

Introduction à la relativité restreinte et générale

Cours d'ouverture, EPF 3eme année

Fabien Besnard

23 janvier 2009

Table des matières

0.1	Avant-Propos	2
1	La physique en 1905	3
1.1	Les référentiels inertiels et la relativité galiléenne	3
1.2	Les transformations de Galilée	4
1.3	Deux petits nuages...	5
2	La relativité restreinte	9
2.1	La solution d'Einstein	9
2.2	Les transformations de Lorentz : Le raisonnement d'Einstein	9
2.3	Loi d'addition des vitesses en cinématique relativiste	14
2.4	Espace-temps de Minkowski	14
2.5	Effets et « paradoxes » en relativité restreinte	19
2.5.1	Dilatation des durées de vie des particules instables	19
2.5.2	Paradoxe des jumeaux de Langevin	20
2.5.3	Le problème de la voiture et du garage	22
2.6	Dynamique relativiste	23
2.6.1	Équivalence masse-énergie, quadrivecteur impulsion-énergie	23
2.6.2	Le PFD en relativité restreinte	26
3	La relativité générale	27
3.1	Motivation	27
3.2	Le principe d'équivalence	27
3.3	« L'inspiration la plus heureuse de ma vie »	29
3.4	« Grossmann, il faut que tu m'aides, sinon je deviens fou. »	34
3.5	Variétés différentiables	41
3.6	La relativité générale en bref	44
3.7	La solution de Schwarzschild	45
3.8	Cosmologie	51
4	Conclusion	55
A	Solutions des exercices	57
B	Une déduction axiomatique des transformations de Lorentz	64
C	Le tenseur énergie-impulsion	66
	Bibliographie	67

0.1 Avant-Propos

L'objectif de ce cours introductif est de mettre en relief les idées principales de la théorie de la relativité d'Einstein, et de donner un bagage mathématique minimal permettant à l'étudiant intéressé d'aller plus loin. L'approche utilisée insistera sur les expériences de pensée fondamentales ayant conduit Einstein à sa théorie, et sera agrémentée d'une perspective historique.

Le premier chapitre rappelle quelques notions de mécanique classique et introduit le problème de l'éther. Le deuxième chapitre est consacré à la relativité restreinte et à la résolution de certains « paradoxes » fameux. L'espace-temps de Minkowski est également introduit. Par manque de temps nous ne pouvons aborder que quelques conséquences physiques, laissant de côté des phénomènes relativistes essentiels, comme l'effet Doppler transverse. Le lecteur pourra approfondir ses connaissances dans [3] ou [7].

Enfin le troisième chapitre est consacré à l'introduction aux prédictions et au formalisme de la relativité générale.

Cette dernière théorie, profonde et difficile, nécessite des développements mathématiques importants. Deux orientations s'offrent alors à l'enseignant : soit rester au niveau de la vulgarisation scientifique (de qualité) en se bornant à une approche heuristique et intuitive des phénomènes, soit viser une compréhension plus précise des concepts géométriques au coeur de la théorie d'Einstein. Nous avons tenté d'ouvrir une voie médiane. Notre but est de nous intercaler entre [10] ou [2] et [8] ou [9] et de permettre aux étudiants d'aborder ces deux derniers ouvrages sans s'effrayer. L'espace-temps très réduit qui nous est imparti nous a contraint à faire des choix drastiques sur le plan mathématique, comme de ne pas aborder la notion de transport parallèle ni celle de dérivée covariante. Nous avons dû sacrifier l'élégance conceptuelle à la rapidité de présentation. Nous ne pouvons qu'encourager le lecteur à combler cette lacune par la lecture des ouvrages de référence déjà mentionnés.

Les applications de la théorie restreinte sont innombrables en physique des hautes énergies, notamment via l'équivalence masse-énergie incarnée par la célèbre formule $E = mc^2$. Les applications pratiques de la théorie générale sont moins nombreuses. On peut citer toutefois le GPS. Cependant, toute la cosmologie moderne, c'est-à-dire l'idée la plus précise que l'homme a de l'univers, est fondée sur la théorie générale. L'expansion de l'univers, le big-bang, les trous noirs, tout ceci ne peut être appréhendé sans l'aide de cette théorie. Il est remarquable que le cheminement intellectuel ayant conduit à un tel bouleversement de la pensée soit finalement assez facile à suivre, et à la portée de l'étudiant motivé, même si sa forme finale est très complexe. Les merveilleux raisonnements d'Albert Einstein font ainsi partie du patrimoine culturel de l'humanité et resteront sans aucun doute comme l'un de ses accomplissements les plus grandioses. À ce titre, ils ne peuvent manquer d'être connus de celui qui est curieux du monde qui l'entoure. L'auteur de ces lignes sera comblé s'il a pu participer à leur diffusion auprès des étudiants de l'EPF.

Chapitre 1

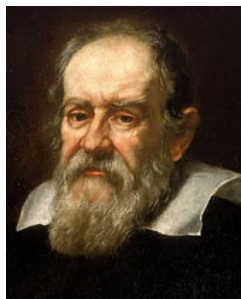
La physique en 1905

1.1 Les référentiels inertiels et la relativité galiléenne

Pour commencer il nous faut revenir au tout début de la mécanique, c'est-à-dire à Galilée. On sait que ce dernier était partisan du système Copernicien selon lequel la Terre tourne autour du Soleil et sur elle-même. L'objection qu'on ne manqua pas de lui faire à l'époque est que si tel était bien le cas, une pierre jetée du haut d'une tour n'aurait pas dû tomber exactement à son pied, à cause de la rotation de la Terre pendant la chute. Voici quelle fut la réponse de Galilée à cette objection :

« Enfermez-vous avec un ami dans la plus grande cabine sous le pont d'un grand navire et prenez avec vous des mouches, des papillons et d'autres petites bêtes qui volent ; munissez-vous aussi d'un grand récipient rempli d'eau avec de petits poissons ; accrochez aussi un petit seau dont l'eau coule goutte à goutte dans un autre vase à petite ouverture placé en dessous. Quand le navire est immobile, observez soigneusement comment les petites bêtes qui volent vont à la même vitesse dans toutes les directions de la cabine, on voit les poissons nager indifféremment de tous les côtés ; les gouttes qui tombent entrent toutes dans le vase placé dessous ; si vous lancez quelque chose à votre ami, vous n'avez pas besoin de jeter plus fort dans une direction que dans une autre lorsque les distances sont égales ; si vous sautez à pieds joints, comme on dit, vous franchirez des espaces égaux dans toutes les directions. Quand vous aurez soigneusement observé cela, bien qu'il ne fasse aucun doute que toutes les choses doivent se passer ainsi quand le navire est immobile, faites aller le navire à la vitesse que vous voulez ; pourvu que le mouvement soit uniforme, sans balancement dans un sens ou l'autre, vous ne remarquerez pas le moindre changement dans tous les effets qu'on vient d'indiquer ; aucun ne vous permettra de vous rendre compte si le navire est en marche ou immobile. »

Galileo Galilei (Dialogue sur les deux grands systèmes du monde)



Galileo Galilei, 1564-1642

Il s'agit d'une formulation de ce qu'on appellera plus tard le principe de relativité du mouvement *uniforme*, ou principe de relativité galiléen. On peut l'énoncer de façon succincte comme ceci :

Principe de relativité galiléen

« Si deux référentiels¹ sont animés d'un mouvement rectiligne uniforme l'un par rapport à l'autre, les lois de la physique sont les mêmes dans les deux référentiels. »

En réalité ce principe n'a pas été énoncé tel quel avant Einstein. Avant lui, on le limitait au cadre des lois de la mécanique, dont il était en fait une conséquence. La mécanique classique (newtonienne) était plutôt fondée sur le principe d'inertie et le PFD. Le principe d'inertie instaure une classe privilégiée de référentiels (les référentiels galiléens ou d'inertie), dans lesquels on pourra appliquer le PFD.

Principe d'inertie

« Il existe des référentiels dits d'inertie, dans lesquels le mouvement d'un corps² sur lequel aucune force ne s'applique est rectiligne uniforme. »

Dans un tel référentiel, les lois de la physique prennent une forme particulière, plus simple que dans les autres référentiels. Le PFD s'y écrit $\vec{F} = m\vec{a}$, alors que dans un référentiel non-inertiel il faut faire apparaître des forces supplémentaires dans le membre de droite, appelées forces d'inertie. Notons que si notre mouvement est accéléré par rapport à un référentiel d'inertie, les forces d'inertie permettent de savoir que nous sommes en mouvement. Par exemple dans une voiture qui accélère, on est plaqué sur son siège par une force d'inertie. Donc le mouvement accéléré est absolu : on n'a pas besoin de regarder dehors pour savoir qu'on bouge, il suffit de faire des expériences de physique (et en particulier de mécanique) à l'intérieur de notre référentiel.

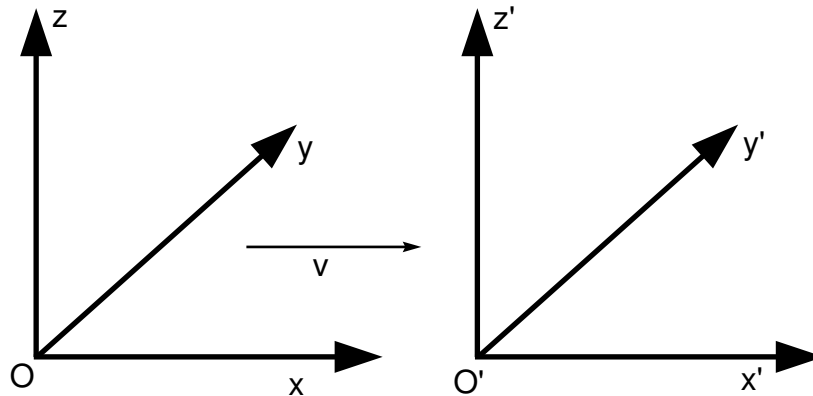
À ce stade on pourrait se demander ce qu'est une force. La réponse serait qu'une force c'est quelque chose qui perturbe le mouvement rectiligne uniforme d'un corps. Le principe d'inertie est donc... une tautologie. En pratique, ça ne pose pas vraiment de problème : il suffit de disposer d'un modèle de référentiel d'inertie. Le premier modèle est le référentiel terrestre, mais son mouvement de rotation diurne engendre des forces d'inertie. Un meilleur exemple est le référentiel héliocentrique. Mais des expériences assez fines pourraient montrer l'existence de forces d'inertie dans ce référentiel. Comme il est logique de supposer qu'un référentiel isolé n'est soumis à aucune force, un tel référentiel serait d'inertie. En fait, la gravitation ayant une portée infinie, aucun référentiel ne peut être complètement isolé du reste de l'Univers. La mécanique classique est donc fondée sur quelque chose qui n'existe pas ! Mais après tout, la géométrie plane parle de droites infinies et de points sans dimension, or personne n'a jamais vu de telles choses. La mécanique classique, comme la géométrie, est fondée sur des idéalizations. Comme elles joueront un grand rôle dans ce qui va suivre, énumérons les ici à nouveau : l'espace absolu (par rapport auquel on peut détecter un mouvement accéléré), le temps absolu (le même pour tous les observateurs), et une classe privilégiée de référentiels : les référentiels d'inertie.

1.2 Les transformations de Galilée

Il est important de savoir passer d'un référentiel à un autre. Soit R et R' deux référentiels tels que R' soit animé d'un mouvement rectiligne uniforme de vitesse \vec{v} par rapport à R . Quitte à utiliser des rotations ou translations (y compris un changement d'origine des dates), on peut supposer qu'à l'instant $t = 0$, les deux référentiels ont la même origine et les mêmes axes, et que la vitesse \vec{v} est dans la direction de l'axe des x .

¹Rappelons qu'un référentiel est un système d'axes de coordonnées lié à un observateur. Nous n'utiliserons que des référentiels dont les axes sont directement orthonormés.

²Ce corps doit être supposé ponctuel. S'il ne l'est pas, le principe s'applique à son centre de gravité.



Il est alors facile de voir que le changement de référentiel s'écrit :

$$\begin{cases} x' &= x - vt \\ y' &= y \\ z' &= z \\ t' &= t \end{cases}$$

En composant ce type de transformation avec des rotations ou des translations, on obtient toutes les transformations possibles qui correspondent à des changements de référentiels dont les axes sont orthonormés directs. L'ensemble de ces transformations forme un groupe appelé groupe de Galilée.

Comme le PFD ne fait intervenir que l'accélération, donc la dérivée seconde par rapport au temps, il est clair que les lois de la mécanique seront les mêmes dans deux référentiels animés l'un par rapport à l'autre d'un mouvement de translation uniforme. Ainsi le principe de relativité galiléen est bien une conséquence des lois de la mécanique. Mais on peut comprendre ce que dit ce principe de façon plus « philosophique ». Il énonce en fait simplement que le mouvement uniforme est relatif, il est toujours mouvement par rapport à quelque chose. Pris dans ce sens, il est naturel, et c'est ce que fera Einstein, de le poser en principe premier et de le généraliser à l'ensemble des lois de la physique.

Remarque : Le groupe de Galilée que nous venons de définir est en réalité seulement une partie du « vrai » groupe de Galilée, appelée « groupe de Galilée propre et orthochrone ». C'est le sous-groupe qui est constitué des transformations qui conservent le sens du temps (orthochrone) et l'orientation de l'espace (propre). Si on ajoute à ce groupe une symétrie P par rapport à un plan quelconque, et le renversement du temps $T : t \mapsto -t$, on obtient, en considérant toutes les compositions possibles, la totalité du groupe de Galilée.

1.3 Deux petits nuages...

William Thomson (Lord Kelvin) était un représentant éminent de la physique de son temps. En 1900, il donna une conférence intitulée « Nineteenth-century clouds over the dynamical theory of heat and light » au cours de laquelle il déclara

« The beauty and clearness of the dynamical theory, which asserts heat and light to be modes of motion, is at present obscured by two clouds. The first came into existence with the undulatory theory of light [...] It involved the question 'How could Earth move through an elastic solid, such as essentially is the luminiferous ether?' The second is the Maxwell-Boltzmann current doctrine regarding the partition of energy.³ »

³« La beauté et la clarté de la théorie dynamique, qui affirme que la chaleur et la lumière sont des modes du mouvement, est à présent obscurcie par deux nuages. Le premier vint à l'existence avec la théorie ondulatoire de la lumière [...] Il

Le second problème deviendra plus tard la tempête « mécanique quantique ». Quant au premier, que nous allons revoir en détail, il aboutira à la physique relativiste qui va nous intéresser dans ce cours d'introduction.

À quoi Thomson faisait-il précisément allusion par « théorie dynamique de la lumière » ? Au cours du XIXe siècle, les physiciens étaient arrivés à la conclusion que la lumière était un phénomène ondulatoire. Or, on ne pouvait imaginer une onde sans un milieu matériel pour la porter. On pensait donc qu'à l'image des ondes sonores qui sont des ondes de pression dans l'air, la lumière était la propagation d'une perturbation mécanique dans un milieu invisible remplissant tout l'espace : l'éther⁴. De plus, comme la théorie de Maxwell avait permis d'identifier la lumière à un phénomène électromagnétique, l'éther devait être constitué de petits dipôles électriques. Ainsi, tout se ramenait au mouvement de charges. Cependant l'éther avait des propriétés étranges... Tout d'abord il devait pénétrer tous les corps, et n'offrir aucune résistance à leur mouvement, sans quoi il aurait ralenti les planètes dans leur course. Mais dans le même temps, il devait se comporter comme un solide ! En effet, les ondes lumineuses sont transversales, et jamais longitudinales, contrairement aux ondes sonores, et cela impose que le milieu de propagation soit solide. Mieux : la vitesse très élevée des ondes lumineuses nécessitait que l'éther soit extrêmement comprimé. Tout cela était bizarre, mais le pire était à venir...

Regardons les équations de Maxwell.

$$\begin{aligned}\nabla \cdot \vec{E} &= \frac{\rho}{\epsilon_0}, & \nabla \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t} \\ \nabla \cdot \vec{B} &= 0, & \nabla \times \vec{B} &= \mu_0 \vec{J} + \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t}\end{aligned}$$

Si vous ne les connaissez pas, pas d'inquiétude : il est inutile d'en expliquer tous les termes. Ce qui nous intéresse c'est que ces équations font intervenir la vitesse de la lumière⁵, c qui est en fait égale à :

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} \quad (1.1)$$

Par conséquent, les équations de Maxwell ne sauraient être valables dans tous les référentiels d'inertie. La théorie de l'électromagnétisme introduit donc un référentiel particulier : celui où l'éther est au repos. Le mouvement par rapport à l'éther peut donc être détecté par une expérience d'électromagnétisme (par exemple avec de la lumière), interne au référentiel. Or ceci viole le principe de relativité galiléen. Einstein se posait cette question en ces termes : que se passerait-il si je me déplaçais à la vitesse de la lumière en tenant devant moi un miroir ? Si je me vois, c'est que la lumière qui quitte mon visage va plus vite que c , ce qui est contraire aux équations de Maxwell. Mais si je ne me vois pas, alors je sais que je bouge sans regarder dehors, ce qui est contraire au principe de relativité de Galilée ! Il se posa cette question pendant des années avant d'entrevoir une réponse.

« Je dois avouer que, tout au début, quand la théorie de la relativité restreinte a commencé à germer en moi, j'ai eu toutes sortes de troubles nerveux. Quand j'étais jeune il m'arrivait de passer des semaines dans un état second, comme quelqu'un qui devait encore à l'époque surmonter l'état de stupéfaction causé par sa rencontre avec ce genre de questions. » (Albert Einstein)

Mais c'est là tout le génie d'Albert que d'avoir pris le principe de relativité galiléen au sérieux. Pour les autres physiciens, s'il existe un état de repos absolu détectable par des expériences d'électromagnétisme⁶,

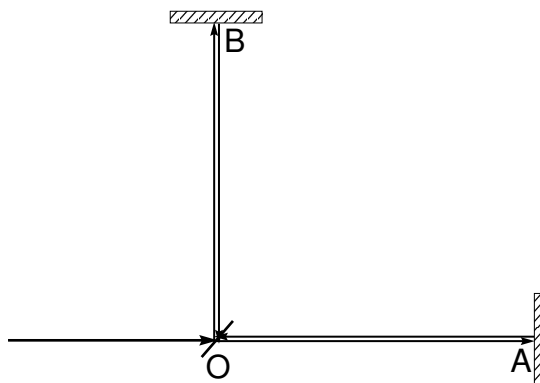
met en jeu la question : 'Comment la Terre pourrait-elle se déplacer dans un solide élastique, telle qu'est essentiellement l'éther lumifère ?' Le second est la doctrine actuelle de Maxwell-Boltzmann sur la répartition de l'énergie. »

⁴Imaginé par Huyghens.

⁵En réalité, c'est en constatant que ses équations avaient pour solution une onde transversale se propageant à la vitesse c définie par la formule (1.1) que Maxwell comprit que la lumière était une onde électromagnétique.

⁶Notons qu'il n'est pas logiquement cohérent de séparer électromagnétisme et mécanique à ce stade, puisque la théorie de l'éther visait précisément à « mécaniser » l'électromagnétisme. Si toute la physique se réduit *in fine* à la mécanique alors il est nécessaire que le principe de relativité galiléen soit étendu à toute la physique.

ainsi soit-il! On tente donc de détecter le mouvement de la Terre par rapport à l'éther. L'effet du « vent d'éther » sur la propagation de la lumière doit être très faible, mais les techniques naissantes d'interférométrie sont très sensibles, assez pour le détecter. En 1887, Albert Michelson et Edward Morley montent leur célèbre expérience. Ainsi que le montre la figure ci-dessous, tracée dans le référentiel terrestre, il s'agit de scinder un rayon lumineux en deux rayons orthogonaux, puis de les réunir après réflexion sur des miroirs. Les trajets parcourus par les deux rayons sont d'égales longueurs.

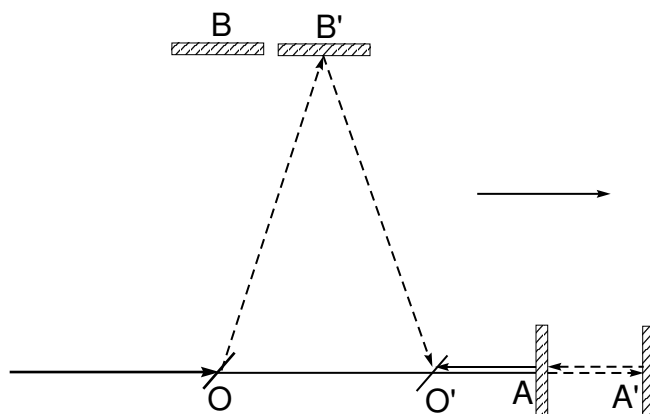


Maintenant supposons que l'axe horizontal de la figure soit parallèle au mouvement de rotation de la Terre (axe Est-Ouest) tandis que l'autre lui est perpendiculaire (axe Nord-Sud). Le rayon horizontal met un temps

$$\Delta t_1 = \frac{L}{c-v} + \frac{L}{c+v} = \frac{2L/c}{1-v^2/c^2}$$

à faire un aller-retour, où l'on a appelé v la vitesse instantanée de rotation de la Terre.

On peut calculer le temps de parcours du rayon Nord-Sud sur la figure ci-dessous, tracée dans le référentiel lié à l'éther.



Le rayon parcourt le triangle $OB'O'$ à la vitesse c , ce qui donne un temps de :

$$\Delta t_2 = \frac{2L/c}{\sqrt{1-v^2/c^2}}$$

Les deux rayons arrivent donc déphasés, ce qui provoque des franges d'interférence. Lorsqu'on fait tourner le dispositif autour d'un axe passant par O , on doit voir ces franges se décaler, et même disparaître

lorsque la position de l'interféromètre est symétrique par rapport à l'axe Nord-Sud. Le problème, c'est que Michelson et Morley n'ont vu... aucun effet !

Il fallait donc supposer que l'éther était emporté par le mouvement de la Terre, de même que l'air est emporté dans l'habitacle d'une voiture, où le son se propage à la même vitesse quelle que soit celle de la voiture. Mais ça ne tenait pas debout car on savait depuis longtemps (c'est le phénomène d'aberration) que la vitesse de la lumière en provenance des étoiles était toujours la même, malgré le mouvement propre de ces étoiles (qui de surcroît se compose avec celui de la Terre autour du Soleil).

La situation était si grave que Lorentz et Fitzgerald firent (indépendamment) l'hypothèse que la vitesse du référentiel terrestre par rapport à celui de l'éther provoquait une contraction des longueurs dans le sens du mouvement, et une dilatation du temps ! C'était audacieux, mais cela attribuait un rôle actif à l'éther et lui conférait une propriété magique de plus... C'est Einstein qui franchit le pas suivant⁷, en jetant l'éther à la poubelle et en modifiant les principes de base de la mécanique. C'est ce que nous allons voir maintenant.

⁷Poincaré était arrivé à une conclusion très similaire, mais il ne l'a pas affirmée aussi clairement et nettement qu'Albert.

Chapitre 2

La relativité restreinte

2.1 La solution d'Einstein

Einstein finit par comprendre que la solution pour se tirer de ces contradictions consistait à énoncer deux principes premiers, ou postulats. Le premier, dont nous avons déjà parlé, est le principe de relativité galiléen étendu à toute la physique.

Postulat 1 : principe de relativité galiléen étendu. Les lois de la physique sont les mêmes pour tous les observateurs inertiels.

Einstein prend Galilée au sérieux : le mouvement uniforme est vraiment relatif, on ne peut le mettre en évidence d'aucune façon. Le second postulat est l'invariance de c .

Postulat 2 : invariance de la vitesse de la lumière. La vitesse de la lumière est la même dans tous les référentiels inertiels.

Ces deux postulats, en tant que tels, n'ont pas à être démontrés par le raisonnement. Mais ils sont aujourd'hui vérifiés avec une très grande précision. Si un jour il advenait que l'un d'entre eux soit réfuté par l'expérience, toute la théorie s'effondrerait. Empressons-nous de dire cependant que la théorie ne fait pas jouer un rôle particulier à la lumière. Il suffit qu'il existe une vitesse maximale, notée c , qu'aucun corps ne puisse dépasser. Cette vitesse est automatiquement invariante, par le premier postulat. Le second postulat pourrait donc s'exprimer de la façon suivante : il existe une vitesse limite pour tout corps ou tout signal. Expérimentalement, cette vitesse limite est alors identifiée à c . De plus, nous verrons que la théorie montre que cette vitesse limite est celle d'un corps de masse nulle. La théorie quantique permettant d'identifier la lumière avec un flux de photons de masse nulle, on retrouve alors le fait que cette vitesse limite est celle de la lumière, mais aussi celle de n'importe quelle particule de masse nulle¹.

Bien entendu, l'invariance d'une vitesse par changement de référentiel est incompatible avec les transformations de Galilée.

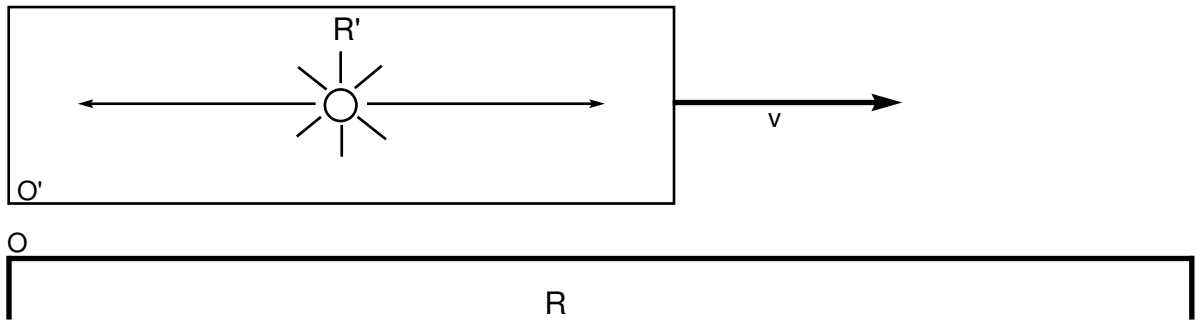
2.2 Les transformations de Lorentz : Le raisonnement d'Einstein

Reprenons ici le raisonnement que fait Einstein en 1905². Il s'agit d'un des chefs d'oeuvre de la pensée humaine, qui, me semble-t-il, devrait faire partie de la culture générale de chacun.

Considérons deux référentiels inertiels R et R' . On peut prendre par exemple pour R un référentiel lié au quai d'une gare et pour R' celui lié à un train. L'origine O de R est située initialement au bout du quai et celle O' de R' à l'arrière du train, de sorte que les deux origines coïncident à l'origine des dates. On appellera Albert l'observateur qui est dans le train, et Isaac celui se trouvant sur le quai.

¹On a longtemps cru que les neutrinos étaient de masse nulle, mais on sait aujourd'hui que ce n'est pas le cas. Les gravitons sont censés posséder une masse nulle, mais on ne les a pas encore observés.

²La version qu'il publie est plus académique dans la forme, mais identique sur le fond.

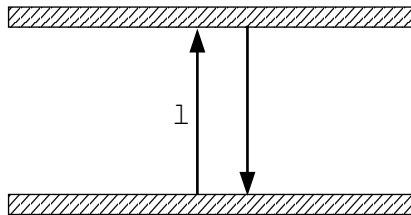


Albert commence par remarquer que la notion de simultanéité n'est pas la même pour lui et pour Isaac. En effet imaginons qu'il y ait deux portes, une à l'avant et une à l'arrière du train, et que ces portes s'ouvrent dès qu'une lumière les éclaire. Disposons une source de lumière au milieu du train. Si le train est à l'arrêt relativement à R , il est clair que les deux portes s'ouvriraient en même temps pour Albert comme pour Isaac. Si le train est en marche avec une vitesse uniforme v , cela ne doit rien changer pour Albert, car sinon il saurait qu'il bouge sans regarder dehors! Mais pour Isaac c'est tout-à-fait différent, car la porte avant du train fuit le rayon lumineux, tandis que la porte arrière va à sa rencontre. Ainsi pour Isaac, la porte arrière s'ouvre avant la porte avant³.

