HAL open science

Relativité générale pour débutants

Michel Le Bellac

To cite this version:

Michel Le Bellac. Relativité générale pour débutants. DEA. 2004. cel-00092961

HAL Id: cel-00092961

https://cel.hal.science/cel-00092961

Submitted on 12 Sep 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Prétirage INLN 2004/17

RELATIVITE GENERALE POUR DEBUTANTS

Michel Le Bellac

Cours donné aux Rencontres non linéaires de Peyresq

Mai 2004

Résumé. Ce cours a pour objectif d'exposer à un public non initié les idées de base de la relativité générale. Il ne suppose aucun prérequis : il contient les notions nécessaires de relativité restreinte et de géométrie différentielle. Les applications traitées sont la cosmologie et les trous noirs.

INSTITUT NON LINEAIRE DE NICE UMR 6638
1361 routes des Lucioles 06560 Valbonne
e-mail : michel.le_bellac@inln.cnrs.fr

Table des matières

1 Introduction **5**
1.1 Bref historique . . . 5
1.2 Plan du cours . . . 6
1.3 Quelques références générales . . . 7

2 Principe d'équivalence **9**
2.1 Référentiels d'inertie . . . 9
2.2 Principe d'équivalence . . . 10
2.3 Décalage vers le rouge gravitationnel . . . 12
2.4 Interprétation géométrique . . . 14
2.5 Effets de marée gravitationnels . . . 15

3 Espace-temps plat **19**
3.1 Photons . . . 19
3.2 Effet Doppler . . . 21
3.3 Métrique de l'espace-temps plat . . . 22
3.4 Tenseur énergie-impulsion . . . 25

4 Cosmologie **29**
4.1 Description qualitative de l'Univers . . . 29
4.2 Coordonnées comobiles . . . 31
4.3 Évolution du facteur de dilatation . . . 34
4.4 Cas de la courbure spatiale non nulle . . . 39
4.5 Le modèle ΛCDM . . . 41

5 Boîte à outils de géométrie différentielle **45**
5.1 Espace tangent à une variété . . . 45
5.2 Champs de tenseurs . . . 48
5.3 Connexions . . . 48
5.4 Métrique et courbure . . . 52
5.5 Adaptation à la relativité générale . . . 56

6 Solutions à symétrie sphérique **59**
6.1 Équation d'Einstein . . . 59
6.2 Métrique de Schwarzschild . . . 61
6.3 Trous noirs . . . 63
6.4 Remerciements. . . . 67

Chapitre 1

Introduction

1.1 Bref historique

Ce cours a pour objectif d'exposer de la façon la plus élémentaire possible les idées de la relativité générale, c'est-à-dire la théorie relativiste de la gravitation. Dans cette introduction, je commencerai par une rapide revue de l'histoire de la relativité générale. En 1905 paraît l'article d'Einstein sur la relativité restreinte, l'un des articles de "l'année miraculeuse". Même si Einstein fut le premier à les interpréter dans le cadre de l'espace-temps, les idées de la relativité restreinte étaient "dans l'air", et il est vraisemblable que Lorentz ou Poincaré (ou un autre) seraient arrivés rapidement à des conclusions identiques. Après avoir établi la relativité restreinte, Einstein commença immédiatement à réfléchir à une théorie relativiste de la gravitation. Il énonça dès 1907 le principe d'équivalence entre gravité et accélération constantes (voir chapitre 2), mais il lui fallut encore huit ans avant d'établir à la fin de 1915 les fondements définitifs d'une théorie géométrique de la gravitation, la relativité générale. Contrairement au cas de la relativité restreinte, il est manifeste que la contribution d'Einstein à la relativité générale est unique, et que sans lui la relativité générale aurait probablement attendu quelques dizaines d'années avant d'être inventée. En effet, la relativité générale est une construction purement intellectuelle, pour laquelle il n'y avait aucune nécessité expérimentale. La théorie de Newton de la gravitation était en accord remarquable avec les mesures très précises de l'astronomie[1], contrairement à la mécanique et à l'électromagnétisme classiques qui commençaient à rencontrer des difficultés. En fait la relativité générale est restée une théorie ésotérique aux yeux de la majorité des physiciens pendant plus d'une cinquantaine d'années, et c'est seulement après la mort d'Einstein en 1955 qu'elle s'est considérablement développée, pour devenir une théorie incontournable de la physique moderne. On peut donner quelques dates clés de son histoire.

- 1916 : Einstein publie dans *Annalen der Physik* son article sur la relativité générale et explique l'avance du périhélie de Mercure. Quelques mois plus tard, Schwarzschild établit la forme (1.2) de la métrique qui porte son nom, et qui généralise la loi de Newton donnant le potentiel gravitationnel $\Phi(r)$ d'une masse ponctuelle M sans structure et avec symétrie sphérique

  $$\Phi(r) = -\frac{GM}{r} \qquad U(r) = -\frac{GMm}{r} \tag{1.1}$$

  où G est la constante de gravitation, r est la distance entre la masse M et le point d'observation et $U(r)$ l'énergie potentielle d'une masse m située à une distance r de M.

  L'expression de la métrique de Schwarzschild, qui sera établie au chapitre 6, est

  $$\mathrm{d}s^2 = \left(1 - \frac{r_S}{r}\right)\mathrm{d}t^2 - \left(1 - \frac{r_S}{r}\right)^{-1}\mathrm{d}r^2 - r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2\right) \tag{1.2}$$

  Dans cette équation, r_S est le rayon de Schwarzschild : $r_S = 2GM/c^2$, c étant la vitesse de la lumière. La métrique de Schwarzschild est l'équivalent einsteinien du potentiel gravitationnel newtonien d'une masse ponctuelle M.

[1] La seule difficulté expérimentale était l'absence d'explication de l'avance du périhélie de Mercure, de 43" (seconde d'arc) par siècle !

- 1919 : une éclipse de Soleil permet de mesurer la déviation d'un rayon lumineux par le Soleil, vérifiant une des prédictions majeures de la relativité générale : l'action de la gravitation sur la lumière. Malgré la précision discutable des expériences et les incertitudes expérimentales, ce résultat rend Einstein célèbre et il devient une vedette médiatique.
- 1929 : l'expansion de l'Univers devient une hypothèse plausible et Hubble établit la loi qui porte son nom (chapitre 4). Einstein abandonne la constante cosmologique qu'il avait introduite de façon *ad hoc* pour rendre compte d'un Univers stationnaire, et qualifie cette introduction de plus grosse erreur de sa vie. En 1932 il publie avec de Sitter le premier modèle "moderne" d'Univers en expansion (voir la section 4.3.1). Puis plus rien d'important ne se passe jusqu'au début des années 1960 ! La relativité générale subit de plein fouet la concurrence de la physique quantique et le nombre d'articles consacrés à la relativité générale dans *Physical Review* est en chute libre.
- 1960 est l'année d'un progrès théorique important : la métrique de Schwarzschild (1.2) semble singulière à $r = r_S$, et cette singularité empêche de comprendre ce qui se passe dans la région $r \leq r_S$. Kruskal et Szekeres proposent indépendamment un système de coordonnées démontrant que la singularité est en fait un artefact du choix des variables (t, r). Cette observation permettra de lancer la physique des trous noirs (chapitre 6).
- 1960 : Pound et Rebka vérifient expérimentalement le principe d'équivalence grâce à une expérience exploitant l'effet Mössbauer.
- 1965 : Penzias et Wilson découvrent par accident le rayonnement cosmologique à 3 K. La découverte de ce rayonnement, prévu par Gamow en 1948, relance la cosmologie sur des bases expérimentales, dans un cadre qui rend indispensable l'utilisation de la relativité générale.
- 1968 : Jocelyn Bell observe le premier pulsar, qui se révèle être une étoile à neutrons. Cette découverte vaudra le prix Nobel aux deux "seniors " de l'équipe[2].
- 1974 : Hulse et Taylor observent le premier pulsar binaire, qui leur permet de vérifier l'émission d'ondes gravitationnelles par les étoiles sur orbite.
- 1976 : l'expérience mesurant le retard gravitationnel d'un écho radar, proposée par Shapiro en 1965, permet la vérification la plus précise ($\sim 10^{-3}$) à cette date de la relativité générale.
- 1998 : deux expériences indépendantes observent l'accélération de l'expansion de l'Univers : $\ddot{a} > 0$ (chapitre 4). La constante cosmologique, qui donne une explication plausible de cette accélération, revient par la grande porte !
- 2003 : le satellite WMAP observe les fluctuations angulaires de température du rayonnement cosmologique à 3 K, ce qui permet de confirmer le jeu de paramètres du modèle standard actuel de la cosmologie (modèle ΛCDM, chapitre 4).

1.2 Plan du cours

Le plan du cours sera le suivant

1. Principe d'équivalence
2. Espace-temps plat
3. Cosmologie
4. Boîte à outils de géométrie différentielle
5. Solutions à symétrie sphérique

Le choix des sujets a été effectué en fonction de leur intérêt physique et de la simplicité de leur traitement théorique. De plus le choix de la cosmologie (chapitre 4) et des trous noirs (chapitre 6) permet d'étudier la relativité générale dans des conditions où elle est essentielle, où elle n'intervient pas seulement comme une petite correction à la gravitation newtonienne. La liste ci-dessous donne un aperçu de la multitude de sujets couverts par la relativité générale, que le caractère limité de ce cours ne permet pas d'aborder ici.

[2] L'"oubli" de Jocelyn Bell n'est malheureusement pas un cas isolé pour les femmes scientifiques : voir Lise Meitner, Rosalind Franklin ...

- Tests classiques (i. e. dans le système solaire) de la relativité générale : précession du périhélie, déviation de la lumière, retard gravitationnel d'un écho radar, entraînement de référentiels d'inertie.
- Ondes gravitationnelles : émission et détection.
- Lentilles gravitationnelles.
- Astrophysique relativiste : pulsars.
- Solutions axisymétriques (métrique de Kerr).
- Théorèmes de singularités : Penrose, Hawking...
- Gravitation quantique.
- etc.

1.3 Quelques références générales

Je donne ci-dessous une bibliographie sommaire qui sera précisée pour chaque chapitre, où les livres seront cités, à titre d'exemple, sous la forme Hartley [2003]

1. Carrol [2004] : S. M. Mc Carrol, *Gravitation*, Addison-Wesley, New-York.
2. Damour [2005] : T. Damour, *Relativité Générale*, contribution à l'ouvrage collectif *Actualité d'Einstein*, EDPSciences/CNRS Editions, Paris.
3. Doubrovine [1983] : B. Doubrovine et al. *Géométrie contemporaine*, Editions Mir, Moscou
4. Eisenstaedt [2002] : J. Eisenstaedt *Einstein et la relativité générale*, CNRS, Paris (sur l'histoire de la relativité générale).
5. Hartle [2003] :J. Hartle *Gravity*, Addison-Wesley, New-York
6. Ludvigsen [2000] : M. Ludvigsen *La relativité générale : une approche géométrique*, Dunod, Paris.
7. Wald [1984] : R. M. Wald, *General Relativity*, The University of Chicago Press, Chicago.
8. Weinberg [1972] : S. Weinberg, *Gravitation and Cosmology*, John Wiley, New-York.

Chapitre 2

Principe d'équivalence

2.1 Référentiels d'inertie

La présentation des cours élémentaires de mécanique est fondée sur la notion de référentiel d'inertie : ce sont les référentiels où les lois de Newton s'expriment simplement. Par exemple la seconde loi de Newton stipule qu'en l'absence de forces, une particule suit un mouvement rectiligne uniforme dans un référentiel d'inertie. Un référentiel en mouvement de translation uniforme par rapport à un référentiel d'inertie est aussi un référentiel d'inertie. Cependant se pose immédiatement la question : comment s'assurer qu'un référentiel est inertiel ? Le plus simple est de s'assurer que l'absence de force entraîne l'absence d'accélération : $\vec{F} = 0 \Longrightarrow \vec{a} = 0$. Mais comment peut-on savoir que $\vec{F} = 0$ autrement qu'en mesurant une accélération nulle ? On tombe manifestement dans un cercle vicieux. Toutefois la situation n'est pas aussi mauvaise qu'il n'y paraît, car on peut s'appuyer sur les propriétés connues des forces. Supposons par exemple que nous voulions nous assurer de l'absence de forces électriques sur un ion O_{16}^{+} de masse m et de charge q. Bien sûr, dans un référentiel d'inertie, on aura $\vec{F} = q\vec{E}$, où $\vec{E}$ est le champ électrique, et on mesurera donc l'accélération correspondante $\vec{a} = q\vec{E}/m$. Cependant dans un référentiel non inertiel d'accélération $\vec{A}$ par rapport à un référentiel d'inertie (en se limitant pour simplifier à des mouvements de translation) on mesure une force $\vec{F}'$ et une accélération $\vec{a}'$ données par

$$\vec{F}\,' = q\vec{E} - m\vec{A} = m\vec{a}\,' \tag{2.1}$$

où $-m\vec{A}$ est la force fictive (ou d'inertie). Il se pourrait que l'on mesure une accélération nulle, non pas parce que la force électrique est nulle, mais parce que cette force est exactement compensée par la force fictive. Il est facile de s'en sortir en utilisant un ion doublement chargé O_{16}^{++}, et cette fois, en négligeant une différence de masse entre les deux ions de l'ordre de 10^{-4}, l'accélération $\vec{a}'$ est

$$m\vec{a}' = \vec{F}' = 2q\vec{E} - m\vec{A} \qquad \vec{F}' - \vec{F} = q\vec{E} \tag{2.2}$$

On peut donc sans problème s'assurer de ce que la force électrique est bien nulle, car on peut modifier *indépendamment* les coefficients de $\vec{E}$ et de $\vec{A}$. Cependant cette procédure ne marche pas dans le cas de forces de gravitation. Si dans un référentiel d'inertie $\vec{F} = m\vec{g}$, dans un référentiel accéléré

$$\vec{F}' = m\vec{g} - m\vec{A} \tag{2.3}$$

et on ne peut pas, comme dans l'exemple précédent, modifier *indépendamment* les coefficients de $\vec{g}$ et de $-\vec{A}$ ... sauf s'il existe deux types de masse ! Comme l'avait parfaitement compris Newton, il est *a priori* possible que la masse gravitationnelle m_g intervenant dans la loi de la gravitation

$$||\vec{F}_g|| = G\,\frac{m_g m_g'}{r^2} \tag{2.4}$$

où G est la constante de gravitation et r la distance entre les deux masses, et la masse d'inertie m_i intervenant dans la force fictive $-m_i\vec{A}$ soient de caractère différent. Dans ce cas le rapport m_i/m_g

pourrait être différent suivant le matériau. Non seulement Newton avait envisagé cette possibilité, mais il réalisa une expérience avec un pendule pour la tester. En effet l'équation du pendule (à l'approximation des petites oscillations) s'écrit avec deux types de masse

$$m_i\ddot{\theta} + m_g g\theta = 0 \tag{2.5}$$

et si l'on construit deux pendules avec des matériaux A et B, le rapport des périodes sera

$$\frac{T^A}{T^B} = \left(\frac{m_i^A}{m_g^A}\frac{m_g^B}{m_i^B}\right)^{1/2} \tag{2.6}$$

Newton n'observa aucun effet de ce type et conclut qu'à la précision de ses expériences le rapport m_i/m_g était indépendant du matériau, ce qui impliquait qu'avec un choix convenable d'unités on pouvait écrire $m_i = m_g = m$; en d'autres termes, il existe un seul type de masse. Des expériences ont été réalisées ultérieurement avec une précision bien plus grande que celle qui était possible pour Newton, et on peut affirmer aujourd'hui que $m_i = m_g$ avec une précision $\sim 10^{-12}$.

