} \Gamma_{\alpha\beta}^{\beta} - \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\alpha} g_{\mu\nu,\alpha}) dg^{\mu\nu}.
 \end{aligned}$$

Si, dans le coefficient de $dg^{\mu\nu}$, on regroupe, d'une part, les deuxième et quatrième termes pour faire apparaître $\Gamma_{\mu\nu}^{\alpha}$; et si, d'autre part, on utilise la relation $\frac{1}{2} g^{\rho\sigma} g_{\rho\sigma,\nu} = \Gamma_{\nu\alpha}^{\alpha}$, on peut écrire

$$\begin{aligned}
 dL_2 &= (\delta_{\nu}^{\lambda} \Gamma_{\mu\beta}^{\beta} - \frac{1}{2} g^{\alpha\lambda} g_{\mu\nu} \Gamma_{\alpha\beta}^{\beta} + \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\lambda} g_{\mu\nu}) dg^{\mu\nu},_{\lambda} \\
 &\quad - (-\Gamma_{\mu\nu}^{\alpha} \Gamma_{\alpha\beta}^{\beta} + g^{\alpha\xi} g_{\xi\nu,\mu} \Gamma_{\alpha\beta}^{\beta} - \Gamma_{\mu\beta}^{\beta} \Gamma_{\nu\alpha}^{\alpha} - \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\alpha} g_{\mu\nu,\alpha}) dg^{\mu\nu}. \quad (C.5)
 \end{aligned}$$

- Calcul de $d\sqrt{-g}$.

$$d\sqrt{-g} = \frac{-dg}{2\sqrt{-g}} = \frac{1}{2} \sqrt{-g} g^{\mu\nu} dg_{\mu\nu} = -\frac{1}{2} \sqrt{-g} g_{\mu\nu} dg^{\mu\nu}. \quad (C.6)$$

• Récapitulation.

Rassemblant les résultats (C.1), (C.3), (C.5) et (C.6), on obtient l'expression de la variation de la densité lagrangienne :

$$dL = (-\Gamma_{\mu\nu}^{\lambda} + \delta_{\nu}^{\lambda} \Gamma_{\mu\beta}^{\beta} - \frac{1}{2} g^{\alpha\lambda} g_{\mu\nu} \Gamma_{\alpha\beta}^{\beta} + \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\lambda} g_{\mu\nu}) \sqrt{-g} dg^{\mu\nu},_{\lambda} - A \sqrt{-g} dg^{\mu\nu}$$

où

$$A = \Gamma_{\mu\alpha}^{\beta} \Gamma_{\nu\beta}^{\alpha} - \Gamma_{\mu\nu}^{\alpha} \Gamma_{\alpha\beta}^{\beta} + g^{\alpha\xi} g_{\xi\nu,\mu} \Gamma_{\alpha\beta}^{\beta} - \Gamma_{\mu\beta}^{\beta} \Gamma_{\nu\alpha}^{\alpha} - \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\alpha} g_{\mu\nu,\alpha} \\ + \frac{1}{2} g_{\mu\nu} g^{\rho\sigma} \Gamma_{\rho\alpha}^{\beta} \Gamma_{\sigma\beta}^{\alpha} - \frac{1}{2} g_{\mu\nu} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\alpha} \Gamma_{\alpha\beta}^{\beta}.$$

Le premier membre des équations d'Euler-Lagrange

$$\frac{\partial}{\partial x^{\lambda}} \frac{\partial L}{\partial (g^{\mu\nu},_{\lambda})} - \frac{\partial L}{\partial g^{\mu\nu}}$$

vaut ainsi

$$\left[(-\Gamma_{\mu\nu}^{\lambda} + \frac{1}{2} \delta_{\nu}^{\lambda} \Gamma_{\mu\beta}^{\beta} + \frac{1}{2} \delta_{\mu}^{\lambda} \Gamma_{\nu\beta}^{\beta} - \frac{1}{2} g^{\alpha\lambda} g_{\mu\nu} \Gamma_{\alpha\beta}^{\beta} + \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\lambda} g_{\mu\nu}) \sqrt{-g} \right]_{,\lambda} + B \sqrt{-g}$$

avec

$$B = \Gamma_{\mu\alpha}^{\beta} \Gamma_{\nu\beta}^{\alpha} - \Gamma_{\mu\nu}^{\alpha} \Gamma_{\alpha\beta}^{\beta} + \frac{1}{2} g^{\alpha\xi} g_{\xi\nu,\mu} \Gamma_{\alpha\beta}^{\beta} + \frac{1}{2} g^{\alpha\xi} g_{\xi\mu,\nu} \Gamma_{\alpha\beta}^{\beta} - \Gamma_{\mu\beta}^{\beta} \Gamma_{\nu\alpha}^{\alpha} - \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\alpha} g_{\mu\nu,\alpha} \\ + \frac{1}{2} g_{\mu\nu} g^{\rho\sigma} \Gamma_{\rho\alpha}^{\beta} \Gamma_{\sigma\beta}^{\alpha} - \frac{1}{2} g_{\mu\nu} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\alpha} \Gamma_{\alpha\beta}^{\beta}.$$

Comme

$$\sqrt{-g},_{\lambda} = \sqrt{-g} \Gamma_{\lambda\xi}^{\xi}, \quad (C.7)$$

ledit membre s'écrit

$$\sqrt{-g} [(-\Gamma_{\mu\nu}^{\lambda} + \frac{1}{2} \delta_{\nu}^{\lambda} \Gamma_{\mu\beta}^{\beta} + \frac{1}{2} \delta_{\mu}^{\lambda} \Gamma_{\nu\beta}^{\beta} - \frac{1}{2} g^{\alpha\lambda} g_{\mu\nu} \Gamma_{\alpha\beta}^{\beta} + \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\lambda} g_{\mu\nu}),_{\lambda} \\ + (-\Gamma_{\mu\nu}^{\lambda} + \frac{1}{2} \delta_{\nu}^{\lambda} \Gamma_{\mu\beta}^{\beta} + \frac{1}{2} \delta_{\mu}^{\lambda} \Gamma_{\nu\beta}^{\beta} - \frac{1}{2} g^{\alpha\lambda} g_{\mu\nu} \Gamma_{\alpha\beta}^{\beta} + \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\lambda} g_{\mu\nu}) \Gamma_{\lambda\xi}^{\xi} \\ + \Gamma_{\mu\alpha}^{\beta} \Gamma_{\nu\beta}^{\alpha} - \Gamma_{\mu\nu}^{\alpha} \Gamma_{\alpha\beta}^{\beta} + \frac{1}{2} g^{\alpha\xi} g_{\xi\nu,\mu} \Gamma_{\alpha\beta}^{\beta} + \frac{1}{2} g^{\alpha\xi} g_{\xi\mu,\nu} \Gamma_{\alpha\beta}^{\beta} - \Gamma_{\mu\beta}^{\beta} \Gamma_{\nu\alpha}^{\alpha} - \frac{1}{2} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\alpha} g_{\mu\nu,\alpha} \\ + \frac{1}{2} g_{\mu\nu} g^{\rho\sigma} \Gamma_{\rho\alpha}^{\beta} \Gamma_{\sigma\beta}^{\alpha} - \frac{1}{2} g_{\mu\nu} g^{\rho\sigma} \Gamma_{\rho\sigma}^{\alpha} \Gamma_{\alpha\beta}^{\beta}].$$