Ainsi deux événements qui sont simultanés pour Albert (l'ouverture des portes), ne le sont pas pour Isaac. La date de ces événements ne peut donc pas être la même pour les deux observateurs. La ligne $t' = t$ des transformations de Galilée, qui signifie que le temps est le même pour tous les observateurs ne peut pas être maintenue. Le temps est relatif à chaque observateur.

Exercice 2.2.1 Se peut-il qu'un observateur voit la porte avant s'ouvrir avant la porte arrière?

Bien que cette conclusion soit déjà bouleversante, nous allons continuer en cherchant à savoir comment le temps se modifie. Pour cela nous allons imaginer une horloge un peu spéciale, constituée d'un rayon de lumière piégé entre deux miroirs parfaits espacés d'une longueur l , comme ci-dessous. Dès que le rayon a fait un aller-retour, l'horloge fait « tic ».



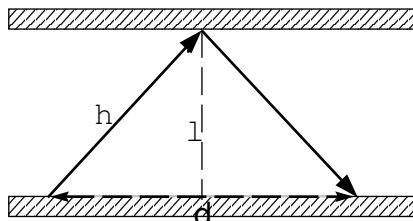
Albert et Isaac disposent chacun d'un exemplaire de cette horloge. Pour Albert, il est clair que l'intervalle de temps entre deux tics d'horloge est :

$$\Delta t' = \frac{2l}{c}$$

D'après le principe de relativité, il en va de même pour Isaac et son horloge.

³Remarquons qu'il ne s'agit aucunement d'une illusion d'optique, ou quoi que ce soit de ce genre : si Isaac veut connaître très précisément l'instant auquel une porte s'est ouverte il peut tenir compte du temps que met la lumière pour lui parvenir, et cet effet « retard » de l'information dû à la finitude de la vitesse de la lumière ne dépend pas de la vitesse du train et n'a rien à voir avec la relativité, contrairement à ce qu'on entend dire parfois.

Maintenant intéressons-nous au temps Δt entre deux tics de l'horloge d'Albert, mais vue depuis le référentiel d'Isaac. D'après Isaac, le trajet d'un rayon lumineux de l'horloge d'Albert est plus long que $2l$, il vaut $2h$.



On a donc

$$\Delta t = \frac{2h}{c}$$

Bien sûr, c'est le même c d'après le second principe.

Comme pendant le temps Δt , le train avance d'une distance

$$d = v\Delta t$$

On peut, par une simple application du théorème de Pythagore, obtenir :

$$\left(\frac{d}{2}\right)^2 + l^2 = h^2$$

Ce qui donne en remplaçant

$$\left(\frac{v\Delta t}{2}\right)^2 + \left(\frac{c\Delta t'}{2}\right)^2 = \left(\frac{c\Delta t}{2}\right)^2$$

D'où en simplifiant, en passant les Δt du même côté et en factorisant

$$c^2\Delta t'^2 = (c^2 - v^2)\Delta t^2$$

Soit encore :

$$\Delta t = \frac{\Delta t'}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.1)$$

On voit donc que si $v \neq 0$, $\Delta t > \Delta t'$, ce qui signifie que, pour Isaac, l'horloge d'Albert bat plus lentement que la sienne ! Une horloge en mouvement (rectiligne uniforme pour l'instant) bat plus lentement qu'une horloge au repos. De plus, l'effet en question ne dépend pas du type d'horloge utilisé, car si c'était le cas, Albert pourrait savoir qu'il bouge en emportant des horloges de construction différente et en observant leur désynchronisation.

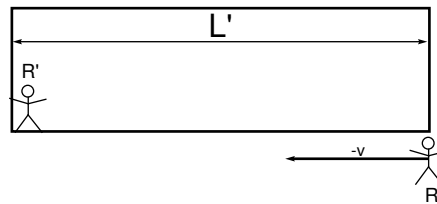
Remarque importante : La formule 2.1 s'applique pour tout intervalle de temps entre deux événements A et B **situés spatialement au même point** pour R' .

Mais n'oublions pas que le mouvement rectiligne uniforme est relatif, et que la situation est donc complètement symétrique. Ainsi, pour Albert, c'est l'horloge d'Isaac qui bat plus lentement, et exactement dans les mêmes proportions.

Supposons maintenant qu'Albert et Isaac veuillent mesurer la longueur du wagon. Pour fixer les idées supposons que le facteur :

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

soit égal à 2 (cela correspond à une vitesse du train environ égale à $0,86c$, ce qui est assez rapide!). Supposons également qu'à l'origine des dates, la position d'Isaac coïncide avec l'extrémité avant du train. L'abscisse d'Isaac au temps $t = 0$ est donc L dans R , et elle est égale à L' au temps $t' = 0$ dans R' , où L et L' sont les longueurs respectives du train dans les deux référentiels.



Du point de vue d'Albert, Isaac se déplace à une vitesse $-v$ vers l'origine, qu'il atteint au bout de disons, 10 secondes. Par conséquent, $L' = 10v$. Cependant, comme $\gamma = 2$, Albert voit que la montre d'Isaac n'a battu que 5 secondes pendant le trajet.

Exercice 2.2.2 Pourquoi peut-on appliquer ce raisonnement ? Peut-on l'itérer et dire que du point de vue d'Isaac, la montre d'Albert n'a battu que 2,5 secondes, et ainsi de suite à l'infini ?

Prenons maintenant le point de vue d'Isaac. Nous venons de voir que sa montre a battu 5 secondes. Le train lui est donc passé devant en totalité pendant un intervalle de temps de 5 secondes à la vitesse v . Il en conclut que $L = 5v$. D'où $L = L'/2$, formule qu'on peut généraliser aisément en prenant des facteurs γ et des temps de parcours quelconques en :

$$L = L'/\gamma = L'\sqrt{1 - \frac{v^2}{c^2}} \quad (2.2)$$

Interprétons cette relation. La longueur L' est la longueur du wagon mesurée dans un référentiel « comobile » : c'est sa longueur propre, ou encore longueur au repos. En effet, si ce type de wagon est fabriqué pour mesurer exactement 10 m de long (dans un référentiel « au repos » c'est-à-dire dans celui d'Isaac si on prend le quai comme référence du repos) alors Albert doit trouver que son wagon mesure exactement 10 m, sinon, encore une fois, il saurait qu'il bouge sans regarder dehors. Mais Isaac lorsqu'il calcule la longueur L trouve moins de 10 m : c'est ce qu'on appelle la contraction des longueurs dans le sens du mouvement, ou encore, contraction de Lorentz, puisque c'est l'effet que ce dernier avait postulé pour expliquer le résultat nul de l'expérience de Michelson et Morley. Mais nous disposons maintenant d'une toute autre explication ne faisant plus intervenir l'éther, mais une modification radicale des concepts de longueur et de durée.

Remarque 1 : La longueur d'un objet est une notion qui ne va pas de soit en relativité restreinte. En effet, aucun corps ne peut être parfaitement rigide dans cette théorie, et ceci en raison de la limite imposée à la vitesse de tout signal. Par conséquent si on tire l'avant du train, les molécules de celui-ci ne peuvent pas se mettre toutes en marche de conserve, car cela signifierait qu'un signal a été transmis instantanément. Ainsi, tout corps doit être considéré comme élastique. Dans ce qui précède, nous avons fait abstraction de ce fait. La formule doit donc s'entendre ainsi : la longueur d'un segment parallèle à l'axe des abscisses est par définition la différence des abscisses de ses extrémités au même instant. Cette notion dépend bien entendu du référentiel choisi, puisque la notion de simultanéité en dépend.

Remarque 2 : On pourrait facilement montrer qu'il n'y a pas de contraction dans la direction perpendiculaire au mouvement en reprenant le trajet de la lumière dans l'horloge-miroir utilisée plus haut. Mais cela ne serait pas très convaincant ! En effet, si on réfléchit bien, on a fait implicitement cette supposition.

Ce point doit donc être justifié autrement. Mais d'autres soupçons pourraient naître sur le raisonnement un peu particulier qu'on vient de faire, et il est nécessaire de démontrer mathématiquement que les principes de relativité et de constance de la vitesse de la lumière (ainsi que l'homogénéité et l'isotropie de l'espace-temps) aboutissent aux effets qu'on vient de décrire. Pour cela nous renvoyons à l'appendice B.

Pour l'instant continuons avec notre train. Supposons qu'un événement se produise dans le coin supérieur avant du train, de coordonnées (x', y', z') à l'instant t' pour Albert (référentiel R'). Ce même événement aura pour coordonnées spatio-temporelles (x, y, z, t) pour Isaac (référentiel R). Nous cherchons la transformation qui fait passer de (x, y, z, t) à (x', y', z', t') et qui va remplacer la transformation galiléenne.

Comme il n'y a pas de modification des longueurs perpendiculairement au mouvement, on aura clairement $y' = y$ et $z' = z$. Par définition de la vitesse et de la longueur du train on a $x = vt + L$. Or on a également par définition $x' = L' = \gamma L = \gamma(x - vt)$. En oubliant le train et en généralisant à n'importe quel événement de coordonnées (x, y, z, t) dans R et (x', y', z', t') dans R' on a déjà obtenu :

$$x' = \gamma(x - vt), \quad y' = y, \quad z' = z$$

Cherchons maintenant t' . Pour cela il suffit de réaliser que la transformation inverse s'obtient en remplaçant v par $-v$. Ainsi on doit avoir :

$$x = \gamma(x' + vt')$$

En remplaçant on obtient :

$$x = \gamma[\gamma(x - vt) + vt']$$

ce qui après quelques calculs donne :

$$t' = \gamma\left(t - \frac{vx}{c^2}\right)$$

D'où la transformation complète :

$$\begin{cases} x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \\ ct' &= \gamma(ct - \frac{v}{c}x) \end{cases} \quad \text{avec } \gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2} \quad (2.3)$$

Une telle transformation s'appelle une transformation spéciale de Lorentz, ou encore « boost ». Elle est spéciale car on a pris des origines de coordonnées ainsi que des axes coïncidants à l'origine des dates dans les deux référentiels. Comme nous l'avons dit au début, pour avoir une transformation générale on compose par des rotations et des translations : on obtient ainsi un groupe de transformations affines de \mathbb{R}^4 appelé groupe de Poincaré. Le sous-groupe constitué des transformations fixant une origine donnée (il est donc constitué d'applications linéaires) est appelé groupe de Lorentz⁴.

Ces transformations résultent à coup sûr de nos principes de départ. Ces principes sont très bien vérifiés expérimentalement, mais il reste deux choses à faire. La première est de comprendre comment la cinématique classique s'insère dans ce nouveau cadre, la seconde est de rebâtir la mécanique du point sur ces nouvelles bases.

Notre première tâche est un simple exercice.

Exercice 2.2.3 Montrer que l'on retrouve une transformation de Galilée lorsqu'on fait tendre v/c vers 0 dans une transformation de Lorentz.

La seconde tâche nous occupera plus tard, et nous ne pourrons que la survoler, mais d'ores et déjà nous pouvons constater que la formule du boost n'a pas de sens si $v > c$. Autrement dit, en théorie de la

⁴Il s'agit en fait du groupe de Lorentz propre et orthochrone, car il n'inclut pas le renversement du temps, ni les renversements de l'orientation spatiale.

relativité, les vitesses supérieures à celles de la lumière n'ont pas de sens⁵. Nous retrouverons ce résultat par des considérations de dynamique.

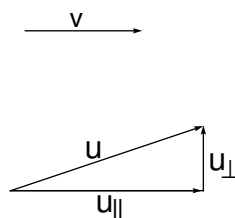
Exercice 2.2.4 Que devient la formule du boost quand v s'approche de c ? Interpréter le résultat.

Exercice 2.2.5 Un astronaute part de la Terre, dans la direction perpendiculaire au plan de l'écliptique, et voyage à la vitesse constante de $0.99c$ pendant 1 an (toutes les données sont énoncées dans le référentiel héliocentrique considéré comme inertiel). Décrire le mouvement de la Terre autour du Soleil du point de vue de l'astronaute pendant la durée du voyage. Décrire le voyage du point de vue du Soleil.

2.3 Loi d'addition des vitesses en cinématique relativiste

On conserve les mêmes notations qu'au paragraphe précédent.

Considérons un mobile animé d'une vitesse \vec{u} dans R . On notera u_{\parallel} (resp. u_{\perp}) la composante de \vec{u} dans la direction et le sens défini par \vec{v} (resp. dans le plan orthogonal à \vec{v}).



Exercice 2.3.1 Montrer que l'on a :

$$u'_{\parallel} = \frac{u_{\parallel} - v}{1 - \frac{u_{\parallel}v}{c^2}} \quad \text{et} \quad u'_{\perp} = u_{\perp} \frac{\sqrt{1 - \frac{v^2}{c^2}}}{1 - \frac{u_{\parallel}v}{c^2}}$$

Exercice 2.3.2 Montrer que l'on retrouve la formule galiléenne d'addition des vitesses lorsque u et v sont petits devant c .

Exercice 2.3.3 Reprendre l'exercice 2.2.5 en utilisant cette formule.

2.4 Espace-temps de Minkowski

« Depuis que les mathématiciens ont envahi la relativité, je ne comprends plus rien à ma théorie ! » (Albert Einstein)

On se place dans un référentiel galiléen R .

Considérons la quantité définie pour deux événements $M_1 = (x_1, y_1, z_1, t_1)$ et $M_2 = (x_2, y_2, z_2, t_2)$ par⁶ :

$$\Delta s^2 = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2$$

On voit facilement que cette quantité, appelée « carré de l'intervalle entre les événements M_1 et M_2 » est invariante par translations et rotations spatiales et par changement de l'origine des dates (translation temporelle).

⁵Il est en fait possible mathématiquement de considérer des particules allant plus vite que la lumière, appelées tachyons. Elles posséderaient néanmoins des propriétés pathologiques, comme une masse imaginaire pure, qui font qu'on considère généralement qu'elles ne peuvent exister. Cependant elles peuvent apparaître formellement comme des intermédiaires de calcul en théorie quantique des champs.

⁶On trouve également la convention inverse pour les signes.

Exercice 2.4.1 Montrer qu'elle est aussi invariante par un boost.

Le Δs^2 est donc invariant par le groupe de Poincaré, exactement comme la longueur au carrée $(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$ est invariante par le groupe de Galilée.

Le signe du Δs^2 donne une information importante. En effet, si $\Delta s^2 < 0$ la distance spatiale séparant les deux événements est supérieure à la distance que parcourt la lumière pendant l'intervalle de temps $|t_2 - t_1|$, autrement dit la lumière ne peut pas se propager de M_1 à M_2 , ni de M_2 à M_1 . Comme aucun signal ne peut aller plus vite que la lumière, aucun signal ne peut se propager de M_1 à M_2 ou inversement. Donc ce qui se passe en M_1 ne peut en aucune manière affecter ce qui se passe en M_2 ou inversement. On dit que M_1 et M_2 ne sont pas causalement reliés, et que l'intervalle au carré entre ces événements est du genre espace.

Si $\Delta s^2 > 0$, au contraire, un signal sub-luminique peut relier ces deux événements et on dit que leur intervalle au carré est du genre temps.

Enfin si $\Delta s^2 = 0$, seul un signal allant à la vitesse de la lumière peut les relier et on dit que leur intervalle au carré est du genre lumière.

Dans les deux derniers cas, on dit que M_1 et M_2 sont causalement reliés. Cette notion est Poincaré-invariante d'après ce que nous venons de voir.

Supposons que M_1 et M_2 soient causalement reliés et que $t_1 \leq t_2$. Alors nous noterons $M_1 \preceq M_2$.

Théorème 2.4.1 *La relation \preceq est Poincaré-invariante.*

Exercice 2.4.2 Démontrer ce théorème.

Exercice 2.4.3 Un joueur de basket marque un panier, puis 3 points supplémentaires pour son équipe s'affichent au tableau de score (tout ça dans le référentiel du terrain, supposé inertiel). Existe-t-il un référentiel d'inertie dans lequel l'affichage du tableau change avant que le ballon ne rentre dans le panier ?

Exercice 2.4.4 Est-il vrai que si un événement a lieu avant un autre dans un référentiel d'inertie, la même chose sera vraie dans tous les référentiels d'inertie ?

Munissons l'espace-vectoriel \mathbb{R}^4 de la forme bilinéaire g de signature $(1, 3)$ dont la matrice dans la base canonique est

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

L'espace \mathbb{R}^4 muni de g est appelé espace de Minkowski et nous le noterons \mathcal{M} . Cet espace représente l'espace-temps de la relativité restreinte lorsqu'on utilise des unités pour lesquelles $c = 1$. (Cela correspond par exemple à mesurer les distances en secondes-lumière. De telles unités sont très pratiques et très utilisées en relativité. On les appelle les unités géométriques⁷.)

Soit $X = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$ et $Y = \begin{pmatrix} y^0 \\ y^1 \\ y^2 \\ y^3 \end{pmatrix}$ deux vecteurs de \mathcal{M} . On a :

$$g(X, Y) = {}^t X \eta Y = \sum_{0 \leq \mu, \nu \leq 3} \eta_{\mu\nu} x^\mu y^\nu = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3$$

En relativité on utilise la convention de sommation d'Einstein qui consiste tout simplement à supprimer le signe \sum et à convenir qu'on somme de 0 à 3 sur tout indice grec répété. Ainsi on écrira :

$$g(X, Y) = \eta_{\mu\nu} x^\mu y^\nu \tag{2.4}$$

et en particulier la forme quadratique $q(X) = g(X, X)$ est donnée par :

⁷Dans les unités géométriques on fixe également la constante de gravitation $G = 1$.

$$q(X) = g(X, X) = \eta_{\mu\nu} x^\mu x^\nu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \quad (2.5)$$

Pour toute formule correcte, un indice répété sur lequel on somme interviendra une fois en haut et une fois en bas. Ceci explique pourquoi on a choisi d'écrire les indices de coordonnées de vecteurs en haut. Avec cette convention, le développement d'un vecteur sur une base s'écrira :

$$\mathbf{v} = v^\mu \mathbf{e}_\mu$$

Ainsi les vecteurs d'une base, ici $(\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ doivent être numérotés avec des indices en bas pour que les coordonnées relativement à cette base aient leurs indices en haut et que la formule de décomposition d'un vecteur sur une base soit « bien équilibrée ».

À des fins qui apparaîtront ultérieurement, il est nécessaire d'aller plus loin dans les développements mathématiques. Une forme linéaire est une application linéaire de \mathcal{M} dans \mathbb{R} . L'ensemble de ces formes linéaires est un espace-vectoriel de même dimension que \mathcal{M} , appelé espace dual et noté \mathcal{M}^* . À la base (\mathbf{e}_μ) vue plus haut, correspond une base duale (ϵ^μ) où la forme ϵ^μ , appelée μ -ème forme coordonnée, est définie par :

$$\epsilon^\mu(\mathbf{v}) = v^\mu$$

Toute forme linéaire ϕ peut se décomposer sur la base duale :

$$\phi = \phi_\mu \epsilon^\mu$$

On peut donc écrire avec les notations d'Einstein :

$$\phi(\mathbf{v}) = \phi_\mu v^\mu$$

De même qu'une forme linéaire est une machine linéaire qui mange un vecteur et recrache un scalaire, une forme bilinéaire, telle que g , est une machine bilinéaire qui mange deux vecteurs et recrache un scalaire. Elle doit donc avoir ses deux indices de coordonnées en bas pour que la formule (2.4) soit bien équilibrée. Mais on peut aussi considérer un endomorphisme de \mathcal{M} : c'est une machine linéaire qui mange un vecteur pour recracher un autre vecteur. Une telle machine, comme par exemple une transformation de Lorentz L , doit donc avoir une coordonnée en haut et une bas de sorte qu'on puisse écrire :

$$L(\mathbf{v}) = L^\nu{}_\mu v^\mu$$

Un vecteur, une forme linéaire, bilinéaire, un endomorphisme : tous ces objets géométriques s'appellent des « tenseurs ». Un tenseur avec k indices en haut et l indices en bas s'appelle un tenseur de type (k, l) : il est dit k fois contravariant et l fois covariant. Cette appellation est liée aux propriétés de transformations de ce tenseur vis-à-vis d'un changement de repère, mais nous ne développerons pas ce point. On dit que $k + l$ est le rang du tenseur. Ainsi η est un tenseur de rang 2 qui est deux fois covariant. Il existe des tenseurs de rang supérieur à deux. Ainsi, un tenseur qui s'écrit :

$$R^\mu{}_{\nu\rho\sigma}$$

est un tenseur de type $(1, 3)$. Géométriquement, on peut dire que c'est une machine linéaire qui mange deux vecteurs et qui recrache un endomorphisme⁸. C'est une forme bilinéaire à valeur endomorphisme. Un tel tenseur comporte 4^4 coordonnées en tout.

On peut multiplier des tenseurs. Ainsi on peut poser :

$$a_{\mu\nu} = \phi_\mu \xi_\nu$$

À gauche on a une forme bilinéaire (une 2-forme), à droite le produit de deux formes linéaires. Cette égalité est effectivement mathématiquement fondée puisque le produit de deux formes linéaires (produit

⁸Mais on pourrait aussi dire que ça mange trois vecteurs et que ça recrache un vecteur.

de fonctions) est bien une forme bilinéaire. On admet qu'on peut ainsi multiplier en général les tenseurs. (Attention, il faut des indices différents pour chaque tenseur que l'on multiplie.)

Enfin, lorsqu'un tenseur comporte un indice en haut et un indice en bas, on peut le « contracter » sur ces deux indices. Par exemple on peut poser, en contractant $R^\mu{}_{\nu\rho\sigma}$ sur le premier et le troisième indice :

$$r_{\nu\sigma} = R^\mu{}_{\nu\mu\sigma}$$

Ceci définit un autre tenseur. On voit qu'en contractant un tenseur de type (k, l) sur deux de ses indices on définit un tenseur de type $(k-1, l-1)$. Un exemple connu est la contraction d'un endomorphisme :

$$\text{Tr}(L) = L^\mu{}_\mu$$

On obtient un scalaire (tenseur de type $(0, 0)$) qui n'est rien d'autre que la trace de l'endomorphisme⁹.

Exercice 2.4.5 Montrer que g est invariante par les transformations orthochrones de Lorentz¹⁰ (rappelons que $c = 1$ dans nos unités), et aussi par renversement du temps.

Théorème 2.4.2 *Les applications linéaires qui conservent g sont exactement les transformations de Lorentz, c'est-à-dire les composées de rotations, de symétries, de boosts, et de renversement du temps. Leur ensemble forme un sous-groupe de $Gl_4(\mathbb{R})$ appelé groupe de Lorentz et noté \mathcal{L} . Si $L \in \mathcal{L}$ alors $\det(L) = \pm 1$.*

Démonstration: Le fait que le sous-ensemble de Gl_4 qui conserve quelque chose soit un groupe est évident. Par ailleurs, une application linéaire L conserve η ssi :

$${}^t L \eta L = \eta \tag{2.6}$$

En prenant le déterminant de chaque côté on obtient $(\det L)^2 = 1$.

En ce qui concerne la première affirmation, on a déjà démontré un sens par l'exercice précédent. Pour la réciproque, il s'agit d'un exercice mathématique que nous laissons au lecteur. Une esquisse de preuve est donnée en appendice. ¶

Exercice 2.4.6 Terminer la démonstration.

Si on regarde l'équation (2.6) en coordonnées, cela donne avec la convention de sommation :

$$\eta_{\alpha\beta} L^\alpha{}_\mu L^\beta{}_\nu = \eta_{\mu\nu}$$

En particulier pour la composante 0,0 de η on obtient :

$$(L^0{}_0)^2 = 1 + (L^1{}_0)^2 + (L^2{}_0)^2 + (L^3{}_0)^2$$

Il en résulte que $(L^0{}_0)^2 \geq 1$, et donc $L^0{}_0 \geq 1$ ou ≤ -1 . On ne pourra donc pas passer continûment d'une transformation de Lorentz L dont la composante 0,0 est positive à une autre L' qui l'aurait négative. De même on ne peut pas cheminer d'une transformation de Lorentz de déterminant égal à 1 à une autre de déterminant égal à -1 . Le groupe \mathcal{L} est ainsi formé de 4 morceaux disjoints, notés \mathcal{L}_+^\uparrow , \mathcal{L}_+^\downarrow , \mathcal{L}_-^\uparrow , et \mathcal{L}_-^\downarrow entre lesquels on ne peut pas cheminer continûment. Le morceau qui contient l'identité est bien sûr \mathcal{L}_+^\uparrow . C'est le seul qui soit un sous-groupe de Gl_4 , et on l'appelle groupe de Lorentz orthochrone propre. Les éléments de \mathcal{L}_+^\uparrow conservent à la fois le sens du temps et l'orientation de l'espace. Ce sont les seuls dont nous aurons vraiment besoin. C'est pourquoi on dira « groupe de Lorentz » au lieu de « groupe de Lorentz orthochrone positif ».

⁹Ainsi, si on interprète $R^\mu{}_{\nu\rho\sigma}$ comme une 2-forme à valeur endomorphisme, le tenseur $r_{\nu\sigma}$ est vu comme la composée avec la trace, c'est-à-dire comme une 2-forme à valeurs scalaires. On peut toujours procéder ainsi et voir les contractions comme des traces.

¹⁰Elle est invariante par le groupe de Lorentz, mais pas par le groupe de Poincaré.

On choisit un référentiel inertiel dont on note les coordonnées (t, x, y, z) . Rappelons qu'avec les unités choisies, $c = 1$. Deux événements M et M' permettent de former le vecteur $\overrightarrow{MM'} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$ avec

$$x^0 = t' - t, \quad x^1 = x' - x, \quad \text{etc.} \dots$$

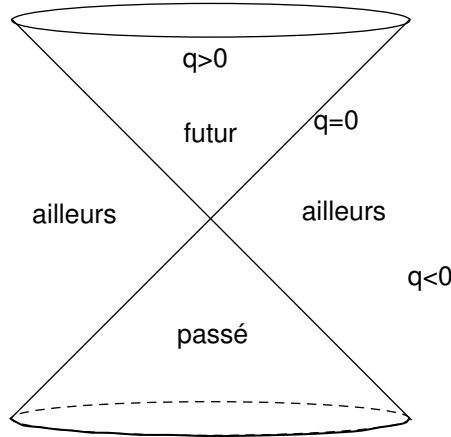
L'intervalle au carré entre les deux événements est donné par :

$$\Delta s^2 = q(\overrightarrow{MM'}) = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$$

Ce Δs^2 est conservé par les transformations de Lorentz de la même façon que la longueur au carré $x^2 + y^2 + z^2$ est conservée par les rotations de l'espace, qui permettent de passer d'un repère orthonormé direct à un autre. Cependant, alors qu'en mécanique classique, le temps et l'espace sont conservés séparément, en relativité restreinte c'est uniquement un mélange des deux qui est conservé et qui constitue la vraie notion géométrique. Pour poursuivre notre analogie avec la géométrie dans l'espace, la quantité x^2 seule n'a pas de sens géométrique et n'est pas conservée par les rotations. Il en va de même pour la quantité $y^2 + z^2$: elle dépend du repère choisi. Cependant, la somme des deux est une quantité géométrique, c'est-à-dire indépendante du repère choisi. On comprend grâce à cette analogie que la distance et la coordonnée de temps n'ont pas de sens géométrique en relativité restreinte : ce sont des quantités qui sont entièrement dépendantes du référentiel.