2.2 Principe d'équivalence

Suivant le raisonnement de la section précédente, l'égalité entre masse d'inertie et masse gravitationnelle implique que l'on ne peut pas distinguer, *au moins localement*, entre une force de gravitation et une force fictive, et on est donc incapable de s'assurer de l'absence de forces de gravitation, alors que l'on peut parfaitement s'assurer de l'absence d'autres types de force, électromagnétiques ou autres : *la gravitation joue donc un rôle particulier*. Au lieu de s'escrimer à définir un référentiel d'inertie où l'on pourrait écrire la loi de Newton pour une force de gravitation, il est plus économique de décider *qu'un référentiel d'inertie est un référentiel en chute libre*. Un ascenseur en chute libre ou un satellite en orbite seront donc des référentiels d'inertie. Avec ce choix, ce n'est pas la pomme qui tombe sur Newton dans un référentiel d'inertie lié à la Terre, c'est Newton qui monte vers la pomme dans le référentiel en chute libre où la pomme est au repos !

Poursuivons notre raisonnement en discutant une expérience simple : un observateur dans une boîte fermée à la surface de la Terre lâche sans vitesse initiale une pomme qui tombe avec une accélération $\vec{g}$. L'observateur verra exactement le même phénomène s'il se trouve dans une fusée dans l'espace, accélérée en direction de sa tête avec une accélération $\vec{A} = -\vec{g}$; il voit la pomme tomber vers ses pieds avec exactement le même mouvement. Sans regarder à l'extérieur, l'observateur est incapable de distinguer entre l'effet d'une force de gravitation et celui d'une accélération en direction opposée[1]. Nous pouvons donc énoncer *le principe d'équivalence*

Principe d'équivalence : aucune expérience à l'intérieur d'une boîte ne permet de distinguer entre une accélération constante et un champ de gravitation constant.

On en déduit une conséquence immédiate pour les photons : dans un référentiel d'inertie, c'est-à-dire, rappelons-le, un référentiel en chute libre, il ne peut pas exister par définition de direction privilégiée, et un photon doit se propager en ligne droite (dans le cas contraire la direction de la déviation indiquerait une direction privilégiée). Dans la figure 2.1, l'ascenseur est lâché quand le photon est émis horizontalement en A, et la trajectoire est une droite dans le référentiel de l'ascenseur. Comme le point d'arrivée B du photon sur la paroi de l'ascenseur ne dépend pas du référentiel, cela veut dire que dans le référentiel de la Terre sa trajectoire est courbée : la lumière est donc sensible à un champ de gravitation. Comme application, il est tentant d'en déduire la déviation d'un rayon lumineux par une masse M, par exemple la déviation de la trajectoire par le Soleil d'un rayon lumineux émis par une étoile et que l'on peut mesurer au cours d'une éclipse totale de Soleil. Effectuons le calcul pour une masse m en supposant la déviation petite, c'est-à-dire à l'approximation du paramètre d'impact (figure 2.2). À cette approximation, on considère que la vitesse v de la masse m varie peu, et l'on peut estimer la variation de l'impulsion transverse Δp_y en intégrant la composante F_y de la force

[1] Ceci rappelle bien évidemment l'énoncé familier concernant les référentiels d'inertie selon lequel un observateur ne peut pas détecter un mouvement uniforme par rapport aux étoiles sans regarder à l'extérieur.

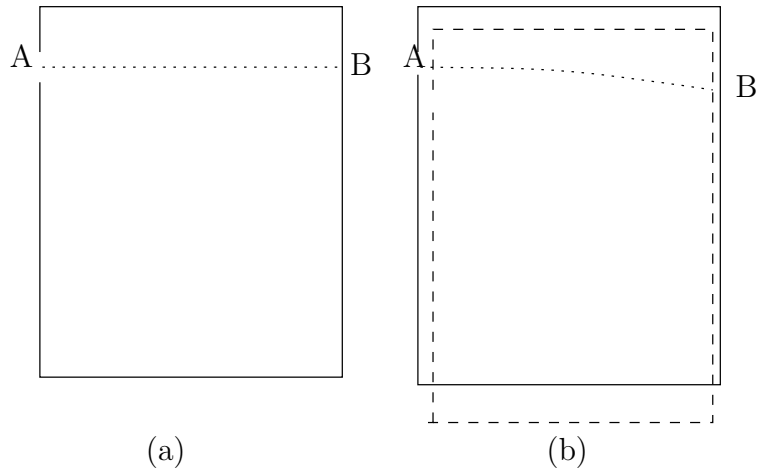

FIG. 2.1 – Trajectoire d'un photon dans un ascenseur. (a) Référentiel en chute libre (b) Référentiel de la Terre.

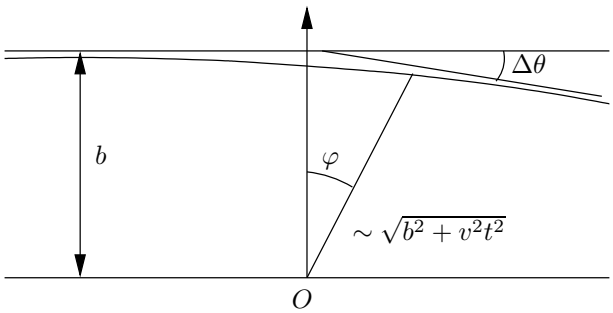

FIG. 2.2 – Déviation d'un rayon lumineux par une masse M placée en O.

$$\Delta p_y = \int_{-\infty}^{+\infty} F_y \, \mathrm{d}t \qquad F_y = -\frac{GmM}{b^2 + v^2t^2} \cos\varphi = -\frac{GmMb}{(b^2 + v^2t^2)^{3/2}}$$

où b est le paramètre d'impact. On obtient donc

$$\Delta p_y = -GmMb \int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{(b^2 + v^2t^2)^{3/2}} = -\frac{2GmM}{bv} \tag{2.7}$$

soit pour l'angle de déviation[2]

$$\Delta\theta = \frac{\Delta p_y}{mv} = \frac{2GM}{bv^2} \tag{2.8}$$

Extrapolant hardiment au cas d'un photon pour lequel $v = c$, on en déduit $\Delta\theta = 2GM/(bc^2)$, ce qui correspond pour une incidence rasante sur le Soleil ($b = R_S =$ rayon du Soleil) à une déviation de 0.87" (seconde d'arc), résultat obtenu initialement par Einstein en 1907. Le résultat de la relativité générale (1915) est le double : 1.75". En fait, dans ce calcul, nous sommes allés au-delà du principe d'équivalence, valable uniquement pour des champs de gravitation constants, et il n'est pas surprenant que notre résultat soit quantitativement incorrect, même si le phénomène est prédit correctement de façon qualitative.

[2] Il est évidemment facile de donner une formule exacte étant donné que les orbites sont des hyperboles

$$\cot\frac{\Delta\theta}{2} = \frac{bv^2}{GM}$$

résultat qui coïncide avec (2.8) pour $\Delta\theta \ll 1$.

2.3 Décalage vers le rouge gravitationnel

Pour simuler un champ de gravitation constant $\vec{g}$, nous allons utiliser le principe d'équivalence et nous placer dans une fusée dont l'accélération vers le haut est constante et égale à $-\vec{g}$. Le calcul qui va suivre est essentiellement un calcul d'effet Doppler. Du plancher de la fusée sont émis des photons avec une période T, aux temps $t = 0, t = T, \ldots, t = nT, \ldots$. Les signaux sont reçus au plafond de la fusée aux temps $t_0, t_1, \ldots, t_n \ldots$ (figure 2.3). La distance plancher-plafond est h et le temps mis par le photon pour aller du plancher au plafond est $\simeq h/c$; pendant ce temps la fusée acquiert une vitesse $v \simeq gh/c$. Le calcul qui va suivre néglige les effets de la relativité restreinte, et il ne sera valable que si $v^2/c^2 \simeq (gh/c^2)^2 \ll 1$: le petit paramètre du problème est en fait gh/c^2.

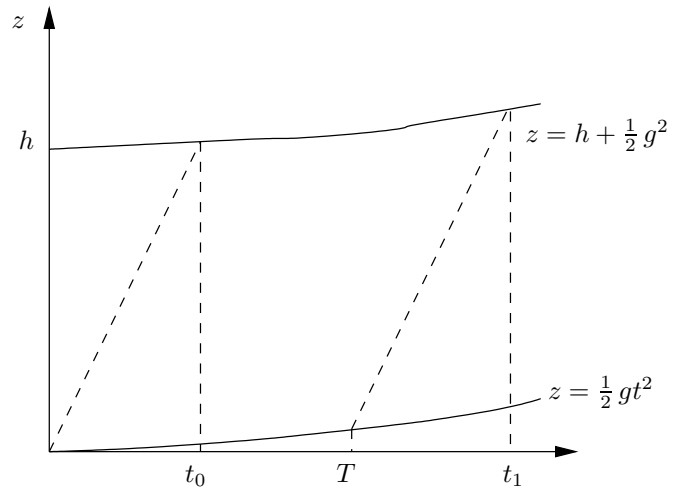

FIG. 2.3 – Décalage vers le rouge gravitationnel. Le plancher de la fusée suit la trajectoire $z = gt^2/2$, le plafond $z = h + gt^2/2$.

La trajectoire du plancher est $z = gt^2/2$ et celle du plafond $z = gt^2/2 + h$. Soit t_n le temps de réception du signal émis à $t = nT$. Ce temps est obtenu en résolvant l'équation du second degré

$$h + \frac{1}{2}\,gt_n^2 = \frac{1}{2}\,g(nT)^2 + c(t_n - nT) \tag{2.9}$$

On peut bien sûr résoudre exactement, mais lorsque $gh/c \ll 1$ il est plus simple de remarquer que

$$t_n = nT + \frac{h}{c}\,(1 + \varepsilon_n) \qquad |\varepsilon_n| \ll 1$$

Dans un calcul au premier ordre en ε_n on peut utiliser

$$t_n - (nT)^2 = (t_n - nT)(t_n + nT) \simeq \frac{2h}{c}\left(nT + \frac{h}{c}\right)$$

et en reportant dans (2.9) on obtient

$$\varepsilon_n = \frac{gnT}{c} + \frac{gh}{c^2} \qquad \Delta T = t_n - t_{n-1} - T = \frac{h}{c}\,(\varepsilon_n - \varepsilon_{n-1}) = \frac{gh}{c^2}\,T$$

Par rapport à l'émission, l'intervalle de temps entre la réception de deux photons est augmenté de ΔT, avec

$$\boxed{\frac{\Delta T}{T} = \frac{gh}{c^2}} \tag{2.10}$$

La période mesurée par l'observateur à l'altitude h est donc plus grande que celle mesurée par l'expéditeur à l'altitude zéro ; c'est le *décalage vers le rouge gravitationnel*. Il en résulte que la fréquence ω du photon reçu est décalée de $\Delta\omega$ par rapport à celle du photon émis

$$\boxed{\frac{\Delta\omega}{\omega} = -\frac{gh}{c^2}} \tag{2.11}$$

On peut donner une déduction plus simple de ce résultat en admettant la formule de l'effet Doppler, que nous démontrerons au prochain chapitre. L'observateur qui reçoit le photon a acquis une vitesse gh/c par rapport à la source pendant le temps de vol du photon, et par effet Doppler la période est allongée de

$$\frac{\Delta T}{T} = \frac{v}{c} = \frac{gh}{c^2}$$

Une dernière façon hardie de voir les choses consiste à écrire qu'entre la source et la réception le photon a perdu une énergie gravitationnelle mgh, avec comme précédemment $m = E/c^2$ et donc, en utilisant la relation de Planck-Einstein $E = \hbar\omega$

$$\frac{\Delta E}{E} = -\frac{1}{E}\left(\frac{E}{c^2}\,gh\right) = \frac{\Delta\omega}{\omega} = -\frac{gh}{c^2}$$

Cet effet fut vérifié expérimentalement par Pound et Rebka en 1960 sur une hauteur de de 20 m.