Dans cette expression, les termes soulignés une fois en ligne droite s'annulent partiellement, les termes soulignés en ondulé s'annulent et les termes soulignés deux fois s'ajoutent. En remarquant que, d'après (C.7),

$$\Gamma_{\mu\beta,\nu}^{\beta} = \Gamma_{\nu\beta,\mu}^{\beta}$$

et en regroupant les termes factorisant $g_{\mu\nu}$, on trouve que le coefficient de $\sqrt{-g}$ vaut

$$\begin{aligned}
 & -\Gamma_{\mu\nu,\lambda}^{\lambda} + \Gamma_{\mu\beta,\nu}^{\beta} - \frac{1}{2} \underbrace{g^{\alpha\lambda} g_{\mu\nu,\lambda} \Gamma_{\alpha\beta}^{\beta}} - 2\Gamma_{\mu\nu}^{\alpha} \Gamma_{\alpha\beta}^{\beta} + \Gamma_{\mu\alpha}^{\beta} \Gamma_{\nu\beta}^{\alpha} + \frac{1}{2} \underbrace{g^{\alpha\xi} g_{\xi\nu,\mu} \Gamma_{\alpha\beta}^{\beta}} + \frac{1}{2} \underbrace{g^{\alpha\xi} g_{\xi\mu,\nu} \Gamma_{\alpha\beta}^{\beta}} \\
 & - \frac{1}{2} g_{\mu\nu} (g^{\alpha\lambda}{}_{,\lambda} \Gamma_{\alpha\beta}^{\beta} + g^{\alpha\lambda} \Gamma_{\alpha\beta,\lambda}^{\beta} - g^{\rho\sigma}{}_{,\lambda} \Gamma_{\rho\sigma}^{\lambda} - g^{\rho\sigma} \Gamma_{\rho\sigma,\lambda}^{\lambda} + g^{\alpha\lambda} \Gamma_{\alpha\beta}^{\beta} \Gamma_{\lambda\xi}^{\xi} - g^{\rho\sigma} \Gamma_{\rho\alpha}^{\beta} \Gamma_{\sigma\beta}^{\alpha}).
 \end{aligned}$$

Les termes soulignés en ondulé se combinent pour générer le terme

$$\Gamma_{\mu\nu}^{\alpha} \Gamma_{\alpha\beta}^{\beta}.$$

L'expression recherchée prend ainsi la forme

$$\begin{aligned}
 & -\sqrt{-g} \{ \Gamma_{\mu\nu,\lambda}^{\lambda} - \Gamma_{\mu\beta,\nu}^{\beta} + \Gamma_{\alpha\beta}^{\beta} \Gamma_{\mu\nu}^{\alpha} - \Gamma_{\mu\alpha}^{\beta} \Gamma_{\nu\beta}^{\alpha} \\
 & - \frac{1}{2} g_{\mu\nu} [-\Gamma_{\alpha\beta}^{\beta} (g^{\alpha\lambda}{}_{,\lambda} + g^{\alpha\lambda} \Gamma_{\lambda\xi}^{\xi}) - g^{\alpha\lambda} \Gamma_{\alpha\beta,\lambda}^{\beta} + \Gamma_{\rho\sigma}^{\lambda} (g^{\rho\sigma}{}_{,\lambda} + g^{\rho\xi} \Gamma_{\xi\lambda}^{\sigma}) + g^{\rho\sigma} \Gamma_{\rho\sigma,\lambda}^{\lambda}] \}.
 \end{aligned}$$

Or, les relations

$$\begin{aligned}
 0 & = g^{\alpha\lambda}{}_{,\lambda} = g^{\alpha\lambda}{}_{,\lambda} + \Gamma_{\xi\lambda}^{\alpha} g^{\xi\lambda} + \Gamma_{\xi\lambda}^{\lambda} g^{\alpha\xi} \\
 0 & = g^{\rho\sigma}{}_{,\lambda} = g^{\rho\sigma}{}_{,\lambda} + \Gamma_{\xi\lambda}^{\rho} g^{\xi\sigma} + \Gamma_{\xi\lambda}^{\sigma} g^{\rho\xi}
 \end{aligned}$$

impliquent respectivement

$$\begin{aligned}
 g^{\alpha\lambda}{}_{,\lambda} + \Gamma_{\xi\lambda}^{\lambda} g^{\alpha\xi} & = -\Gamma_{\xi\lambda}^{\alpha} g^{\xi\lambda} \\
 g^{\rho\sigma}{}_{,\lambda} + \Gamma_{\xi\lambda}^{\sigma} g^{\rho\xi} & = -\Gamma_{\xi\lambda}^{\rho} g^{\xi\sigma}.
 \end{aligned}$$

Finalement, le premier membre des équations d'Euler-Lagrange s'écrit

$$\begin{aligned}
 & -\sqrt{-g} [\Gamma_{\mu\nu,\lambda}^{\lambda} - \Gamma_{\mu\beta,\nu}^{\beta} + \Gamma_{\alpha\beta}^{\beta} \Gamma_{\mu\nu}^{\alpha} - \Gamma_{\mu\alpha}^{\beta} \Gamma_{\nu\beta}^{\alpha} \\
 & - \frac{1}{2} g_{\mu\nu} (\Gamma_{\alpha\beta}^{\beta} \Gamma_{\xi\lambda}^{\alpha} g^{\xi\lambda} - g^{\alpha\lambda} \Gamma_{\alpha\beta,\lambda}^{\beta} - \Gamma_{\rho\sigma}^{\lambda} \Gamma_{\xi\lambda}^{\rho} g^{\xi\sigma} + g^{\rho\sigma} \Gamma_{\rho\sigma,\lambda}^{\lambda})],
 \end{aligned}$$

c'est-à-dire

$$-\sqrt{-g} (R_{\mu\nu} - \frac{R}{2} g_{\mu\nu}). \quad (\text{C.8})$$

Ainsi nous avons établi que

$$\frac{\partial}{\partial x^{\lambda}} \frac{\partial L}{\partial (g^{\mu\nu}{}_{,\lambda})} - \frac{\partial L}{\partial g^{\mu\nu}} = -\sqrt{-g} (R_{\mu\nu} - \frac{R}{2} g_{\mu\nu}). \quad (\text{C.9})$$