On dira qu'un 4-vecteur v est du genre temps si $q(v) > 0$, du genre lumière (ou nul) si $q(v) = 0$, et du genre espace si $q(v) < 0$. L'équation des vecteurs du genre lumière :

$$q(v) = 0 \Leftrightarrow (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = 0$$



est l'équation d'un cône en dimension 4, appelé cône de lumière. On le notera \mathcal{C} . Si un événement M correspond à l'émission d'une onde lumineuse, le cône de lumière issu de M est $M + \mathcal{C} = \{M' | \overrightarrow{MM'} \in \mathcal{C}\}$ représente l'ensemble des trajectoires de tous les rayons lumineux issus de M . En particulier, \mathcal{C} est l'ensemble des trajectoires de tous les rayons lumineux issus de l'origine. Si on coupe \mathcal{C} par un hyperplan $H : x^0 = k > 0$ on obtient l'équation d'une sphère dans l'hyperplan H :

$$(x^1)^2 + (x^2)^2 + (x^3)^2 = k^2$$

ce qui donne, si on revient aux unités usuelles :

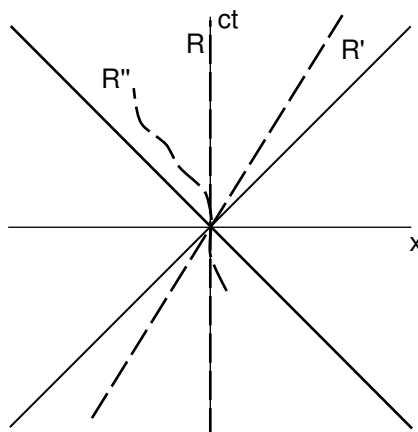
$$x^2 + y^2 + z^2 = c^2 T^2$$

C'est l'équation de la sphère atteinte par la lumière, qu'on appelle aussi le front d'onde, au bout du temps $T = k/c$.

Exercice 2.4.7 Montrer physiquement puis mathématiquement que \mathcal{C} est Lorentz-invariant.

On décompose habituellement \mathcal{C} en deux parties : le cône du futur $\mathcal{C}^+ = \mathcal{C} \cap \{x^0 \geq 0\}$ et le cône du passé $\mathcal{C}^- = \mathcal{C} \cap \{x^0 \leq 0\}$. Comme le groupe de Lorentz (orthochrone) ne contient pas les renversements du temps, ces deux parties sont Lorentz-invariantes.

L'espace de Minkowski permet de visualiser les situations, ce qui est toujours une bonne chose. Ainsi on trace souvent des « diagrammes d'espace-temps », où l'on ne figure que deux coordonnées : ct et x . Dans la figure ci-dessous nous avons tracé le diagramme d'espace-temps d'un observateur inertiel R . Sa ligne d'univers, autrement dit sa trajectoire spatio-temporelle n'est rien d'autre que l'axe du temps. Nous avons représenté la ligne d'univers d'un autre observateur inertiel R' qui rencontre R à l'origine : c'est une droite de genre temps. Toute droite inclinée à plus de 45 degrés par rapport à l'axe du temps représenterait un objet se propageant plus vite que la lumière, ce qui est interdit. La ligne courbe représente un observateur R'' qui est accéléré, et donc non-inertiel. Notons que le vecteur tangent de la ligne d'univers de R'' doit partout être du genre temps pour que cette trajectoire corresponde au mouvement autorisé d'un corps massif.



Si on ignore les coordonnées y et z , le boost peut s'écrire sous la forme matricielle suivante :

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma v/c \\ -\gamma v/c & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \quad (2.7)$$

2.5 Effets et « paradoxes » en relativité restreinte

2.5.1 Dilatation des durées de vie des particules instables

Exercice 2.5.1 Une particule instable parcourt dans le laboratoire la distance $l = 5,19 \text{ m}$ en $\Delta t = 20 \times 10^{-9} \text{ s}$ avant de se désintégrer. Quelle est sa vitesse v , supposée uniforme ? En déduire sa durée de vie propre τ_0 . Connaissant τ_0 , quelle prédiction ferait la mécanique classique pour la distance l parcourue à la vitesse v par la particule avant de se désintégrer ?

Cet exercice montre qu'on peut vérifier directement la dilatation du temps en laboratoire en utilisant des particules instables. Cet effet est en fait vérifié quotidiennement dans les accélérateurs de particules.

Dans le même ordre idée, la dilatation du temps a permis d'expliquer pourquoi on détectait des muons produits dans la haute atmosphère au niveau du sol alors que selon un calcul classique, ils auraient dû se désintégrer en vol. Pour être plus précis, on devrait dire qu'on en observait plus que l'on aurait dû au

niveau du sol, puisque la loi de désintégration est une loi de probabiliste. Il convient donc de raisonner sur la durée de vie moyenne des particules. Cette expérience fut réalisée en 1941 par Rossi et Hall. Il trouvèrent un excellent accord avec la prédiction de la relativité restreinte.

2.5.2 Paradoxe des jumeaux de Langevin

C'est sûrement le plus célèbre « paradoxe » de la relativité restreinte. On considère deux jumeaux que nous nommerons Alice et Bob. Alice reste sur Terre tandis que Bob effectue un aller-retour vers une planète située à 9 années-lumière de la Terre. Il effectue son voyage aller et son voyage retour à la vitesse uniforme de $9/10$ de c .

Considérons d'abord les choses du point de vue d'Alice. La durée du voyage est, par définition, de $2 \times 9 \div \frac{9}{10} = 20$ ans.

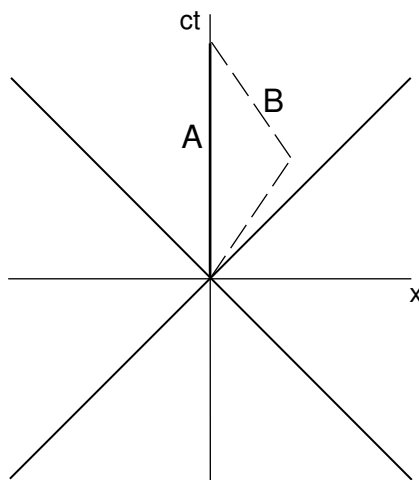
En appliquant la formule des transformations spéciales de Lorentz elle trouve que la durée du voyage aller mesurée par Bob est :

$$T_a = \gamma^{-1} \times 9 \simeq 3,9 \text{ ans}$$

Exercice 2.5.2 Vérifiez de deux façons différentes cette formule.

Il en va de même pour la durée du voyage de retour, et par conséquent, d'après Alice, lorsque Bob reviendra il aura vieilli d'environ 8 ans tandis qu'elle, aura vieilli de 20 ans. C'est curieux, mais le paradoxe n'est pas là !

En effet, pourquoi ne pas regarder les choses du point de vue de Bob ? Selon lui c'est Alice qui voyage. Et par symétrie, il trouve que c'est elle qui a le plus vieilli lorsqu'ils se retrouvent ! Il y a là une contradiction. Qu'en est-il en réalité ? En réalité les deux situations ne sont pas symétriques, ce qu'on voit bien en faisant un diagramme d'espace-temps.



Même si on néglige le problème de l'accélération (qui est infinie lorsque Bob rebrousse chemin, mais qu'on peut rendre finie en arrondissant le point anguleux de sa trajectoire), on voit que Bob utilise deux référentiels inertiels différents, tandis qu'Alice n'en utilise qu'un seul. On ne peut donc pas faire le calcul « dans le référentiel de Bob », car ce dernier n'est pas un seul et même référentiel d'inertie pendant toute la durée du voyage.

Signalons que cet effet a été mesuré à l'aide de deux horloges atomiques. Les résultats ont montré un très bon accord avec la théorie de la relativité restreinte.

L'effet Langevin illustre le fait que le temps mesuré par un observateur en mouvement dépend de sa trajectoire. Ce temps, qu'on appelle temps propre, est dans le cas de Bob :

$$\Delta\tau = \Delta\tau_{aller} + \Delta\tau_{retour} = \sqrt{1 - \frac{v_{aller}^2}{c^2}} \Delta t_{aller} + \sqrt{1 - \frac{v_{retour}^2}{c^2}} \Delta t_{retour}$$

On peut généraliser à une trajectoire en ligne brisée, formée de n morceaux parcourus à vitesse rectiligne uniforme. La formule est alors :

$$\Delta\tau = \sum_{i=1}^n \sqrt{1 - \frac{v_i^2}{c^2}} \Delta t_i$$

où v_i est la vitesse sur le i -ème morceau, et Δt_i le temps de parcours de ce morceau, exprimé dans le référentiel R . On somme donc l'intervalle de temps qui serait mesuré par une horloge emportée par le mouvement sur chaque intervalle.

En faisant tendre vers zéro les intervalles Δt_i , on définit pour toute courbe rectifiable¹¹ :

$$\tau - \tau_0 = \Delta\tau = \int_{t_0}^{t_1} \sqrt{1 - \frac{v(t)^2}{c^2}} dt \quad (2.8)$$

Cette formule définit le temps propre τ au choix arbitraire d'une origine τ_0 près. C'est le temps mesuré par une horloge emportée par le mouvement.

Une autre expression pour le temps propre est possible. Il suffit en effet d'utiliser l'intervalle infinitésimal au carré :

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = \eta_{\mu\nu} dx^\mu dx^\nu$$

En factorisant on obtient :

$$ds^2 = c^2 dt^2 \left(1 - \frac{dx^2 + dy^2 + dz^2}{c^2 dt^2}\right) = c^2 dt^2 \left(1 - \frac{v^2}{c^2}\right)$$

où v est la vitesse instantanée. Si la courbe est telle que sa tangente est toujours du genre temps, alors le ds^2 sera positif. On aura alors :

$$d\tau := \frac{1}{c} ds = \sqrt{1 - \frac{v^2}{c^2}} dt$$

d'où en intégrant :

$$\tau - \tau_0 = \int \sqrt{1 - \frac{v^2}{c^2}} dt = \frac{1}{c} \int ds \quad (2.9)$$

La seconde formule permet de calculer τ pour une courbe qui ne serait pas paramétrée par le temps coordonnée t .

La quantité ds quant à elle, qui a la dimension d'une longueur, est égale au temps propre si on utilise les unités géométriques. On voit qu'il s'agit d'un « élément de longueur spatio-temporel », généralisant l'élément euclidien $dl = \sqrt{dx^2 + dy^2 + dz^2}$ qui sert à mesurer la longueur des courbes en géométrie euclidienne. Il y a toutefois une différence importante, c'est que ds n'est défini que si $ds^2 \geq 0$. Si deux événements A et B sont causalement reliés, on peut alors définir la distance minkowskienne $d(A, B)$ par :

$$d(A, B) = \int_{[AB]} ds$$

Exercice 2.5.3 En généralisant le paradoxe des jumeaux, montrez que si ABC est un triangle d'événements deux à deux causalement reliés, alors :

$$d(A, C) \geq d(A, B) + d(B, C)$$

¹¹C'est-à-dire toute courbe dont on peut mesurer la longueur en l'approchant par des lignes brisées. En particulier toute courbe possédant un vecteur vitesse à chaque instant est rectifiable.

Cet exercice montre que la distance minkowskienne vérifie une inégalité triangulaire exactement inverse de l'inégalité usuelle pour les distances spatiales¹².

À l'aide de cette pseudo-distance, on peut dire que si A et B sont causalement reliés, le segment $[AB]$ est la courbe du genre temps ou du genre lumière de **longueur maximale** qui les relie. Le segment $[AB]$ est donc tel que la quantité $\int_{[AB]} ds$ soit extrémale. Une telle courbe s'appelle une géodésique. Plus précisément, une géodésique est une courbe telle que pour chacun de ses points A, B , on ait :

$$\delta \int ds = 0$$

où le δ signifie qu'on considère la variation au premier ordre (c'est une sorte de dérivée), et l'intégrale s'étend sur l'arc de courbe entre A et B . On peut montrer que les segments du genre lumière vérifient également cette condition. Nous avons ainsi obtenu :

Proposition 2.5.1 *Les géodésiques de l'espace-temps de Minkowski sont les droites du genre temps ou du genre lumière.*

Dans l'espace euclidien \mathbb{R}^3 , les géodésiques sont aussi les droites (toutes les droites), et elles réalisent le minimum de la distance entre deux quelconques de leurs points. Sur une sphère, les géodésiques sont les grands cercles. Un arc de grand cercle réalise le minimum de la distance entre ses extrémités si son ouverture est inférieure à π . Nous reviendrons sur cette notion de géodésique plus tard.

Remarque : On peut définir des géodésiques plus générales dans l'espace de Minkowski. Ce sont les courbes telles que

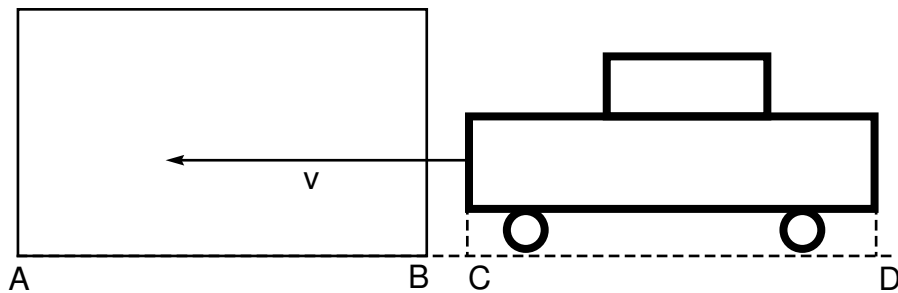
$$\delta \int \sqrt{|ds^2|} = 0$$

En plus des segments de droites du genre temps ou lumière, on trouve ceux du genre espace. Ces derniers ne correspondent pas au mouvement d'un corps matériel.

Exercice 2.5.4 Soient deux horloges de fabrication identique et synchronisées. On emporte l'une dans un voyage à vitesse constante $v = 1000$ km/h autour de la Terre, tandis que l'autre reste immobile. En supposant inertiel le référentiel géocentrique, que constatera-t-on en comparant les deux horloges à la fin du voyage ? Faire l'application numérique.

2.5.3 Le problème de la voiture et du garage

Imaginez une voiture (segment $[CD]$) qui rentre exactement en longueur dans une place de garage (segment $[AB]$ de même longueur au repos que $[CD]$). Il y a une porte à l'avant et une à l'arrière du garage. On ouvre ces portes, et on fait arriver la voiture dans le garage à très grande vitesse. Est-ce qu'on peut fermer les portes du garage de sorte que la voiture soit entièrement à l'intérieur de celui-ci (pour un bref instant...)?



¹²Ce n'est donc pas une distance au sens de la théorie mathématique des espaces métriques.

Du point de vue du garage, la voiture subit une contraction de Lorentz, donc elle rentre facilement. La réponse est donc oui.

Du point de vue de la voiture, la réponse est non ! En effet, c'est le garage qui subit une contraction, il est donc trop petit !

Exercice 2.5.5 Quelle est la bonne réponse ? (Indication, faire un diagramme d'espace-temps pour comprendre où est la difficulté.)

2.6 Dynamique relativiste

2.6.1 Équivalence masse-énergie, quadrivecteur impulsion-énergie

En physique classique la masse est une propriété d'une particule qui ne dépend pas du référentiel de l'observateur. Admettons simplement qu'en relativité la masse devienne dépendante de la vitesse de la particule par rapport au référentiel inertiel de référence, et cherchons son expression $m(v)$.

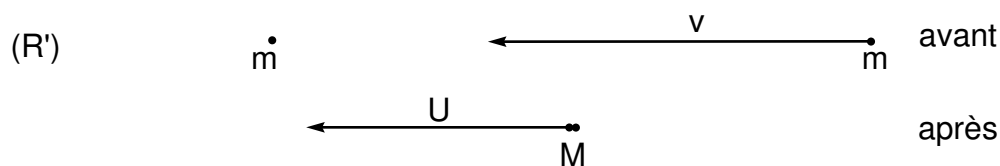
Pour cela considérons deux particules identiques. Dans le référentiel R du centre de masse, elles se déplacent l'une vers l'autre à la même vitesse (en valeur absolue), qu'on notera U . Supposons que le choc soit inélastique : l'objet constitué des deux particules est alors au repos après le choc. On notera m la masse des particules et M la masse de l'objet combiné.



Dans le référentiel R' de l'une des deux particules, l'autre se déplace vers elle à la vitesse uniforme v , que nous obtenons par la formule d'addition relativiste des vitesses :

$$v = \frac{2U}{1 + \frac{U^2}{c^2}}$$

L'objet combiné est lui animé d'une vitesse U après le choc.



En cherchant U en fonction de v on obtient :

$$U^2 - \frac{2c^2}{v}U + c^2 = 0$$

On trouve alors :

$$U = \frac{c^2}{v} \left(1 \pm \sqrt{1 - \frac{v^2}{c^2}} \right)$$

Quand v tend vers 0, l'expression avec un $+$ tend vers l'infini, ce qui est absurde. On doit prendre :

$$U = \frac{c^2}{v} \left(1 - \sqrt{1 - \frac{v^2}{c^2}}\right) \quad (2.10)$$

Nous admettons la conservation de la masse relativiste $m(v)$ et de la quantité de mouvement définie par $p = m(v)v$. (C'est-à-dire qu'on cherche $m(v)$ tel que ces deux quantités se conservent lors d'un choc inélastique.)

On a alors dans le référentiel R' :

$$m(v) + m(0) = M(U) \quad (2.11)$$

et

$$m(v)v + 0 = M(U)U \quad (2.12)$$

d'où l'on déduit

$$m(v) = m(0) \frac{U}{v - U} \quad (2.13)$$

En remplaçant par l'expression (2.10), on obtient après un peu d'algèbre :

$$m(v) = \gamma m_0 \quad (2.14)$$

où l'on a posé $m_0 = m(0)$. C'est la masse d'un corps dans un référentiel comobile¹³. On l'appelle masse au repos ou masse propre.

Si on fait un développement limité de $m = m(v)$ en v/c au voisinage de 0 on trouve :

$$m = m_0 + \frac{1}{c^2} \left(\frac{mv^2}{2}\right) + \mathcal{O}\left(\frac{v^4}{c^4}\right)$$

En multipliant par c^2 on obtient

$$mc^2 = \gamma m_0 c^2 = m_0 c^2 + \frac{1}{2} m v^2 + c^2 \mathcal{O}\left(\frac{v^4}{c^4}\right) \quad (2.15)$$

Ainsi on a un terme constant, puis l'expression classique de l'énergie cinétique. Cette dernière est donc une composante de la masse relativiste, au facteur de conversion $1/c^2$ près. Ceci unit la masse et l'énergie, de façon totalement inédite par rapport à la mécanique classique. L'énergie mécanique classique d'une particule, hors de tout potentiel, étant égale à son énergie cinétique plus une constante arbitraire, on peut très bien poser :

$$mc^2 = E_{classique} + \text{termes correctifs tendant vers 0 pour les petites vitesses}$$

Ceci suggère qu'en relativité l'énergie d'une particule soit définie par :

$$E = mc^2 = \gamma m_0 c^2 \quad (2.16)$$

Plusieurs remarques s'imposent. Tout d'abord il y a une certaine part d'arbitraire dans la méthode que nous avons suivie. Ce à quoi nous pouvons répondre que nous sommes en train de découvrir une nouvelle mécanique, et qu'il n'y a pas de recette pour cela. Nos deux seuls guides sont la cohérence avec les postulats de départ, et la conformité avec l'expérience. On ne peut pas arriver aux lois de la dynamique relativiste à partir des deux seuls postulats et du raisonnement pur. Il serait possible d'explicitier les hypothèses supplémentaires que nous devons faire, mais ceci est fort long et compliqué. À ce sujet on pourra consulter [3] et les références qui y sont citées. Contentons nous d'admettre que nos formules sont valables dans un cadre beaucoup plus général que celui dans lequel on les a obtenues, ce qu'on peut

¹³On parle aussi parfois de référentiel propre.

montrer par un raisonnement plus complet ou par l'expérience. Par exemple, ces relations s'appliquent à tous les corps, et ne dépendent pas de leur structure interne. La deuxième remarque est que la formule $E = mc^2$ donne par différence :

$$\Delta E = \Delta mc^2 \quad (2.17)$$

Ceci montre qu'un système qui change de masse change d'énergie, et réciproquement. Cela est clair si ce système acquiert ou perd de la vitesse, et donc de l'énergie cinétique, par la relation (2.15). Mais c'est aussi vrai s'il change de masse au repos, ce qui est le cas dans certaines réactions nucléaires. Ainsi la théorie de la relativité restreinte fait entrevoir la possibilité théorique de convertir de la masse au repos en énergie, mais ne donne aucune indication sur les moyens pour y parvenir.

Enfin, une question : à quelle masse « classique » la masse relativiste correspond-elle ? Est-ce la masse inertielle, qui s'oppose à l'accélération d'un corps ? Est-ce la masse grave, qui entre dans la formule de l'attraction des corps ? La réponse est oui aux deux questions. On peut s'en rendre compte qualitativement pour la première par le raisonnement suivant, que l'on affinera un peu plus tard : si on fait tendre v vers c , la masse relativiste tend vers l'infini. Il est donc de plus en plus difficile d'accélérer un corps dont la vitesse tend vers c , et la formule $a = F/m$ montre que cette accélération finit par s'annuler, ce qui explique dynamiquement pourquoi on ne peut pas faire aller un corps plus vite que la lumière. Pour la seconde question, l'explication viendra seulement avec la théorie de la relativité générale.

Nous avons défini plus haut la quantité de mouvement relativiste $p = mv = \gamma m_0 v$. On l'appelle également impulsion. La quantité vectorielle associée est naturellement $\vec{p} = m\vec{v}$, mais pour avoir une quantité géométrique, c'est-à-dire une quantité qui se transforme par une transformation de Lorentz quand on change de référentiel inertielle, il faut lui associer une quatrième composante. On pose :

$$P = \begin{pmatrix} E/c \\ \vec{p} \end{pmatrix} = \begin{pmatrix} E/c \\ mv_x \\ mv_y \\ mv_z \end{pmatrix} \quad (2.18)$$

On appelle P le quadrivecteur énergie-impulsion. La raison pour laquelle on écrit E/c plutôt que mc est que cette dernière quantité est indéfinie pour les particules de masse propre nulle, telles que les photons, tandis que E/c a toujours un sens. En effet, pour ces derniers E s'obtient par la relation $E = h\nu$, issue de la physique quantique.

Exercice 2.6.1 Montrer que P est bien un quadrivecteur. Pour cela, il est nécessaire et suffisant de montrer que si R et R' sont deux référentiels inertiels, et si P et P' sont définis par (2.18) dans leurs référentiels respectifs, alors $P' = \Lambda P$, où Λ est la transformation de Lorentz permettant de passer R à R' . On reconnaît ici la formule permettant de changer les coordonnées d'un quadrivecteur lorsqu'on change de référentiel, ce qui montre que P définit bien un quadrivecteur.

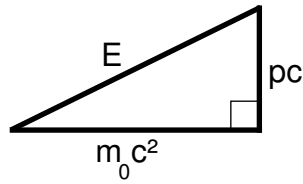
Il suit de l'exercice précédent que $q(P)$, la pseudo-norme au carré de P , est invariante de Lorentz. Elle vaut :

$$\eta_{\mu\nu} P^{\mu\nu} = \frac{E^2}{c^2} - p^2 = \gamma^2 m_0^2 c^2 - \gamma^2 m_0^2 v^2 = m_0^2 c^2$$

d'où la relation fort utile :

$$E^2 = p^2 c^2 + m_0^2 c^4 \quad (2.19)$$

qui peut se retenir en appliquant Pythagore dans le triangle ci-dessous.



2.6.2 Le PFD en relativité restreinte

Le principe d'inertie reste valable en relativité restreinte. Par conséquent, pour une particule libre, l'impulsion \vec{p} est constante. Toute variation d'impulsion peut être attribuée à l'action d'une force de sorte que l'on ait :

$$\vec{f} = \frac{d\vec{p}}{dt}$$

C'est la version relativiste du PFD. Néanmoins il faut connaître sa loi de transformation par un changement de référentiel inertiel. Pour cela formons le quadrivecteur force, qui est par définition :

$$F = \frac{dP}{d\tau} \tag{2.20}$$

Comme P est un quadrivecteur et $d\tau$ un invariant, F est bien un quadrivecteur. Il se transforme donc de Lorentz, ce qui permet d'obtenir par identification les lois de transformation de \vec{f} .

Exercice 2.6.2 Montrer que les composantes de F sont $(\frac{\gamma}{c} \vec{f} \cdot \vec{v}, \gamma \vec{f})$.

La remarque importante ici est que la force n'est pas Lorentz-invariante, tandis qu'en mécanique Newtonienne, la force est invariante par un changement de référentiel inertiel. En particulier si on a une relation telle que $\vec{f}_1 + \vec{f}_2 = \vec{0}$, elle ne sera plus vraie en général après changement de référentiel. Par conséquent, le principe de l'action et de la réaction n'est pas valable en relativité, car il ne peut même pas être exprimé de façon sensée!

Pour plus de détails et d'exercices sur la dynamique relativiste, on pourra consulter [3].

Chapitre 3

La relativité générale

3.1 Motivation

Plusieurs indices conduisent à la nécessité d'aller plus loin que la relativité restreinte dans la refonte de la mécanique. Le premier est tout simplement que, contrairement aux équations de Maxwell, les équations du champ de gravitation ne sont compatibles avec les transformations de Lorentz, et doivent donc être modifiées. Plusieurs chercheurs se sont d'ailleurs attaqués à ce problème en même temps qu'Einstein, en suivant cette piste.

Mais ce dernier était motivé par un deuxième indice qui est un principe d'élégance et de simplicité : pourquoi conserver un statut spécial aux référentiels d'inertie ? S'il est vrai que le mouvement uniforme est relatif, n'est-il pas vrai également que le mouvement en général l'est ? Imaginons que dans l'univers il n'y ait que la Terre et le Soleil. Qu'est-ce qui nous permettrait de dire que c'est la Terre et non le Soleil qui tourne ? Si l'on veut donner un sens au mouvement absolu, il faut qu'il y ait un espace absolu, un cadre fixe, par rapport auquel mesurer ce mouvement. Mais ce cadre, qu'on a cru trouver un moment avec l'éther, s'est évanoui. De toute façon, Einstein, fortement influencé en cette matière par les idées de Mach, considère l'espace absolu comme une absurdité philosophique. Il se trouve que ce reproche conceptuel fut adressé en son temps à Newton, qui y répondit par une expérience de pensée célèbre. Considérons un seau d'eau en rotation. La surface de l'eau s'incurve. Dans le référentiel du seau, c'est l'univers qui tourne, et l'eau n'a aucune raison de s'incurver. Pourtant elle s'incurve ! C'est donc bien que les deux référentiels ne sont pas équivalents : on peut affirmer que le seau tourne. Mais s'il tourne c'est bien par rapport à quelque chose ! Enlevons les étoiles et les planètes du ciel, et considérons le seau en rotation tout seul dans l'univers. D'après Newton, le seau tourne par rapport à l'espace absolu, et l'eau s'incurve. Mach a critiqué ce point de vue : d'après lui, si le seau est seul dans l'univers, il n'y a pas de sens à dire qu'il tourne, et l'eau ne doit pas s'incurver. Par conséquent, c'est l'influence des masses distantes de l'univers qui fait s'incurver l'eau. Autrement dit, pour Mach, l'origine des forces d'inertie est dans l'action gravitationnelle de l'univers tout entier. C'est un deuxième indice qu'il pourrait exister un lien entre inertie et gravitation, et qu'une théorie qui cherche à mettre tous les observateurs sur un pied d'égalité doit aussi s'occuper de la gravitation. Enfin le troisième indice, qui pointe lui aussi vers un lien entre inertie et gravitation, est le plus crucial. Nous y consacrons le paragraphe suivant.