Les corrections de décalage vers le rouge gravitationnel trouvent une application dans le GPS. Dans l'espace à trois dimensions, il faut quatre satellites pour déterminer une position GPS. Nous allons nous limiter à une caricature de GPS en prenant une seule dimension d'espace, et il suffira donc de deux satellites. Dans le plan (x, ct), les satellites sont supposés suivre des trajectoires (lignes d'Univers) S_A et S_B (figure 2.4) et l'utilisateur une trajectoire U. U reçoit les signaux des deux satellites au temps t et au point x, et il a accès aux positions x_A et x_B des satellites au moment de l'émision du signal, et aux temps d'émission t_A et t_B des signaux reçus car le signal GPS est codé[3]. Les deux droites de pente ± 1 issues des points (x_A, ct_A) et (x_B, ct_B) se coupent en (x, ct)

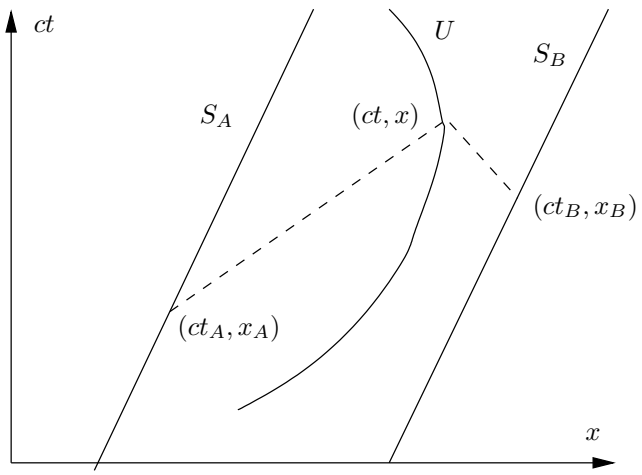

FIG. 2.4 – Schéma du GPS. Les lignes d'Univers S_A et S_B des deux satellites sont des droites, celle de l'observateur U est courbe.

$$\begin{aligned} c(t - t_A) &= x - x_A \\ c(t - t_B) &= -x + x_B \end{aligned}$$

[3]Si l'utilisateur connaissait exactement le temps, il lui suffirait d'un seul satellite (trois satellites à trois dimensions) pour repérer sa position. Le problème est qu'il n'est pas simple de transporter une horloge atomique, par exemple si l'on est en randonnée !

Le système a pour solution immédiate

$$\begin{aligned} ct &= \frac{1}{2}\left[c(t_A + t_B) + (x_B - x_A)\right] \\ x &= \frac{1}{2}\left[c(t_B - t_A) + (x_B + x_A)\right] \end{aligned} \tag{2.12}$$

Si l'on veut une précision d'un mètre[4], il faut une précision sur le temps de 3 ns, et ceci montre que les effets relativistes ne sont pas négligeables. En effet, pour un satellite de période 12 heures, et qui se trouve donc sur une orbite circulaire de rayon $R_s = 4.2\,R_T$, de vitesse $v/c \simeq 1.3 \times 10^{-5}$, l'effet de la relativité restreinte $\sim v^2/c^2$ est $\sim 10^{-10}$. En ce qui concerne le décalage vers le rouge gravitationnel, si l'on compare une horloge sur Terre et une horloge dans le satellite à une horloge hypothétique située dans un référentiel d'inertie, l'effet le plus important est en $GM_T/(R_T c^2) \simeq 1.6 \times 10^{-10}$. En une minute la dérive est $\sim 10^{-9}$ s, et il faut donc remettre les horloges à l'heure toutes les minutes pour une précision du mètre.

2.4 Interprétation géométrique

Revenons à l'expérience démontrant le décalage vers le rouge gravitationnel, en essayant de lui trouver une interprétation qui sera justifiée et affinée par la suite. Nous allons partir de l'idée que les deux observateurs ont en commun un *paramètre de temps* t, mais que leurs deux horloges ne mesurent pas t. En fait le temps mesuré par les deux horloges est un temps τ, appelé *temps propre*, et relié à t par

$$\tau = t\left(1 + \frac{\Phi}{c^2}\right) \tag{2.13}$$

où Φ est le potentiel gravitationnel[5] au point où se trouve l'horloge : une masse test m possède une énergie gravitationnelle $U = m\Phi$, et on suppose $|\Phi/c^2| \ll 1$. Dans ce cas la différence entre les temps τ_A (plancher) et τ_B (plafond) mesurés par les deux horloges est

$$\tau_B - \tau_A = \Delta T = T\left(1 + \frac{\Phi_B}{c^2}\right) - T\left(1 + \frac{\Phi_A}{c^2}\right) = T\,\frac{\Phi_B - \Phi_A}{c^2} = T\,\frac{gh}{c^2}$$

en accord avec (2.10). Une première idée pour interpréter (2.13) consisterait à proposer que le fonctionnement des horloges est affecté par la gravitation. Ce n'est pas l'idée retenue par Einstein, qui fait porter la responsabilité de (2.13) *sur la géométrie*. Pour Einstein, la distance pseudo-euclidienne (ou de Minkowski) au carré $\mathrm{d}s^2$ (voir le chapitre 3 pour plus d'explications) entre deux événements séparés par un intervalle de temps $\mathrm{d}t$ et d'espace $\mathrm{d}\vec{r}$ est affectée par la gravitation, et lorsque $\Phi/c^2 \ll 1$ on peut écrire

$$\mathrm{d}s^2 = \left(1 + \frac{2\Phi(\vec{r})}{c^2}\right) c^2 \mathrm{d}t^2 - \left(1 - \frac{2\Phi(\vec{r})}{c^2}\right) \mathrm{d}\vec{r}^{\,2} \tag{2.14}$$

Nous verrons au chapitre 6 comment cette équation peut se déduire de la métrique de Schwarzschild. Pour mieux comprendre le raisonnement d'Einstein, on peut utiliser une analogie due à Hartle : dans une projection de Mercator, les distances entre Paris et Montréal et entre Lagos et Bogota sont les mêmes. Or il est connu qu'un avion met plus de temps à voler de Lagos à Bogota que de Paris à Montréal. On peut donner deux interprétations.

1. Les règles raccourcissent quand la latidude croît (la gravité affecte les horloges).
2. Les règles restent les mêmes, mais la géométrie est celle d'une sphère (la géométrie est affectée par la gravité).

Lorsque des événements se passent au même point ($\mathrm{d}\vec{r} = 0$), on dit que l'intervalle séparant les deux événements est un *intervalle de temps propre* : $\mathrm{d}s^2 = \mathrm{d}\tau^2$. Suivant (2.14), la relation entre le paramètre de temps t et le temps propre τ est donc

$$\Delta\tau = \Delta t\left(1 + \frac{2\Phi(\vec{r})}{c^2}\right)^{1/2} \simeq \Delta t\left(1 + \frac{\Phi(\vec{r})}{c^2}\right) \tag{2.15}$$

[4] Cette précision semblera modeste d'ici quelques années : on nous promet pour bientôt une précision du centimètre !

[5] Il est important de noter qu'en relativité la valeur absolue de l'énergie potentielle a une signification physique : il n'y a pas de constante arbitraire dans la définition de l'énergie.

ce qui permet de retrouver (2.13). En l'absence de gravité, si $(c\Delta t, \Delta\vec{r})$ est l'intervalle d'espace-temps entre l'émission et la réception d'un photon

$$\Delta s^2 = c^2\Delta t^2 - \Delta\vec{r}^{\,2} = 0$$

car un photon se propage à la vitesse c, et dans un diagramme $(ct, \vec{r})$ sa ligne d'Univers, c'est-à-dire son temps en fonction de sa position, est une droite de pente unité. En présence de gravité, nous aurons toujours par définition $\mathrm{d}s^2 = 0$, et si nous nous limitons à une dimension d'espace, z, nous avons d'après (2.14) et compte tenu de $|\Phi/c^2| \ll 1$

$$\frac{\mathrm{d}z}{\mathrm{d}t} \simeq c\left(1 + \frac{2\Phi}{c^2}\right)$$

Cette équation semble indiquer que la vitesse du photon n'est plus c, mais c'est une illusion car t est seulement un *paramètre de temps* et z un paramètre d'espace. Nous pouvons par exemple définir z' par

$$z' = \int^z \left(1 - \frac{2\Phi(v)}{c^2}\right) \mathrm{d}v \qquad \frac{\mathrm{d}z'}{\mathrm{d}z} = 1 - \frac{2\Phi(z)}{c^2}$$

et donc

$$\frac{\mathrm{d}z'}{\mathrm{d}t} = \frac{\mathrm{d}z'}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}t} \simeq c$$

et avec ce nouveau paramètre d'espace les lignes d'Univers des photons sont des droites !

La forme (2.14) de la métrique nous a permis de rendre compte du décalage vers le rouge gravitationnel dans un cas où $\mathrm{d}\vec{r} = 0$. Calculons maintenant l'intégrale s_{AB} de la métrique sur la trajectoire d'une particule massive de vitesse $\ll c$: comme nous le verrons au chapitre suivant, cette intégale est la "distance" minkowskienne ou le temps propre entre le point de départ A et le point d'arrivée B. Nous avons

$$\begin{aligned} s_{AB} &= \int_A^B \mathrm{d}t \left[\left(1 + \frac{2\Phi(\vec{r})}{c^2}\right)c^2 - \left(1 - \frac{2\Phi(\vec{r})}{c^2}\right)\left(\frac{\mathrm{d}\vec{r}}{\mathrm{d}t}\right)^2\right]^{1/2} \\ &\simeq c\int_A^B \mathrm{d}t \left[\left(1 + \frac{2\Phi(\vec{r})}{c^2}\right) - \frac{1}{c^2}\left(\frac{\mathrm{d}\vec{r}}{\mathrm{d}t}\right)^2\right]^{1/2} \\ &\simeq c\int_A^B \mathrm{d}t \left[1 - \frac{1}{c^2}\left(\frac{1}{2}\left(\frac{\mathrm{d}\vec{r}}{\mathrm{d}t}\right)^2 - \Phi(\vec{r})\right)\right] \end{aligned} \tag{2.16}$$

On reconnaît en facteur de $-1/c^2$ le lagrangien de la particule, et suivant le principe de moindre action la minimisation de s_{AB} par rapport à $\vec{r}(t)$ donne tout simplement les équations du mouvement

$$\frac{\delta s_{AB}}{\delta\vec{r}(t)} = 0 \Longleftrightarrow \frac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2} = -\vec{\nabla}\Phi(\vec{r})$$

Autrement dit la particule suit une trajectoire qui extrémise le temps propre s_{AB}. On voit de (2.16) qu'en général cet extremum sera un maximum, car l'intégrale du deuxième terme dans le crochet n'est autre que l'action, qui est en général minimisée par la trajectoire. Nous verrons ultérieurement que cet énoncé est équivalent à l'affirmation selon laquelle la trajectoire d'une particule soumise uniquement à la gravitation est une géodésique de l'espace-temps. Nous voyons donc que l'interprétation géométrique nous donne *à la fois* le décalage vers le rouge gravitationnel et les lois de Newton.

2.5 Effets de marée gravitationnels

Le défaut principal (mais il est inévitable !) de notre définition d'un référentiel d'inertie est que cette définition est forcément locale si le champ de gravitation n'est pas uniforme. En effet, si le champ de gravitation n'est pas uniforme, on va observer des effets de marée. Prenons l'exemple d'un satellite en

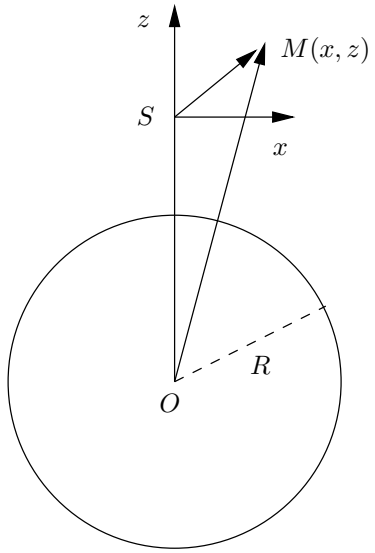

FIG. 2.5 – Effets de marée gravitationnels. Le centre du satellite S est à une distance d du centre O de la Terre. La masse m a pour coordonnées (x, z) dans un référentiel d'origine S.

chute libre au-dessus de la Terre, et comparons la chute de son centre masse situé en S à une distance d du centre de la Terre à celle d'objets dans le satellite (figure 2.5). Si les coordonnées d'un objet dans le satellite sont (x, z), en choisissant une géométrie à deux dimensions pour simplifier, l'énergie potentielle gravitationnelle d'un objet de masse m est

$$\begin{aligned} \Phi(x,z) &= \frac{GmM}{[x^2+(d+z)^2]^{1/2}} \\ &\simeq -\frac{GmM}{d}\left[1-\frac{z}{d}-\frac{x^2}{2d^2}+\frac{z^2}{d^2}\right] \end{aligned} \tag{2.17}$$

où M est la masse de la Terre et nous avons utilisé $|x|, |z| \ll d$ pour effectuer un développement limité. Si la masse était située en S, la force sur cette masse aurait pour composantes $F_x^0 = 0$, $F_z^0 = -GmM/d^2$. Pour une masse en (x, z) , nous avons

$$F_x - F_x^0 = F_x = -GmM\,\frac{x}{d^3} \tag{2.18}$$

$$F_z - F_z^0 = GmM\,\frac{2z}{d^3} \tag{2.19}$$

Ces équations montrent qu'au cours de la chute du satellite x diminue et z augmente : il y contraction suivant la direction horizontale et dilatation suivant la direction verticale. Ceci se comprend qualitativement : si une personne est en chute libre verticale, l'accélération de ses pieds est plus grande que celle de son centre de masse, et celle de sa tête plus faible. Elle a donc tendance à être dilatée verticalement dans le référentiel de sa chute. Si elle étend les bras, l'accélération de ses mains est dirigée vers le centre de forces et elle a tendance à être contractée horizontalement. C'est un *effet de marée* typique, sur lequel nous faisons les deux remarques suivantes.

- Un exemple extrême d'effet de marée est donné par la chute d'un astronaute sur une étoile en effondrement gravitationnel (chapitre 6) : l'astronaute est dilaté de la tête aux pieds et contracté dans la direction horizontale, et finalement mis en pièces !
- Si $\vec{r}$ est la position de l'origine d'un référentiel d'inertie, et si $\vec{\xi}$ est la position d'un objet dans ce référentiel, les équations du mouvement pour $\vec{\xi}$ sont données par une généralisation facile de ce qui précède

$$\frac{\mathrm{d}^2\xi^i}{\mathrm{d}t^2} = -\delta^{ij}\,\frac{\partial^2\Phi}{\partial x^i\partial x^k}\Big|_{\vec{r}}\,\xi^k \tag{2.20}$$

En résumé deux idées importantes ressortent de cette première approche.

1. Les forces de gravitation apparaissent comme des effets de marée et on abandonne l'idée de donner une valeur absolue à la force de gravitation.
2. Si elles sont soumises uniquement à des forces de gravitation, les particules suivent des géodésiques de l'espace-temps.

Cependant cette première approche heuristique de la relativité générale souffre pour le moment d'un défaut essentiel : nous n'avons pas considéré le cas de particules massives dont la vitesse peut être proche de celle de la lumière. Avant d'examiner ce qui se passe en relativité générale, il nous faut revenir au chapitre suivant à la relativité restreinte.

Bibliographie.