On en déduit les équations d'Einstein dans le vide :

$$R_{\mu\nu} - \frac{R}{2} g_{\mu\nu} = 0.$$

Pour retrouver les équations d'Einstein en présence d'une constante cosmologique Λ , il est facile de voir, à partir de (C.6) et de (C.9), qu'il faut ajouter à la densité lagrangienne L le terme

$$-2\Lambda \sqrt{-g}. \quad (\text{C.10})$$

Annexe D

Quelques résultats de mécanique statistique

D.1 Potentiels thermodynamiques

Soit un système constitué d'un mélange de divers types de particules, en quantité N_i pour chaque type i . On appelle *potentiel thermodynamique* une fonction d'état s'exprimant à l'aide d'une variable mécanique (volume V ou pression P) et d'une variable thermique (entropie S ou température T) et dont les dérivées partielles permettent de calculer toutes les autres grandeurs thermodynamiques¹.

variables	potentiel thermodynamique	différentielle
V, S, N_i	énergie interne E	$dE = T dS - P dV + \sum \mu_i dN_i$
P, S, N_i	enthalpie $W = E + PV$	$dW = T dS + V dP + \sum \mu_i dN_i$
V, T, N_i	énergie libre de Helmholtz $F = E - TS$	$dF = -S dT - P dV + \sum \mu_i dN_i$
P, T, N_i	énergie libre de Gibbs $\Phi = E - TS + PV$	$d\Phi = -S dT + V dP + \sum \mu_i dN_i$

Les quantités μ_i sont appelées *potentiels chimiques*.

À l'équilibre thermique, F et Φ sont minimaux, la première de ces grandeurs par rapport à toute variation de l'état à T et V constants, et la seconde dans le cas de T et P constants.

On peut montrer que

$$\Phi = \sum_i N_i \mu_i$$

et donc

$$F = \sum_i N_i \mu_i - PV.$$

D.2 Densité et énergie d'un gaz parfait

Nous ne considérerons ici que le cas particulier du gaz parfait constitué d'un seul type de particules élémentaires obéissant à la statistique de Bose-Einstein ou de Fermi-Dirac².

1. L. Landau et E. Lifchitz, *Physique théorique, Tome 5, Physique statistique*, Mir, 1967, §§ 14, 15, 24, 86.

2. *Ibid.*, § 55.

La distribution des valeurs absolues de l'impulsion p est donnée par

$$dN_p = \frac{gV}{2\pi^2\hbar^3} \frac{p^2}{\exp\left(\frac{\varepsilon - \mu}{kT}\right) \pm 1} dp \quad (\text{D.1})$$

(signe supérieur pour la statistique de Fermi-Dirac et signe inférieur pour celle de Bose-Einstein), où $\varepsilon = mc^2$ est l'énergie et g le nombre d'états quantiques de spin (auquel cas $g = 2s + 1$, s étant le spin des particules considérées) ou d'hélicité [$g = 2$ pour les photons (polarisation gauche ou droite), $g = 1$ pour les neutrinos (on n'a observé expérimentalement que des neutrinos polarisés à gauche et des antineutrinos polarisés à droite)]. On a

$$p^2 c^2 = \varepsilon^2 - m_0^2 c^4$$

donc

$$\begin{aligned} dN_\varepsilon &= \frac{gV}{2\pi^2 (\hbar c)^3} \frac{\varepsilon \sqrt{\varepsilon^2 - m_0^2 c^4}}{\exp\left(\frac{\varepsilon - \mu}{kT}\right) \pm 1} d\varepsilon \\ N &= \frac{gV}{2\pi^2 (\hbar c)^3} \int_{m_0 c^2}^{+\infty} \frac{\varepsilon \sqrt{\varepsilon^2 - m_0^2 c^4}}{\exp\left(\frac{\varepsilon - \mu}{kT}\right) \pm 1} d\varepsilon. \end{aligned}$$

Posant

$$z = \frac{\varepsilon}{kT},$$

on obtient

$$\frac{N}{V} = \frac{g}{2\pi^2} \left(\frac{kT}{\hbar c}\right)^3 \int_{\frac{m_0 c^2}{kT}}^{+\infty} \frac{z \sqrt{z^2 - \left(\frac{m_0 c^2}{kT}\right)^2}}{\exp\left(z - \frac{\mu}{kT}\right) \pm 1} dz. \quad (\text{D.2})$$

Cette relation, tout à fait générale, donne implicitement μ en fonction de N/V et de T . Si $\mu = 0$, elle permet d'exprimer N/V en fonction de T .

L'énergie totale est donnée par

$$\begin{aligned} E &= \int_{m_0 c^2}^{+\infty} \varepsilon dN_\varepsilon \\ &= \frac{gV}{2\pi^2 (\hbar c)^3} \int_{m_0 c^2}^{+\infty} \frac{\varepsilon^2 \sqrt{\varepsilon^2 - m_0^2 c^4}}{\exp\left(\frac{\varepsilon - \mu}{kT}\right) \pm 1} d\varepsilon, \end{aligned}$$

soit, en utilisant la variable z ,

$$E = \frac{gV}{2\pi^2} \frac{(kT)^4}{(\hbar c)^3} \int_{\frac{m_0 c^2}{kT}}^{+\infty} \frac{z^2 \sqrt{z^2 - \left(\frac{m_0 c^2}{kT}\right)^2}}{\exp\left(z - \frac{\mu}{kT}\right) \pm 1} dz. \quad (\text{D.3})$$

D.3 Quatre intégrales utiles

Soit la fonction zêta de B. Riemann :

$$\zeta(x) = \sum_{n=1}^{\infty} \frac{1}{n^x} \quad (x > 1).$$

En particulier, $\zeta(3) = 1.20206$.

On peut démontrer les quatre résultats suivants.

$$\begin{aligned} \int_0^{+\infty} \frac{z^2}{e^z - 1} dz &= \Gamma(3) \zeta(3) \\ &= 2 \zeta(3); \end{aligned} \quad (\text{D.4})$$

$$\int_0^{+\infty} \frac{z^2}{e^z + 1} dz = \frac{3}{2} \zeta(3); \quad (\text{D.5})$$

$$\begin{aligned} \int_0^{+\infty} \frac{z^3}{e^z - 1} dz &= \Gamma(4) \zeta(4) \\ &= \frac{\pi^4}{15}; \end{aligned} \quad (\text{D.6})$$

$$\int_0^{+\infty} \frac{z^3}{e^z + 1} dz = \frac{7\pi^4}{120}. \quad (\text{D.7})$$

D.4 Application au gaz de photons (V et T fixés)

Le cas des photons³ est intéressant car, d'une part, il n'existe qu'*un seul type de particules* et, d'autre part, *leur nombre n'est pas fixé d'avance*. Comme nous l'avons vu dans la section D.1, à l'équilibre thermique pour T et V donnés l'énergie libre de Helmholtz F est minimale, donc