3.2 Le principe d'équivalence

« C'est alors, considérant ces faits, qu'il me vint à l'esprit que si l'on supprimait totalement la résistance du milieu, tous les corps descendraient avec la même vitesse. » (Galilée)

Cette observation de Galilée, tout étudiant a pu la vérifier : dans un tube à vide, la plume et la bille de plomb tombent exactement en même temps. Ce fait prend, dans le cadre de la mécanique classique, la

forme d'une coïncidence : la masse grave est égale à la masse inerte. La masse grave est celle qui apparaît dans la loi d'attraction des corps :

$$F = G \frac{M m_g}{r^2}$$

où M est la masse de la Terre, disons, et m_g la masse grave d'une particule test, par exemple une petite bille de plomb. Mais si l'on veut calculer la trajectoire de la bille, il faut appliquer le PFD :

$$F = m_i a$$

où cette fois m_i est la masse inerte de la bille : c'est ce qui mesure sa résistance au mouvement. Il n'y a a priori aucun rapport entre ces deux masses. Par exemple, si l'on considère la loi de Coulomb :

$$F = \frac{1}{4\pi\epsilon_0} \frac{Qq}{r^2}$$

on voit que la masse grave est remplacée par la charge de la particule test. Ainsi, si nous cherchons l'accélération de la bille, nous obtenons dans le premier cas :

$$a = GM \frac{m_g}{m_i} \frac{1}{r^2}$$

et dans le second

$$a = \frac{Q}{4\pi\epsilon_0} \frac{q}{m_i} \frac{1}{r^2}$$

Ces deux formules sont similaires, mais ce n'est que parce qu'on a, empiriquement :

$$m_i = m_g$$

que la première se simplifie et devient :

$$a = GM \frac{1}{r^2}$$

On voit qu'aucune propriété de la bille n'intervient dans cette formule, autrement dit, tous les corps tombent de la même façon dans le champ de gravitation, ce qui n'est pas du tout le cas dans un champ électrique. Répétons-le, cette constatation vient de la coïncidence :

$$m_i = m_g$$

qui n'a aucune explication en mécanique classique. Cependant, elle est très bien vérifiée expérimentalement (la précision est de l'ordre de 10^{-12} aujourd'hui). Une telle coïncidence semble curieuse, et signale que quelque chose ne va pas dans la théorie. On n'a peut-être pas pris les choses par le bon bout... La méthode d'Einstein va ici encore consister à changer de point de vue, et à considérer l'observation de Galilée comme un point de départ, un principe premier, ou postulat, c'est le principe d'équivalence.

Principe d'équivalence : « Le mouvement d'une particule test dans un champ de gravitation est indépendant des propriétés (masse, composition chimique, etc...) de cette particule. »

Einstein va appliquer ce principe sous une forme particulière.

3.3 « L'inspiration la plus heureuse de ma vie »

« J'étais assis sur une chaise au bureau des brevets à Berne quand, soudain, il me vint une idée : lorsqu'un homme est en chute libre, il ne ressent pas son propre poids. J'étais épaté. Cette expérience de pensée toute simple me fit grande impression et me conduisit à la théorie de la gravitation. » (Albert Einstein)

Imaginons-nous à l'intérieur d'un ascenseur en chute libre. Nous sommes alors en état d'apesanteur, exactement comme un astronaute dans une capsule spatiale en orbite. Si nous lâchons un objet sans vitesse initiale, celui-ci reste en place. Autrement dit, un observateur qui serait né dans cet ascenseur et n'aurait jamais eu accès au monde extérieur, pourrait très bien formuler le principe d'inertie et découvrir les lois de la mécanique, et il le pourrait avec plus de facilité même que nous n'en avons eu sur Terre. L'idée d'Einstein est d'étendre le principe de relativité à cet observateur, et même à n'importe quel observateur.

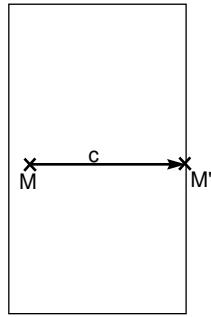
Principe de relativité générale : « La forme des lois de la physique est la même dans tous les référentiels. »¹

Ce principe nous impose de donner une forme particulière, dite « généralement covariante », aux lois de la physique². Par ailleurs, ce principe semble contradictoire avec l'existence des forces d'inertie. Comment Einstein fait-il pour expliquer ces dernières ? L'expérience de l'ascenseur montre qu'on peut faire disparaître (localement) le champ de gravité en utilisant un observateur en chute libre. Mais on peut aussi faire apparaître un champ de gravité ! En effet, transportons notre ascenseur loin de tout corps. Le champ de gravité est alors négligeable. Imprimons une accélération uniforme $-g$ à notre ascenseur à l'aide d'un câble. À l'intérieur de l'ascenseur, l'observateur se retrouve plaqué au sol, exactement comme il le serait à la surface de la Terre sous l'action de son poids. Tout se passe comme si on venait « d'allumer » un champ de gravitation. Aucune expérience (locale) ne pourra distinguer les deux situations de l'intérieur de l'ascenseur. Autrement dit, rien ne permet de distinguer (localement) entre un champ de gravitation et une force d'inertie. L'existence ou non d'un champ de gravitation est uniquement une question de choix de référentiel ! Il va donc falloir formuler la théorie de sorte que le champ de gravitation ne soit pas vraiment une force, mais une propriété mathématique de l'espace-temps. Nous préciserons plus tard ce qui n'est pour l'instant qu'une intuition vague, et nous expliquerons pourquoi nous insistons tellement sur le caractère local des expériences de pensée précédentes.

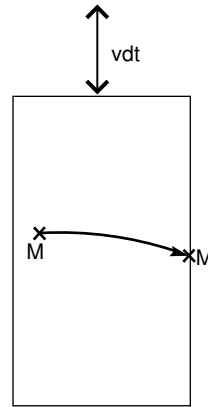
Nous ne pouvons pour l'instant rien calculer, mais nous pouvons quand même donner quelques conséquences qualitatives des principes que nous venons d'énoncer. Revenons à notre observateur dans l'ascenseur en chute libre à la surface de la Terre, et munissons-le d'une lampe. Qu'observe-t-il ? Rien de particulier, la lumière doit se propager en ligne droite pour satisfaire le principe de relativité générale. Mais pour qu'il en soit ainsi, elle doit se propager selon une courbe du point de vue d'un observateur terrestre qui n'est pas en chute libre et pour lequel existe un champ de gravitation ! Autrement dit, le champ de gravitation doit courber les rayons lumineux.

¹On voit que la théorie de la relativité est bien mal nommée, puisque son principe fondamental affirme une universalité !

²Nous laissons ici de côté les phénomènes quantiques.



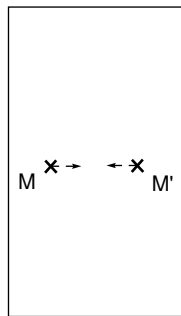
Dans le référentiel de l'ascenseur, la lumière se propage en ligne droite



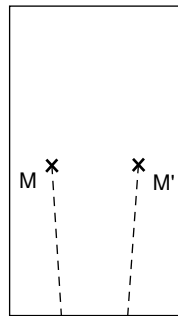
Dans le référentiel terrestre, le point M' a chuté de vdt pendant que la lumière se propageait de M à M' .

Allons plus loin. Pour l'observateur en chute libre, l'espace-temps est celui de Minkowski, et la lumière s'y déplace selon des géodésiques du genre lumière. Une petite bille de plomb s'y déplace selon une géodésique du genre temps³. L'intérêt de cette formulation est qu'elle est géométrique : elle ne dépend pas d'un choix particulier de coordonnées. Nous pouvons alors avoir l'intuition qu'une telle formulation a déjà la forme « généralement covariante » que l'on cherche et va donc aussi être valable pour l'observateur terrestre. Si c'est le cas, alors pour ce dernier, les géodésiques ne seront plus des lignes droites, autrement dit l'espace-temps n'est plus celui de Minkowski, il est courbé. C'est donc que les effets du champ de gravitation se manifestent par une courbure de l'espace-temps.

Voyons une autre expérience qui appuie cette intuition. Si nous lâchons sans vitesse non pas une mais deux petites billes, séparées par une petite distance d , que voyons-nous ? Depuis l'extérieur de l'ascenseur, on voit que les deux billes convergent vers le centre de la Terre. Ainsi, dans l'ascenseur on doit voir les deux billes se rapprocher l'une de l'autre.



Dans le référentiel de l'ascenseur, les deux billes se rapprochent l'une de l'autre.



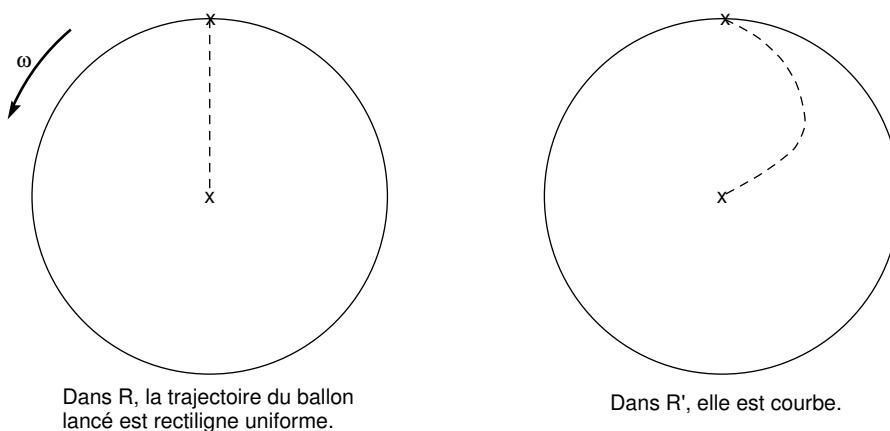
Dans le référentiel terrestre, elles convergent vers le centre de la Terre.

³Ne perdons pas de vue que nous parlons ici de la trajectoire dans l'espace-temps, c'est-à-dire de la ligne d'univers de la bille.

Exercice 3.3.1 Écrire l'équation différentielle vérifiée par la distance x entre les deux billes en utilisant la théorie Newtonienne.

Si on trace un diagramme d'espace-temps dans le référentiel de l'ascenseur, les lignes d'univers des deux billes sont des courbes convergeant l'une vers l'autre. Ceci est une autre manifestation de la courbure de l'espace-temps, qu'on appelle « déviation géodésique » et sur laquelle nous reviendrons. Notons que ceci semble contredire notre point de départ, à savoir que le référentiel de l'ascenseur est Lorentzien ! Il ne l'est en effet qu'en première approximation. Si l'ascenseur est caractérisée par une taille⁴ ϵ , son référentiel est Lorentzien au premier ordre en ϵ . Nous dirons qu'il est localement Lorentzien. La convergence des géodésiques des deux billes est un effet du second ordre en ϵ .

Voyons une dernière expérience de pensée pour laquelle nous imaginons un disque en rotation uniforme, une sorte de manège⁵. Ce manège est très loin de tout corps, il est isolé. Nous devons dire par rapport à quoi s'effectue la rotation : elle s'effectue par rapport à un observateur extérieur dont nous noterons R le référentiel, tandis que le référentiel lié au manège sera noté R' .



Comme R est inertiel, un objet lâché sans vitesse y demeure immobile. Il n'en va pas de même dans R' ! Un tel objet décrira une courbe. Si on lance un ballon vers le centre du cercle, il décrira une trajectoire incurvée, et si l'on veut vraiment atteindre le centre il faudra faire preuve d'une certaine habileté.

Exercice 3.3.2 Quelle est la courbe tracée par le ballon dans R' ? (Répondre en supposant que la vitesse de rotation est toujours petite devant c , et en utilisant la cinématique classique.)

Enfin, tout habitant de ce manège pourra ressentir une force le poussant vers l'extérieur, et s'il veut tenir debout, devra compenser en se penchant vers le centre. Mais le principe de relativité générale est très démocratique : les habitants de R' ont le droit d'y faire de la physique et de découvrir les mêmes lois de la Nature que celle qui ont lieu dans R . La différence est que ce que R appelle force d'inertie, les habitants de R' l'appelleront par un autre nom, comme par exemple « gravitation ». Après tout, la force qu'ils ressentent et qui les pousse vers l'extérieur n'est pas très différente de celle qui nous plaque au sol et que nous appelons « pesanteur »⁶. En effet, les corps se meuvent tous de la même façon dans ce champ de force. La seule différence est que la force est de plus en plus grande à mesure que l'on s'éloigne du centre, alors que sur Terre, c'est l'inverse.

En faisant d'autres expériences, les habitants du manège découvriront un phénomène très étrange. En mesurant le périmètre d'un cercle, ils se rendront compte que celui-ci n'est pas égal à 2π fois le rayon ! En effet, considérons un cercle centré sur le centre du manège. Utilisons la relativité restreinte, qui est

⁴spatio-temporelle, dans les unités géométriques

⁵Nous négligeons ici le fait déjà mentionné que la notion de solide n'a pas cours en relativité.

⁶Les agences spatiales envisagent d'ailleurs, pour les futurs très longs vols, d'induire une telle pesanteur artificielle en imprimant aux véhicules un mouvement de rotation. Stanley Kubrick a utilisé cette idée dans « 2001, l'odyssée de l'espace ».

valable dans R . Appelons d le diamètre de ce cercle et p son périmètre, le tout dans R . Comme d est constamment perpendiculaire au mouvement, il ne subit par de contraction de Lorentz. Le périmètre du cercle⁷ est donc égal dans R à $p = \pi d$.

Considérons maintenant les choses dans R' . On sait que $d' = d$, mais bien sûr nous ne pouvons pas utiliser les transformations de Lorentz pour calculer p' , puisque R' n'est pas inertiel. Néanmoins, imaginons que les habitants de R' utilisent une petite règle de 1 cm (mesuré radialement) pour évaluer le périmètre du cercle. Vue de R , cette règle est contractée lorsqu'on la tient le long du cercle. Les habitants de R' devront donc la reporter un plus grand nombre de fois que « la normale », et ils en déduiront que $p' > \pi d'$. Ils pourront alors conclure... que la géométrie de leur monde n'est pas euclidienne ! C'est une géométrie courbée, et la seule origine possible de cette courbure est leur champ de « gravitation ».

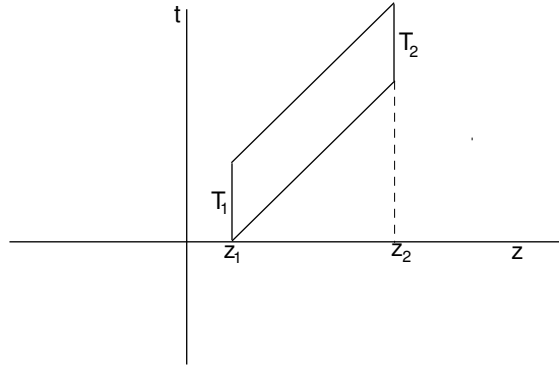
Voyons un autre phénomène. Si on place une horloge à une certaine distance du centre du manège, on constate en utilisant R qu'elle sera ralentie par rapport à une même horloge placée au centre. Ce ralentissement relatif entre les deux horloges sera perçu par les habitants de R' comme un effet du champ de gravitation. De plus, plus l'horloge est loin du centre, plus elle est ralentie. En utilisant l'équivalence locale entre champ de gravitation et référentiel accéléré, nous pouvons généraliser cet effet à n'importe quel champ de gravitation, et nous obtenons une prédiction qualitative importante : la dilatation du temps par le champ gravitationnel⁸.

L'analyse précise des référentiels en rotation uniforme est assez complexe. Nous avons repris ici essentiellement l'argument d'Einstein (voir [6]). Notons que le calcul de la circonférence d'un cercle en rotation a donné lieu à des débats houleux. Selon Ehrenfest, celui-ci devait subir une contraction de Lorentz (problème connu sous le nom de paradoxe d'Ehrenfest), alors que nous avons vu que pour Einstein, il n'en est rien. Néanmoins, il semble que la lecture de l'article d'Ehrenfest ait mis Einstein sur la voie de la géométrie non euclidienne, et finalement c'est ce qui importe. L'argument du manège manque certainement de rigueur (on peut montrer qu'un référentiel tournant ne peut exister en relativité restreinte), mais il conduit à deviner que la métrique dépend de la position dans un champ de gravitation, et à envisager que l'espace-temps puisse être courbé, même si dans le cas du disque tournant, c'est seulement l'espace qui est courbé.

Il est toutefois possible de donner un argument plus solide en faveur d'une courbure de l'espace-temps par le champ de gravitation. Celui-ci est dû à Schild (voir [9]). Supposons que de la lumière soit émise verticalement depuis une altitude z_1 par rapport à la surface de la Terre, et qu'elle soit reçue à une altitude $z_2 > z_1$. La lumière doit perdre de l'énergie en travaillant contre le champ gravitationnel pour passer de z_1 à z_2 . Si tel n'était pas le cas, on pourrait produire de l'énergie gratuitement en utilisant l'énergie lumineuse reçue en z_2 pour créer de la matière que l'on ferait chuter en z_1 , pour la transformer à nouveau en rayonnement réémis vers z_2 et ainsi de suite. Si la lumière grimpe sans perte d'énergie, on gagnerait à chaque cycle une énergie égale à $mg(z_2 - z_1)$ (cet argument de physique classique était bien connu d'Einstein). En vertu de la relation $E = h\nu$, on peut conclure que la fréquence ν_2 de la lumière reçue en z_2 est inférieure à la fréquence ν_1 de la lumière émise. Supposons maintenant que l'émission en z_1 soit maintenue pendant une durée T_1 exactement égale à une période. Alors la réception d'une période en z_2 durera $T_2 > T_1$ puisque $\nu_2 < \nu_1$. Si on trace dans l'espace-temps de Minkowski le diagramme correspondant à la situation, on obtient un parallélogramme puisque deux côtés opposés sont inclinés à 45 degrés par rapport à l'horizontale, et sont donc parallèles.

⁷On pourrait se dire que le périmètre du cercle subit une contraction de Lorentz. Ce n'est pas le cas. En effet, prenons un ruban sans mouvement dans R dont nous entourons exactement le cercle, en supposant que cela n'induit pas de frottement. Alors le diamètre du ruban est égal à celui du cercle, et son périmètre aussi. Mais comme le ruban est immobile, il ne subit aucune contraction. Il est clair que dans tout ce raisonnement nous ne devons nous servir des effets de relativité restreinte que dans la mesure où l'on peut assimiler sur de tous petits intervalles de temps un mouvement de rotation uniforme et un mouvement de translation uniforme. Tout cela mériterait bien sûr d'être rendu plus rigoureux, mais nous ne cherchons ici que des indications qualitatives.

⁸Nous verrons que l'effet dépend du potentiel gravitationnel.



On en conclut que les côtés verticaux ont des longueurs égales sur le dessin, et représentent donc des durées égales. On aboutit à l'absurde : on ne peut pas représenter cette expérience dans l'espace-temps de Minkowski. Le champ de gravitation doit modifier la géométrie de l'espace-temps.

Remarquons que l'on peut modifier l'argument pour tenir de l'influence du champ de gravitation sur la propagation de la lumière. Si celle-ci n'est pas représentée sur la figure par une droite inclinée à 45 degrés mais par une courbe plus générale, on obtient un « parallélogramme » dont deux côtés opposés sont des courbes parallèles, et l'argument tient encore.

Pour terminer ce paragraphe, nous allons essayer de déterminer un ordre de grandeur à partir duquel la relativité générale va jouer un rôle dominant pour les phénomènes de gravitation. La première approche consiste à calculer la vitesse de libération d'un corps de masse M et de rayon R en gravitation Newtonienne. Cette vitesse est :

$$v = \sqrt{\frac{2GM}{R}}$$

Si $v \approx c$, il est clair qu'une théorie relativiste de la gravitation est nécessaire. Autrement dit, lorsque

$$\frac{GM}{Rc^2} \approx 1$$

nous sommes dans le régime de la relativité générale. Une seconde approche confirme cette heuristique. Le rapport de l'énergie potentielle gravitationnelle du corps sur son énergie de masse est :

$$\frac{GM^2/R}{Mc^2} = \frac{GM}{Rc^2}$$

Nous retrouvons ce même rapport. Le tableau suivant⁹ donne l'ordre de grandeur de ce paramètre pour différents objets.

Objet	GM/Rc^2
Noyau atomique	10^{-38}
Atome	10^{-43}
Homme	10^{-25}
Terre	10^{-9}
Soleil	10^{-6}
Galaxie	10^{-7}
Naines blanches	10^{-4}
Étoile à neutrons	10^{-1}
Univers	10^0
Trou noir	1

⁹Le tableau et les calculs précédents sont issus de [8]

Il ne faudrait pas en conclure que les effets de la relativité générale au voisinage de la Terre sont totalement négligeables, comme on l'a longtemps cru. En effet, une expérience dont la sensibilité est de 10^{-9} pourra les mettre en évidence, et de telles expériences existent. Par ailleurs, dans certains cas, les effets sont cumulatifs et peuvent au cours du temps dépasser largement ce seuil, c'est ce qui se passe avec le système GPS sur lequel nous reviendrons.

Nous sommes allés très loin en partant de principes et d'expériences de pensée simples. Nous avons pu nous convaincre que la théorie de la relativité générale doit englober la gravitation et traiter d'espace-temps courbés, et que la trajectoire d'une particule libre doit être une géodésique dans un tel espace-temps. Comme c'est la matière qui engendre la gravitation, c'est elle qui doit dire à l'espace-temps comment se courber, et en retour ce dernier doit dire à la matière comment se déplacer. Mais nous devons maintenant voir si toutes ces belles idées peuvent s'incorporer dans une théorie précise, qui redonnera pour les champs faibles la relativité restreinte, et pour les petites vitesses la théorie de la gravitation Newtonienne.

Remarque : On commet fréquemment l'erreur de penser la gravitation en terme de courbure purement spatiale. Ainsi on utilise parfois l'image d'une boule de billard figurant une étoile, placée sur une toile tendue qui se courbe localement sous son poids. En réalité la courbure purement spatiale au voisinage d'une étoile typique ne rend que très partiellement compte des effets gravitationnels. Au voisinage d'une planète elle est même négligeable. Il ne faut donc pas perdre de vue que c'est l'espace-temps qui est courbé, et pas seulement l'espace. Bien sûr, notre capacité de visualisation se limite aux surfaces dans l'espace, et pour donner une idée intuitive de la courbure on doit forcément commencer par ce cas, c'est ce que nous ferons au paragraphe suivant, en espérant ne pas engendrer de confusion.

3.4 « Grossmann, il faut que tu m'aides, sinon je deviens fou. »

« Je m'occupe exclusivement du problème de la gravitation, et je crois maintenant que je surmonterai toutes les difficultés avec l'aide d'un ami mathématicien¹⁰. Il y a au moins une chose certaine, c'est que je n'avais jamais travaillé aussi dur de ma vie, et que j'ai acquis un grand respect des mathématiques, dont j'avais jusqu'à présent, dans mon innocence, considéré les aspects les plus subtils comme un luxe superflu ! À côté de ce problème, la première théorie de la relativité est un jeu d'enfant. » Albert Einstein, Lettre à Sommerfeld (1912)

Bien sûr nous ne pourrions pas, en quelques heures, rentrer dans les détails de ce qu'Einstein a mis des années à assimiler. Notre but est seulement de présenter la forme mathématique achevée de la théorie de la relativité générale, d'en expliquer le fonctionnement sur des exemples, et de montrer comment cette forme mathématique traduit bien les idées et incorpore les principes que nous avons énoncés plus haut. Il nous faut pour cela faire une incursion dans les mathématiques pures, en particulier dans la théorie des variétés riemanniennes et le calcul tensoriel. Ces outils prennent des mois, voire des années, à présenter complètement, c'est pourquoi nous ne ferons que survoler les idées générales, en partant d'un exemple précis.

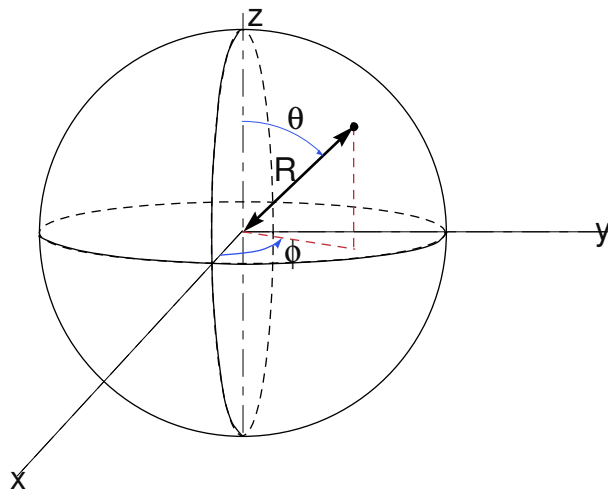
Variétés et métrique, l'exemple de la sphère

Considérons la sphère Σ de rayon R centrée sur l'origine d'un repère orthonormé direct $(O, \vec{i}, \vec{j}, \vec{k})$ de l'espace euclidien \mathbb{R}^3 .

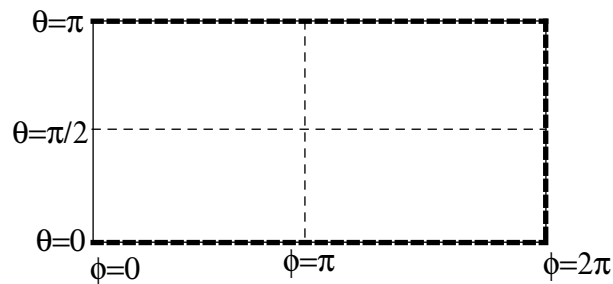
Imaginons que cette sphère soit habitée par des créatures absolument confinées à sa surface. De plus, contrairement aux Terriens, ils ne reçoivent aucune information de l'espace extérieur : pas d'étoiles ni de Soleil dans leur ciel. S'ils veulent se repérer, il leur faut introduire des coordonnées, par exemple les coordonnées sphériques. Ces coordonnées sont représentées mathématiquement par une application bijective $f : [0; 2\pi[\times]0; \pi[\rightarrow \Sigma$ définie comme suit :

$$f(\phi, \theta) = (R \cos(\phi) \sin(\theta); R \sin(\phi) \sin(\theta); R \cos(\theta))$$

¹⁰Il s'agit de Marcel Grossmann



Les habitants de Σ , les sigmiens, pourront alors dresser une carte de leur planète, c'est-à-dire représenter sa surface par un rectangle, comme ci-dessous.



En réalité cette carte ne couvre pas toute la planète : en effet, les pôles N et S ne sont pas représentés car l'application f cesse d'être bijective si on les inclut. On sait bien que, quelle que soit la méthode qu'on emploie, on ne pourra jamais représenter toute la surface d'une sphère par un bout de plan. On voit qu'il faudra au moins deux cartes. On peut par exemple prendre deux autres points antipodaux A et B , et considérer des nouvelles coordonnées sphériques $(\alpha; \beta)$ en faisant jouer à ces points le rôle des pôles Nord et Sud.

On obtient ainsi deux cartes, c'est-à-dire des bijections :

$$f : T \longrightarrow \Sigma \setminus \{N; S\}, \quad g : T' \longrightarrow \Sigma \setminus \{A; B\}$$

où T et T' sont deux rectangles ouverts¹¹ que l'on a pas précisés. Muni de ces deux cartes, les sigmiens pourront se repérer n'importe où sur la planète, à condition de savoir identifier un même lieu présent sur les deux cartes à la fois, autrement dit, à condition de savoir passer des coordonnées (ϕ, θ) aux coordonnées (α, β) à chaque fois que cela est possible. Dans notre exemple, il est clair que l'application :

$$g^{-1} \circ f : (\phi; \theta) \longmapsto (\alpha; \beta)$$

qu'on appelle **changement de carte** est différentiable, car f et g sont les restrictions d'applications différentiables de \mathbb{R}^2 dans \mathbb{R}^3 aux domaines T et T' . On dit que ces deux cartes munissent Σ d'une structure de variété différentiable de dimension 2 (voir paragraphe suivant).