Hartle [2003], chapitres 1 à 3 et 6 ; N. Ashby, *Relativity and the Global Positioning System*, Physics Today mai 2002, p. 41 ; F. Wilczek, *Total Relativity : Mach 2004*, Physics Today avril 2004 (sur la relativité et le principe de Mach).

Chapitre 3

Espace-temps plat

Ce chapitre examine l'espace-temps plat de la relativité restreinte, non encore "déformé" par la gravité. Il s'agit donc de l'espace-temps idéal (et théorique !) d'un Univers entièrement vide de matière et d'énergie, où le temps est homogène et l'espace homogène et isotrope : c'est une arène neutre où se déroulent les processus physiques. On ne pourra donner ici qu'un survol de la relativité restreinte, et la plupart des résultats seront rendus plausibles, et non démontrés en détail. Le lecteur est renvoyé à la bibliographie pour des exposés plus complets.

N.B. Dans ce chapitre et les suivants nous utiliserons un système d'unités où la vitesse de la lumière $c = 1$ et nous utiliserons également la convention de sommation sur les indices répétés

$$\sum_i x^i y^i = x^i y^i$$

3.1 Photons

Rappelons qu'un *événement* est quelque chose qui se passe à un temps déterminé t en un point déterminé $\vec{r}$: un événement est caractérisé par une coordonnée de temps t, souvent notée $x^0 = t$, et trois coordonnées d'espace $\vec{r} = (x, y, z) = (x^1, x^2, x^3)$. En relativité restreinte, on a l'habitude de définir un référentiel d'inertie en quadrillant l'espace-temps par un réseau de règles rigides et d'horloges synchronisées, mais l'utilisation de règles rigides est à proscrire en relativité générale, du moins sur de grandes distances, en raison de la courbure de l'espace. En conséquence nous allons essayer de tout mesurer en utilisant uniquement des photons.

Un observateur sera représenté par un physicien qui transporte avec lui une horloge, mesurant son temps propre, et un émetteur/détecteur de photons capable de mesurer les fréquences des photons reçus et émis. Entre d'autres termes, un observateur = une horloge et un radar.

Nous allons maintenant énoncer l'hypothèse de base de la relativité restreinte en l'illustrant au moyen de l'explosion d'une étoile : un événement O (explosion de l'étoile) se produit en émettant une bouffée de photons et de particules massives (débris). Nous allons faire l'hypothèse fondamentale que *tous les photons, indépendamment de leur couleur, arrivent au même instant à un observateur*, mais pas en général avec la même couleur que celle de l'émission. Si tel n'était pas le cas, l'explosion semblerait durer dans le temps et un événement ponctuel ne serait pas vu comme ponctuel par un observateur. Les particules massives, au contraire, n'arrivent pas au même instant et sont détectées *après* les photons. Les courbes d'espace-temps formées des positions successives de l'observateur, des photons ou des particules et les instants correspondants, sont appelées *lignes d'Univers* de l'observateur, des photons ou des particules. Les propriétés précédentes doivent évidemment être valables quel que soit l'observateur. Pour en rendre compte, on va tracer dans l'espace-temps à partir de O un cône de sommet O, $N(O)$, qui est une *structure géométrique absolue*, indépendante de l'observateur, et qui est appelée *cône de lumière* de O. Les lignes

d'Univers des photons passant par le point O à $t = 0$ suivent les génératrices de ce cône. La ligne d'Univers de l'observateur coupe $N(O)$ en un point unique O', qui définit la position de l'observateur et le temps auquel il reçoit les photons émis par l'explosion (figure 3.1). Les lignes d'Univers des particules massives émises au moment de l'explosion sont entièrement contenues à l'intérieur du cône futur $N^+(O)$ de O. Comme $N(O)$ est une structure géométrique absolue, il ne définit aucune direction privilégiée, et tout observateur passant par O voit $N(O)$ avec une parfaite symétrie sphérique.

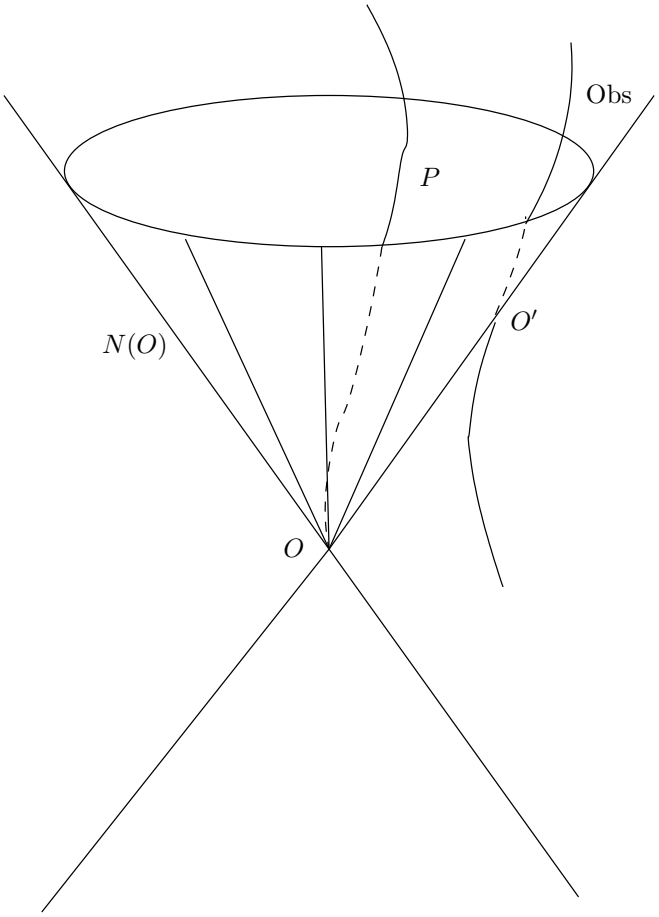

FIG. 3.1 – Cône de lumière $N(O)$ du point O. La ligne d'Univers d'une particule massive passe par P, celle de l'observateur coupe $N(O)$ en O'. Les lignes d'Univers des photons sont les génératrices du cône.

Considérons deux observateurs, Alice (A) et Bob (B)[1] qui échangent des photons. Si les fréquences des photons reçus sont identiques à celles des photons émis, on dira qu'Alice et Bob suivent des *lignes d'Univers parallèles*, et les deux observateurs peuvent alors aisément synchroniser leurs horloges. L'échange de photons permet à Alice de définir les coordonnées de l'événement P "réflection du photon par le miroir de Bob"[2] (figure 3.2). Alice mesure les coordonnées de l'événement P comme étant $(t, x, c = 1)$

$$t = \frac{1}{2}(t_2 + t_1) \qquad x = \frac{1}{2}(t_2 - t_1) \tag{3.1}$$

Mais ce résultat est indépendant du fait que le photon rebondit sur le miroir de Bob ou sur celui d'un troisième observateur, Chiara (C), qui ne suit pas nécessairement une ligne d'Univers parallèle à celles

[1] Ces deux héros de de l'informatique quantique sont apparus pour la première fois dans la discussion des systèmes de cryptographie à clé secrète.

[2] Pour une version littéraire de cet échange de photons, on pourra se reporter à la nouvelle de D. Buzzatti, "Les sept messagers", où les photons sont remplacés par des cavaliers.

d'Alice et de Bob (figure 3.2). Autrement dit les coordonnées (t, x) attribuées par Alice à l'événement P sont indépendantes du partenaire qui réfléchit le photon.

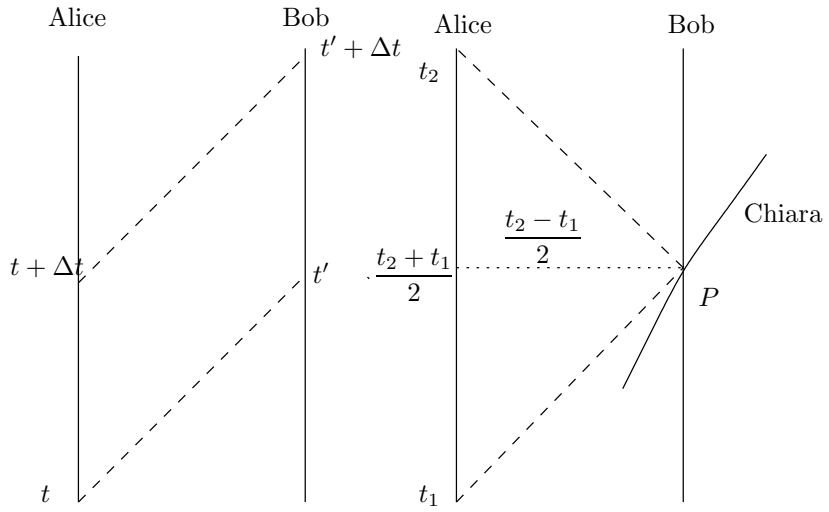

FIG. 3.2 – Mesure des coordonnées d'un événement P. Alice et Bob suivent des lignes d'Univers parallèles. Les lignes d'Univers des photons sont en tirets.

3.2 Effet Doppler

Nous allons maintenant établir la formule de l'effet Doppler en prenant l'exemple de la mesure de la vitesse d'une voiture par un radar. La voiture croise le radar au temps $t = 0$, suivant les horloges du radar et de la voiture ; ceci est toujours possible par un choix convenable de l'origine des temps des deux horloges. Le radar émet un photon[3] au temps t, *suivant l'horloge du radar*. Si la voiture était immobile, elle recevrait ce photon au temps t, mais comme elle s'éloigne elle le recevra à un temps Kt *suivant l'horloge de la voiture*, avec $K > 1$ (figure 3.3)[4]. Mais la situation radar-voiture est symétrique, car les directions x et $-x$ sont équivalentes : il n'y a pas de direction privilégiée dans un référentiel d'inertie. Si le photon est réfléchi au temps Kt suivant l'horloge de la voiture, il atteindra le radar au temps K^2t *suivant l'horloge du radar*. Pour le radar, l'événement P : le photon est réfléchi par la voiture, a pour coordonnées, avec $t_1 = t$ et $t_2 = K^2t$ dans (3.1)

$$\begin{array}{lrcl} \text{temps} & \frac{1}{2}(t_2 + t_1) & = & \frac{1}{2}t(K^2 + 1) \\ \text{espace} & \frac{1}{2}(t_2 - t_1) & = & \frac{1}{2}t(K^2 - 1) \end{array}$$

Effectuant le rapport espace/temps, la vitesse de la voiture mesurée par le radar est donc

$$v = \frac{K^2 - 1}{K^2 + 1} \tag{3.2}$$

ou encore

$$K = \sqrt{\frac{1 + v}{1 - v}} \tag{3.3}$$

[3]En pratique un signal électromagnétique codé.
[4]En raison de l'homogénéité temporelle, K ne peut pas dépendre du temps.

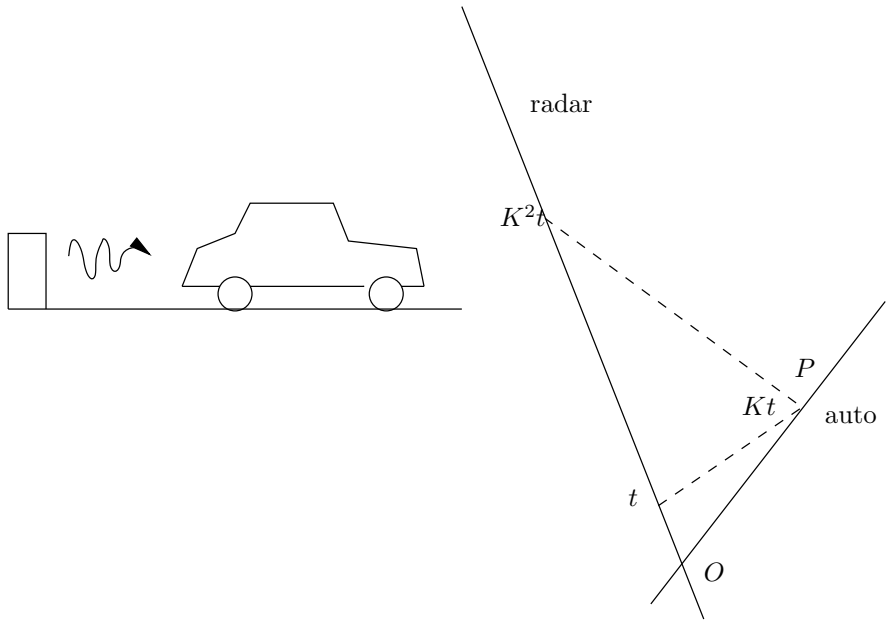

FIG. 3.3 – Mesure de la vitesse d'une voiture avec un radar. En traits pleins les lignes d'Univers du radar et de l'auto, en tirets celles des photons. Le signal est reçu par l'auto en P.

Une autre façon d'exprimer ce résultat est de dire que la fréquence du photon mesurée par la voiture, ω_{rec} est reliée à la fréquence ω_{em} d'émission par le radar par

$$\omega_{\text{rec}} = \frac{1}{K}\,\omega_{\text{em}} = \omega_{\text{em}}\sqrt{\frac{1-v}{1+v}} \tag{3.4}$$

Cette formule n'est évidemment pas autre chose que celle de l'effet Doppler relativiste.

3.3 Métrique de l'espace-temps plat

Pour formaliser ces considérations géométriques, on introduit une métrique sur l'espace-temps ; si

$$x^\mu = (x^0, \vec{x}) \quad \text{et} \quad y^\mu = (y^0, \vec{y})$$

sont les coordonnées spatio-temporelles de deux événements (l'exposant 0 désigne la coordonnée de temps), on pourra aussi considérer x^μ et y^μ comme des *quadrivecteurs*[5], les quadrivecteurs joignant l'origine à ces points. On définira le *produit scalaire de Minkowski* de ces deux quadrivecteurs par

$$\boxed{x \cdot y = x^0 y^0 - \vec{x}\cdot\vec{y} = x^0 y^0 - x^i y^i} \tag{3.5}$$

Par définition des quadrivecteurs sont des objets à quatre composantes tels que *leur produit scalaire (3.5) soit indépendant du référentiel d'inertie*. Cette condition permet de déterminer la loi de transformation des quadrivecteurs quand on passe d'un référentiel à un autre, ou transformation de Lorentz, mais nous n'aurons pas besoin de sa forme explicite. Il suffira de savoir qu'elle laisse par construction le produit scalaire invariant. De façon équivalente, nous introduisons un *tenseur métrique*[6] $\eta_{\mu\nu}$, le tenseur

[5] Ce qui ne sera pas possible en Relativité Générale, où les coordonnées ne sont pas des vecteurs.