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{T,V} = 0. \quad (\text{D.8})$$

Pour obtenir N et E , il suffit ainsi d'annuler le potentiel chimique μ et de poser $m_0 = 0$ dans les relations (D.2) et (D.3) particularisées aux bosons :

$$\frac{N}{V} = \frac{g}{2\pi^2} \left(\frac{kT}{\hbar c} \right)^3 \int_0^{+\infty} \frac{z^2}{e^z - 1} dz$$

avec $g = 2$ pour les photons. Compte tenu de (D.4), on a donc

$$\frac{N}{V} = \frac{2}{\pi^2} \zeta(3) \left(\frac{kT}{\hbar c} \right)^3. \quad (\text{D.9})$$

On calcule de même

$$E = \frac{gV}{2\pi^2} \frac{(kT)^4}{(\hbar c)^3} \int_0^{+\infty} \frac{z^3}{e^z - 1} dz$$

3. La linéarité des équations de l'électrodynamique exprime le fait que les photons n'interagissent pas entre eux. Le « gaz de photons » peut, par conséquent, être assimilé à un gaz parfait.

donc, par (D.6),

$$E = \frac{4\sigma}{c} T^4 V \quad (\text{D.10})$$

(loi de Stefan-Boltzmann), où

$$\sigma = \frac{\pi^2 k^4}{60 \hbar^3 c^2}$$

est la constante de Stefan-Boltzmann, égale à $5.67 \times 10^{-8} \text{ kg}/(\text{s}^3 \text{ K}^4)$. D'après (D.9) et (D.10), l'énergie moyenne par photon vaut ainsi

$$\begin{aligned} \frac{E}{N} &= \frac{\pi^4}{30 \zeta(3)} kT \\ &\simeq 2,70 kT. \end{aligned}$$

D'autre part, on peut démontrer qu'un gaz parfait ultra-relativiste de particules élémentaires vérifie la relation

$$PV = \frac{E}{3}. \quad (\text{D.11})$$

La pression de radiation d'un gaz de photons est donc donnée par⁴

$$P = \frac{4\sigma}{3c} T^4.$$

Calculons enfin l'entropie du gaz de photons. On a $\Phi = N\mu = 0$. Or, $\Phi = E - TS + PV$ donc, par (D.10) et (D.11),

$$\begin{aligned} S &= \frac{E + PV}{T} \\ &= \frac{16\sigma}{3c} T^3 V. \end{aligned} \quad (\text{D.12})$$

Conséquence intéressante, on constate, en comparant (D.9) et (D.12), que *l'entropie d'un gaz de photons est proportionnelle au nombre de particules*. L'entropie par photon vaut

$$\begin{aligned} \frac{S}{N} &= \frac{2\pi^4}{45 \zeta(3)} k \\ &\simeq 3,60 k. \end{aligned}$$

4. Remarquons que l'on peut formellement écrire une équation du type $PV = N\kappa T$ avec

$$\kappa = \frac{2\pi^2 \hbar^3 c^2 \sigma}{3k^3 \zeta(3)}.$$

Annexe E

Physique et philosophie

Introduction à la *Critique de la raison pure* de Kant ¹

« Quel dommage que les enfants ne puissent pas parler durant les deux premières années de leur vie, car, s'ils le pouvaient, ils balbutieraient sans doute la philosophie kantienne. » ²

E.1 Origine et signification de la *Critique de la raison pure*

Le siècle des Lumières marque le triomphe de la pensée rationnelle héritée du XVII^e siècle, et cela aussi bien dans le contexte politique et social qui prépara la Révolution française que dans le domaine scientifique. La science nouvelle héritée de Galilée et de Newton, le concept de *loi scientifique* remportent chaque jour de nouveaux succès, tant dans le domaine de la connaissance pure qu'appliquée. Peu à peu, une séparation s'amorce entre des techniciens de la science, manipulateurs d'outils mathématiquement évolués (essentiellement l'analyse infinitésimale) et d'autres penseurs soucieux de la légitimer (comment ce type de savoir scientifique est-il possible?) ou de lui fixer un cadre (tout peut-il être rationalisé?).

Dans son *Traité de la nature humaine*, David Hume (1711-1776) est ainsi amené à critiquer le fondement de la pensée rationnelle à l'origine des lois scientifiques : la relation de la cause et de l'effet. Pour cela, il considère l'exemple d'une bille de billard roulant sur une table et heurtant une autre bille initialement au repos ³. Elle lui confère du mouvement, c'est là un fait d'expérience tellement commun que l'on ne songe même plus à s'en étonner. Les lois mathématiques des collisions ont d'ailleurs été parmi les premières à recevoir une formulation par Descartes, Huygens et Leibniz. Mais derrière ces banalités apparentes se cachent deux difficultés. *Primo*, qui sait si, au bout de n expériences, les billes ne vont pas se comporter différemment? Par exemple, après la collision elles pourraient se trouver au repos, ou bien une des deux billes pourrait s'envoler ou encore disparaître. Comme on ne peut démontrer le général à partir du particulier, *la démarche inductive, pourtant aux fondements de la science, n'est pas logiquement justifiable*. Les lois de la physique ne sont pas rationnellement nécessaires au sens où l'est une démonstration mathématique.

1. Le texte de cette annexe n'a pas la prétention d'atteindre la rigueur de pensée des spécialistes en philosophie des sciences. Le lecteur intéressé par un discours plus fouillé pourra consulter, par exemple : L. Halleux, *Le transcendantal en physique contemporaine*, mémoire de licence en philosophie, Université de Liège, 2003.

2. P. Deussen, cité par E. Schrödinger, *La Nature et les Grecs*, Seuil, 1992, p. 185.

3. D. Hume, *Abrégé du Traité de la nature humaine*, 1740, traduction française de D. Deleule, Aubier, Paris, 1971, pp. 49-51. *Traité de la nature humaine*, 1739, traduction française de A. Leroy, Aubier, Paris, 1946, pp. 163-165, 166-167, 240, 252.

Secundo, et cette remarque est encore plus grave, le *principe de causalité* lui-même est hors de la portée de l'expérience, incapable de le réfuter ou de le prouver. La notion de connexion causale est une sorte de croyance, spontanée (ce qui la rend d'autant plus subtile à décoder) mais dénuée de tout appui empirique.