¹¹privés de leurs côtés

Nous allons maintenant nous demander comment mesurer des distances sur la sphère. Pour cela nous disposons d'abord du point de vue « aérien » de \mathbb{R}^3 . Soit une courbe sur la sphère, qui a partout un vecteur vitesse. Une telle courbe peut se représenter par une application différentiable $\gamma : I \rightarrow \Sigma$, où I est un intervalle de \mathbb{R} . La longueur de la portion de courbe comprise entre $\gamma(t_1)$ et $\gamma(t_2)$ est naturellement définie par :

$$l(\gamma)_{t_1}^{t_2} = \int_{t_1}^{t_2} \|\gamma'(t)\| dt \quad (3.1)$$

Considérons à titre d'exemple le quart de méridien défini par :

$$\begin{aligned} \gamma &: [\pi/4; 3\pi/4] \longrightarrow \Sigma \\ \gamma(t) &= (R \cos(\phi) \sin(t); R \sin(\phi) \sin(t); R \cos(t)) \end{aligned}$$

On a :

$$\gamma'(t) = (-R \cos(\phi) \cos(t); -R \sin(\phi) \cos(t); R \sin(t))$$

d'où :

$$\|\gamma'(t)\| = R \sqrt{\cos^2(\phi) \cos^2(t) + \sin^2(\phi) \cos^2(t) + \sin^2(t)} = R$$

Ainsi la longueur totale du quart de méridien est :

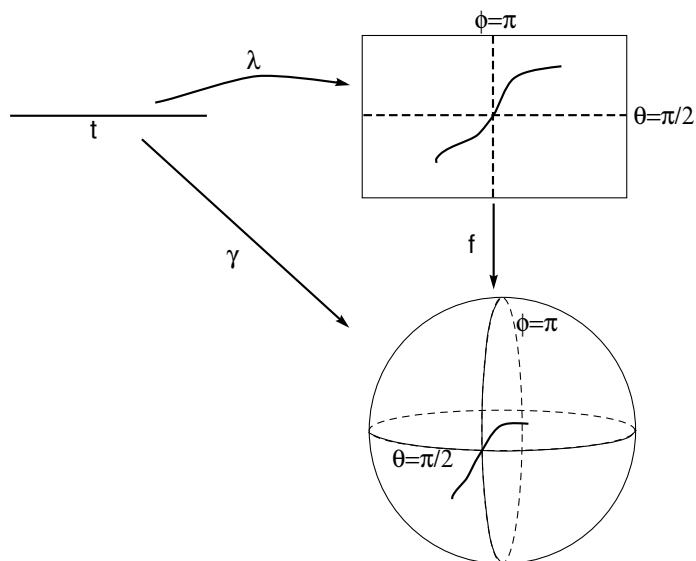
$$\int_{\pi/4}^{3\pi/4} R dt = \frac{\pi}{2} R$$

comme il se doit.

Exercice 3.4.1 Calculer la longueur d'un parallèle.

Nous voulons adopter maintenant un point de vue « piéton », autrement dit un point de vue intrinsèque à la sphère. Considérons que nous sommes dans le domaine de la carte f définie plus haut. Une courbe γ est donnée dans cette carte par une application :

$$\lambda : t \mapsto (\phi(t), \theta(t))$$



En posant $\gamma = f \circ \lambda$ on peut écrire, en repassant au point de vue aérien :

$$\gamma(t) = f(\lambda(t)) = (R \cos(\phi(t)) \sin(\theta(t)); R \sin(\phi(t)) \sin(\theta(t)); R \cos(\theta(t)))$$

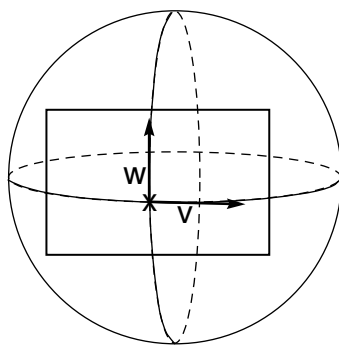
En appelant f_1, f_2, f_3 les composantes de f , cela peut se réécrire plus simplement (et en colonne pour faciliter la visualisation) :

$$\gamma(t) = \begin{pmatrix} f_1(\phi(t), \theta(t)) \\ f_2(\phi(t), \theta(t)) \\ f_3(\phi(t), \theta(t)) \end{pmatrix}$$

d'où

$$\gamma'(t) = \begin{pmatrix} \frac{\partial f_1}{\partial \theta} \theta'(t) + \frac{\partial f_1}{\partial \phi} \phi'(t) \\ \frac{\partial f_2}{\partial \theta} \theta'(t) + \frac{\partial f_2}{\partial \phi} \phi'(t) \\ \frac{\partial f_3}{\partial \theta} \theta'(t) + \frac{\partial f_3}{\partial \phi} \phi'(t) \end{pmatrix} = \theta'(t) \begin{pmatrix} \frac{\partial f_1}{\partial \theta} \\ \frac{\partial f_2}{\partial \theta} \\ \frac{\partial f_3}{\partial \theta} \end{pmatrix} + \phi'(t) \begin{pmatrix} \frac{\partial f_1}{\partial \phi} \\ \frac{\partial f_2}{\partial \phi} \\ \frac{\partial f_3}{\partial \phi} \end{pmatrix} = \phi'(t)v(t) + \theta'(t)w(t)$$

Nous pouvons interpréter géométriquement les vecteurs v et w que nous avons introduits à la dernière égalité. En effet, on obtient $\gamma'(t) = w(t)$ lorsque $\theta(t) = t$ et $\phi(t) = k = cte$. Autrement dit, $w(t)$ dirige la tangente au méridien $\phi = k$, au point $\gamma(t)$. De même $v(t)$ dirige la tangente au parallèle passant par ce point. Ces deux vecteurs forment une base du plan tangent à la sphère au point $\gamma(t)$.



L'équation $\gamma'(t) = \phi'(t)v(t) + \theta'(t)w(t)$ ne fait donc qu'exprimer le vecteur tangent quelconque $\gamma'(t)$ dans la base $(v(t), w(t))$ du plan tangent à Σ au point $\gamma(t)$. On peut calculer sa norme en utilisant cette décomposition :

$$\|\gamma'(t)\|^2 = (\phi'(t)v(t) + \theta'(t)w(t)) \cdot (\phi'(t)v(t) + \theta'(t)w(t)) = [\phi'^2 \|v\|^2 + 2\theta' \phi' v \cdot w + \theta'^2 \|w\|^2](t)$$

Ceci peut encore s'écrire :

$$\|\gamma'(t)\|^2 = g\left(\begin{pmatrix} \dot{\phi} \\ \dot{\theta} \end{pmatrix}, \begin{pmatrix} \dot{\phi} \\ \dot{\theta} \end{pmatrix}\right) \quad (3.2)$$

en introduisant les notations $\dot{\theta} := \theta'(t)$, $\dot{\phi} = \phi'(t)$ et la forme bilinéaire g , dont la matrice dans la base (v, w) est :

$$g_{\mu\nu} = \begin{pmatrix} v \cdot v & v \cdot w \\ v \cdot w & w \cdot w \end{pmatrix}$$

avec $\mu, \nu = 1, 2$.

Seulement les sigmiens n'ont pas de point de vue aérien. Ils disposent seulement de cartes, que leurs explorateurs ont pu dresser. Pour eux la courbe dont on cherche à mesurer la longueur est représentée sur la carte par la courbe $\lambda(t) = (\phi(t), \theta(t))$, et le vecteur $(\dot{\phi}, \dot{\theta})$ n'est rien d'autre que le vecteur vitesse $\lambda'(t)$. Seulement, s'ils intègrent $(\dot{\phi}^2 + \dot{\theta}^2)^{1/2}$, ils ne trouveront pas la bonne longueur, puisque celle-ci est donnée par :

$$l(\lambda) = l(\gamma) = \int_{t_1}^{t_2} \sqrt{g\left(\begin{pmatrix} \dot{\phi} \\ \dot{\theta} \end{pmatrix}, \begin{pmatrix} \dot{\phi} \\ \dot{\theta} \end{pmatrix}\right)} dt \quad (3.3)$$

Explicitons ce calcul dans le cas très simple où γ est un parallèle, i.e. λ est une droite horizontale. Calculons d'abord $g_{\mu\nu}$.

Exercice 3.4.2 Montrer que

$$g_{\mu\nu}(\phi, \theta) = R^2 \begin{pmatrix} \sin^2(\theta) & 0 \\ 0 & 1 \end{pmatrix}$$

Comment interpréter la dépendance en θ ?

Le fait que les termes anti-diagonaux soient nuls est particulier à la carte que nous avons choisie, et exprime simplement que les parallèles et les méridiens se coupent orthogonalement. La longueur totale du parallèle

$$\begin{aligned} \lambda : [0; 2\pi[&\longrightarrow [0; 2\pi[\times]0; \pi[\\ t &\longmapsto (t; \theta_0) \end{aligned}$$

est donc :

$$l = \int_0^{2\pi} R[1^2 \times \sin^2(\theta_0) + 0^2 \times 1]^{1/2} dt = 2\pi R |\sin \theta_0|$$

La longueur des différents parallèles est d'autant plus petite que leur latitude est élevée, ce qui est évident du point de vue aérien, mais a de quoi surprendre les sigmiens, puisque ces parallèles sont représentés par des segments qui semblent égaux sur la carte. Plus exactement, ces segments seraient de même longueur si la « métrique » était euclidienne, autrement dit si l'élément de longueur était donné par

$$ds^2 = dx^2 + dy^2$$

comme dans un plan euclidien rapporté à des axes orthonormés avec des coordonnées x et y , tandis qu'ici l'élément de longueur est :

$$ds^2 = R^2(d\theta^2 + \sin^2(\theta)d\phi^2) \quad (3.4)$$

ce qu'on peut encore écrire

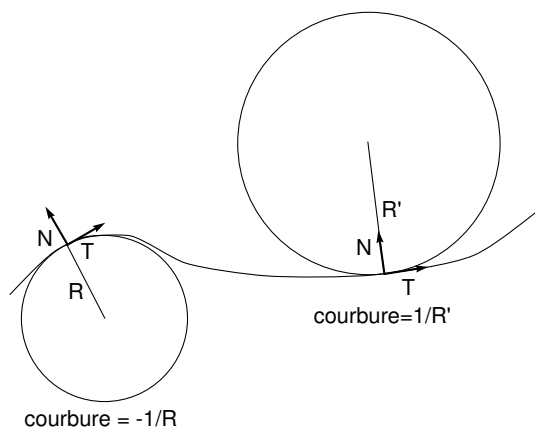
$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (3.5)$$

en posant $dx^1 = d\phi$ et $dx^2 = d\theta$.

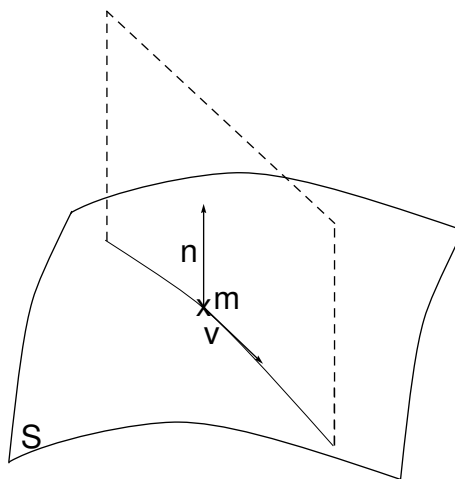
La formule (3.4) constitue un moyen pour les sigmiens de se rendre compte qu'ils ne vivent pas dans un espace plat¹² ou « euclidien ». Comment pourraient-ils se rendre compte qu'en outre, ils vivent dans un espace dont la courbure est, en un sens intuitif évident de notre point de vue aérien, constante ? Comment les sigmiens ont-ils accès à la courbure, et d'ailleurs, comment définir cette notion de courbure ?

Tout d'abord, rappelons que la courbure d'une courbe plane est définie comme plus ou moins l'inverse du rayon du cercle osculateur à la courbe en m , le signe étant induit par la position du centre du cercle sur la normale orientée par \vec{n} .

¹²Cet « espace » est ici une surface.

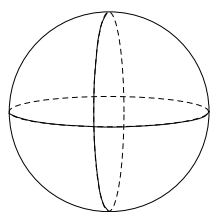


Considérons maintenant une surface S quelconque dans \mathbb{R}^3 . En un point $m \in S$ on peut considérer un vecteur normal unitaire \vec{n} à S en m . Pour chaque vecteur unitaire tangent \vec{v} en m , on définit le plan (m, \vec{v}, \vec{n}) . L'intersection de ce plan avec S est une courbe, dont on appelle $k(\vec{v})$ la courbure.

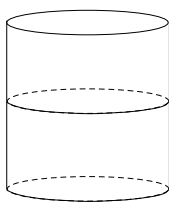


Si on fait tourner \vec{v} autour de \vec{n} , on trouvera une courbure minimale k_{min} et une maximale k_{max} . La courbure gaussienne de S en m est définie par $k = k_{min}k_{max}$.

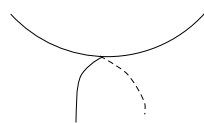
Si $k_{min} \neq k_{max}$ les directions correspondantes sont uniques (les extrema de $k(\vec{v})$ ne sont atteints qu'une fois chacun) et s'appellent les directions principales. Prenons trois exemples. Pour la sphère Σ , les courbes qu'on obtient sont toutes des grands cercles de rayon R . La courbure $k(\vec{v})$ est constante et égale à $1/R$. On a donc $k_{min} = k_{max} = 1/R$, et la courbure gaussienne vaut $k = 1/R^2$, et elle ne dépend pas du point m . Si on considère maintenant un cylindre dont la base est un cercle de rayon r , on a $k_{min} = 0$ (correspond à la génératrice du cylindre passant par m) et $k_{max} = 1/r$. On voit que la courbure gaussienne est nulle. Enfin si on prend un parabolôïde hyperbolique (une selle de cheval), on voit que la courbure gaussienne est négative.



$k > 0$



$k = 0$



$k < 0$

Euler a remarqué et démontré que les directions principales sont toujours orthogonales (c'est une application de la réduction des formes quadratiques aux axes principaux). Ce qui est encore plus remarquable, c'est que la courbure gaussienne ne dépend pas du « point de vue aérien ». Cela signifie qu'il y a un moyen de la définir à partir de la métrique de la surface. C'est si remarquable que Gauss, qui a démontré ce fait, l'a appelé « Theorema egregium » (Théorème remarquable). La démonstration est toutefois trop longue pour être donnée ici (on pourra consulter [1]).

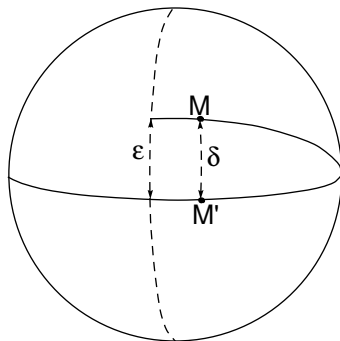
On peut déduire du Theorema Egregium que les sigmiens disposent nécessairement d'un moyen de calculer par eux-mêmes la courbure gaussienne de leur monde, sans faire appel au point de vue aérien. Pour voir comment ils peuvent s'y prendre, il nous faut d'abord définir la notion de géodésique, ce qui correspond pour eux à la notion euclidienne de ligne droite. Une géodésique est une courbe réalisant le minimum de la distance entre deux points. Les géodésiques de la sphère sont les grands cercles, c'est-à-dire les intersections de la sphère avec les plans passant par l'origine. Intuitivement on comprend que ce sont les courbes « les moins courbées » puisque leur rayon de courbure, R , est le plus grand possible. Par exemple les méridiens sont des géodésiques, tandis que les parallèles n'en sont pas, sauf l'équateur. Si on regarde deux méridiens, par exemple $\phi = 0$ et $\phi = \epsilon$, avec $\epsilon > 0$, on voit qu'ils convergent aux pôles. Les sigmiens peuvent s'en rendre compte en suivant ces deux méridiens jusqu'à leur point de rencontre. La convergence des géodésiques leur prouve que leur espace est courbé.

Plus précisément, considérons deux mobiles, initialement situés en $(\pi/2; 0)$ et $(\pi/2 - \epsilon; 0)$ en coordonnées sphériques, et suivant leurs grands cercles respectifs à vitesse constante et égale à 1 m/s , de sorte que chaque courbe soit paramétrée par la longueur d'arc. L'équation horaire du premier mobile est :

$$M(t) : \begin{cases} \lambda(t) &= 0 \\ \phi(t) &= \frac{t}{R} \end{cases}$$

où l'on a introduit la latitude $\lambda = \pi/2 - \theta$, pour simplifier les calculs. L'équation horaire du second mobile est plus compliquée, mais on peut montrer qu'on a :

$$M'(t) : \begin{cases} \lambda(t) &= \arcsin(\sin(\epsilon) \cos(t/R)) \\ \phi(t) &= \arctan\left(\frac{\tan(t/R)}{\cos \epsilon}\right) \end{cases}$$



Soit $\delta(t)$ la distance à la surface de la sphère séparant les deux mobiles à l'instant t .
On a :

$$\delta(t) = R \arccos(\overrightarrow{OM} \cdot \overrightarrow{OM'}) \quad (3.6)$$

d'où

$$\delta(t) = R \arccos(\cos(\epsilon) \cos^2(t/R) + \sin^2(t/R))$$

Au deuxième ordre en ϵ on a :

$$\delta(t) = R\epsilon \cos(t/R) + o(\epsilon^2)$$

d'où l'équation différentielle, satisfaite pour les petites valeurs de ϵ :

$$\ddot{\delta}(t) + \frac{1}{R^2} \delta = 0 \quad (3.7)$$

Cette équation, qui donne l'accélération relative que ressentent les deux mobiles, fait directement intervenir la courbure de Gauss $1/R^2$, ce qui est un moyen pour les sigmiens de la calculer. Cette équation s'appelle l'équation de la déviation géodésique.

Il existe d'autres moyens d'avoir accès à la courbure. L'ensemble des points situés à une même distance intrinsèque ϵ d'un point m à la surface de Σ , est un cercle de rayon ϵ pour les sigmiens. D'un point de vue aérien on voit que c'est aussi un cercle, mais dont le rayon est plus petit que ϵ . Par conséquent le rapport entre le périmètre l du cercle et son rayon intrinsèque ϵ sera plus petit que 2π . On a en fait la formule :

$$K = \lim_{\epsilon \rightarrow 0} \frac{6}{\epsilon^2} \left(1 - \frac{l}{2\pi\epsilon}\right) \quad (3.8)$$

où K est la courbure de Gauss, ici égale à $1/R^2$, mais cette formule est valable pour toute surface. Elle permet de calculer directement la courbure de Gauss en mesurant des rayons et des périmètres.

Enfin on peut aussi utiliser des triangles dont les côtés sont des segments géodésiques. La différence entre la somme des angles intérieurs du triangles et π est alors égale à l'intégrale de la courbure de Gauss sur le triangle. Ainsi, dans le cas d'une courbure de Gauss positive, comme pour la sphère, la somme des angles d'un triangle est toujours plus grande que π .

3.5 Variétés différentiables

Résumons et généralisons les leçons tirées de l'exemple de la sphère. Nous commençons par définir la notion de variété différentiable. La définition technique est celle-ci :

Définition 3.5.1 *Une variété différentiable de dimension n est un espace topologique M , muni d'une collection de cartes locales $\phi_i : U_i \rightarrow \mathbb{R}^n$, où les U_i sont des ouverts qui recouvrent M , les ϕ_i des homéomorphismes¹³ sur leurs images, tels que les applications de changement de cartes soient différentiables.*

Sans rentrer dans le détail des explications de tous les termes de cette définition disons, qu'en gros, une variété différentiable est un truc qui ressemble localement à \mathbb{R}^n et qui est muni d'une collection complète de cartes bien fichues. Une sphère, un ellipsoïde, un tore, sont des exemples de variétés différentiables de dimension 2. Dans le cas de la sphère on a vu qu'on devait utiliser au moins deux cartes, mais on peut bien sûr en utiliser davantage.

On peut généraliser très facilement les exemples précédents. Ainsi l'hypersphère de rayon 1 de \mathbb{R}^4 , notée S^3 et définie par $\|\overrightarrow{OM}\|^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$, où x_1, \dots, x_4 sont les coordonnées canoniques, est

¹³C'est-à-dire des applications bijectives continues et dont la réciproque est aussi continue.

un exemple de variété de dimension 3. Cependant il est important de comprendre que la définition d'une variété différentiable est intrinsèque : une telle variété n'est pas nécessairement une sous-variété de \mathbb{R}^k . Il n'y pas forcément d'espace ambiant dans lequel une variété est plongée. Ce point de vue intrinsèque est important en relativité générale, puisqu'il n'y a pas de « cinquième dimension » dans laquelle notre espace-temps serait plongé¹⁴.

L'hypothèse de différentiabilité est d'une grande importance technique pour ce qu'on va faire ensuite. Comme contre-exemple on peut citer un cube. Ce n'est pas une variété différentiable car les changements de cartes ne sont pas différentiables « dans les coins ». Le fait d'avoir une structure différentiable permet de parler d'applications différentiables entre deux variétés M et M' . Une application $F : M \rightarrow M'$ sera dite différentiable ssi elle l'est « vue à travers les cartes », i.e. ssi $\psi_i \circ \phi \circ \phi_j^{-1}$ est différentiable pour toute carte ψ_j de M et ϕ_i de N telle que la composée soit définie.

Nous avons maintenant besoin d'un outil pour mesurer les distances localement.

Définition 3.5.2 *Soit M une variété différentiable. Une métrique sur M est la donnée, pour tout point $m \in M$, d'une forme bilinéaire symétrique $g(m)$. De plus on demande que l'application $m \mapsto g(m)$ soit différentiable.*

Si pour tout m la forme $g(m)$ est définie positive on dit que (M, g) est une variété riemannienne. Si pour tout m la forme $g(m)$ est de signature $(1, n - 1)$, on dit que (M, g) est une variété de Lorentz.

Cette définition appelle un commentaire immédiat. La forme bilinéaire symétrique $g(m)$ sera donnée dans une carte locale par une matrice¹⁵ $(g_{ab}(m))$, dont les valeurs dépendent du point m , et la différentiabilité demandée équivaut à ce que les fonctions $g_{ab}(m)$ soient toutes C^∞ . La mesure locale des distances sera donnée par

$$ds^2 = g_{ab} dx^a dx^b \quad (3.9)$$

Cela signifie, dans le cas riemannien, que la distance au carré entre le point m de coordonnées (x^a) et le point infinitésimalement proche m' de coordonnées $(x^a + dx^a)$ est donnée par la formule ci-dessus. Dans le cas Lorentzien, il s'agit de l'intervalle au carré. Dans une autre carte, les points m et m' auront d'autres coordonnées, et la matrice de g sera différente, mais le ds^2 restera le même. Mathématiquement on dit que g est un tenseur, c'est-à-dire un objet géométrique, indépendant des coordonnées. Nous avons déjà rencontré les tenseurs, que nous avons définis à l'aide d'un espace vectoriel et de son dual. Ici nous avons une variété, ce qui est différent. Néanmoins nous pouvons définir pour tout point m de M un espace vectoriel, appelé espace tangent. L'espace tangent $T_m M$ est l'ensemble de tous les vecteurs tangents au point m , c'est-à-dire l'ensemble des tous les vecteurs vitesse en m de toutes les courbes passant par M . Lorsque M est de dimension n , l'espace tangent $T_m M$ est un espace vectoriel de dimension n . Cela se comprend aisément lorsque M est une sous-variété de \mathbb{R}^k , comme par exemple la sphère dans \mathbb{R}^3 . Définir précisément l'espace tangent dans les cas où M n'est pas plongée dans \mathbb{R}^k est possible, mais demande des développements mathématiques hors de la portée de ce cours¹⁶. Une fois $T_m M$ défini, on peut dire qu'une métrique est la donnée pour tout m d'une application bilinéaire symétrique :

$$g(m) : T_m M \times T_m M \longrightarrow \mathbb{R}$$

La métrique est donc un objet qui mange deux vecteurs et recrache un réel, de façon bilinéaire et symétrique. Ceci est tout-à-fait similaire à ce que nous avons déjà vu pour l'espace de Minkowski, à ceci près que l'espace vectoriel n'est pas fixe : il y en a un pour chaque point m de M .

Citons un théorème important pour les variétés riemmaniennes aussi bien que lorentziennes.

Théorème 3.5.1 (*platitude locale*)

Pour tout point $m \in M$, où M est une variété de Riemann (resp. de Lorentz) de métrique g , il existe un ouvert assez petit $U \ni m$ et des coordonnées locales (x^i) sur U telles que $g(m)$ soit donnée par la matrice

¹⁴Les théorèmes de plongement de Nash et Whitney permettent toujours de se ramener au cas des sous-variétés, mais d'une façon artificielle et non unique. Ainsi cette vision des choses ne convient pas en relativité générale.

¹⁵Nous utilisons les indices romains a, b, c, \dots , qui parcourent tous les nombres $1, 2, \dots, n$.

¹⁶Cela ne pose pas de problème en pratique puisque, dans un domaine de carte $U \subset \mathbb{R}^n$, un vecteur tangent au point $m \in U$ n'est rien d'autre qu'un vecteur de \mathbb{R}^n dont l'origine est fixée en m .

identité (resp. par la matrice $\mathbf{diag}(1; -1; \dots; -1)$) dans ces coordonnées, et que de plus, les dérivées $\partial g / \partial x^i$ soient toutes nulles en m .

Nous verrons plus loin quelle est l'interprétation physique de ce théorème en relativité générale. La longueur d'une courbe $t \mapsto \gamma(t)$ est donnée, dans le cas riemannien, par

$$l(\gamma) = \int_{t_1}^{t_2} \sqrt{g(\gamma(t))[\dot{\gamma}(t), \dot{\gamma}(t)]} dt = \int_{t_1}^{t_2} \sqrt{g_{ab} \frac{dx^a}{dt} \frac{dx^b}{dt}} dt \quad (3.10)$$

Dans le cas Lorentzien on pose :

$$l(\gamma) = \int_{t_1}^{t_2} \sqrt{\left| g_{ab} \frac{dx^a}{dt} \frac{dx^b}{dt} \right|} dt \quad (3.11)$$

Cette quantité s'interprète comme une longueur ou un temps propre, suivant les cas.

Définition 3.5.3 Une géodésique est une courbe γ telle que la variation au premier ordre de la quantité $l(\gamma)$ entre deux quelconques de ses points est nulle.

Dans le cas riemannien, une géodésique va donc être une courbe qui minimise localement la distance¹⁷. Dans le cas Lorentzien, une géodésique du genre temps réalise un maximum local de temps propre entre deux de ses points (penser au paradoxe des jumeaux dans l'espace de Minkowski).

On peut montrer que le système d'équations différentielles vérifié par une géodésique parcourue à vitesse unité est donnée en carte locale par :

$$\ddot{x}^a + \Gamma_{bc}^a \dot{x}^b \dot{x}^c = 0 \quad (3.12)$$

où les Γ_{bc}^a s'appellent les symboles de Christoffel¹⁸, et sont définis de façon unique à partir de la métrique par les conditions :

$$\partial_c g_{ab} = \Gamma_{ac}^d g_{db} + \Gamma_{bc}^d g_{ad} \quad \text{et} \quad \Gamma_{bc}^a = \Gamma_{cb}^a \quad (3.13)$$

où ∂_c est la dérivée partielle par rapport à la coordonnée x^c . On peut montrer que la solution du système précédent est donnée par :

$$\Gamma_{bc}^a = \frac{1}{2} g^{ad} (\partial_b g_{dc} + \partial_c g_{db} - \partial_d g_{bc}) \quad (3.14)$$

où g^{ab} est le tenseur contravariant dont l'expression en coordonnées est simplement donnée par la matrice inverse de g_{ab} . Cette dernière formule n'étant pas toujours très pratique car nécessitant l'inversion de la matrice de g_{ab} on lui préfère souvent l'utilisation directe du système (3.13).

Exercice 3.5.1 Calculer les symboles de Christoffel et l'équation des géodésiques dans le cas de la sphère, pour les coordonnées sphériques.

Exercice 3.5.2 Calculer les symboles de Christoffel pour l'espace de Minkowski. Faire de même, mais en coordonnées sphériques (on pourra se référer à l'exercice 3.7.1). Vérifier que, dans les deux cas, les géodésiques sont bien des droites.