[6] Il existe deux conventions pour le tenseur $\eta_{\mu\nu}$, la convention (3.6) et la convention

$$\eta_{\mu\nu} = \text{diagonal } (-1, 1, 1, 1)$$

La convention utilisée dans cet exposé minimise le nombre de signe moins, car on utilise beaucoup plus d'intervalles du genre temps $\mathrm{d}s^2 > 0$ (avec la convention (3.6) !) que d'intervalle du genre espace. La convention opposée à (3.6) est avantageuse en théorie quantique des champs, car on tombe directement sur la métrique euclidienne lorsque l'on effectue une rotation de Wick.

de Minkowski

$$\boxed{\eta_{\mu\nu} = \text{diagonal } (1, -1, -1, -1)} \tag{3.6}$$

et nous récrivons (3.5)

$$x \cdot y = \eta_{\mu\nu} x^\mu y^\nu = x_\mu y^\nu \quad \text{avec} \quad x_\mu = \eta_{\mu\nu}\, x^\nu$$

Les composantes x^μ sont appelées conventionnellement composantes *contravariantes* du quadrivecteur et les x_μ sont les composantes *covariantes*. Le produit scalaire d'un quadrivecteur par lui-même, $x^2 = x_\mu x^\mu$ est la "longueur au carré", ou norme carrée (de Minkowski) de ce vecteur[7]. On écrit en général cette longueur au carré de Minkowski pour un vecteur infinitésimal $\mathrm{d}x^\mu$

$$\boxed{\mathrm{d}s^2 = (\mathrm{d}x^0)^2 - \mathrm{d}\vec{x}^2 = \mathrm{d}t^2 - \mathrm{d}\vec{x}^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu} \tag{3.7}$$

Si un photon est émis à l'origine[8] O $(t = 0, \vec{x} = 0)$ et est reçu au point x^μ, alors $|x^0| = ||\vec{x}||$ et

$$(x^0)^2 - \vec{x}^{\,2} = x_\mu x^\mu = 0 \tag{3.8}$$

Ceci n'est autre que l'équation du cône de lumière $N(O)$ défini dans la section 3.1. Pour un autre observateur qui verra l'événement réception du photon avec des coordonnées différentes x'^μ, l'équation du cône de lumière sera $x'^0 - \vec{x}'^2 = 0$. L'intervalle d'espace-temps x^μ entre le point d'émission et le point de réception d'un photon vérifie donc $x_\mu x^\mu = x^2 = 0$, et on dira que le quadrivecteur x^μ est du *genre lumière*. Si $x_\mu x^\mu = x^2 > 0$, on dira que x^μ est *du genre temps*, et si $x^2 < 0$, x^μ est du *genre espace*.

Appliquons ces notions au cas d'un mouvement uniforme sur une droite, $x = vt$, $|v| < 1$: par exemple une particule massive est émise en O au temps $t = 0$ et est détectée en P au temps t et à la position x (figure 3.4). La norme carrée du quadrivecteur $x^\mu = (t, x)$ est donnée par

$$x_\mu x^\mu = \tau^2 = t^2 - v^2 t^2 = t^2(1 - v^2) = \frac{t^2}{\gamma^2}$$

avec $\gamma = (1 - v^2)^{-1/2}$, soit $t = \gamma\tau$. Mais τ^2 est un produit scalaire, qui est indépendant du système de coordonnées. Si un observateur se déplace avec une vitesse v en suivant la particule, les coordonnées de P pour cet observateur seront $(t', x' = 0)$, et on aura

$$t'^2 = t^2 - x^2 = \tau^2$$

Le temps mesuré par une horloge liée à la particule est par définition t', et c'est le temps propre de la particule : $t' = \tau$. L'expression $t = \gamma\tau$ montre que *le temps propre est toujours le plus court*.

Afin de montrer la cohérence de ce résultat avec notre raisonnement précédent sur l'effet Doppler, examinons la situation décrite dans la section 3.2 dans le référentiel où Alice est au repos (figure 3.5). D'après (3.3) Bob mesure sur sa montre la réception du photon au temps

$$\Delta\tau' = K\Delta t = \Delta t\sqrt{\frac{1+v}{1-v}}$$

Pour Alice, les coordonnées de P sont données par l'intersection des droites

$$x = t - \Delta t \qquad x = vt$$

soit

$$t = \frac{\Delta t}{1-v} \qquad x = \frac{v\Delta t}{1-v}$$

Nous obtenons donc

$$\Delta\tau^2 = \frac{1}{(1-v)^2}\,(1 - v^2)\Delta t^2 = \frac{1+v}{1-v}\,\Delta\tau'^2$$

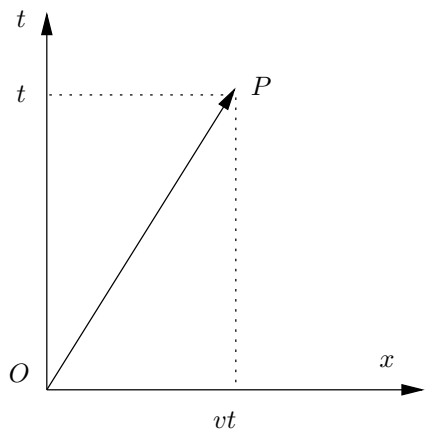

FIG. 3.4 – Mouvement uniforme d'une particule de vitesse v, qui quitte l'origine au temps $t = 0$ et arrive en $P : (t, x = vt)$ au temps t.

Le temps propre $\Delta\tau$ calculé à l'aide de l'invariance du produit scalaire est donc bien le même que celui $\Delta\tau'$ obtenu par le raisonnement donnant l'effet Doppler.

Examinons maintenant la ligne d'Univers $x^\mu(\tau)$ d'une particule massive paramétrée à l'aide de son temps propre. Nous appellerons *quadrivitesse* de la particule le quadrivecteur $u^\mu(\tau)$ tangent à la ligne d'Univers au point P de temps propre τ

$$\boxed{u^\mu(\tau) = \frac{\mathrm{d}x^\mu(\tau)}{\mathrm{d}\tau}} \tag{3.9}$$

u^μ est bien un quadrivecteur car c'est le quotient d'un quadrivecteur x^μ par un scalaire τ. Pendant un temps (propre) infinitésimal $\mathrm{d}\tau$

$$\mathrm{d}x^\mu = u^\mu \mathrm{d}\tau$$

mais comme

$$\mathrm{d}x^2 = \mathrm{d}\tau^2 = (u_\mu u^\mu)\mathrm{d}\tau^2$$

on trouve $u^2 = 1$: *la quadrivitesse est un vecteur unitaire de genre temps*. Il est intéressant de donner l'expression de la quadrivitesse en fonction de la vitesse $\vec{v}$ dans un référentiel d'inertie (figure 3.4)

$$\frac{\mathrm{d}x^0}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma \qquad \frac{\mathrm{d}\vec{r}}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = \gamma\,\vec{v}$$

soit

$$\boxed{u^\mu = \gamma\,(1, \vec{v})} \tag{3.10}$$

La quadrivitesse d'une particule libre dans un référentiel d'inertie obéit à $\mathrm{d}u^\mu/\mathrm{d}\tau = 0$: τ est ce que l'on appelle un *paramètre affine* pour la trajectoire. Ce paramètre affine n'est pas unique, $\sigma = a\tau + b$ est aussi affine. Si l'on effectue un changement de paramétrisation

$$y^\mu(\sigma) = x^\mu(\tau(\sigma)) \qquad w^\mu = \frac{\mathrm{d}y^\mu}{\mathrm{d}\sigma} = u^\mu(\tau(\sigma))\,\frac{\mathrm{d}\tau}{\mathrm{d}\sigma}$$

alors

$$\frac{\mathrm{d}^2 y^\mu}{\mathrm{d}\sigma^2} = \frac{\mathrm{d}w^\mu}{\mathrm{d}\sigma} = \frac{\mathrm{d}^2\tau}{\mathrm{d}\sigma^2}\,u^\mu \tag{3.11}$$

[7] Toutes mes excuses : x peut désigner soit une coordonnée, soit le quadrivecteur x. J'espère que le contexte lèvera toute ambiguïté.

[8] Pour un photon émis en x^μ et reçu à l'origine, on a bien évidemment $x^0 < 0$.

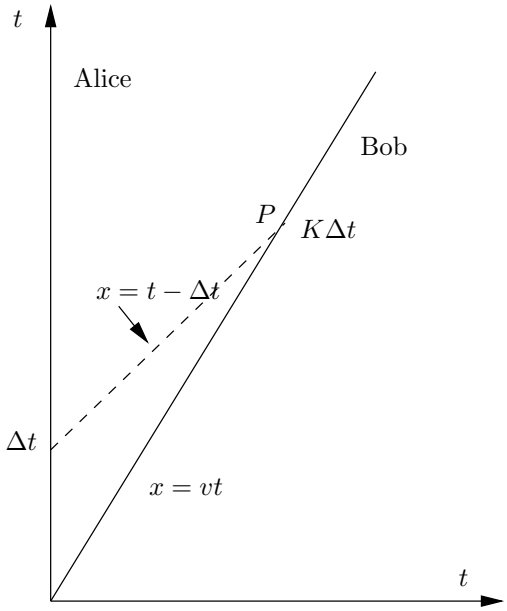

FIG. 3.5 – Autre déduction de l'effet Doppler. On a choisi un référentiel où Alice est au repos.

Si le changement de variable $\tau \to \sigma$ n'est pas linéaire, $\mathrm{d}w^\mu/\mathrm{d}\sigma \neq 0$, et σ n'est pas un paramètre affine. Les choses en vont différemment pour un photon : prenons par exemple un photon se propageant suivant Ox et dont la ligne d'Univers est $x^0 = x$, ce que l'on peut écrire

$$x^\mu = \lambda u^\mu \quad \text{avec} \quad u^\mu = (1, 1, 0, 0)$$

On a $\mathrm{d}x^\mu/\mathrm{d}\lambda = u^\mu$, $u^2 = 0$ et $\mathrm{d}u^\mu/\mathrm{d}\lambda = 0$: λ est un paramètre affine. Mais on pourrait aussi bien prendre $x^\mu = \sigma^3\, u^\mu$, et on vérifie que σ n'est pas un paramètre affine.

Pour terminer cette section, nous observons que le produit scalaire de deux quadrivitesses a une interprétation intéressante. Soit Alice et Bob se croisant en O avec des quadrivitesses u_A et u_B. Dans le référentiel au repos d'Alice, $u_A = (1, \vec{0})$. Soit v_{AB} la valeur absolue de la vitesse relative de A et B, par exemple la vitesse de Bob mesurée par Alice. On a alors

$$u_A \cdot u_B = u_A^0 u_B^0 - \vec{u}_A \cdot \vec{u}_B = u_B^0 = \frac{1}{\sqrt{1 - v_{AB}^2}}$$

et donc

$$u_A \cdot u_B = \frac{1}{\sqrt{1 - v_{AB}^2}} \tag{3.12}$$

Cette équation permet de calculer la vitesse relative v_{AB} quel que soit le référentiel où l'on connaît u_A et u_B.

3.4 Tenseur énergie-impulsion

En multipliant la quadri-vitesse u^μ d'une particule (massive) par sa masse m, on obtient la *quadri-impulsion* p^μ, quadri-vecteur formé de l'énergie (composante de temps) et de l'impulsion (composantes d'espace)

$$p^\mu = m u^\mu \qquad p^0 = E = \gamma m \qquad \vec{p} = \gamma m \vec{v} \tag{3.13}$$

où nous avons utilisé (3.10) ; la masse étant un scalaire, le produit $m u^\mu$ est bien un quadri-vecteur. On notera la relation importante $\vec{v} = \vec{p}/E$. Pour des faibles vitesse, $v \ll c$, on peut effectuer un développement

limité en fonction de v/c

$$E = mc^2 + \frac{p^2}{2m} + \cdots \qquad \vec{p} = \frac{m\vec{v}}{\sqrt{1-(v/c)^2}} = m\vec{v} + \cdots$$

où nous avons rétabli c. Le premier terme de E est l'énergie de masse mc^2, le second l'énergie cinétique non relativiste habituelle.

Nous allons maintenant discuter la notion de courant en relativité. Considérons un ensemble de particules instables de vie moyenne τ dans le référentiel où elles sont au repos, et supposons pour simplifier que ces particules se désintègrent *exactement* au bout d'un temps τ. Si ces particules se déplacent à une vitesse v dans un référentiel d'inertie $\mathcal{R}$, un observateur dans ce référentiel trouvera que les particules se désintègrent au bout d'un temps $\gamma\tau$, en raison de la relation entre le temps propre de la particule et le temps mesuré par l'observateur. Les particules parcourent donc dans $\mathcal{R}$, non pas une distance $v\tau$, mais une distance[9] $\gamma v\tau$. Supposons que soit disposée dans dans $\mathcal{R}$ une densité linéaire de charge ρ^* suivant la direction de la vitesse des particules, une particule verra, pendant un temps τ, $\gamma\rho^* v\tau$ charges. En résumé, une particule voit les charges défiler à la vitesse $-v$, et elle observe une densité de charges qui n'est pas ρ^*, mais $\gamma\rho^*$. Un observateur lié à la particule définira donc une densité de charge $\rho = \gamma\rho^*$ et une densité de courant $\vec{\jmath} = -\gamma\rho^*\vec{v} = -\rho\vec{v}$. Si la densité de charges est ρ^* dans un référentiel où ces charges sont au repos, cet observateur de quadri-vitesse u^μ dans ce référentiel définira un quadri-courant j^μ par

$$\boxed{j^\mu = \rho^* u^\mu} \tag{3.14}$$

Ce quadri-courant obéit à la *loi de conservation*

$$\boxed{\partial_\mu j^\mu = \partial_0 j^0 + \vec{\nabla}\cdot\vec{\jmath} = 0} \tag{3.15}$$

qui est bien connue des équations de Maxwell. Il est très important de noter qu'en relativité *une loi de conservation est obligatoirement une loi de conservation locale*. En effet, supposons que nous partions d'une situation où la charge totale est nulle, mais qu'au temps $t = 0$ une charge négative soit créée au point O, tandis qu'une charge positive est créée en un autre point, ce qui conserve globalement la charge. Cependant, étant donné la relativité de la simultanéité, il est facile de trouver des référentiels où la charge positive, par exemple, est créée *avant* la charge négative, et la charge n'est pas conservée. Autrement dit, la charge contenue dans un volume ne peut varier que parce qu'elle passe par la frontière de ce volume[10], ce qui n'est pas vrai par exemple pour la conservation de l'impulsion en mécanique des fluides galiléenne.