Analysant la situation, Hume constate que le principe de causalité n'est en définitive que l'expression d'un instinct reposant uniquement sur trois critères : la *contiguïté* dans le temps et dans l'espace, l'*antériorité* dans le temps et la *conjonction constante* entre la cause et l'effet. Mais il ne voit là aucune *nécessité logique* : cette dernière échappe à l'expérimentation. Comme l'écrit E. Yakira :⁴

« le sens de la notion de cause (ce que nous en entendons) est un lien nécessaire entre la cause et son effet ; mais sa référence (l'objet réel dont nous parlons effectivement) est la conjonction constante de la cause et de l'effet. »

Hume en conclut que la relation de cause à effet ne repose que sur *la force de l'habitude*. Sa pensée relève du *scepticisme* :

« on peut voir dans l'œuvre de Hume un défi incontournable à la science et à la philosophie : la pensée, disait-il, s'impose des critères de vérité et de rationalité qu'elle ne peut réaliser car elle est incommensurable à l'Être, le fossé qui sépare pensée et Être est infranchissable. »⁵

La lecture de Hume va bouleverser la pensée de E. Kant (1724-1804) en le tirant, selon la formule célèbre, « de son sommeil dogmatique », allusion à l'état d'esprit qui règne alors sur le monde philosophique allemand. Kant se demande aussi comment la science est possible. Il reconnaît la pertinence de la critique de Hume (la raison ne peut accéder à l'Être) mais, contrairement à ce dernier, ne peut se contenter d'une réaction sceptique, et voici pourquoi. La science, produit de la pensée humaine, fonctionne sur base de quelques règles *nécessaires et universelles* dans la mesure où elles déterminent systématiquement et *a priori* toute connaissance scientifique possible. Ainsi, le principe selon lequel *tous les changements arrivent suivant la loi de liaison de la cause et de l'effet* est constitutif de la nature de la science humaine. Le premier enjeu de la *Critique de la raison pure* (1781, 1787), que l'on peut qualifier d'*épistémologique*, consiste ainsi à réhabiliter la science rationnelle dans ses prétentions à établir une certaine forme de connaissance nécessaire, parce que *a priori*.

Étudions un peu plus en détail l'argumentation de Kant. Mathématicien et physicien de formation, il s'était d'abord interrogé sur l'apparente *différence de statut* entre ces deux disciplines : pourquoi la mathématique paraît-elle absolument vraie et nécessaire tandis que la physique semble prêter le flanc aux critiques de Hume ?

On connaissait au moins une classe de propositions vraies *a priori* et universelles, ce sont les jugements tautologiques, c'est-à-dire ceux dont le contraire implique une contradiction. Par exemple : « Tous les corps sont étendus », « Les hommes chauves sont des hommes », « Les chats sont des félins ». Kant les appelle *analytiques*. Dans chacun de ces énoncés en effet, le prédicat est implicitement contenu dans la définition du sujet, il résulte obligatoirement de son analyse.

« C'est précisément aussi la raison pour laquelle toutes les propositions analytiques sont des jugements *a priori*, lors même que leurs concepts sont empiriques, par exemple : l'or est un métal jaune ; car pour savoir cela, en dehors de mon concept d'or qui impliquait que ce corps est jaune et que c'est un métal, je n'ai besoin d'aucune expérience supplémentaire, car c'est

4. E. Yakira, *La causalité de Galilée à Kant*, Presses Universitaires de France, 1994, pp. 108-109.

5. *Ibid.*, p. 111.

précisément cela qui constituait mon concept, et il me suffisait de l'analyser, sans avoir à me mettre en quête de rien d'autre qui lui soit antérieur. »⁶

Avant Kant, on pensait que la nécessité mathématique ne reposait que sur son caractère analytique, donc tautologique.

D'autre part, les jugements attribuant à un sujet une propriété nouvelle sont dits *synthétiques*. Avant Kant, il semblait clair que de tels énoncés ne peuvent être connus qu'*a posteriori*, c'est-à-dire empiriquement. Par exemple : « Tous les corps sont pesants ». Le caractère *a posteriori* résulte du fait qu'il est toujours possible de *concevoir* que l'attribut ne s'applique pas au sujet, en l'occurrence, dans notre exemple, qu'il existe des corps non pesants. Science expérimentale, la physique semblait dès lors devoir renoncer à toute prétention à l'universalité et à la nécessité, ainsi que Hume l'avait noté.

On a donc le tableau de classification des énoncés ci-dessous :⁷

		Forme logique	
		<i>analytique</i>	<i>synthétique</i>
Ce sur quoi s'appuie la prétention à la vérité ou à la fausseté	<i>a priori</i>	« Un cercle est rond »	?
	<i>a posteriori</i>	—	« Le satellite naturel de la Terre est rond »

(L'impossibilité mentionnée au niveau des propositions analytiques *a posteriori* résulte du fait que les propositions analytiques sont nécessairement des jugements *a priori*.)

Le grand mérite de Kant a été de compléter ce tableau en découvrant l'existence de *jugements synthétiques a priori*. C'est la *mathématique* qui lui en a suggéré la possibilité. Qu'elle fût *vraie a priori* n'était, nous l'avons dit, pas nouveau pour l'époque. Mais admettre qu'elle est constituée de propositions synthétiques revient à accepter le point de vue de Kant, selon lequel *la mathématique ne se réduit pas à une tautologie*, c'est-à-dire à un jeu purement formel et logique. La raison en est la suivante : c'est l'homme qui se donne lui-même les objets mathématiques en les offrant à sa propre intuition (les concepts mathématiques sont ce que Kant appelle des *concepts sensibles purs*⁸) :

« Le premier qui démontra le *triangle isocèle* (qu'il s'appelât *Thalès* ou de tout autre nom) eut une illumination ; car il trouva qu'il ne devait pas s'attacher à ce qu'il voyait dans la figure, ou même au simple concept qu'il en avait, pour en apprendre en quelque sorte les propriétés, mais qu'il devait produire cette figure par ce qu'il y pensait et présentait (par construction) *a priori* d'après les concepts eux-mêmes, et que, pour connaître sûrement une chose *a priori*, il ne devait attribuer à cette chose que ce qui résultait nécessairement de ce qu'il y avait mis lui-même, conformément à son concept. »⁹

« La définition réelle serait donc celle qui rend clair non seulement un concept, mais en même temps sa *réalité objective*. Les définitions mathématiques, qui présentent l'objet dans l'intuition conformément au concept, sont de cette dernière espèce. »¹⁰

6. E. Kant, *Prolégomènes à toute métaphysique future qui pourra se présenter comme science*, traduction de L. Guillermit, Librairie Philosophique J. Vrin, 1993, p. 27.

7. K. Popper, *La connaissance objective*, Aubier, 1991, p. 160.

8. E. Kant, *Critique de la raison pure*, édition publiée sous la direction de F. Alquié, traduit de l'allemand par A. J.-L. Delamarre et F. Marty à partir de la traduction de J. Berni, folio essais numéro 145, Gallimard, 1980, p. 192.

9. *Ibid.*, pp. 42-43.

10. *Ibid.*, p. 280.