La courbure d'une variété de Riemann ou de Lorentz est codée dans le tenseur suivant, appelé tenseur de Riemann et défini à partir des symboles de Christoffel :

$$R^a{}_{bcd} = \partial_c \Gamma_{bd}^a - \partial_d \Gamma_{bc}^a + \Gamma_{bd}^e \Gamma_{ec}^a - \Gamma_{bc}^e \Gamma_{ed}^a \quad (3.15)$$

¹⁷Pour comprendre pourquoi cette minimisation n'est que locale, penser à une courbe représentant les 3/4 d'un méridien sur la sphère. Cette courbe ne réalise pas le minimum de la distance entre ses extrémités puisqu'on peut trouver un chemin plus court en prenant le quart de méridien restant.

¹⁸Ces symboles ne sont pas des tenseurs.

On peut définir le tenseur de Riemann autrement que par cette formule indigeste en faisant appel à la notion de dérivée covariante, que nous ne définirons pas dans ce cours, et qui donne une expression plus simple et naturelle pour la courbure.

Faisons quelques commentaires sur cette notion de courbure. Considérons deux vecteurs tangents et orthonormés $x = (x^a)$ et $y = (y^b)$ en un point m . Ces deux vecteurs engendrent un plan $P \subset T_m M$. Pour tout vecteur unitaire w du plan P , on peut considérer l'unique géodésique de M passant par m et ayant w comme vecteur tangent. En suivant cette géodésique sur une longueur $\epsilon > 0$ assez petite, on obtient un bout de géodésique partant du point m . En faisant varier le vecteur w dans P , on trace ainsi un bout de surface $S_\epsilon \subset M$, contenant le point m . Le tenseur de Riemann permet de calculer la courbure de Gauss de S_ϵ . Plus précisément le scalaire :

$$K = R^a{}_{bcd} g_{ae} x^e y^b x^c y^d$$

est la courbure de Gauss de S_ϵ . Moralité : le tenseur de Riemann contient les courbures de Gauss de toutes les surfaces géodésiques S_ϵ contenues dans M . En particulier, si $n = 2$, le tenseur de Riemann est entièrement déterminé par la courbure de Gauss.

Le tenseur de Riemann intervient directement dans l'équation de la déviation géodésique, donnant l'accélération relative de deux géodésiques voisines. Les termes de cette équation sont un peu long à définir, aussi nous renvoyons à [9] pour son expression précise.

En contractant le tenseur de Riemann sur le premier et le troisième indice, on obtient un tenseur appelé « courbure de Ricci », qui est une sorte de moyenne du tenseur de Riemann.

$$Ric_{ab} = R^d{}_{adb} \quad (3.16)$$

Ce tenseur est symétrique. Il a donc $n(n+1)/2$ composantes indépendantes. Pour en avoir une interprétation concrète, considérons un vecteur unitaire tangent x^a . Alors la quantité $Ric_{ab} x^a x^b$ est égale à $n-1$ fois la moyenne de la courbure de Gauss de toutes les surfaces géodésiques S_ϵ contenant le vecteur x^a .

Une dernière courbure reste à définir, la courbure scalaire, donnée par :

$$R = g^{ab} Ric_{ab} \quad (3.17)$$

On voit qu'il s'agit (à un facteur près) d'une moyenne des courbures de Gauss de toutes les surfaces géodésiques passant par le point considéré. On peut également dire que si la courbure scalaire est positive en un point m , alors le volume d'une boule de rayon petit autour du point m sera plus petit que le volume d'une boule de même rayon dans \mathbb{R}^n . Le contraire est vrai si $R(p) < 0$.

3.6 La relativité générale en bref

Entre 1907 et 1916, Einstein va, avec l'aide de son ami le mathématicien Marcel Grossmann, mettre en forme les intuitions exposées plus haut. Nous donnons ici le résultat final de ces recherches qui durèrent près de dix ans.

L'espace-temps est une variété de Lorentz de dimension 4, dont la métrique g vérifie l'équation :

$$R_{\mu\nu} - \frac{1}{2}(R + 2\Lambda)g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \quad (3.18)$$

où $R_{\mu\nu}$ est le tenseur de Ricci, R la courbure scalaire, Λ une constante appelée « constante cosmologique » et $T_{\mu\nu}$ le tenseur énergie-impulsion dont la forme est donnée par le contenu matériel de l'espace-temps. Nous renvoyons à l'appendice C pour plus d'informations sur $T_{\mu\nu}$. Ce sont les équations de champ de la relativité générale : la matière (donc l'énergie) dit à l'espace-temps comment se courber. On écrit parfois cette équation sous la forme « courbure=matière » dans les ouvrages de vulgarisation. Il faut cependant être conscient que la courbure en question est une sorte d'écart à la moyenne, donnée par le tenseur $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$ dans le membre de gauche. Quant à la matière du membre de droite, il s'agit plus exactement de la densité d'énergie-matière. La constante Λ a une longue histoire. Nous en reparlerons

à propos de la cosmologie. Les dernières observations astronomiques tendent à montrer qu'elle est très petite, mais non nulle. Ses effets sont donc négligeables sur les courtes distances, et ne se font sentir qu'à l'échelle de l'univers. On fait souvent passer le terme $\Lambda g_{\mu\nu}$ du côté droit de l'équation. On l'interprète alors comme une densité d'énergie présente même quand $T_{\mu\nu} = 0$, autrement dit comme une densité d'énergie du vide. Le nom à la mode pour cette forme d'énergie est « énergie noire ».

Notons que (3.18) est une égalité entre tenseurs symétriques de rang deux, ce qui donne $4 \times 5/2 = 10$ équations en tout, qui sont des équations différentielles non linéaires du second ordre portant sur les composantes de la métrique $g_{\mu\nu}$, à partir desquelles on calcule le tenseur de Riemann, puis la courbure de Ricci et la courbure scalaire. Ce système d'équations différentielles, qu'on appelle « équations d'Einstein » prend la place de l'équation de Poisson :

$$\Delta\phi = 4\pi G\rho$$

qui détermine le potentiel gravitationnel ϕ en fonction de la densité de matière ρ en théorie Newtonienne. On voit ainsi en comparant les deux théories que c'est le tenseur métrique $g_{\mu\nu}$ qui joue le rôle du potentiel gravitationnel en relativité générale. Notons cependant une différence importante. Si on fait évoluer dans le temps le membre de droite de l'équation de Poisson, le potentiel s'adapte instantanément puisque le membre de gauche ne fait pas intervenir le temps. Cela traduit l'existence d'une action instantanée à distance dans la théorie de Newton. Au contraire, en relativité générale, si les sources de gravitation (l'énergie-matière) évoluent, le champ se modifiera de proche en proche, et on peut montrer que la vitesse de cette propagation est celle de la lumière. Il y a ici un parallèle très fort avec l'électromagnétisme de Maxwell. Les équations d'Einstein sont en quelque sorte les équations de Maxwell du champ de gravitation, mais elles sont bien plus complexes.

Une particule-test de masse négligeable¹⁹ (donc ne provoquant elle-même qu'une courbure négligeable de l'espace-temps) se meut selon une géodésique, c'est-à-dire vérifie l'équation :

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma_{\nu\rho}^\mu \frac{dx^\nu}{d\tau} \frac{dx^\rho}{d\tau} = 0 \quad (3.19)$$

C'est l'équation du mouvement de la relativité générale. L'espace-temps dit à la matière comment se mouvoir.

Maintenant il faut interpréter tout cela et notamment comprendre dans quelles situations on retrouve la théorie Newtonienne, et quels nouveaux phénomènes sont prédits par la relativité générale. C'est ce que nous ferons dès le prochain paragraphe. En attendant, terminons celui-ci par un tableau donnant l'interprétation physique de certains concepts mathématiques.

Mathématiques	Physique
carte locale	référentiel
équation entre tenseurs	loi généralement covariante
métrique	potentiel gravitationnel
courbure	force de marée
géodésiques	lignes d'univers des particules test
théorème de platitude locale	principe d'équivalence

3.7 La solution de Schwarzschild

Einstein fut fort surpris lorsqu'un mathématicien, Karl Schwarzschild, lui annonça qu'il avait trouvé une solution exacte aux équations (3.18), quelques mois seulement après la publication de celles-ci. Pensez-donc, un système de dix équations différentielles non linéaires du second ordre! On ne peut bien sûr pas les résoudre en toute généralité, mais Schwarzschild fut capable de trouver une solution dans le cas

¹⁹Mais non-nulle... Une particule de masse nulle se meut sur une géodésique de genre lumière. Il faudrait dans ce cas remplacer le temps propre, qui vaut 0, par un autre paramètre s . Tout autre choix de paramètre s' tel que (3.19) soit vérifiée serait de la forme $s' = as + b$ où a et b sont des constantes.

particulier du champ engendré par une masse M ponctuelle seule dans l'univers, en régime permanent. Le champ est alors statique (ce qui signifie que dans une certaine carte, il n'y a pas d'évolution dans le temps), à symétrie sphérique²⁰, « tend vers zéro à l'infini » (ce qui signifie qu'on retrouve l'espace-temps de Minkowski à l'infini) et où le tenseur énergie-impulsion est nul (l'univers est vide) sauf à l'origine, où il est infini. La constante Λ est également prise égale à zéro. Cette solution est, en coordonnées sphériques :

$$ds^2 = (1 - 2m/r)c^2 dt^2 - (1 - 2m/r)^{-1} dr^2 - r^2(d\theta^2 + \sin^2 \theta d\phi^2) \quad (3.20)$$

où $m = GM/c^2$

On peut montrer que c'est la seule solution à symétrie sphérique des équations d'Einstein dans le vide. Par conséquent la métrique de Schwarzschild fournit une très bonne approximation du potentiel gravitationnel existant au voisinage d'une étoile ou d'une planète dont on considère que toute la masse est concentrée au centre, tant que l'on peut considérer que la symétrie sphérique est respectée et que l'influence d'éventuels autres corps est négligeable. Notons qu'elle n'est valable qu'à l'extérieur de l'étoile ou planète, car à l'intérieur de celle-ci, le tenseur énergie-impulsion n'est pas nul.

Exercice 3.7.1 Montrer que la métrique de Lorentz en coordonnées sphériques s'écrit :

$$ds^2 = c^2 dt^2 - [dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2)] \quad (3.21)$$

Remarque : On pourrait penser que dans les conditions énoncées, l'espace-temps est forcément plat puisqu'il n'y a pas d'énergie/matière, donc pas de source de champ gravitationnel. Ce n'est pourtant pas le cas, car la solution de Schwarzschild ne couvre pas tout l'espace-temps : la source du champ est cachée à l'origine des coordonnées sphériques, qui ne fait pas partie de la carte locale dans laquelle la solution est exprimée.

Discutons maintenant de cette solution. On peut constater en comparant (3.20) et (3.21) que la métrique de Schwarzschild redonne bien celle de Lorentz quand $r \rightarrow \infty$. En prenant des coupes à $t = cte$, on obtient des sous-variétés spatiales à 3 dimensions, dont la métrique est ($dt = 0$) :

$$ds^2 = -(1 - 2m/r)^{-1} dr^2 - r^2(d\theta^2 + \sin^2 \theta d\phi^2)$$

Celle-ci étant indépendante de t , on conclut que la solution est bien statique.

Mettons en garde le lecteur contre un piège : les coordonnées t et r n'ont pas le sens immédiat de temps et de distance à l'origine. Tout d'abord, t est le temps-coordonnée, qui n'a pas plus de sens physique qu'en relativité restreinte. Pour comprendre le lien entre r et la distance, considérons deux événements : $p_1 = (t, r_1, \theta, \phi)$ et $p_2 = (t, r_2, \theta, \phi)$, et calculons la longueur du « rayon » $p_1 p_2$. Le long de ce rayon on a $dt = d\theta = d\phi = 0$ et $ds^2 = -dl^2 = -(1 - 2m/r)^{-1} dr^2$, d'où la longueur :

$$l = \int_{r_1}^{r_2} \frac{dr}{\sqrt{1 - 2\frac{GM}{rc^2}}}$$

Exercice 3.7.2 Montrer que pour r_1 et r_2 grands devant $r_S = \frac{2GM}{c^2}$ (voir l'interprétation de r_S plus bas) on a le DL :

$$l \approx r_2 - r_1 + \frac{GM}{c^2} \log\left(\frac{r_2}{r_1}\right) \quad (3.22)$$

La coordonnée r n'a donc pas plus d'interprétation en terme de distance que n'en a la longitude sur une carte de la Terre. Néanmoins, dans certaines conditions, les corrections sont assez faibles pour être négligées, de même que sur une carte de la Terre en projection de Mercator on ne commet pas d'erreur trop importante en assimilant la longitude à une distance tant qu'on ne s'éloigne pas trop de l'équateur. Pour savoir si l'approximation est justifiée, on se rapportera à (3.22). Remarquons également un phénomène étrange. Lorsque $r \rightarrow r_S := \frac{2GM}{c^2}$, l'une des composantes de la métrique s'annule, tandis qu'une autre tend

²⁰On peut démontrer, ce qui est d'ailleurs très curieux, que la symétrie sphérique du champ implique que celui-ci est statique. C'est absolument faux en électromagnétisme par exemple.

vers l'infini. On appelle r_S le rayon de Schwarzschild²¹. Pendant longtemps on a négligé ce phénomène. La raison en est que, pour des étoiles ordinaires, les événements correspondant à $r = r_S$ sont à l'intérieur de l'étoile, et donc dans une zone où la solution de Schwarzschild n'est plus valide. Pour le Soleil r_S vaut environ 3 km, et pour la Terre 9 mm. Pour que r_S soit plus grand que le rayon de l'étoile (en assimilant la coordonnée r au rayon), il faudrait que la densité de celle-ci soit gigantesque, et on a longtemps cru que des objets aussi denses n'existaient pas. On sait aujourd'hui qu'ils existent bel et bien, et on les appelle des « trous noirs ». La surface $r = r_S$ s'appelle « horizon des événements ». On peut montrer qu'aucun signal provenant des régions où $r < r_S$ ne peut s'en échapper. Pour analyser correctement un tel trou noir, il faut faire un changement de coordonnées, car celles de Schwarzschild sont singulières pour $r = r_S$, ainsi que nous l'avons remarqué plus haut. Ce phénomène est très semblable à ce qui se passe en coordonnées sphériques lorsqu'on s'approche des pôles : physiquement rien ne distingue les pôles des autres points, et pourtant en coordonnées sphériques la métrique devient dégénérée. C'est parce qu'on cherche à utiliser une carte en dehors de son domaine de validité, en un endroit où les distances sont infiniment écrasées puisqu'un point est représenté par une droite. En $r = r_S$, les distances sont infiniment distordues, mais cela ne correspond pas à un phénomène physique, ainsi qu'on s'en rendrait compte en changeant de carte. Ce point met en évidence une caractéristique fondamentale de la relativité générale : c'est une théorie locale. Les équations sont locales, et les solutions sont a priori locales et représentent un morceau de l'espace-temps avec sa carte locale. Pour obtenir tout l'espace-temps, on est obligé de recoller des solutions, et de faire des hypothèses sur la topologie globale de l'espace-temps, qui n'est pas déterminée par la théorie.

Pour interpréter la coordonnée de temps, considérons un observateur à l'infini spatial. Pour un tel observateur, un intervalle infinitésimal de temps propre $d\tau$ est égal à dt . Par conséquent, en fixant l'origine du temps propre à l'origine du temps coordonnée, on obtient :

$$\tau_\infty = t$$

Le temps coordonnée est le temps propre, celui qui est mesuré par l'horloge H , d'un observateur à l'infini spatial. Supposons qu'un modèle identique d'horloge H' se trouve ailleurs dans l'espace et que (t, r, θ, ϕ) et $(t + dt, r, \theta, \phi)$ soient les coordonnées des événements correspondants à deux tics successifs de H' . L'intervalle de temps propre

$$\Delta\tau_r = \left(1 - 2\frac{GM}{rc^2}\right)^{1/2} \Delta t$$

est par définition l'intervalle de temps mesuré par H' , tandis que $\Delta t = \Delta\tau_\infty$ est l'intervalle de temps entre les deux tics de H' mesuré par H . Ainsi on a :

$$\Delta\tau_\infty = \frac{\Delta\tau_r}{\sqrt{1 - 2\frac{GM}{rc^2}}} \quad (3.23)$$

Si par exemple H et H' sont des horloges qui battent la seconde, le battement de l'horloge H' mesuré par H est plus long qu'une seconde, et il d'autant plus long que r est petit, c'est-à-dire que H' est soumis à un fort potentiel gravitationnel. Pour deux horloges H^1 et H^2 ayant pour coordonnées radiales respectives r_1 et r_2 , on trouve :

$$\frac{\Delta\tau_\infty^1}{\Delta\tau_\infty^2} = \left(\frac{1 - 2\frac{GM}{r_2c^2}}{1 - 2\frac{GM}{r_1c^2}}\right)^{1/2} \quad (3.24)$$

où $\Delta\tau_\infty^{1,2}$ est la durée de la « seconde de $H^{1,2}$ » mesurée en « secondes de H ».

Ce phénomène s'appelle la dilatation gravitationnelle du temps. Nous l'avons obtenu à partir de la solution de Schwarzschild, mais il est en fait beaucoup plus général. Nous pouvons en déduire le phénomène de

²¹Notons que le mot rayon est mal choisi, d'après ce que l'on vient de dire !

rougissement gravitationnel (gravitational redshift) des ondes²². En effet, comme horloges H, H^1, H^2 on peut prendre une onde électromagnétique de période fixée, par exemple l'onde émise lors d'une certaine transition d'un certain élément chimique. L'intervalle de temps battu par une telle horloge est T , la période de l'onde. La formule (3.23) s'écrit alors :

$$\frac{T_{obs}}{T_{em}} = \frac{1}{\sqrt{1 - 2\frac{GM}{rc^2}}} \quad (3.25)$$

où T_{obs} est la période de l'onde mesurée par une horloge située à l'infini, et T_{em} est la période de l'onde émise. On voit que l'onde observée est décalée vers les grandes longueurs d'ondes, on dit qu'elle est décalée vers le rouge. On peut ainsi observer depuis la Terre, dont on néglige l'influence gravitationnelle, le spectre d'émission d'une étoile suffisamment dense pour que le paramètre GM/Rc^2 soit appréciable, par exemple une naine blanche. Le décalage du spectre vers le rouge dû au champ gravitationnel est donné par cette formule. Si on connaît la masse de l'étoile par une autre méthode, cela donne un moyen de vérifier expérimentalement une prédiction de la relativité générale. De telles vérifications furent longues à mettre au point, car les difficultés observationnelles sont grandes, et c'est seulement dans les années 60 que les premières observations assez précises apparurent. Elles donnent un accord avec la relativité générale dans les limites de précision expérimentales, qui sont d'environ 5%. Cependant une expérience directe, en laboratoire et dans le champ de la Terre est possible, en se basant sur la formule (3.24). La première expérience de cette nature fut celle de Pound et Rebka en 1960. Elle consiste à mesurer le décalage dans le spectre d'émission de deux corps identiques situés à deux altitudes différentes (la différence d'altitude était de 22 m, et l'effet de l'ordre de 10^{-15} seulement!).

Exercice 3.7.3 Montrer que, au premier ordre en $h = z_2 - z_1$, où $z_{1,2}$ sont les altitudes respectives de deux horloges identiques situées à la verticale l'une de l'autre dans un laboratoire, on a :

$$\frac{\Delta\tau_\infty^1}{\Delta\tau_\infty^2} = 1 + \frac{gh}{c^2}$$

avec g l'accélération de la pesanteur (on néglige la rotation de la Terre). Application : $h = 22$ m.

Malgré la faiblesse de l'effet recherché, l'expérience donna un accord avec la théorie de l'ordre de 1%. Aujourd'hui cet effet est pris en compte de façon cruciale dans le système GPS. En effet, ce système permet d'estimer la position d'un point P à la surface de la Terre en triangulant sa localisation à l'aide de 3 ou 4 satellites. La distance entre P et les satellites est estimée grâce au temps mis par une onde électromagnétique pour la parcourir. Ce temps doit être connu avec une grande précision. Il est mesuré à l'aide d'horloges atomiques embarquées par les satellites qui envoient des signaux périodiques servant à resynchroniser en permanence l'horloge interne du récepteur GPS. Pour estimer la distance parcourue par l'onde il suffit que le récepteur connaisse l'heure précise à laquelle celle-ci a été émise, faire la différence avec l'heure de réception et multiplier par c . Mais ceci ne peut fonctionner que si le récepteur et le satellite mesurent le même temps, ce qui n'est pas le cas, puisque le récepteur au sol est plus près du centre de la Terre, et que la « seconde au sol » est un peu plus longue que la « seconde du satellite ». Le quotient entre les deux « secondes » peut être calculé par la formule (3.24). L'effet est très faible, puisque l'horloge au sol retarde d'environ une demi-nanoseconde par seconde. Mais au bout de 24 heures on peut constater un décalage d'environ 0.00005 secondes entre les deux horloges, ce qui multiplié par c donne... dans les 15 km! En fait il faut tenir compte également de la vitesse du satellite²³, qui donne au contraire un ralentissement de l'horloge embarquée par rapport à celle au sol, en vertu de la relativité restreinte. Cet effet est non-négligeable mais plus faible que l'effet gravitationnel. Le GPS constitue ainsi un excellent test de la relativité générale²⁴!

²²Le rougissement se distingue de la dilatation du temps dans un espace-temps général, car il peut dépendre du chemin suivi par l'onde. Dans l'espace-temps de Schwarzschild les deux phénomènes sont équivalents.

²³La formule (3.24) suppose les deux horloges sans vitesse relative.

²⁴Le calcul présenté ici est très simplifié. Il faut en outre tenir compte de : la non-rotundité de la Terre, la non-circularité des orbites des satellites, etc...

Exercice 3.7.4 Retrouver les affirmations précédentes par le calcul en utilisant la solution de Schwarzschild. Données : la vitesse du satellite est d'environ 4 km/s et son altitude environ 20000 km.

Les deux autres effets de la relativité générale à l'échelle du système solaire et qui ont conduit aux tests classiques de cette théorie sont l'avance du périhélie de Mercure et la déviation de la lumière au voisinage du Soleil. Nous citons ici quelques extraits de [7] auquel nous renvoyons pour plus de détails.

Nous nous intéressons au « problème à un corps » c'est-à-dire au mouvement d'une particule test dans le champ d'un corps, ici le Soleil, de masse M . Pour cela nous devons en principe déterminer les géodésiques de la métrique de Schwarzschild. Nous nous intéressons aux orbites planes, dans le plan équatorial $\theta = \pi/2$ en coordonnées de Schwarzschild²⁵. On trouve après calcul que ces dernières vérifient l'équation :

$$\frac{d^2u}{d\phi^2} + u = \frac{m}{h^2} + 3mu^2 \quad (3.26)$$

où $m = GM/c^2$, $u = 1/r$, et $h = r^2\dot{\phi}$ est une constante (qui s'identifie au moment cinétique de la particule test divisée par sa masse). On peut comparer cette équation avec la formule de Binet en théorie Newtonienne :

$$\frac{d^2u}{d\phi^2} + u = \frac{m}{h^2} \quad (3.27)$$

on voit apparaître un terme supplémentaire qui fait de (3.26) une équation non-linéaire. Cependant, pour une planète dans le système solaire, le terme supplémentaire est très petit devant m/h^2 . C'est pour Mercure qu'il est le plus grand, et le rapport des deux termes, $3h^2/r^2$, est alors de l'ordre de 10^{-7} . On peut dans ce cas utiliser une approche perturbative et considérer que l'orbite de Mercure est l'orbite képlérienne classique perturbée par un petit terme. On sait en mécanique classique que l'effet du terme perturbateur est d'empêcher les orbites de se refermer exactement. On peut alors assimiler la trajectoire de la planète à une suite d'ellipses dont les grands axes tournent très légèrement à chaque passage au plus près du Soleil : c'est le phénomène de précession. Toutes les planètes connaissent ce phénomène et on pouvait l'expliquer classiquement par la perturbation de l'orbite képlérienne par les autres planètes. Pourtant, pour Mercure, une fois prises en compte les contributions des autres planètes, il restait une avance inexplicée de 43 secondes d'arc par siècle. Cela donne déjà une idée du niveau stupéfiant de précision atteint par l'astronomie classique ! On chercha²⁶ à rendre compte de cette avance par l'effet d'une planète inconnue, située entre le Soleil et Mercure, baptisée Vulcain. Une méthode similaire avait déjà porté ses fruits avec la découverte de Neptune en 1846 par Urbain Le Verrier. Mais on ne trouva jamais Vulcain ! Aucune explication convaincante de l'avance du périhélie de Mercure ne fut donnée avant Einstein. En utilisant la méthode de relativité générale esquissée plus haut, on trouve une précession de l'orbite de :

$$2\pi\epsilon \approx \frac{24\pi^3 a^2}{c^2 T^2 (1 - e^2)}$$

où T est la période de l'orbite, a le demi-grand axe et e l'excentricité. Pour Mercure on obtient après application numérique... 42.98 secondes d'arc par siècle. Einstein a raconté qu'après avoir effectué ce calcul, il fut transporté par une joie tellement intense qu'il resta dans un état second pendant plusieurs jours...

Terminons avec la déviation de la lumière. L'analogie de (3.26) pour une géodésique de genre lumière est :

$$\frac{d^2u}{d\phi^2} + u = 3mu^2 \quad (3.28)$$

²⁵N'oublions pas que les géodésiques vivent dans l'espace-temps. Il s'agit donc ici des géodésiques dont la projection spatiale est plane.

²⁶On était alors dans les années 1850.

S'il n'y a pas de masse au centre, $m = 0$, et on trouve une équation linéaire dont la solution générale est :

$$u = \frac{1}{D} \sin(\phi - \phi_0) \quad (3.29)$$

où D est une constante. C'est l'équation d'une droite en coordonnées polaires (r, ϕ) , et la constante D est la distance entre la droite et l'origine. L'angle ϕ_0 est l'angle que fait la droite avec l'axe polaire. Pour trouver une solution approchée de (3.28), on utilise le fait que mu est petit dans le cas qui nous préoccupe. On cherche alors une solution sous la forme :

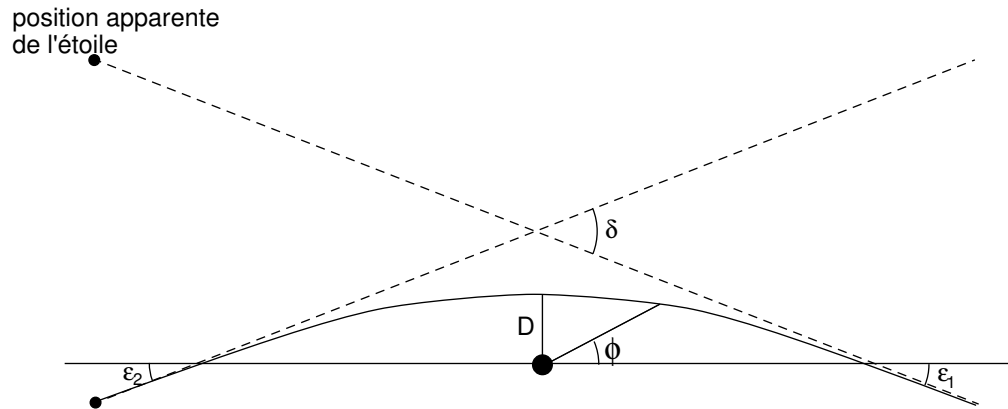
$$u = u_0 + 3mu_1$$

où u_0 est la solution (3.29) de l'équation linéarisée. On choisit $\phi_0 = 0$ pour simplifier. En négligeant tous les termes d'ordre $(mu)^2$, on obtient pour u_1 :

$$u_1'' + u_1 = \frac{\sin^2 \phi}{D^2}$$

On peut vérifier que $(1 + C \cos(\phi) + \cos^2(\phi))/3D^2$ est solution. On trouve alors comme solution générale approchée de (3.28) :

$$u \approx \frac{\sin \phi}{D} + \frac{m(1 + C \cos \phi + \cos^2 \phi)}{D^2}$$



On constate qu'un observateur à l'infini reçoit la lumière sous un angle apparent ϵ_1 alors que l'angle d'incidence est ϵ_2 . On trouve ces angles en cherchant les asymptotes (exo!). La déviation $\delta = \epsilon_1 + \epsilon_2$ est environ :

$$\delta \approx \frac{4m}{D} = \frac{4GM}{Dc^2}$$

Pour mesurer la déviation on compare la position apparente de l'étoile à un moment où celle-ci est proche du Soleil (ce qui ne peut s'observer que lors d'une éclipse totale) et à un moment où elle ne l'est pas. Une célèbre expédition pour mesurer cette déviation pendant une éclipse totale de Soleil fut conduite en 1919 par Eddington. Un accord correct avec la théorie compte tenu des imprécisions de mesure fut obtenu, ce qui fit grand bruit à l'époque et transforma instantanément Einstein en célébrité. La déviation de la lumière par les masses est aussi à l'origine d'un effet « lentille ». On utilise aujourd'hui abondamment ces lentilles gravitationnelles en astronomie pour détecter des corps invisibles autrement, ou pour estimer des masses.