Poursuivons l'étude des lois de conservation. Une surface $t =$cste est une *3-surface de genre espace* de l'espace-temps, le vecteur orthogonal à cette surface étant le vecteur $n^\mu = (1, \vec{0})$. On peut généraliser en prenant une surface orthogonale à un vecteur unitaire de genre temps $n^\mu, n^2 = 1$. Une surface orthogonale au vecteur de genre espace $n^\mu = (0, 1, 0, 0)$ par exemple sera une *3-surface de genre temps*, et en général une telle surface sera orthogonale à un vecteur unitaire de genre espace $n^\mu, n^2 = -1$. On peut tracer un volume ΔV dans une 3-surface de genre temps ou de genre espace. Soit ΔN le nombre de charges dans ΔV. La quantité ΔN est indépendante du référentiel, c'est un scalaire de Lorentz. Le seul scalaire que l'on puisse construire avec le courant j^μ et le vecteur n^μ caractérisant la 3-surface est $j \cdot n = j_\mu n^\mu$. Lorsque $n^\mu = (1, \vec{0})$

$$(j_\mu n^\mu)\Delta V = \rho^*\Delta V = \Delta N$$

Si le vecteur n^μ est du genre espace, par exemple $n^\mu = (0, 1, 0, 0)$, alors

$$\Delta N = (j_\mu n^\mu)\Delta t\Delta y\Delta z = j^x\Delta t\Delta y\Delta z = \frac{\Delta N}{\Delta y\Delta z\Delta t}\,\Delta t\Delta y\Delta z$$

[9] Cette propriété est utilisée dans les accélérateurs de particules. Si l'on produit un faisceau de mésons-π de vie moyenne $\sim 10^{-8}$ s, on pourrrait s'attendre qu'un méson-π dont la vitesse est proche de celle de la lumière puisse parcourir au maximun un distance ~ 3 m. En fait il peut parcourir des distances bien plus grandes, ce qui permet d'éloigner la zone d'expériences de celle de production de plusieurs centaines de mètres, sans perte appréciable de mésons-π.

[10] Cette observation permet de démontrer (3.15) en utilisant le théorème de la divergence.

et $j^x = \Delta N/(\Delta A \Delta t)$ est bien le flux de charges à travers la surface $\Delta A = \Delta y \Delta z$. On a donc dans tous les cas

$$\Delta N = (j_\mu n^\mu)\Delta V \tag{3.16}$$

Si au lieu de ΔN, qui est un scalaire, on a affaire à une quantité qui est un quadrivecteur, par exemple une quadri-impulsion, on doit introduire au lieu d'un courant vectoriel j^μ un courant tensoriel $T^{\mu\nu}$. La relation correspondant à (3.16) est alors

$$\Delta p^\mu = (T^{\mu\nu} n_\nu)\, \Delta V \tag{3.17}$$

où $T^{\mu\nu}$ est le *tenseur énergie-impulsion*. Si $n^\mu = (1, \vec{0})$, cette équation donne pour la composante de temps et les trois composantes d'espace

$$\begin{aligned} \Delta p^0 &= \Delta E = T^{00}\Delta V \\ \Delta p^i &= T^{i0}\Delta V \end{aligned}$$

$T^{00} = \epsilon$ est la densité d'énergie, $T^{i0} = \pi^i$ la densité d'impulsion suivant la direction i. Prenons l'exemple d'une boîte de particules de même vitesse v dans $\mathcal{R}$, de masse m et de densité ρ^* dans le référentiel où elles sont au repos. Nous aurons alors pour les densités d'énergie et d'impulsion

$$\begin{aligned} \epsilon &= T^{00} = (m\gamma)(\rho^*\gamma) = m\rho^* u^0 u^0 \\ \pi^i &= T^{i0} = (m\gamma v^i)(\rho^*\gamma) = m\rho^* u^0 u^i \end{aligned}$$

car la densité dans $\mathcal{R}$ est $\rho = \rho^*\gamma$.

Il nous faut maintenant interpréter T^{0i} et T^{ij}. Suivant le même raisonnement que dans le cas scalaire, considérons une surface de genre temps orthogonale à $n^\mu = (0, 1, 0, 0)$

$$\Delta p^\mu = T^{\mu x}\Delta t \Delta y \Delta z = T^{\mu x}\Delta t \Delta A$$

$T^{\mu x} = \Delta p^\mu/(\Delta A \Delta t)$ est le flux de p^μ à travers ΔA, et en particulier $T^{0x} = \Delta p^0/(\Delta A \Delta t)$ est le flux d'énergie dans la direction x. Mais, en raison de la conservation locale de l'énergie-impulsion

$$(\text{flux d'energie}) \times \Delta A \Delta t = (\text{densite d'energie}) \times v^x \Delta A \Delta t = (\text{densite d'impulsion})^x \Delta A \Delta t$$

où nous avons utilisé $v^x = p^x/E$. En divisant par $\Delta A \Delta t$ nous obtenons $T^{0x} = T^{x0}$. D'autre part

$$T^{ix} = \frac{\Delta p^i}{\Delta A \Delta t} = \frac{\Delta p^i/\Delta t}{\Delta A}$$

qui est une force par unité de surface. Plus généralement

$$\Delta F^i = T^{ij} n_j \Delta A$$

est la composante i de la force $\vec{F}$ exercée sur une surface de normale $\vec{n}$. À la limite non-relativiste, les composantes T^{ij} ne sont autres que les composantes du tenseur des pressions, familier en mécanique des fluides. En résumé, le tenseur énergie-impulsion $T^{\mu\nu}$ est un tenseur symétrique : $T^{\mu\nu} = T^{\nu\mu}$. La conservation de l'énergie-impulsion se traduit par la généralisation de (3.15)

$$\boxed{\partial_\mu T^{\mu\nu} = 0} \tag{3.18}$$

Un cas particulier important est celui du fluide parfait. Dans ce cas le tenseur énergie-impulsion ne peut dépendre que de la vitesse d'ensemble u^μ du fluide, car il n'existe pas d'autre direction disponible[11] et aussi du tenseur de Minkowski $\eta_{\mu\nu}$. Dans le référentiel où le fluide est au repos, on doit avoir

$$T^{\mu\nu} = \mathrm{diag}\,(\rho, \mathcal{P}, \mathcal{P}, \mathcal{P})$$

[11] Le flux de chaleur définit une direction privilégiée dans le référentiel où le fluide est au repos, mais précisément il n'y a pas de flux de chaleur dans un fluide parfait.

car la densité d'énergie est $\rho = m\rho^*$ et $T^{ij} = \mathcal{P}\delta^{ij}$, $\mathcal{P}$ étant la pression. Si l'on écrit la forme la plus générale possible de $T^{\mu\nu}$ avec des coefficients arbitraires A et B

$$T^{\mu\nu} = Au^\mu u^\nu + B\eta^{\mu\nu}$$

on obtient dans le référentiel au repos où $u^\mu = (1, \vec{0})$

$$T^{00} = A + B = \rho \qquad T^{ij} = -B\,\delta^{ij} = \mathcal{P}\,\delta^{ij}$$

ce qui donne

$$\boxed{T^{\mu\nu} = (\rho + \mathcal{P})u^\mu u^\nu - \mathcal{P}\eta^{\mu\nu}} \tag{3.19}$$

Il est facile de vérifier que l'équation de conservation (3.18) donne l'équation de conservation de la masse et l'équation d'Euler à la limite des faibles vitesses[12]. Nous allons voir au chapitre 6 que le tenseur énergie-impulsion joue un rôle fondamental dans l'écriture de l,équation d'Einstein

courbure locale de l'espace-temps = tenseur énergie-impusion

ou en formule

$$R_{\mu\nu} - \frac{1}{2}\,g_{\mu\nu}R = -8\pi G T_{\mu\nu} \tag{3.20}$$

Le tenseur de Ricci $R_{\mu\nu}$, la courbure R et la métrique $g_{\mu\nu}$, qui seront définies de façon précise au chapitre 5, sont des caractéristiques de la géométrie.

Bibliographie.

Hartle [2003], chapitres 4 et 5. Ludvigsen [2000], chapitres 2 à 8. Weinberg [1972], chapitre 2.

[12] L'équation de conservation de la masse suit de

$$u^\mu \partial_\mu T^{\mu\nu} = 0$$

et celle d'Euler de

$$(\eta_{\rho\nu} - u_\rho u_\nu)\partial_\mu T^{\mu\nu} = 0$$

Le tenseur $(\eta_{\rho\nu} - u_\rho u_\nu)$ est le projecteur, au sens de la métrique de Minkowski, sur la surface orthogonale à u^μ. Les deux équations ci-dessus sont les projections de $\partial_\mu T^{\mu\nu} = 0$ sur u^μ et sur la surface perpendiculaire à u^μ.

Chapitre 4

Cosmologie

Les considérations de géométrie dont nous aurons besoin dans ce chapitre de cosmologie sont suffisamment intuitives pour qu'il ne soit pas nécessaire de faire appel aux résultats de géométrie différentielle du chapitre 5.

4.1 Description qualitative de l'Univers

Matière sombre. Il est aujourd'hui admis que la matière visible (étoiles, nuages de gaz ...) ne représente qu'une très faible partie de la masse de l'Univers, au plus quelques pour cents. Une composante essentielle est la *matière sombre*. Cette matière sombre est mise en évidence par l'étude de la rotation de nuages de gaz autour du centre de galaxies. Par exemple la figure 4.1 montre la vitesse de rotation $v(r)$ de nuages de gaz dans la galaxie d'Andromède en fonction de la distance r au centre de la galaxie.

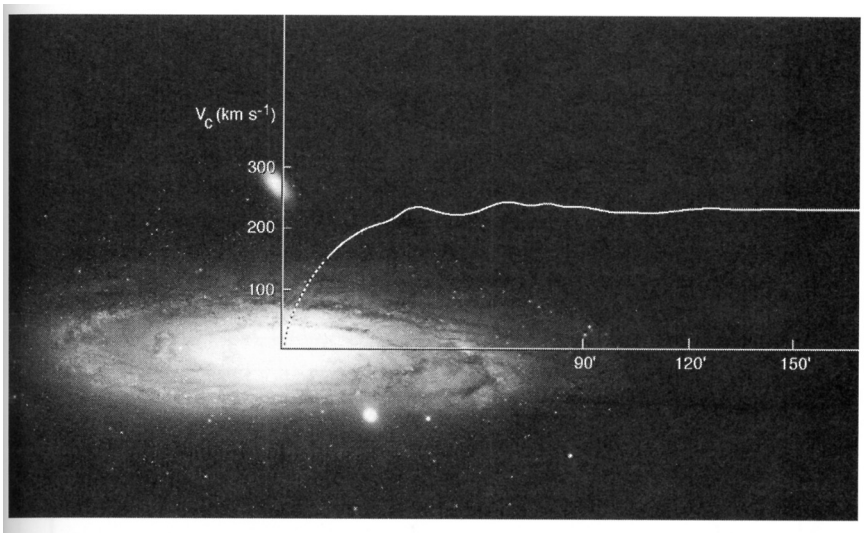

FIG. 4.1 – Vitesse de rotation de nuages de gaz en fonction de la distance au centre de la galaxie, ici la galaxie d'Andromède. L'échelle horizontale est en minutes d'arc et donne la taille angulaire des objets. D'après Hartle [2003].

Si le nuage est situé à une distance r du centre de la galaxie, et si $M(r)$ est la masse contenue à l'intérieur de l'orbite du nuage, alors d'après la loi de Newton

$$\frac{GM(r)}{r^2} = \frac{v^2(r)}{r} \tag{4.1}$$

Si l'essentiel de la masse était concentrée au voisinage de $r = 0$, comme on pourrait le déduire de l'observation de la matière visible, on aurait $v(r) \propto r^{-1/2}$. En fait on trouve $v(r) \sim$cste, ce qui indique la présence d'une matière non visible importante. L'existence de cette matière sombre a reçu récemment une confirmation indépendante grâce à l'utilisation de l'effet de lentille gravitationnelle : voir Hartle [2003], chapitre 11 et l'article de Koopmans et Blanford.

Isotropie et homogénéité de l'Univers = principe cosmologique. Si l'on fait abstraction des fluctuations locales (galaxies...), l'observation montre que l'Univers est homogène et isotrope (en fait l'isotropie en tout point entraîne l'homogénéité). Il faut moyenner sur des distances suffisamment grandes pour que l'homogénéité soit valable et l'échelle de transition entre la répartition homogène et la structure granulaire est de l'ordre de 3×10^7 année-lumière. Un autre argument convaincant est l'isotropie du rayonnement cosmologique[1] à 3 K : les résultats des satellites COBE et WMAP montrent qu'en dehors d'une asymétrie due à l'effet Doppler provenant du mouvement de notre galaxie[2] avec une vitesse de l'ordre $10^{-3}\,c$, ce rayonnement est isotrope avec des fluctuations qui ne dépassent pas 10^{-5}. En fait la qualité de la courbe de rayonnement de corps noir du rayonnement cosmologique est bien meilleure que celle de toute courbe réalisée au laboratoire ! Le point de départ de la cosmologie standard est donc l'hypothèse que *l'Univers est homogène et isotrope à grande échelle*, hypothèse qu'Einstein fut le premier à formuler en 1917, à l'époque une pure spéculation théorique sans la moindre base expérimentale.