« On devrait sans doute penser d'abord que la proposition $7 + 5 = 12$ est une proposition simplement analytique, qui résulte d'après le principe de contradiction du concept d'une somme de sept et cinq. Mais, si on y regarde de plus près, on trouve que le concept de la somme de 7 et 5 ne contient rien de plus que la réunion de deux nombres en un unique, ce par quoi on ne pense absolument pas ce qu'est ce nombre unique, qui comprend les deux autres. Le concept de douze n'est aucunement déjà pensé du fait que je pense simplement cette réunion de sept et de cinq, et j'aurai beau prolonger l'analyse de mon concept d'une telle somme possible, je n'y trouverai cependant pas le nombre douze. On doit aller au-delà de ces concepts, en s'aidant de l'intuition qui correspond à l'un des deux, par exemple ses cinq doigts, ou (comme Segner dans son arithmétique) cinq points, et ainsi ajouter progressivement les unités du nombre cinq donné dans l'intuition au concept du nombre sept. Je prends en effet d'abord le nombre 7, et en m'aidant pour le concept de 5 des doigts de ma main comme intuition, j'ajoute les unités, que je prenais auparavant ensemble, pour composer le nombre 5, progressivement sur cette image que j'ai, au nombre 7, et je vois ainsi résulter le nombre 12. Que 7 *devait* être ajouté à 5, je l'avais bien pensé dans le concept d'une somme $= 7 + 5$, mais non que cette somme soit égale au nombre 12. La proposition arithmétique est donc toujours synthétique; ce que l'on réalise d'autant plus clairement que l'on prend des nombres quelque peu plus grands, parce qu'il saute alors aux yeux que nous pouvons tourner et retourner notre concept comme nous voulons : sans nous aider de l'intuition, au moyen de la simple analyse de nos concepts, nous ne pourrions jamais trouver la somme.

Un principe quelconque de la géométrie n'est pas davantage analytique. Que la ligne droite soit entre deux points la plus courte, c'est une proposition synthétique. Car mon concept de droit ne contient rien qui se rapporte à la quantité, mais seulement une qualité. Le concept de plus court vient donc entièrement s'ajouter, et ne peut être tiré par aucune analyse du concept de la ligne droite. Il faut donc s'aider de l'intuition, au moyen de laquelle seulement la synthèse est possible. »¹¹

Afin d'illustrer encore cette notion de *synthèse* si fondamentale pour l'intelligence des phénomènes selon Kant, considérons l'exemple d'un tronc de cylindre droit à base circulaire. Observé latéralement, il se présente comme un rectangle. Vu d'en haut, il apparaît circulaire. On peut dire que notre cerveau possède l'*intuition du cylindre, comme synthèse du rectangle et du cercle*. Il serait impossible à un mathématicien bidimensionnel d'effectuer cette synthèse, même s'il possédait une intuition parfaite du rectangle et du cercle envisagés séparément. Il ne pourrait d'ailleurs établir un cours complet de géométrie tridimensionnelle qu'en utilisant la méthode analytique.

Tâchons maintenant de pénétrer davantage la pensée critique. Kant constate que les deux types d'opérations mentales sur lesquelles se fonde toute connaissance (la *sensibilité*, qui recueille, et l'*entendement*, qui organise) sont nécessairement conformes à des *a priori*. Ce sont

— les intuitions pures de la sensibilité, *l'espace et le temps* ;

« Espace et temps ne sont pas des propriétés des choses, mais des conditions de la connaissance que nous pouvons avoir d'elles, et ces conditions sont *a priori*. Le temps est la condition de toute succession et de toute simultanéité - ce n'est nullement la succession ou la simultanéité qui engendrent le temps. Grâce à cet *a priori* nous pouvons comprendre pourquoi il peut exister une arithmétique nécessaire et universelle en tant que science de la durée, et une géométrie nécessaire et universelle comme science de l'espace. »¹²

11. *Ibid.*, pp. 75-76.

12. J. Hersch, *L'étonnement philosophique*, folio essais numéro 216, Gallimard, 1993, p. 209.

— les concepts purs de l'entendement, les *catégories* (par exemple, celle de causalité et dépendance). L'entendement opère dans les limites de l'expérience sensible, et donc sous les conditions de l'espace et du temps. Par conséquent, les catégories n'ont pas d'autre contenu que celui qui leur est donné par les intuitions sensibles, elles ne s'appliquent qu'à des objets d'expérience.

Kant en déduit l'existence de jugements synthétiques *a priori* en physique, tel le principe de causalité déjà mentionné :

« Lorsque je dis « cause », je comprends ce que je dis, mais je n'apprends rien. Mais lorsque je dis « à tout changement il y a une cause », je comprends et j'apprends quelque chose qui n'était pas dit dans « cause » (ni, d'ailleurs, dans « changement ») : cette proposition est un jugement synthétique *a priori*. Or, de la catégorie de causalité naît effectivement une telle proposition, qui est un principe universel et nécessaire de l'expérience. Plus précisément, ce principe constate que « tous les changements arrivent suivant la loi de liaison de la cause et de l'effet ». »¹³

Ces découvertes constituent la réponse de Kant au scepticisme. Comme l'écrit E. Yakira,

« la rationalité n'est plus quelque chose que la science cherche à reproduire, mais à établir sinon à créer. »¹⁴

Cette victoire comporte un revers : c'est que, justement, la connaissance ne peut s'aventurer en-dehors du champ des choses telles que récoltées au travers de l'intuition. Un concept auquel ne correspond aucune intuition est bien une *pensée*, mais non une *connaissance*. Je puis penser ce que je veux mais je ne puis connaître que ce qui m'est donné dans l'intuition. Et comme le seul mode d'intuition, pour l'homme, est l'intuition sensible, il en résulte que nous ne pouvons rien connaître au-delà de l'expérience, elle-même inéluctablement inscrite dans un cadre spatio-temporel. Ainsi, la raison ne peut légitimement atteindre d'autre réalité que la réalité sensible. Hors des objets du monde, nos concepts ne saisissent rien. Tous les progrès de l'expérience ne nous permettent jamais de sortir des limites de l'expérience, qui fixent ainsi les limites de notre connaissance.

Kant appelle *phénomènes* les objets de l'expérience, « Das Ding für mich ». Par opposition, il qualifie de *noumènes* les objets d'une intuition non sensible (et donc en-dehors du champ de la connaissance humaine), par exemple intellectuelle. L'expression « Das Ding an sich » pour le noumène est peut-être quelque peu équivoque, dans la mesure où elle tendrait à faire croire que le noumène est plus proche de la vérité que le phénomène. Mais le message kantien dépasse cette interprétation :

« La chose en soi n'est pas chez Kant une autre chose : elle est un autre point de vue sur la chose... »¹⁵

Cependant, ce que l'on ne peut connaître peut cependant être pensé, et la raison humaine est ainsi faite qu'elle ne peut s'empêcher de s'aventurer au-delà du champ de l'expérience en recherchant *l'inconditionné* ; par exemple, une causalité ultime. Mais cela n'a pas de sens car toute cause doit – selon la logique de l'entendement – provenir d'une cause supérieure, et une telle régression n'a pas de fin. Le texte qui ouvre la *Critique* est très explicite à cet égard :¹⁶

13. E. Yakira, *op. cit.*, p. 115.

14. *Ibid.*, p. 122.