Terminons avec la métrique de Schwarzschild en nous demandant ce qui se passe si le rayon de l'étoile est plus petit que le rayon critique pour lequel on a une singularité de coordonnées. Dans ce cas, entre

l'étoile et le rayon critique R_S , il y a de l'espace vide, et la solution devrait être valable. Les physiciens ont pendant très longtemps négligé ce cas, correspondant à des densités incroyables de matière. Mais la possibilité théorique, puis la découverte, des étoiles à neutrons, formidablement denses et proche de la limite où le rayon de Schwarzschild est plus grand que celui de l'étoile, les obligèrent à se pencher sur ce cas. Sans rentrer dans les détails, on peut dire que la surface $R = R_S$ s'interprètent comme un « horizon des événements ». Rien de ce qui pénètre en deçà ne peut jamais en revenir. Un tel astre est pour cette raison appelé « trou noir ». Longtemps envisagée avec scepticisme, l'existence des trous noirs ne fait aujourd'hui plus de doute. Il en existe même probablement de gigantesques au centre de la plupart des galaxies. En utilisant un certain changement de coordonnées, on peut examiner ce qui se passe pour $r < R_S$. À l'origine des coordonnées, on a une vraie singularité. Cela signifie que la courbure y est infinie. On peut se dire dans un premier temps qu'il s'agit là d'un artefact mathématique, qui ne veut rien dire physiquement. Pourtant, un observateur suivant une géodésique entrant à l'intérieur de l'horizon finit nécessairement sa course sur la singularité en un temps propre fini τ . Autrement dit, pour un tel observateur, on ne peut pas dire ce qui se passe après τ , la théorie « bugue » en quelque sorte. De plus, on ne peut pas s'en sortir en disant que la singularité est cachée à l'intérieur d'un petit astre très dense, car aucun astre d'aucune sorte, en tout cas pas en l'état actuel des connaissances physiques, ne peut garder sa cohésion face aux énormes forces de marée régnant à proximité de la singularité. Si un trou noir se forme, toute la matière présente à l'intérieur de l'horizon doit finir dans la singularité (d'où une densité infinie de matière). Il y a là un problème qu'on ne sait pas (encore) résoudre.

Exercice 3.7.5 Deux horloges atomiques sont synchronisées au niveau du sol. On emporte l'une dans un voyage en avion autour du globe à la vitesse constante de 1000 km/h et à l'altitude constante de 10 km. Que constate-t-on à l'arrivée ? (On négligera le décollage et l'atterrissage et on utilisera la métrique de Schwarzschild)

3.8 Cosmologie

Nous ne pouvons parler que très brièvement de cette application très importante de la relativité générale : à savoir l'utilisation de la théorie pour la description de l'univers tout entier, et l'étude de sa dynamique. Il existe de nombreux modèles cosmologiques, et nous ne discuterons que du plus populaire, celui de FLRW (Friedmann-Lemaître-Robertson-Walker). Ce modèle est fondé sur deux hypothèses. La première est l'existence d'un temps cosmologique. Cela signifie qu'on suppose qu'il existe une décomposition de la variété d'espace-temps en espace+temps. En relativité générale, une telle décomposition ne peut pas avoir un sens intrinsèque. On lui donne cependant un sens en considérant qu'elle est liée à une classe particulière d'observateurs, définie précisément. Ceci ne brise pas le principe de relativité, car cette classe d'observateurs n'a pas de privilège d'un point de vue théorique, seulement d'un point de vue pratique. En effet la métrique d'espace-temps prendra une forme plus simple pour cette classe d'observateurs, et elle aura une interprétation physique claire. Pour définir ces observateurs, on suppose que les galaxies sont comme les particules d'un gaz en chute libre, c'est-à-dire soumis au seul champ de gravitation. Elles ont donc un mouvement d'ensemble auquel s'ajoute un mouvement aléatoire. Cette hypothèse est justifiée par l'observation. Les observateurs qu'on considère sont ceux dont le mouvement coïncide avec le mouvement d'ensemble des galaxies. Le temps indiqué par un de ces observateurs constitue le temps cosmique t . La deuxième hypothèse, le principe cosmologique, est que chaque tranche spatiale $t = cte$ est homogène et isotrope. On peut alors montrer que la métrique la plus générale satisfaisant à ces hypothèses est²⁷ :

$$ds^2 = dt^2 - R(t)^2 \left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right] \quad (3.30)$$

où k est une constante égale à 1, 0 ou -1 . Si l'on fait une coupe spatiale $t = cte$, on obtient une variété de dimension 3 de courbure constante égale²⁸ à $K = kR(t)^{-2}$.

²⁷Les unités utilisées ici sont les unités géométriques, pour lesquelles $G = c = 1$.

²⁸Plus précisément, la courbure de Ricci est égale à $2Kg_{ab}$, et pour celle de Riemman on a $R_{abcd} = K(g_{ac}g_{bd} - g_{ad}g_{bc})$.

Dans le cas $k = 1$, les géodésiques spatiales issues d'un même point se recoupent toujours. Cela signifie que l'univers est spatialement borné (on dit aussi fermé) : on peut faire le tour de l'univers ! En faisant une hypothèse de simplicité sur la topologie globale de l'univers, on peut montrer que celui-ci est alors nécessairement isométrique à une 3-sphère de rayon $R(t)$, c'est-à-dire la généralisation de dimension 3 de la sphère S^2 . L'hypothèse de simplicité en question s'exprime mathématiquement par la requête que l'univers soit simplement connexe. Sans rentrer dans les détails, on peut comprendre ce qu'est un espace simplement connexe en regardant la sphère S^2 , qui l'est, et le tore (une surface en forme de chambre à air ou de doughnut) qui ne l'est pas.

Si $k = 0$, l'univers est globalement spatialement plat, et sa géométrie spatiale est euclidienne. Les observations actuelles semblent favoriser ce cas. Mais attention : l'espace-temps, lui, n'est pas plat. En faisant l'hypothèse de simple connexité, on peut montrer qu'il est spatialement non-borné (on dit aussi ouvert). Mais il existe des variétés spatialement fermées de courbure nulle. Leur topologie (du type chambre à air, ou bretzel avec plus de trous !) est plus compliquée, mais rien n'exclut a priori que notre univers ait une telle forme. Il faut bien comprendre que la relativité générale ne dit rien sur la topologie globale de l'univers, puisque c'est une théorie purement locale. Enfin si $k = -1$, l'univers est globalement de géométrie hyperbolique. Il est difficile d'en donner une image concrète car aucune surface de \mathbb{R}^3 n'est de courbure constante négative. Dans un tel espace, la somme des angles intérieurs d'un triangle est toujours inférieure strictement à π . Là encore, si on suppose que l'univers est simplement connexe, il est forcément non-borné spatialement (ouvert).

Le facteur $R(t)^2$ s'interprète comme un facteur d'échelle qui représente une dilatation ou une contraction globale des distances spatiales. Si on prend un triangle ABC formé par trois galaxies génériques à un temps t , et qu'on considère ces trois mêmes galaxies à un temps t' , l'évolution des distances mutuelles ne peut être qu'une contraction ou une dilatation uniforme, en vertu du principe cosmologique. La possible évolution dans le temps du facteur d'échelle s'interprète comme une contraction ou expansion de l'univers. Dans le cas d'une expansion, par exemple, les distances moyennes entre toutes les galaxies augmentent avec le temps. Cependant, cette expansion n'est pas une fuite à partir d'un point particulier, ce sont toutes les distances qui augmentent²⁹. Ceci se comprend très clairement dans le cas $k = 1$ puisque la variation des distances mutuelles entre les galaxies correspond à la variation du rayon $R(t)$ de la 3-sphère. Les équations du champ (3.18) appliquées à la métrique (3.30) et au tenseur énergie-impulsion du « fluide cosmique » donnent deux équations sur $R(t)$:

$$3\frac{\dot{R}^2 + k}{R^2} - \Lambda = 8\pi\rho$$

$$\frac{2R\ddot{R} + \dot{R}^2 + k}{R^2} - \Lambda = -8\pi p$$

où ρ et p sont respectivement la densité et la pression du fluide cosmique (dépendant du temps). Comme nous l'avons vu, le fluide cosmique aujourd'hui est un gaz de galaxie, et sa pression peut être considérée comme nulle. Cependant, cela n'a pas toujours été le cas (voir plus bas). Si l'on fait cette hypothèse de pression nulle, on peut intégrer facilement les équations précédentes et l'on trouve :

$$\dot{R}^2 = \frac{C}{R} + \frac{1}{3}\Lambda R^2 - k \quad (3.31)$$

où $C = 8\pi R^3 \rho$ est une constante > 0 . C'est l'équation de Friedmann en l'absence de pression.

Dans le cas où $k = 0$, regardons les différentes solutions.

1. Si $\Lambda > 0$ on trouve comme solution :

$$R^3 = \frac{3C}{2\Lambda} [\cosh(\omega t) - 1], \text{ avec } \omega = \sqrt{3\Lambda} \quad (3.32)$$

²⁹Ce qui se passe à une échelle plus petite que l'échelle intergalactique peut être très différent, car aux plus petites échelles l'univers n'est ni isotrope ni homogène. La métrique de FLRW n'est applicable que dans la mesure où l'on peut considérer une galaxie comme ponctuelle.

Le facteur d'échelle R est alors une fonction strictement croissante, et même exponentiellement croissante du temps : l'univers est en expansion continue, et l'expansion ne fait que s'accélérer.

2. Si $\Lambda < 0$ on trouve :

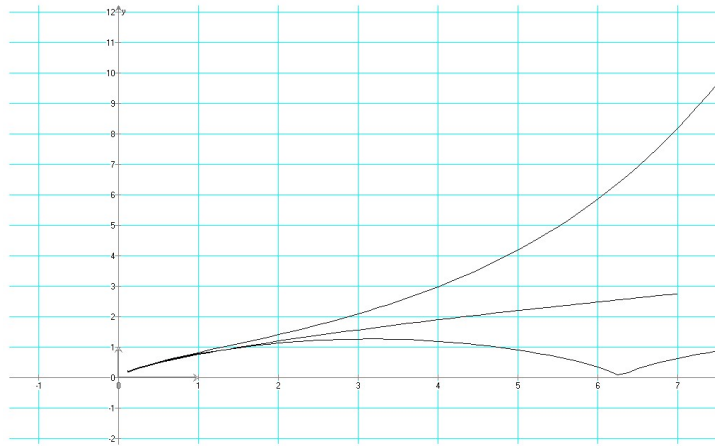
$$R^3 = \frac{3C}{-2\Lambda} [1 - \cos(\omega t)] \text{ avec } \omega = \sqrt{-3\Lambda} \quad (3.33)$$

On obtient une fonction périodique pour le facteur d'échelle. Si on observe une phase d'expansion, celle-ci est freinée, puis lui succède une phase de contraction.

3. Enfin si $\Lambda = 0$ on trouve :

$$R^3 = \frac{9}{4} C t^2$$

C'est une expansion de plus en plus ralentie, mais le facteur d'échelle tend tout-de-même vers l'infini.



R en fonction de t pour $\Lambda > 0$ (courbe du haut), $\Lambda = 0$ (courbe du milieu), $\Lambda < 0$, (courbe du bas), dans le cas où $k = 0$.

Dans le cas $k = -1$, les solutions ont des comportements qualitativement semblables aux solutions précédentes. Dans le cas $k = 1$, si $\Lambda \leq 0$, on obtient un facteur d'échelle périodique. Le cas $k = 1$ et $\Lambda > 0$ mérite d'être étudié. Dans ce cas il existe une valeur critique Λ_c de la constante cosmologique pour laquelle le facteur d'échelle est constant. C'est le modèle statique d'Einstein. Si $\Lambda > \Lambda_c$, l'expansion initiale est freinée, atteint une sorte de plateau puis s'accélère à nouveau. Enfin si $\Lambda < \Lambda_c$, l'expansion est suivie d'une contraction.

Notons que dans tous ces cas, lorsque R est très petit, ρ est très grand. Le gaz de galaxies est alors très condensé, et le modèle n'est plus valide car on ne peut plus considérer la pression comme nulle. Il faut alors utiliser d'autres modèles.

Quelles leçons tirer de tout ceci, et quelle est la situation observationnelle ? La leçon la plus évidente est que tous ces modèles sont dynamiques à l'exception d'un seul, et résultent d'une compétition entre l'attraction mutuelle des galaxies, qui tend à faire s'effondrer l'univers sur lui-même, et la constante cosmologique dont l'action est inverse.

Il existe un seul modèle statique qui requiert un ajustement parfait de la constante cosmologique à une valeur critique, et on peut montrer qu'il est instable. C'est donc le moins réaliste de tous, et c'est pourtant le premier qui fut considéré, tout simplement parce qu'à l'époque (vers 1917), tout le monde croyait que l'univers était statique et éternel. À ce titre, l'histoire de la constante cosmologique est exemplaire : Einstein l'avait d'abord prise égale à 0 pour rendre plus simples et élégantes ses équations de champ. Puis, se rendant compte qu'il fallait absolument une constante cosmologique pour obtenir une possibilité

d'univers statique, il changea d'avis et introduisit cette constante. S'il ne l'avait pas fait, il aurait prédit que l'univers doit être en expansion ou en contraction. Or Hubble découvrit en 1929 que l'univers est en expansion. Einstein a alors déclaré qu'il avait commis la plus grande erreur de sa vie en introduisant cette constante ! Cependant, il semble aujourd'hui que l'expansion s'accélère ! Pour cela, il faut que Λ soit strictement positive. Bien que très petite, la constante cosmologique détermine le destin de l'univers, et elle contribuerait pour environ 70% à l'énergie totale de ce dernier. Finalement Einstein n'a pas commis d'erreur... Si l'expansion s'accélère effectivement, alors quelle que soit la valeur³⁰ de k , l'univers continuera sa dilution sans fin. Voilà pour son avenir, mais si nous regardons vers le passé alors nous voyons que l'univers se contracte³¹ à mesure que t s'approche de 0. En raccordant le modèle FLRW avec un modèle plus réaliste pour un univers jeune dans lequel la pression de radiation est grande, on peut poursuivre l'exploration des temps reculés. On constate alors que rien n'arrête la diminution du facteur d'échelle, qui tend vers 0 avec t . La métrique présente alors une singularité³², nommée « big-bang ». Si on calcule le temps propre mesuré par une particule cosmique depuis le big-bang jusqu'à nous, on trouve un temps de l'ordre de 15 milliards d'années. Ainsi la relativité générale, alliée au principe cosmologique, prédit que dans un passé lointain, l'univers était dans un état incroyablement dense et donc chaud. Le rayonnement émis à cet époque³³ constitue le rayonnement dit « fossile » et son observation en 1964 par Penzias et Wilson est un magnifique succès de la cosmologie relativiste.

³⁰Les observations actuelles penchent pour $k = 0$, i.e. pour une valeur assez petite de la courbure pour être compatible avec cette hypothèse.

³¹Encore une fois, cette contraction ne se fait pas en direction d'un point particulier. Par ailleurs l'univers peut se contracter tout en étant ouvert.

³²L'existence d'une telle singularité est largement indépendante des modèles, voir plus loin.

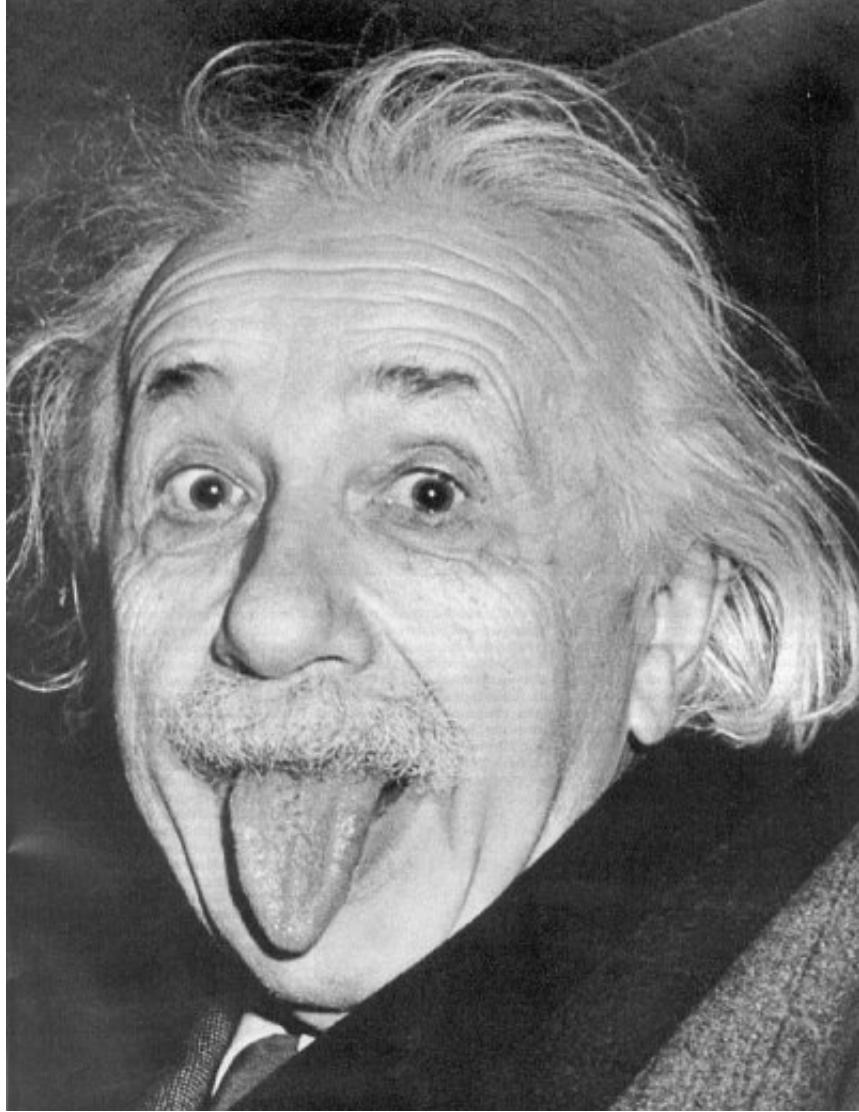
³³Plus précisément, ce rayonnement fut émis au moment dit de « recombinaison » où les atomes d'hydrogène se sont formés, soit environ 400 000 ans après le big-bang.

Chapitre 4

Conclusion

La relativité restreinte a bouleversé les concepts d'espace et de temps. La relativité générale est allée encore plus loin, en interprétant la gravitation comme une courbure de l'espace-temps. Elle nous force à réinterpréter les anciens concepts de la physique à l'aide de la géométrie des variétés de Lorentz. Mais elle ne peut pas être le fin mot de l'histoire. En effet, nous avons déjà rencontré deux singularités : le big-bang, et la singularité présente au centre d'un trou noir. Il ne s'agit pas là de cas particulier. En effet, au début des années 1970, Stephen Hawking et Roger Penrose démontrèrent que les solutions de la relativité générale présentaient toujours des singularités, du moment qu'il y avait de la matière quelque part dans l'univers ! Ainsi, cette théorie prédit sa propre perte, elle est nécessairement incomplète. Par ailleurs, une autre révolution d'ampleur comparable en physique eut lieu au début du XXe siècle : la théorie quantique. Or la physique quantique fournit un cadre dans lequel toute théorie physique devrait pouvoir s'exprimer : on a ainsi pu « quantifier » l'électromagnétisme, et établir également des théories quantiques pour deux autres interactions, la faible et la forte. On a même pu donner un cadre commun à ces trois forces, le modèle standard. La possibilité de fabriquer ainsi une théorie quantique des champs, et les formidables succès expérimentaux de cette dernière ont incité les physiciens à appliquer cette technique au champ de gravitation, en le traitant comme n'importe quel autre champ, oubliant peut-être un peu vite les principes fondamentaux sur lesquels la relativité générale est bâtie. Mais malgré les efforts acharnés de nombreux physiciens depuis plus de 30 ans¹, la gravitation refuse obstinément de se laisser « quantifier ». De son côté, la relativité générale offre également un cadre conceptuel pour toute théorie physique, cadre dans lequel la théorie quantique ne rentre pas ! Ainsi nous avons deux théories, chacune merveilleusement bien vérifiée dans son domaine d'application, chacune apportant son lot de bouleversements conceptuels. Il reste juste à les faire cohabiter. Il reste juste un petit nuage...

¹Notons cependant qu'une approche un peu différente de « l'approche quantique standard » semble aujourd'hui très prometteuse : la gravité quantique à boucles.



Annexe A

Solutions des exercices

Exercice 2.2.1

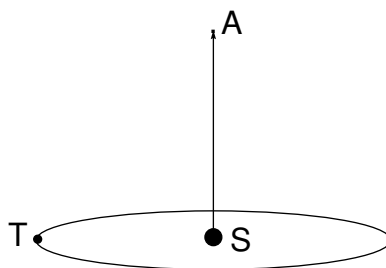
Oui. Il suffit que pour cet observateur le train « recule ». Par exemple on prend un observateur dans un autre train allant dans le même sens que le premier à une vitesse plus grande.

Exercice 2.2.2

Notons A et B l'avant et l'arrière du wagon. Albert mesure l'intervalle de temps entre deux événements : « Isaac est en A » et « Isaac est en B ». Ces deux événements sont situés spatialement au même point pour Isaac, il peut donc appliquer la formule de dilatation des durées.

Au contraire, ces deux mêmes événements ne sont pas situés au même point de l'espace pour Albert. Ainsi, Isaac ne peut pas appliquer cette même formule, et on ne peut pas itérer le raisonnement. Isaac devrait faire appel à une transformation de Lorentz (ici un boost) pour calculer l'intervalle de temps mesuré par Albert entre ces deux événements. C'est un très bon exercice de vérifier qu'on trouve bien 10 secondes.

Exercice 2.2.5



Commençons par le point de vue du Soleil. Au bout d'une révolution complète de la Terre, l'astronaute A est arrivé au bout de son voyage : disons qu'il atteint un objet O immobile par rapport au soleil et situé à 0,99 années-lumière du Soleil. Cependant, le calendrier de bord de l'astronaute est ralenti d'un facteur $\gamma \approx 7$. Ainsi, lorsque A est en O , il s'est passé environ $1/7$ d'année pour A . L'événement « A est en O » est simultané avec l'événement « la Terre a complété une révolution », du point de vue du Soleil. Du point de vue de l'astronaute, l'objet O est atteint au bout de $1/7$ d'année, comme nous venons de le voir. De plus, en regardant la Terre, il voit une sorte de grande horloge animée d'un mouvement uniforme de vitesse v . Par conséquent cette horloge est ralentie d'un facteur γ . Aussi, du point de vue de l'astronaute la Terre a tourné d'environ $1/49$ de tour lorsqu'il atteint O .

Montrons que l'on peut retrouver tous ces résultats par des transformations de Lorentz. Nous choisissons comme unité de distance l'année-lumière, et comme unité de temps l'année, de sorte que les vitesses soient exprimées en fraction de c . Notons R le référentiel héliocentrique et R' le référentiel de l'astronaute.

Nous considérons seulement une dimension d'espace orientée dans le sens du mouvement de l'astronaute. L'événement « départ de l'astronaute » a pour coordonnées $(0; 0)$ dans les deux référentiels. L'événement « A est en O » a pour coordonnées dans $R : (x, t) = (0, 99; 1)$, et dans $R' : (x'; t')$. On a naturellement $x' = 0$ puisque A est l'origine de R' . Cherchons t' en appliquant la formule du boost :

$$t' = \gamma(t - \frac{vx}{c^2})$$

On obtient $t' = \frac{1}{\sqrt{1-0,99^2}}(1 - 0,99^2) = \sqrt{1-0,99^2} = 1/\gamma \approx 1/7$.

Considérons maintenant le mouvement de la Terre. Soit l'événement de coordonnées $(0; \frac{\theta}{2\pi})$ dans R . Il correspond à une rotation de la Terre d'un angle θ autour du Soleil à partir de sa position initiale. En appliquant un boost on voit que cet événement a pour coordonnées $(0; \gamma \frac{\theta}{2\pi})$ dans R' . Si on cherche θ pour que cet événement soit simultané avec l'arrivée de A en O dans R' , on voit qu'il faut prendre $\theta/2\pi = 1/\gamma^2 \approx 1/49$.

Nous constatons que l'assimilation du système Terre-Soleil à une horloge localisée à l'origine de R est justifiée par le fait que le mouvement de la Terre est perpendiculaire au mouvement. Ainsi, les coordonnées y et z de la Terre n'ont aucune influence sur la transformation en x, t , et on peut appliquer la formule de dilatation des durées pour les événements situés spatialement au même point dans R .

Exercice 2.3.1

La trajectoire du mobile dans R est donnée par $x(t), y(t), z(t)$, et l'on a $\vec{u} = \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \\ \frac{dz}{dt} \end{pmatrix}$.

Calculons

$$\vec{u}' = \begin{pmatrix} \frac{dx'}{dt'} \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{dy'}{dt'} \\ \frac{dz'}{dt'} \end{pmatrix} = \vec{u}_{\parallel} + \vec{u}_{\perp}$$

On a :

$$\frac{dx'}{dt'} = \frac{dx'}{dt} \frac{dt}{dt'}$$

Or par la formule du boost de R vers R' on a :

$$\frac{dx'}{dt'} = \gamma(\frac{dx}{dt} - v) \frac{dt}{dt'}$$

et par le boost inverse on obtient :

$$t = \gamma(t' + \frac{vx'}{c^2})$$

D'où :

$$\frac{dx'}{dt'} = \gamma(\frac{dx}{dt} - v)\gamma(1 + \frac{dx'}{dt'} \frac{v}{c^2})$$

En réorganisant on obtient le résultat. En procédant de la même manière on trouve dy'/dt' et dz'/dt' et l'on déduit la formule.

Exercice 2.3.3 : Dans R la vitesse de la Terre se décompose en $u_{\parallel} = 0$ et $u_{\perp} = U = cte$. En appliquant les formules, on trouve $u'_{\parallel} = -v$ et $u'_{\perp} = U/\gamma \approx U/7$. La vitesse de révolution de la Terre autour du Soleil est donc ralentie d'un facteur 7 pour R' .

Exercice 2.4.2 : On voit clairement que cette relation est invariante par rotation et translation. Il faut vérifier qu'elle l'est par un boost, et pour cela, on se ramène au cas où M_1 et M_2 ont pour coordonnées respectives $(x_1, 0, 0, t_1)$ en $(x_1, 0, 0, t_2)$ dans R , et où le boost est selon l'axe des x . Alors on trouve :

$$t'_2 - t'_1 = \gamma(t_2 - t_1) - \frac{v}{c^2}(x_2 - x_1)$$

en utilisant l'hypothèse :

$$c^2(t_2 - t_1)^2 \geq (x_2 - x_1)^2$$

et le fait que $v \leq c$ et $\gamma \geq 1$, on trouve que $t'_2 - t'_1$ est ≥ 0 si $t_2 - t_1 \geq 0$.