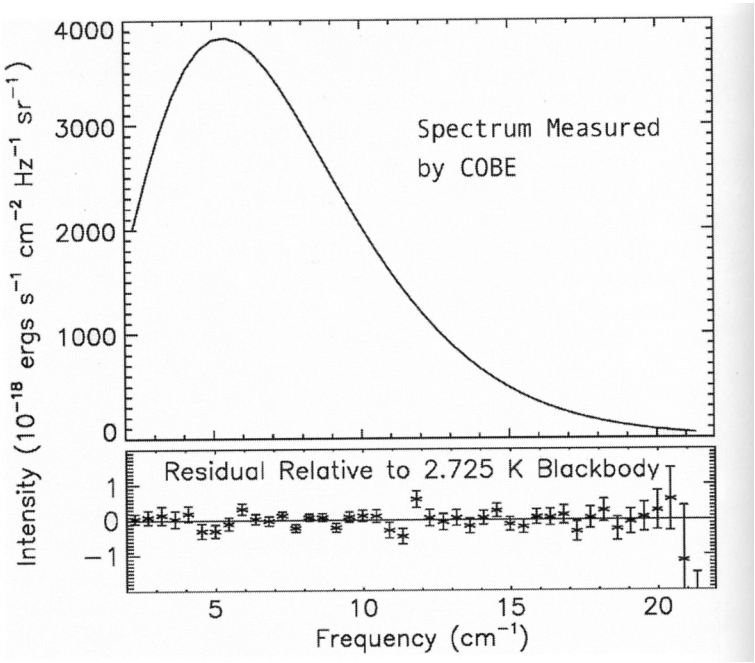

FIG. 4.2 – Courbe de Planck pour le rayonnement cosmologique mesurée par le satellite COBE (1992). D'après Hartle [2003].

Expansion de l'Univers. Comme on le sait depuis Hubble (1927), le décalage Doppler de la lumière émise par les galaxies est proportionnel à leur distance : pour une galaxie s'éloignant de la nôtre avec une vitesse v, le décalage Doppler est suivant (3.4) pour $v/c \ll 1$

$$\frac{v}{c} = \frac{\Delta\lambda}{\lambda} \equiv z \tag{4.2}$$

[1] Environ 380 000 ans après le Big Bang, les photons se sont découplés de la matière et se sont retrouvés hors équilibre thermodynamique. Depuis cette époque, leur longueur d'onde a varié en proportion du facteur de dilatation $a(t)$ (voir (4.17)), et ils forment aujourd'hui le rayonnement cosmologique à 3 K.

[2] La vitesse de notre groupe local de galaxies est de 600 km/s par rapport au fond cosmologique. La vitesse du système solaire par rapport au centre de notre galaxie est de 270 km/s.

où λ est la longueur d'onde. Une définition plus précise de z sera donnée ultérieurement. La loi établie par Hubble est $v = H_0 d$, où H_0 est la *constante de Hubble*, avec la valeur numérique

$$H_0 = 72 \pm 7\,\mathrm{km.s^{-1}/Mpc} \qquad t_H = \frac{1}{H_0} = 4.3 \times 10^{17}\,\mathrm{s} = 13.6 \times 10^9 \mathrm{ans} \tag{4.3}$$

Rappelons que 1 parsec (pc) vaut 3.26 année-lumière (a.l). Le temps t_H est une première approximation pour l'âge de l'Univers. En effet, en supposant que les galaxies ont eu une vitesse uniforme depuis le Big Bang, $d = vt_H$, et tenant compte d'autre part de la loi de Hubble $d = v/H_0 \implies t_H = 1/H_0$. Pour vérifier la loi de Hubble, on a besoin d'une mesure de distance. Celle-ci est fournie en utilisant des "bougies standard", c'est-à-dire des étoiles dont la luminosité L est connue (ou supposée telle !). En effet, si f est le flux recueilli par un détecteur sur la Terre, ce flux est relié à L et à la distance d par

$$f = \frac{L}{4\pi d^2} = \frac{LH_0^2}{4\pi z^2 c^2} \tag{4.4}$$

d'après la loi de Hubble. Nous verrons ultérieurement une version plus précise de cette formule ; z est connu par le décalage Doppler, f est mesuré expérimentalement, et si l'on connaît L, on en déduit d.

Il convient de faire quelques mises en garde contre des confusions possibles.

- Il ne faut pas interpréter le Big Bang comme une explosion à partir d'un centre ! L'Univers n'a pas de centre, chaque galaxie voit les autres galaxies s'éloigner suivant la loi de Hubble. Une image classique (à deux dimensions) est celle d'un ballon que l'on gonfle : des points marqués sur le ballon s'éloignent les uns des autres, sans que l'on ait un point central sur le ballon.
- Comme on le verra au chapitre 5, on ne sait pas *a priori* comparer la vitesse de deux galaxies éloignées dans un espace courbe. Pour ce faire il faudrait "transporter parallèlement" le vecteur vitesse d'une des deux galaxies au point où se trouve la seconde galaxie. Or le transport parallèle d'un vecteur dépend de la courbe choisie dans un espace de courbure non nulle. Les questions qui sont simples dans un espace plat le sont beaucoup moins dans un espace courbe !

4.2 Coordonnées comobiles

L'hypothèse d'isotropie et d'homogénéité permet de décomposer l'espace-temps M en une famille de 3-surfaces de genre espace, paramétrées par un scalaire que l'on peut appeler le temps t. Chacune de ces 3-surfaces est une variété Σ_t à trois dimensions homogène et isotrope : $M = \mathbb{R}^+ \times \Sigma$. On a donc un feuilletage de l'espace-temps en tranches t =cste. Il existe seulement trois types de variétés à trois dimensions homogènes et isotropes

1. L'espace plat.
2. La sphère S^3 à courbure constante > 0.
3. L'hyperboloïde à courbure constante < 0.

Ce résultat est intuitif, et il peut être montré rigoureusement. Remarquons que nous ne disons rien sur les propriétés topologiques *globales* (sauf dans le cas de la sphère) : par exemple dans le cas de l'espace plat, rien n'interdit de choisir le tore T^3, le tore étant une surface de courbure nulle. Nous allons nous limiter pour le moment au cas de l'espace plat, en revenant dans la section 4.4 sur les deux autres cas ; l'espace plat est le cas le plus simple, et le bonus est que c'est, semble-t-il, le cas physiquement pertinent !

Considérons les galaxies se trouvant au temps t sur la variété Σ_t : par homogénéité, ce temps, appelé *temps comobile*, peut être choisi identique pour toutes les galaxies. Au voisinage de chaque galaxie, on mesure au temps t une densité de matière $\rho(t)$ qui est la même pour toutes les galaxies. L'isotropie entraîne que la métrique est de la forme[3]

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2) \tag{4.5}$$

[3] Anticipons sur les résultats du chapitre 5 : chaque galaxie suit une géodésique G de l'espace-temps et on choisit comme paramètre le temps propre $t = \tau$. On a donc pour la métrique

$$g_{tt} =< \frac{\partial}{\partial t}, \frac{\partial}{\partial t} >= u^2 = 1$$

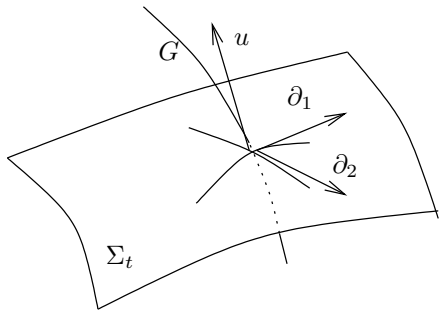

FIG. 4.3 – Variété Σ_t et géodésique dune galaxie G. On a représenté la quadrivitesse u et deux vecteurs tangents aux axes de coordonnées ∂_1 et ∂_2 (chapitre 5)

Les coordonnés (x, y, z) repèrent la positions dans l'espace d'une galaxie et elle sont *indépendantes du temps* : ce sont des *coordonnées comobiles*. Le facteur $a(t)$ est le *facteur de dilatation* ou le *facteur d'échelle*. La métrique (4.5) est la généralisation la plus simple de la métrique de Minkowski (3.7). Une tranche d'espace-temps $t =$cste est muni de la métrique euclidienne ordinaire pour $t =$cste

$$-\mathrm{d}s^2 = a^2(t)(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2) = a^2(t)\mathrm{d}S^2 \tag{4.6}$$

Cette métrique sera évidement modifiée dans le cas de l'hyperboloïde et de la sphère. La métrique (4.5) correspond à un *espace* plat, mais pas à un *espace-temps* plat !

Nous allons maintenant généraliser le résultat de Hubble, qui n'est valable que pour des galaxies pas trop éloignées, en établissant la formule du *décalage vers le rouge cosmologique*. Considérons deux galaxies dont la différence de coordonnées comobiles est $(\Delta x, \Delta y, \Delta z)$. La distance qui sépare ces deux galaxies au temps t est d'après (4.6)[4]

$$d(t) = a(t)\left[\Delta x^2 + \Delta y^2 + \Delta z^2\right]^{1/2} = a(t)d_{\mathrm{com}} \tag{4.7}$$

Si $a(t)$ croît avec t, les deux galaxies s'éloignent : l'Univers est en expansion, ce que nous allons désormais supposer : $\dot{a}(t) > 0$. Il sera commode d'utiliser des coordonnées sphériques (r, θ, φ) dans l'espace et de récrire (4.5)

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)\left[\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2)\right] = \mathrm{d}t^2 - a^2(t)\left[\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega^2\right] \tag{4.8}$$

Soit deux galaxies, de coordonnées comobiles $r = 0$ (nous) et $r = r_{\mathrm{com}}$, et un photon se propageant entre les deux galaxies. Dans le cas d'un photon, on doit avoir $\mathrm{d}s^2 = 0$, mais à cause du facteur $a(t)$ dans la métrique, les lignes d'Univers d'un photon ne sont pas des droites (figure 4.4). Si les temps d'émission par la galaxie de deux photons succcessifs à $r = r_{\mathrm{com}}$ sont t_e et $t_e + \delta t_e$ et les temps de réception sur la Terre sont t_0 et $t_0 + \delta t_0$, on tire de $\mathrm{d}s^2 = 0$ la relation $\mathrm{d}t^2 = a^2(t)\mathrm{d}r^2$ et

$$r_{\mathrm{com}} = \int_{t_e}^{t_0} \frac{\mathrm{d}t}{a(t)} = \int_{t_e+\delta t_e}^{t_0+\delta t_0} \frac{\mathrm{d}t}{a(t)} \tag{4.9}$$

car

$$\frac{\partial}{\partial t} = \frac{\mathrm{d}}{\mathrm{d}\tau}\Big|_{\text{le long de } G}$$

Les coordonnées spatiales sont telles que les vecteurs tangents $\partial/\partial x^i$ sont orthogonaux à $\partial/\partial t$. La métrique est donc de la forme

$$\mathrm{d}s^2 = \mathrm{d}t^2 - g_{ij}\mathrm{d}x^i\mathrm{d}x^j$$

Géométriquement, les géodésiques suivies par les galaxies sont orthogonales à Σ_t, sinon elles définiraient une direction privilégiée sur Σ_t. Cela veut dire que l'on n'a pas de termes en $\mathrm{d}t\,\mathrm{d}x^i$.

[4]Il convient de préciser la signification de la distance $d(t)$: en théorie, il faudrait disposer d'un grand nombre d'observateurs disposés entre notre galaxie et la galaxie lointaine, qui mesurent chacun la distance entre eux-mêmes et leur plus proche voisin au même temps t. La somme de ces distances donne $d(t)$. Inutile de dire que ce n'est pas très pratique...

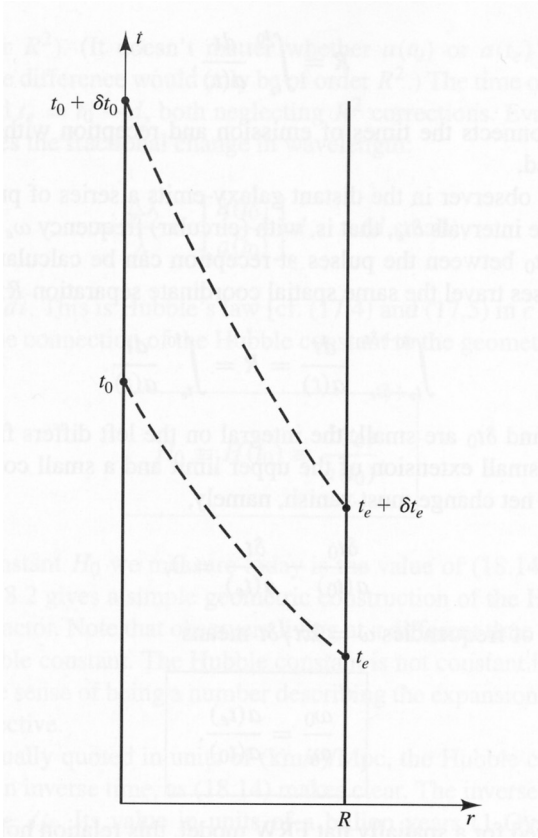

FIG. 4.4 – Propagation de photons entre deux galaxies et décalage vers le rouge gravitationnel, $R = r_{\text{com}}$.

soit pour des δt petits

$$\frac{\delta t_0}{a(t_0)} = \frac{\delta t_e}{a(t_e)}$$

Ceci donne pour le décalage de fréquence entre photons émis (ω_e) et reçus (ω_0)

$$\boxed{\frac{\omega_0}{\omega_e} = \frac{a(t_e)}{a(t_0)} \left(< 1\right)} \tag{4.10}$$

Cette équation donne l'expression générale du décalage vers le rouge cosmologique. On définit le facteur z par

$$\boxed{1 + z = \frac{\lambda_0}{\lambda_e} = \frac{\omega_e}{\omega_0} = \frac{a(t_0)}{a(t_e)}} \tag{4.11}$$

On verra dans la section 4.4.1 que ces équations sont aussi valables pour un espace courbe. Dans ce raisonnement, toute référence à la notion hasardeuse de vitesse relative de deux galaxies a disparu. En fait il est bien préférable de définir la distance à une galaxie lointaine par la donnée de z, qui est une donnée *observationnelle* non ambiguë, alors que le temps d'émission d'un photon (t_e dans (4.11)) ou bien la distance dépendent du modèle d'Univers choisi et de la définition de cette distance (voir la section 4.4.2). On dira par exemple que le découplage des photons du rayonnement cosmologique s'est produit à $z \simeq 1100$, ou que les objets les plus lointains que l'on a pu observer aujourd'hui ont $z \sim 6$.

Montrons que la loi de Hubble des équations (4.2) et (4.3) est une approximation de (4.11) en considérant deux galaxies proches. Le temps mis par le photon pour aller d'une galaxie à l'autre est $\Delta t \simeq a(t_0)R = d$, où t_0 représente l'instant actuel, mesuré à partir du Big Bang, et donc le facteur de dilatation aujourd'hui.