15. J.-C. Billier, *Kant et le kantisme*, Armand Colin, 1998, p. 41.

16. E. Kant, *Critique de la raison pure, op. cit.*, pp. 31-32.

« La raison humaine a cette destinée particulière, dans un genre de ses connaissances, d'être accablée de questions qu'elle ne peut écarter ; car elles lui sont proposées par la nature de la raison elle-même, mais elle ne peut non plus y répondre, car elles dépassent tout pouvoir de la raison humaine.

Ce n'est pas sa faute si elle tombe dans cet embarras. Elle part de principes dont l'usage est inévitable dans le cours de l'expérience, et en même temps suffisamment garanti par elle. Avec leur aide, elle s'élève toujours plus haut (comme le comporte aussi bien sa nature), vers des conditions plus éloignées. Mais, s'apercevant que, de cette manière, son œuvre doit toujours rester inachevée, puisque les questions ne cessent jamais, elle se voit contrainte de se réfugier dans des principes qui dépassent tout usage possible d'expérience, et qui pourtant paraissent si peu suspects que la raison humaine commune elle-même se trouve en accord avec eux. Mais, par là, elle se précipite dans l'obscurité et des contradictions, d'où elle peut certes conclure que cela doit tenir à des erreurs cachées quelque part, mais sans pouvoir les découvrir, parce que les principes dont elle se sert, comme ils vont au-delà des limites de toute expérience, ne connaissent plus désormais de pierre de touche prise à l'expérience. Le champ de bataille de ces combats sans fin, voilà ce qu'on nomme *Métaphysique*. »

Ainsi apparaissent les *idées transcendantales*, c'est-à-dire des concepts rationnels inéluçables auxquels ne correspond aucun objet donné dans les sens. Ces idées sont au nombre de trois : l'*âme*, le *monde* et *Dieu*. Elles relèvent de la *métaphysique* et pas de la *science*.

« ... nous avons conclu qu'il ne doit y avoir que trois sortes de raisonnements dialectiques, lesquels se rapportent aux trois sortes de raisonnements par lesquels la raison peut aller de principes à des connaissances et qu'en tout sa fonction propre est de s'élever de la synthèse conditionnée à laquelle l'entendement reste toujours attaché à une synthèse inconditionnée qu'il ne peut jamais atteindre. »

« Or, tous les concepts purs en général ont à s'occuper de l'unité synthétique des représentations, mais les concepts de la raison pure (les idées transcendantales) s'occupent de l'unité synthétique inconditionnée de toutes les conditions en général. Par suite, toutes les idées transcendantales se laissent ramener sous trois classes, dont la *première* contient l'*unité* absolue (inconditionnée) du *sujet pensant* ; la *seconde*, l'*unité* absolue de la *série des conditions du phénomène* ; la *troisième*, l'*unité* absolue de la *condition de tous les objets de la pensée en général*.

Le sujet pensant est l'objet de la *psychologie* ; l'ensemble qui comprend tous les phénomènes (le monde), celui de la *cosmologie*, et la chose qui contient la condition suprême de la possibilité de tout ce qui peut être pensé (l'être de tous les êtres), l'objet de la *théologie*. La raison pure nous fournit donc l'idée d'une doctrine transcendantale de l'âme (*psychologia rationalis*), d'une science transcendantale du monde (*cosmologia rationalis*), enfin d'une connaissance transcendantale de Dieu (*theologia transcendentalis*). »¹⁷

Dans le chapitre intitulé « L'antinomie de la raison pure », Kant entend montrer que la raison est impuissante à traiter rationnellement la question cosmologique. Pour ce faire, il expose quatre conflits des idées transcendantales. Chacun est discuté en double page : la thèse est démontrée à gauche... et l'antithèse à droite !

– Première antinomie.

Thèse : *Le monde a un commencement dans le temps, et il est aussi, quant à l'espace, renfermé dans des limites.*

Antithèse : *Le monde n'a ni commencement ni limites dans l'espace, mais il est infini aussi bien par rapport au temps que par rapport à l'espace.*

17. *Ibid.*, pp. 347-348.

– Deuxième antinomie.

Thèse : *Toute substance composée dans le monde est faite de parties simples, et il n'existe absolument rien que le simple ou ce qui en est composé.*

Antithèse : *Aucune chose composée, dans le monde, n'est faite de parties simples, et il n'existe absolument rien de simple dans le monde.*

– Troisième antinomie.

Thèse : *La causalité suivant les lois de la nature n'est pas la seule d'où puissent être dérivés les phénomènes du monde dans leur ensemble. Il est encore nécessaire d'admettre, pour les expliquer, une causalité par liberté.*

Antithèse : *Il n'y a pas de liberté, mais tout dans le monde arrive suivant les lois de la nature.*

– Quatrième antinomie.

Thèse : *Au monde appartient quelque chose qui, soit comme sa partie, soit comme sa cause, est un être absolument nécessaire.*

Antithèse : *Il n'existe nulle part aucun être absolument nécessaire, ni dans le monde, ni hors du monde, comme sa cause.*

Le second enjeu de la *Critique*, que l'on peut qualifier de *métaphysique*, consiste donc à tracer une ligne de démarcation nette entre la physique (la connaissance des phénomènes) et la métaphysique, cette dernière ne pouvant être considérée comme une science. Loin de déforer les prétentions de la pensée rationnelle, Kant les renforce, au même titre qu'une personne connaissant à fond les tables de multiplication et leur utilisation est plus fiable qu'une autre ayant étudié le calcul intégral mais l'utilisant erronément ou à mauvais escient :

« On croira percevoir, dans une rapide inspection de cette œuvre, que l'utilité en est toute *négative* : ne jamais nous risquer avec la raison spéculative au-delà des limites de l'expérience, et c'est bien aussi en fait sa première utilité. Mais elle devient *positive*, dès qu'on se rend compte que les principes, avec lesquels la raison spéculative se risque hors de ses limites, ont en fait pour conséquence inévitable, non pas l'*élargissement*, mais, à y regarder de plus près, le *rétrécissement* de l'usage de notre raison, en menaçant réellement d'étendre surtout les limites de la sensibilité, à laquelle ils appartiennent proprement, et ainsi de repousser l'usage pur (pratique) de la raison. »¹⁸

En résumé, la philosophie développée par Kant, et qui porte le nom d'*idéalisme transcendantal*, est basée sur les deux observations suivantes :

- 1) l'objet (comme objet des sens) se règle sur la nature de notre faculté d'intuition, et non l'inverse ;
- 2) l'objet (comme objet de connaissance scientifique) se règle sur les concepts *a priori* de notre entendement, et non l'inverse.