Exercice 2.4.3 :

Non, c'est impossible. En effet si M_1 est l'événement « le ballon entre dans le panier » et $M_2 =$ « le score augmente » alors $M_1 \preceq M_2$, ainsi qu'il apparaît clairement dans le référentiel du panier par exemple (les deux événements sont causalement reliés, et l'un a lieu avant l'autre). Donc ceci sera vrai dans tous les référentiels.

Exercice 2.4.4 : Non, c'est faux. En effet si les deux événements ne sont pas causalement reliés, on peut trouver un référentiel où l'un a lieu avant l'autre, et un autre dans lequel on a l'inverse. Nous avons déjà vu des exemples dans le paragraphe 2.2.

Exercice 2.4.6 :

Considérons $L \in \mathcal{L}$. Notons e_0, e_1, e_2, e_3 les vecteurs de la base canonique. Par conservation de g , $(L(e_0), L(e_1), L(e_2), L(e_3))$ est encore une base g -orthonormée. Nous allons montrer qu'il existe une transformation de Lorentz f telle que $f(L(e_0)) = e_0, \dots, f(L(e_3)) = e_3$. Alors il sera clair que $L = f^{-1}$ est encore une transformation de Lorentz. Nous utiliserons à plusieurs reprises qu'une transformation de Lorentz préserve l'orthogonalité au sens de g , puisqu'elle préserve g .

Écrivons $L(e_0) = \alpha_0 e_0 + \vec{v}$, où l'on a mis une flèche pour indiquer que \vec{v} appartient à l'espace $E = \text{Vect}\{e_1, e_2, e_3\}$. On a $|\alpha_0|^2 - \|\vec{v}\|^2 = 1$, où la norme désigne la norme euclidienne de l'espace E . Si $\vec{v} = 0$ alors $L(e_0) = \pm e_0$. Dans ce cas, quitte à composer avec un renversement du temps, on peut se ramener au cas où $L(e_0) = e_0$. Mais alors, les vecteurs $L(e_1), L(e_2), L(e_3)$ qui sont nécessairement g -orthogonaux à e_0 vont former une base orthonormée de E . Il existe alors une rotation de cette espace qui amène $L(e_1)$ sur e_1 , etc. . . et le tour est joué. Si $\vec{v} \neq 0$, on complète par un procédé du type Gram-Schmidt la famille $(e_0, \frac{\vec{v}}{\|\vec{v}\|})$ en base g -orthonormée de \mathcal{M} . Dans cette base, $L(e_0)$ a pour coordonnées $(\alpha_0, \|\vec{v}\|, 0, 0)$, et $e_0 = (1, 0, 0, 0)$. On voit alors que le boost de direction \vec{v} et de vitesse $\|\vec{v}\|/\alpha_0 < 1$ amène $L(e_0)$ sur e_0 . Nous sommes alors ramené au cas précédent.

Exercice 2.5.3 : On traite dans un premier temps le cas où les trois événements sont séparés mutuellement par des intervalles du genre temps. On considère alors un corps animé d'un mouvement rectiligne uniforme se propageant de A en C . Dans le référentiel de ce corps, en prenant A comme origine, on a un diagramme d'espace-temps tout-à-fait similaire à celui tracé pour résoudre le paradoxe des jumeaux. Le résultat suit très simplement, et le cas général s'obtient par un passage à la limite : si $[AB]$ est du genre lumière, par exemple, alors on considère deux suites d'événements A_n et B_n tels que $[A_n B_n]$ soit du genre temps et converge vers $[AB]$, et on applique l'inégalité précédente. On peut alors passer à la limite, ce qui préserve les inégalités larges.

Exercice 2.5.4 : Le temps mesuré par l'horloge au sol est de $t = 40$ h. Le temps mesuré par l'horloge en vol est :

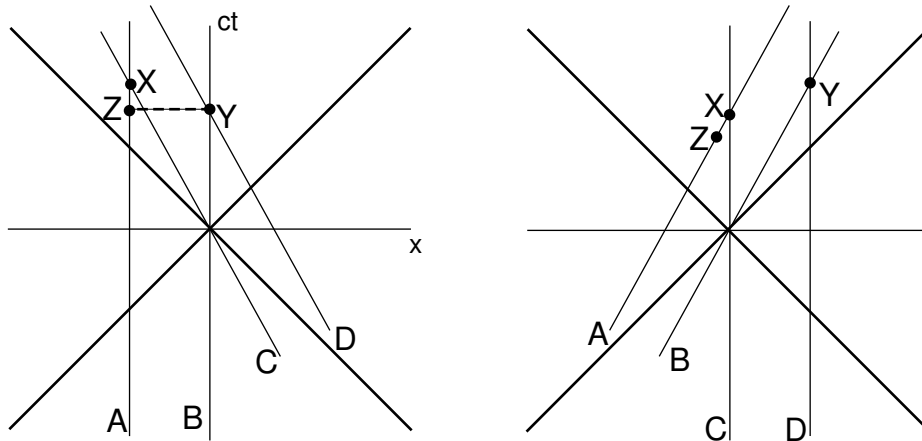
$$\tau = \int_0^{40} \sqrt{1 - v^2/c^2} dt$$

Comme la vitesse est constante, on obtient :

$$\tau = \gamma^{-1} \times 40 \text{ h}$$

L'application numérique donne une différence de l'ordre de 60 nanosecondes entre t et τ , ce qui peut être mesuré par une horloge atomique. Cependant, un autre effet, de plus grande ampleur, s'ajoute à la dilatation due à la vitesse. Cet effet dont nous reparlerons est d'origine gravitationnelle.

Exercice 2.5.5 : En traçant les diagrammes d'espace-temps dans le référentiel R du garage puis dans celui R' de la voiture, on voit que le problème vient de la relativité de la simultanéité. En effet, si les portes se ferment simultanément dans R , la porte A se ferme avant la porte B dans R' . Faisons le calcul explicite.



On a représenté les lignes d'univers des portes A et B et des extrémités C et D de la voiture.

On suppose qu'à l'instant $t = t' = 0$ les points B et C coïncident.

L'événement X : « $A = C$ » intervient à l'instant $t_1 = L/v$ dans R , où L est la longueur du garage, égale à la longueur au repos de la voiture, et v la valeur absolue de la vitesse de la voiture. Dans R' cet événement a pour coordonnée de temps $t'_1 = L/(\gamma v) = t/\gamma$.

Considérons maintenant l'événement Y : « $B = D$ ». Cet événement a pour coordonnées respectives dans les deux repères : $t_2 = L/(\gamma v)$, et $t'_2 = L/v$. Or comme $\gamma > 1$, on a $t_2 < t_1$ tandis que $t'_1 < t'_2$. Supposons que les portes se ferment simultanément à l'instant t_2 dans R .

Considérons maintenant l'événement Z : « fermeture de la porte A ». Cet événement a pour coordonnées $(-L, L/(\gamma v))$ dans R , par hypothèse. Pour chercher ses coordonnées dans R' on effectue un boost (attention la voiture se dirige vers les x négatifs). En appelant t''_2 la coordonnée de temps de la fermeture de la porte en A dans R' on trouve $t''_2 = L/v - \gamma v L/c^2$. On trouve que $t''_2 < t'_1$ (exercice). Résumons tout ceci dans un tableau :

événement	coord. de temps dans R	coord. de temps dans R'
X : A=C	L/v	$L/\gamma v$
Y : B=D, porte B se ferme	$L/\gamma v$	L/v
Z : porte A se ferme	$L/\gamma v$	$L/v - \gamma(vL/c^2)$
ordre des événements	Y et Z puis X	Z, puis X, puis Y

Le récit d'un observateur de R serait le suivant : les deux portes se sont fermées au moment où l'arrière de la voiture franchissait la porte arrière (porte B) du garage. À cet instant la voiture était entièrement contenue dans le garage. Puis l'avant de la voiture a percuté la porte avant (A) du garage et l'a défoncée. Un observateur de R' dirait plutôt : la porte A du garage s'est fermée, puis la voiture l'a enfoncé, puis l'arrière de la voiture a franchi la porte B au moment où celle-ci se fermait.

On voit que ces récits se contredisent seulement sur l'ordre dans lequel se sont produits des événements qui ne sont pas causalement reliés. Il n'y a donc pas de paradoxe. Il faut bien comprendre que la phrase « la voiture est entièrement contenue dans le garage » porte sur la position de l'avant et de l'arrière de la voiture *au même instant*, ce qui n'a pas de sens physique.

Exercice 3.3.1 : En appelant r la distance au centre de la Terre on a $x = r\theta$, avec θ une constante très petite, et on trouve $\ddot{x} = -\frac{GM}{r^3}x$ qui est une équation qu'on peut envisager comme linéaire en x et à coefficient constant sur un intervalle de temps assez court pour que les variations de $1/r^3$ soient négligeables.

Exercice 3.3.2 : C'est une spirale d'Archimède !

Voir l'article consacré à cette courbe sur <http://www.mathcurve.com>

Exercice 3.5.1 : On trouve :

$$\Gamma_{\phi\phi}^{\theta} = -\sin\theta \cos\theta; \Gamma_{\phi\theta}^{\phi} = \Gamma_{\theta\phi}^{\phi} = \frac{1}{\tan\theta}$$

Exercice 3.5.2 : En coordonnées cartésiennes (3.13) est un système d'équations linéaires dont le second membre est nul, et on sait que ce système a une unique solution, qui est donc nécessairement la solution nulle. (On peut aussi appliquer directement la formule 3.14, puisque toutes les dérivées de la métrique sont nulles). L'équation des géodésiques est donc :

$$\ddot{x}^{\mu} = 0$$

qui s'intègre en

$$x^{\mu} = a^{\mu}t + b^{\mu}$$

On trouve bien une droite.

En coordonnées sphériques on a :

$$ds^2 = c^2 dt^2 - [dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)]$$

Autrement dit $g_{\mu\nu}$ est donnée, en numérotant les coordonnées dans l'ordre t, r, θ, ϕ , par la matrice diagonale :

$$g_{\mu\nu} = \begin{pmatrix} c^2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2\theta \end{pmatrix}$$

L'inverse de cette matrice étant diagonale, on déduit de (3.14) les équations :

$$\Gamma_{bc}^a = \frac{1}{2}g^{aa}(\partial_b g_{ac} + \partial_c g_{ab} - \partial_a g_{bc})$$

où il n'y a pas de sommation sur a . Les seules dérivées partielles non nulles étant $\partial_{\theta}g_{\phi\phi}$, $\partial_r g_{\phi\phi}$, et $\partial_r g_{\theta\theta}$, on en déduit immédiatement que les seuls symboles susceptibles d'être non nuls sont : $\Gamma_{\theta\theta}^r, \Gamma_{\phi\phi}^{\theta}, \Gamma_{\phi\phi}^r, \Gamma_{r\theta}^{\theta} = \Gamma_{\theta r}^{\theta}, \Gamma_{\phi r}^{\phi} = \Gamma_{r\phi}^{\phi}, \Gamma_{\phi\theta}^{\phi} = \Gamma_{\theta\phi}^{\phi}$.

On trouve :

$$\Gamma_{\theta\theta}^r = \frac{1}{2}g^{rr}(-\partial_a(-r^2)) = -r$$

$$\Gamma_{\phi\phi}^a = -\frac{1}{2}g^{aa}\partial_a(-r^2 \sin^2\theta)$$

d'où : $\Gamma_{\phi\phi}^{\theta} = -\sin\theta \cos\theta$, $\Gamma_{\phi\phi}^r = -r \sin^2\theta$.

$$\Gamma_{r\theta}^{\theta} = \frac{1}{2r^2}\partial_r(r^2) = \frac{1}{r}$$

Enfin on a :

$$\Gamma_{\phi c}^{\phi} = \frac{1}{2r^2 \sin^2\theta}\partial_c(r^2 \sin^2\theta)$$

d'où $\Gamma_{\phi\theta}^{\phi} = \frac{\cos\theta}{\sin\theta}$, $\Gamma_{\phi r}^{\phi} = \frac{1}{r}$.

Finalement les seuls symboles non-nuls sont :

$$\Gamma_{\theta\theta}^r = -r; \Gamma_{\phi\phi}^{\theta} = -\sin\theta \cos\theta; \Gamma_{\phi\phi}^r = -r \sin^2\theta; \Gamma_{r\theta}^{\theta} = \Gamma_{\phi r}^{\phi} = \frac{1}{r}; \Gamma_{\phi\theta}^{\phi} = \frac{\cos\theta}{\sin\theta}$$

et ceux qui s'en déduisent par symétrie des deux indices du bas.

Soit $x^\mu(s)$ une géodésique paramétrée par s . La première des équations de (3.12) donne :

$$\ddot{t} = 0 \Rightarrow t = as + b$$

si $a = 0$ on a une géodésique du genre espace entièrement incluse dans une tranche spatiale $t = cte$. Laisant ce cas de côté, on peut toujours modifier le paramètre et supposer que la géodésique est paramétrée par t . Regardons les autres équations :

$$\begin{aligned} \ddot{r} + \Gamma_{\theta\theta}^r \dot{\theta}\dot{\theta} + \Gamma_{\phi\phi}^r \dot{\phi}\dot{\phi} &= 0 \Leftrightarrow \ddot{r} - r\dot{\theta}^2 - r\sin^2\theta\dot{\phi}^2 = 0 \\ \ddot{\theta} - \sin\theta\cos\theta\dot{\phi}^2 + \frac{2}{r}\dot{r}\dot{\theta} &= 0 \\ \ddot{\phi} + \frac{2}{r}\dot{\phi}\dot{r} + \frac{2\cos\theta}{\sin\theta}\dot{\phi}\dot{\theta} &= 0 \end{aligned}$$

Pour vérifier que ces équations différentielles ont bien des droites pour solutions, on peut vérifier que le système est équivalent à $\ddot{x} = \ddot{y} = \ddot{z} = 0$, avec x, y, z exprimés en fonction de r, θ, ϕ . Cela ne présente pas de difficulté particulière mais c'est un peu fastidieux. Vérifions seulement que les trois équations précédentes impliquent $\ddot{z} = 0$.

On a :

$$z = r \cos \theta, \quad \dot{z} = \dot{r} \cos \theta - r \dot{\theta} \sin \theta, \quad \ddot{z} = \ddot{r} \cos \theta - r \ddot{\theta} \sin \theta - 2\dot{r}\dot{\theta} \sin \theta - r\dot{\theta}^2 \cos \theta$$

D'où :

$$\ddot{z} = (r\dot{\theta}^2 + r\sin^2\theta\dot{\phi}^2) \cos \theta - (\sin\theta\cos\theta\dot{\phi}^2 - \frac{2}{r}\dot{r}\dot{\theta})r \sin \theta - 2\dot{r}\dot{\theta} \sin \theta - r\dot{\theta}^2 \cos \theta = 0$$

Exercice 3.7.2 : On peut en fait calculer exactement cette intégrale en posant $y = \sqrt{\frac{r}{r-k}}$. On trouve :

$$l = \sqrt{r_2(r_2 - r_S)} - \sqrt{r_1(r_1 - r_S)} + \frac{r_S}{2} \ln\left(\frac{2r_1 - 2\sqrt{r_1(r_1 - r_S)} - r_S}{2r_2 - 2\sqrt{r_2(r_2 - r_S)} - r_S}\right)$$

On note au passage que le résultat est bien défini seulement pour r_1 et $r_2 > r_S$ et que l'intégrale est convergente quand r_1 tend vers r_S . En faisant un DL en r_2/r_S et r_1/r_S à l'ordre 0, on trouve le résultat (il faut faire un DL à l'ordre 2 sous le log pour obtenir le terme d'ordre 0 dans le résultat final).

En particulier pour $r_1 = r_S$ et $r_2 = r \gg r_S$ on obtient le DL :

$$l = r + \frac{r_S}{2} \ln\left(\frac{4r}{r_S}\right)$$

Pour la Terre, $r_S \approx 9$ mm. La différence $l - r$ pour r de l'ordre de 1000 km est de l'ordre de la dizaine de centimètres.

Exercice 3.7.4 : On prend des coordonnées angulaires θ, ϕ telles que le mouvement du satellite soit $r = cte$, $\theta = cte = \pi/2$ et $\phi = kt$ où k est une constante. On a alors pour le satellite :

$$ds^2 = \left(1 - \frac{r_S}{r} dt^2 - r^2 k^2\right) dt^2 = c^2 d\tau^2$$

d'où

$$d\tau_{sat} = \left[1 - \frac{r_S}{r} - v^2/c^2\right]^{\frac{1}{2}} d\tau_{sol}^\infty$$

avec v la vitesse du satellite. On en déduit :

$$\frac{d\tau_{sat}^\infty}{d\tau_{sol}^\infty} = \sqrt{\frac{1 - \frac{r_S}{r_{sol}}}{1 - \frac{r_S}{r_{sat}} - v^2/c^2}}$$

Comme $r_S/r_{sat/sol}$ est de l'ordre de 10^{-9} et $(v/c)^2$ de l'ordre de 10^{-10} , on peut ne garder que les termes d'ordre 1 en ces quantités et on trouve :

$$\begin{aligned}\frac{d\tau_{sat}^\infty}{d\tau_{sol}^\infty} &= \left(1 - \frac{1}{2}r_s/r_{sol}\right)\left(1 + \frac{1}{2}r_s/r_{sat} + v^2/2c^2\right) + \dots \\ &= 1 - \frac{r_s}{2}\left(\frac{1}{r_{sol}} - \frac{1}{r_{sat}}\right) + v^2/2c^2 + \dots\end{aligned}$$

En utilisant la formule de l'exercice 3.7.2 pour la différence $l - r$, on voit que la différence entre r_s/r et r_s/r est en $(r_s/r)^2 \ln(r/r_s)$ qui est négligeable devant r_s/r . On peut donc négliger ici la différence entre r et la distance au centre de la Terre. En utilisant les données on trouve après application numérique :

$$\frac{d\tau_{sat}^\infty}{d\tau_{sol}^\infty} \approx 1 - (0,52 - 0,08) \times 10^{-9}$$

Autrement dit, la seconde du satellite est environ une demi-nanoseconde plus courte que celle du sol. On peut s'apercevoir tout au long du calcul que la formule incorpore l'effet de la relativité restreinte, et que ce dernier est de signe opposé à celui du champ de gravité.

Exercice 3.7.5 :

En développant au premier ordre en $GM/Rc^2 \approx 10^{-9}$ la formule (3.24) on trouve :

$$\frac{\Delta\tau^1}{\Delta\tau^2} \approx 1 + 10^{-9} \frac{h}{R}$$

Sur la durée du voyage, la différence de temps enregistrée est donc de :

$$40 \times 3600 \times 10^{-9} \times \frac{10}{6400} \approx 2 \times 10^{-7} s$$

Cet effet est donc de l'ordre de 200 nanosecondes mesurées en moins par l'horloge au sol. L'effet net, combiné avec celui de la relativité restreinte est de l'ordre de $200 - 60 = 140$ nanosecondes. Bien que les calculs précédents soient très grossiers, ils donnent l'ordre de grandeur correct de l'effet mesuré en 1971 par Hafele et Keating.

Annexe B

Une déduction axiomatique des transformations de Lorentz

On trouve dans de nombreux ouvrages des déductions des transformations de Lorentz quelque peu hasardeuses sur le plan mathématique. En particulier, on écrit souvent que ces transformations sont nécessairement linéaires (on devrait dire affines si l'on parle des transformations les plus générales qui sont celles de Poincaré) en vertu de l'homogénéité de l'espace-temps sans réellement le justifier. Un autre argument est qu'une transformation de Poincaré doit envoyer un mouvement rectiligne uniforme sur un mouvement rectiligne uniforme d'après le principe d'inertie et que cela impose le caractère affine de ces transformations. En effet, le théorème fondamental de la géométrie affine stipule que toute transformation qui conserve l'alignement est affine. Seulement, on oublie dans ce cas que toutes les droites de l'espace-temps ne correspondent pas à des mouvements physiquement possibles. Il est alors peu clair que l'utilisation du principe d'inertie soit pleinement justifié. Si l'on se restreint aux droites du genre-temps, qui correspondent effectivement aux mouvements physiquement possibles de particules matérielles, on ne peut plus appliquer le théorème fondamental de la géométrie affine tel quel. On peut montrer que le résultat tient toujours mais cela n'est pas en général indiqué dans la littérature.

Nous allons donner ci-dessous une déduction de la forme générale des transformations de Poincaré en nous fondant sur trois principes : le principe de relativité et le principe d'invariance de c , à la base de la théorie de la relativité, et l'hypothèse que l'espace-temps est homogène.

Soient R et R' deux référentiels inertiels, tel que R' soit animé d'une vitesse v par rapport à R . On appelle L l'application de \mathbb{R}^4 dans \mathbb{R}^4 qui aux coordonnées (x_0, x_1, x_2, x_3) d'un événement P dans R associe les coordonnées (x'_0, \dots, x'_3) de ce même événement dans R' . Nous appellerons L^μ avec $\mu = 0, \dots, 3$ les différentes composantes de L , i.e. :

$$L(x^0, \dots, x^3) = \begin{pmatrix} L^0(x^0, \dots, x^3) \\ \vdots \\ L^3(x^0, \dots, x^3) \end{pmatrix}$$

Première étape : L est de classe \mathcal{C}^1 . Il est difficile d'imaginer que L ne soit pas de classe \mathcal{C}^1 , cependant donnons un argument. Soit ϕ un champ scalaire de classe \mathcal{C}^1 , c'est-à-dire une application $\phi : \mathbb{R}^4 \rightarrow \mathbb{R}$ (on peut penser que ϕ représente par exemple un potentiel défini sur l'espace-temps). Si ϕ représente ce champ dans R , alors $\phi \circ L^{-1}$ représente ce champ dans R' . En vertu du principe de relativité, $\phi \circ L^{-1}$ est aussi de classe \mathcal{C}^1 . Comme ceci est vrai pour tout ϕ , cela impose que L soit de classe \mathcal{C}^1 .

Deuxième étape : L est affine. Soit $\psi : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ un champ vectoriel exprimé dans R . Soit ψ' ce champ vectoriel exprimé dans R' . Pour tout $p' \in \mathbb{R}^4$ on a la formule :

$$\psi'(p') = d_p L(\psi(p))$$

où $p = L^{-1}(p')$ et $d_p L$ est l'application différentielle de L , encore appelée application linéaire tangente à L en p . Cette application, en coordonnées, n'est rien d'autre que la matrice jacobienne de L . Ainsi on a :

$$\psi'(p') = \frac{\partial L^\mu}{\partial x^\nu}(p) u^\nu$$

où l'on a noté u^ν les composantes de $\psi(p)$ et utilisé la convention de sommation d'Einstein. D'après l'homogénéité de l'espace-temps, la façon dont le vecteur $\psi(p)$ défini au point p voit ses composantes se transformer en passant de R à R' ne peut dépendre que de celles-ci et pas du point p . Ainsi les dérivées partielles de L doivent être indépendantes de p . Il existe donc des constantes $\Lambda^\mu{}_\nu$ telles que :

$$\frac{\partial L^\mu}{\partial x^\nu} = \Lambda^\mu{}_\nu$$

d'où par intégration :

$$L^\mu = \Lambda^\mu{}_\nu x^\nu + k^\mu$$

où les k^μ sont des constantes. En d'autres termes, L est une application affine dont la partie linéaire est Λ , donnée en coordonnées par la matrice $\Lambda^\mu{}_\nu$.

Troisième étape : Λ conserve la forme η . Quitte à composer ultérieurement par une translation, supposons que $k^\mu = 0$ et concentrons-nous maintenant sur la partie linéaire Λ . Un événement $p = (x_0, \dots, x_3)$ appartient au cône de lumière ssi $q(p) = x_0^2 - (x_1^2 + x_2^2 + x_3^2) = 0$.

D'après l'invariance de c , si p appartient au cône de lumière, il en va de même pour Λp . Ainsi :

$$q(p) = 0 \Rightarrow q(\Lambda p) = 0 \tag{B.1}$$

Notons $B_{\mu\nu}$ les composantes de la forme quadratique $q \circ \Lambda$, i.e. :

$$q \circ \Lambda(x) = B_{\mu\nu} x^\mu x^\nu$$

En appliquant (B.1) à $p = (1, -1, 0, 0)$, puis $(1, 0, -1, 0)$ et $(1, 0, 0, -1)$ on trouve que $B_{0i} = 0$ et $B_{ii} = -B_{00}$, pour $i = 1, 2, 3$. En utilisant enfin $p = (\sqrt{2}, -1, -1, 0)$ on trouve $B_{12} = 0$ et de même on trouve $B_{ij} = 0$ pour $i \neq j$. Ainsi, en posant $\alpha = B_{00}$, on a :

$$q \circ \Lambda = \alpha q$$

Or la forme quadratique q détermine de façon unique la forme bilinéaire η par l'identité de polarisation :

$$\eta(u, v) = \frac{1}{2}(q(u+v) - q(u) - q(v))$$

Ainsi on trouve que (matriciellement) :

$${}^t \Lambda \eta \Lambda = \alpha \eta$$

Donc Λ conserve η à un facteur près, qui correspond à une homothétie, i.e. un changement d'unités de longueurs et de durée. En fixant les unités, on peut supposer que $\alpha = 1$, et ainsi on a montré que Λ conserve η . La forme générale de Λ suit alors du théorème 2.4.2.

Annexe C

Le tenseur énergie-impulsion

Ce qui suit est un résumé du résumé du chapitre 5 de [9], p. 131, destiné à donner une idée du contenu du tenseur énergie-impulsion.

En tant qu'objet mathématique, $T_{\mu\nu}$ est une forme bilinéaire qui peut s'appliquer à deux vecteurs tangents. Soit u^ν la quadrivitesse d'un observateur. Alors, pour tout point m de l'espace-temps on a :

$$T_{\mu\nu}u^\nu = -\text{densité de quadri-impulsion en } m$$

le tout mesuré dans le référentiel de Lorentz local de l'observateur au point m . De plus :

$$T_{\mu\nu}u^\mu u^\nu = \text{densité d'énergie-masse en } m$$

Prenons un exemple. Supposons que l'on puisse négliger les effets gravitationnels dans un certain référentiel muni de coordonnées (t, x, y, z) et considérons un observateur éventuellement accéléré passant en (x, y, z) à l'instant t , muni d'une petite boîte. En calculant $T_{\mu\nu}u^\nu$ en (t, x, y, z) , on trouve moins la somme de toutes les impulsions de toutes les particules se trouvant dans la boîte, divisé par le volume de cette boîte, telle que mesurée par l'observateur en ce point et à cet instant.

Enfin on peut montrer que $T_{\mu\nu}$ est un tenseur symétrique.

Le tenseur énergie-impulsion d'un fluide parfait dans le référentiel de son centre d'inertie est :

$$T_{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & -p & 0 & 0 \\ 0 & 0 & -p & 0 \\ 0 & 0 & 0 & -p \end{pmatrix}$$

où ρ est la densité et p la pression.

Bibliographie

- [1] M. Audin, *Geometry*, Springer
- [2] S. Bergia, *Einstein, le père du temps moderne*, Les génies de la science
- [3] M. Bertin, J.P. Faroux, J. Renault, *Mécanique 1*, Dunod
- [4] J. Eisenstaedt, *Einstein et la relativité générale*, CNRS éditions
- [5] A. Einstein, *La théorie de la relativité restreinte et générale*, Dunod
- [6] A. Einstein, *Les fondements de la théorie de la relativité générale*, in « Sur les épaules des géants », S. Hawking, Dunod.
- [7] Ray d'Inverno, *Introducing Einstein's relativity*, Clarendon Press
- [8] R. Hakim, *Gravitation Relativiste*, InterÉditions
- [9] C. Misner, K. Thorne, J. Wheeler, *Gravitation*, Freeman
- [10] J. Schwartz, M. McGuinness, *Introducing Einstein*, Icon Books