Par une analogie évidente, on a appelé *révolution copernicienne* ce renversement de perspective. La structuration de nos perceptions spatio-temporelles (les phénomènes) par nos concepts *a priori* (les catégories) est ainsi constitutive de toute science possible. Il importe de distinguer l'idéalisme transcendantal de l'*idéalisme dogmatique* (Berkeley : il n'existe aucune réalité extérieure à la pensée).

18. *Ibid.*, p. 50.

E.2 Lexique

Catégories. Formes *a priori* (ou encore : concepts purs) de l'entendement. Elles sont réparties en quatre groupes : de la *quantité* ; de la *qualité* ; de la *relation* (dont celle de la cause et de l'effet) ; de la *modalité*.

Entendement. Voir *Sensibilité et entendement*.

Espace et temps. Formes *a priori* de la sensibilité, c'est-à-dire conditions *nécessaires* et *universelles* de l'expérience sensible. Kant affirme ainsi leur idéalité transcendante, entraînant leur réalité *empirique* (ils n'existent pas objectivement).

Esthétique. Science des règles de la sensibilité en général.

Esthétique transcendante. L'esthétique est *transcendante* lorsqu'elle étudie les formes *a priori* de la sensibilité.

Idéalisme. Doctrine pour laquelle tout ce qui existe, ou du moins tout ce dont nous pouvons connaître l'existence, doit être d'une façon ou d'une autre de nature mentale.¹⁹

Idée transcendante. « J'entends par idée un concept nécessaire de la raison auquel aucun objet qui lui corresponde ne peut être donné dans les sens. Ainsi les concepts purs de la raison ... sont des *idées transcendantes*. »²⁰

Intuition. « De quelque manière et par quelque moyen qu'une connaissance puisse se rapporter à des objets, le mode par lequel elle se rapporte immédiatement à des objets ... est l'*intuition*. »²¹

Logique. Science des règles de l'entendement en général.

Logique transcendante. « [...] science de l'entendement pur et de la connaissance rationnelle par laquelle nous pensons des objets tout à fait *a priori*. Une telle science, qui déterminerait l'origine, l'étendue et la valeur objective de telles connaissances, devrait porter le nom de *logique transcendante* ; car elle a affaire seulement aux lois de l'entendement et de la raison, mais uniquement en tant qu'elle se rapporte à des objets *a priori*, et non, comme la logique générale, indifféremment aux connaissances empiriques aussi bien que pures. »²²

Noumène, objet transcendantal, chose en soi. Ces termes sont parfois pris par Kant comme synonymes en tant qu'ils s'opposent tous trois au *phénomène*, que les formes de l'espace et du temps ne les concernent pas et que, par conséquent, ils ne peuvent être connus par les catégories de l'entendement. Cependant si cette synonymie peut être acceptée pour la lecture de certains textes, elle ne saurait être maintenue dans une étude plus approfondie.²³

Phénomène. C'est ce qui est objet d'expérience, c'est donc le réel non tel qu'il est en soi, mais tel qu'il est par rapport à notre faculté de connaître. Il s'oppose donc à la fois à l'*appareance* et au *noumène*.²⁴

19. B. Russell, *Problèmes de philosophie*, Introduction et traduction de l'anglais par F. Rivenc, Bibliothèque philosophique Payot, Editions Payot, 1989, p. 59.

20. E. Kant, *Critique de la raison pure*, *op. cit.*, p. 342.

21. *Ibid.*, p. 87.

22. *Ibid.*, p. 123.

23. Kant, *la raison pure. Extraits de la critique*, choisis et présentés par F. Khodoss, Presses Universitaires de France, 1990. Vocabulaire, pp. 217-230.

24. *Ibid.*

Principe. Les principes sont les règles de l'usage objectif des catégories.²⁵ Kant énonce les jugements *a priori* issus de ces dernières, dont le fameux « *Tous les changements arrivent suivant la loi de liaison de la cause et de l'effet* ».

Raison. Ce mot a chez Kant un sens étroit et un sens large généralement faciles à distinguer par le contexte. Au sens étroit, la Raison est la faculté de lier des connaissances en un système. Au sens large, c'est la faculté de former des principes *a priori*. (Elle est donc Raison *pure*.) Si ces principes déterminent ce qui doit être, la Raison pure est pratique, si ces principes sont des principes de connaissance, la Raison pure est théorique. Dans les deux cas elle s'oppose à l'expérience.²⁶

Scepticisme. Toute attitude philosophique qui nie la possibilité de la certitude. Doctrine d'après laquelle l'homme ne peut atteindre la vérité, dans un domaine ou sur un sujet déterminé. *Scepticisme scientifique, moral*.²⁷

Sensibilité et entendement. « La capacité de recevoir (la réceptivité) des représentations grâce à la manière dont nous sommes affectés par des objets s'appelle *sensibilité*. C'est donc au moyen de la sensibilité que des objets nous sont donnés, et elle seule nous fournit des *intuitions*; mais c'est par l'entendement qu'ils sont *pensés*, et c'est de lui que proviennent les *concepts*. »²⁸

« Sans la sensibilité, nul objet ne nous serait donné; sans l'entendement, nul ne serait pensé. Des pensées sans contenu sont vides; des intuitions sans concepts sont aveugles. »²⁹

Temps. Voir *Espace et temps*.

Transcendental. Cet adjectif désigne toujours une connaissance (forme, concept, synthèse, déduction, etc.) et s'oppose à *empirique*. Il désigne ce par quoi une connaissance *a priori* est possible. Ainsi l'*exposition transcendantale* du concept d'espace montre comment ce concept rend possible cette connaissance *a priori* qu'est la géométrie. La *déduction transcendantale* des catégories montre que ces catégories s'appliquent nécessairement à tout objet de connaissance, etc. Par suite transcendantal devient un des synonymes de *a priori, pur*. Il importe de rappeler que l'*a priori* chez Kant, s'il est indépendant de l'expérience, lui est relatif.³⁰

25. E. Kant, *Critique de la raison pure*, *op. cit.*, p. 206.

26. Kant, *la raison pure. Extraits de la critique*, choisis et présentés par F. Khodoss, *op. cit.*, pp. 217-230.

27. *Le Petit Robert 1*, Dictionnaires Le Robert, 1991.

28. E. Kant, *Critique de la raison pure*, *op. cit.*, pp. 87-88.

29. *Ibid.*, p. 118.

30. Kant, *la raison pure. Extraits de la critique*, choisis et présentés par F. Khodoss, *op. cit.*, pp. 217-230.