Selon une convention habituelle en cosmologie, l'indice 0 étiquette aujourd'hui, par exemple H_0 est la constante de Hubble aujourd'hui et $a_0 \equiv a(t_0)$. Écrivant

$$a(t_e) \simeq a(t_0 - d) \simeq a(t_0) - d\,\dot{a}(t_0)$$

on déduit de (4.11)

$$1 + z \simeq \frac{a(t_0)}{a(t_0) - d\,\dot{a}(t_0)} \simeq 1 + d\,\frac{\dot{a}(t_0)}{a(t_0)} = 1 + dH_0$$

soit

$$z \simeq H_0 d \qquad H_0 \equiv \frac{\dot{a}(t_0)}{a(t_0)} = \frac{\dot{a}_0}{a_0} \tag{4.12}$$

L'interprétation du décalage vers le rouge cosmologique comme provenant de l'effet Doppler et de la "fuite des galaxies" n'est correcte que pour des galaxies suffisamment proches. Dans le cas général on doit utiliser (4.11).

Nous avons vu que l'inverse de la constante de Hubble (aujourd'hui) donne une indication sur l'âge de l'Univers t_0 : si la vitesse d'expansion a été uniforme depuis le Big Bang, alors $H(t) = H_0$ et l'âge de l'Univers est bien l'inverse de la constante de Hubble : $t_0 = t_H = 1/H_0$. Cependant il est intuitivement évident (et ce sera montré ci-dessous), que la gravité ne peut que ralentir l'expansion de l'Univers : une pierre lancée vers le haut est toujours freinée par la gravité, même si on la lance avec une vitesse supérieure à la vitesse de libération et $\ddot{a}(t) < 0$. La figure 4.5 montre l'évolution du facteur de dilatation $a(t)$: on voit que $t_0 \leq t_H$. On verra que la prédiction simple et remarquable $\ddot{a}(t) < 0$ *n'est pas vérifiée expérimentalement.*

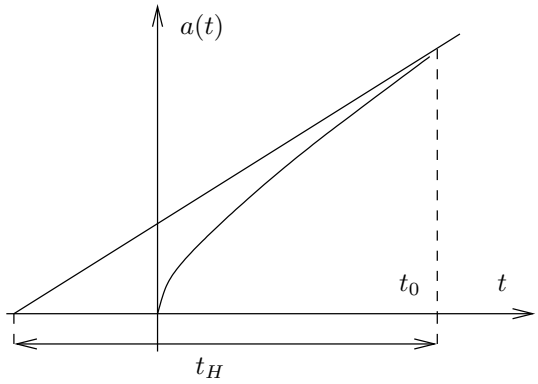

FIG. 4.5 – Le facteur de dilatation $a(t)$ en l'absence d'énergie du vide. La tangente au point $t = t_0$ donne la vitesse d'expansion aujourd'hui. Si cette vitesse était restée constante, l'âge de l'Univers serait t_H.

4.3 Évolution du facteur de dilatation

4.3.1 Équation de Friedmann

En première approximation, l'Univers est un fluide parfait dont les molécules sont les galaxies. Un flux de chaleur ou de particules serait incompatible avec l'isotropie, car il fixerait une direction privilégiée, et on est en droit de négliger tout phénomène irréversible. Le deuxième principe (E = énergie, S = entropie, $\mathcal{P}$ =pression, V = volume, T = température absolue)

$$\mathrm{d}E = T\mathrm{d}S - \mathcal{P}\mathrm{d}V$$

devient

$$\mathrm{d}(\Delta E) = -\mathcal{P}\mathrm{d}(\Delta V) \tag{4.13}$$

Des coordonnées comobiles $(\Delta x, \Delta y, \Delta z)$ définissent un covolume $\Delta V_{\rm com} = \Delta x \Delta y \Delta z$ où le nombre de galaxies est constant, tandis que le volume ΔV est

$$\Delta V = a^3(t)\,\Delta x \Delta y \Delta z = a^3(t)\,\Delta V_{\rm com} \tag{4.14}$$

Si $\rho(t)$ est la densité d'énergie (qui inclut bien évidemment l'énergie de masse), nous avons d'après (4.13) et (4.14)

$$\frac{\rm d}{{\rm d}t}\left[\rho(t) a^3(t) \Delta V_{\rm com}\right] = -\mathcal{P}(t)\,\frac{\rm d}{{\rm d}t}\left[a^3(t)\Delta V_{\rm com}\right]$$

soit

$$\frac{\rm d}{{\rm d}t}\left[\rho(t) a^3(t)\right] = -\mathcal{P}(t)\,\frac{\rm d}{{\rm d}t}\left[a^3(t)\right] \tag{4.15}$$

Trois cas de figure limites sont possibles.

1. *Univers dominé par la matière.* La pression du gaz de galaxies est négligeable par rapport à l'énergie de masse et (4.15) devient

$$\rho(t) a^3(t) = \text{cste} \qquad \text{ou} \qquad \rho(t) = \rho(t_0)(1+z)^3 \tag{4.16}$$

   Cette équation exprime simplement que la densité est divisée par $(a(t_0)/a(t))^3 = (1+z)^3$ quand le facteur de dilatation passe de $a(t)$ à $a(t_0)$.

2. *Univers dominé par le rayonnement.* Dans ce cas (voir le rayonnement du corps noir), on a $\mathcal{P} = \rho/3$ et $\rho \propto T^4$, équations qui sont aussi valides pour tout gaz de particules ultra-relativistes, quelle que soit leur statistique. On en déduit

$$\frac{\dot\rho}{\rho} = -3\,\frac{\dot a}{a}$$

   soit

$$\rho(t) = \rho(t_0)\left[\frac{a(t_0)}{a(t)}\right]^4 = \rho(t_0)(1+z)^4 \qquad T(t) = T(t_0)\left[\frac{a(t_0)}{a(t)}\right] = T(t_0)(1+z) \tag{4.17}$$

   Cette équation montre que la longueur d'onde des photons du rayonnement du corps noir cosmologique croît en proportion de $a(t)$ après le découplage rayonnement-matière.

3. *Univers dominé par le vide.* Dans ce cas ρ doit être indépendant du temps[5] : la densité d'énergie du vide ne dépend pas de l'expansion de l'Univers. On a donc

$$\rho(t)\frac{{\rm d}a^3(t)}{{\rm d}t} = -\mathcal{P}(t)\frac{{\rm d}a^3(t)}{{\rm d}t}$$

   soit

$$\mathcal{P} = -\rho \tag{4.18}$$

Nous reviendrons ultérieurement sur cette "énergie du vide". Pour le moment nous prenons en compte uniquement les paramètres classiques, matière et rayonnement, et nous allons essayer d'établir l'équation d'évolution du paramètre d'échelle $a(t)$ en utilisant un raisonnement newtonien discutable, mais qui a le mérite de la simplicité. Choisissons une origine arbitraire O dans l'Univers et considérons une galaxie G de masse m située à une distance $d(t)$ de O. En admettant la validité du théorème de Gauss pour une distribution de matière infinie, son énergie potentielle est

$$U(d) = -G\frac{mM}{d} \qquad M = \frac{4\pi}{3}\,d^3\rho$$

soit

$$U(d) = -\frac{4\pi}{3}\,Gmd^2\rho \tag{4.19}$$

[5] Ceci est une simplification qui est contestée dans certains modèles. Avec l'hypothèse simplificatrice $\rho_v(t) = \text{cste}$, l'effet de l'énergie du vide est équivalent à celui d'une constante cosmologique. Dans un référentiel d'inertie local, le seul tenseur disponible est le tenseur de Minkowski $\eta_{\mu\nu}$ et si tous les observateurs voient le même vide, on doit avoir $T^v_{\mu\nu} \propto \eta_{\mu\nu}$. Ceci rajoute un terme $\Lambda g_{\mu\nu}$ dans l'équation d'Einstein : voir(4.42), où Λ est la constante cosmologique.

Écrivons $d(t) = a(t)R$, où R est la coordonnée comobile radiale de la galaxie, l'origine O ayant une coordonnée comobile $R = 0$, d'où la vitesse de la galaxie $v(t) = R\,\dot{a}(t)$. L'énergie de la galaxie est alors

$$\begin{aligned} E &= \frac{1}{2}\,mR^2\,\dot{a}^2(t) - \frac{4\pi}{3}mR^2G\rho(t)\,a^2(t) \\ &= \frac{1}{2}\,mR^2\left(\dot{a}^2(t) - \frac{8\pi}{3}\,G\rho(t)\;a^2(t)\right) \end{aligned}$$

Cette énergie mécanique doit être constante, et après redéfinition convenable de $a(t)$ par un facteur multiplicatif, on obtient *l'équation de Friedmann*

$$\boxed{\dot{a}^2(t) - \frac{8\pi}{3}\,G\rho(t)\;a^2(t) = -k \qquad k = 1, 0, -1} \tag{4.20}$$

où k peut prendre les valeurs $k = -1$: énergie mécanique positive, $k = 0$: énergie mécanique nulle et $k = 1$: énergie mécanique négative. Pour $k = -1\,(E > 0)$ et $k = 0$ $(E = 0)$ l'expansion se poursuit indéfiniment, tandis que pour $k = 1\,(E < 0)$ l'expansion finit par s'arrêter et on revient à la singularité initiale ("Big Crunch"). La raison de la convention de signe pour k apparaîtra à la section 4.4 : k caractérise la courbure de l'espace, $k = +1$ pour la sphère, $k = 0$ pour l'espace plat et $k = -1$ pour l'hyperboloïde. De l'équation (4.20) on déduit, en la combinant avec (4.15)

$$\boxed{\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(3\mathcal{P} + \rho)} \tag{4.21}$$

ce qui montre que $\ddot{a} < 0$ si $(3\mathcal{P} + \rho) > 0$, en particulier si la pression et l'énergie sont positives : s'il n'y a que de la matière et/ou du rayonnement, l'expansion est ralentie par la gravité, même dans le cas où elle se poursuit indéfiniment. L'observation $\ddot{a} > 0$ implique donc une nouvelle physique !

L'Univers spatialement plat correspond à $k = 0$. De l'équation (4.20) prise à $t = t_0$ on déduit la densité d'énergie aujourd'hui en tenant compte de ce que $H_0 = \dot{a}_0/a_0$

$$\boxed{\rho_0 = \rho_c = \frac{3H_0^2}{8\pi G}} \tag{4.22}$$

La densité ρ_c est la *densité critique* : pour $\rho \leq \rho_c$, l'expansion se poursuit indéfiniment, pour $\rho > \rho_c$ l'Univers finit dans le Big Crunch. Les cosmologistes ont l'habitude de définir les rapports des différents types de densité d'énergie à la densité d'énergie critique (m = matière, r = rayonnement, v = vide)

$$\Omega_m = \frac{\rho_m^0}{\rho_c} \qquad \Omega_r = \frac{\rho_r^0}{\rho_c} \qquad \Omega_v = \frac{\rho_v^0}{\rho_c} \tag{4.23}$$

et $\Omega_m + \Omega_r + \Omega_v = 1$ pour un Univers spatialement plat ($\rho_0 = \rho_c$).

Pour fixer les idées, prenons le cas particulier d'un Univers plat dominé par la matière, $\Omega_m = 1$, $\Omega_r = \Omega_v = 0$, ce qui est en fait le premier modèle d'Univers en expansion proposé en 1932 par Einstein et de Sitter. Les deux équations (4.17) et (4.20) deviennent

$$\dot{a}^2 - \frac{8\pi G}{3a} = 0 \qquad \rho a^3 = 1$$

d'où l'on tire

$$\sqrt{a}\,\mathrm{d}a \propto \mathrm{d}t \qquad a^{3/2} \propto t$$

soit finalement une loi en $t^{2/3}$

$$a(t) = a_0\left(\frac{t}{t_0}\right)^{2/3} \tag{4.24}$$

et donc $t_0 = 2t_H/3$, ce qui confirme la courbe de la figure 4.5. Pour un Univers dominé par le rayonnement on trouve $a(t) \propto t^{1/2}$ et pour un Univers dominé par le vide $a(t) \propto \exp(Ht)$: si l'énergie du vide est non nulle, elle finit toujours par l'emporter ! Un exemple simple (non réaliste) avec $\Omega_m = \Omega_r = \Omega_v = 1/3$ est donné dans la figure 4.6.

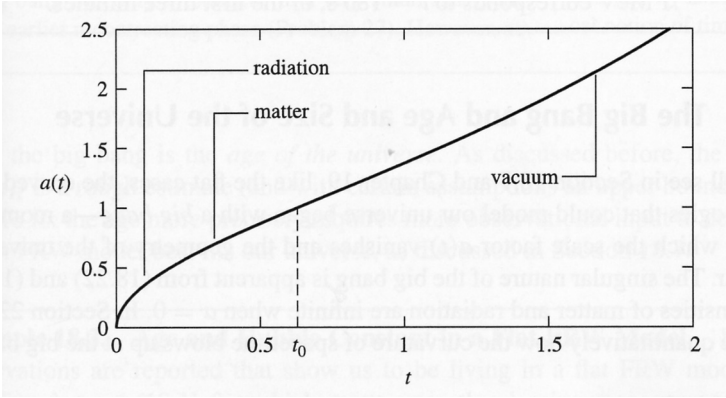

FIG. 4.6 – Le paramètre d'échelle $a(t)$ en fonction du temps avec $\Omega_m = \Omega_r = \Omega_v = 1/3$. D'après Hartle [2003]

4.3.2 Le problème de l'horizon

En raison de la vitesse finie de la lumière on ne peut observer qu'une fraction finie de l'univers : c'est le problème de *l'horizon*. On aurait envie de dire que l'horizon est simplement ct_0, mais c'est un peu plus compliqué en raison de l'expansion. Il est commode pour tracer les figures d'utiliser le "temps conforme" η défini par

$$\mathrm{d}\eta = \frac{\mathrm{d}t}{a(t)} \tag{4.25}$$

En effet, avec le "temps" η, les lignes d'Univers des photons sont des droites, car la métrique devient

$$\mathrm{d}s^2 = a^2(t)[\mathrm{d}\eta^2 - (\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega^2)] \tag{4.26}$$

Dans les coordonnées (η, r), les lignes d'Univers des photons sont des droites à 45°. L'horizon $r_H(t)$ est donné par (voir aussi (4.9))

$$r_H(t) = \int_0^t \frac{\mathrm{d}t'}{a(t')} \tag{4.27}$$

La distance physique $d_H(t)$ à l'horizon au temps d'observation t est

$$\boxed{d_H(t) = a(t) r_H(t) = a(t) \int_0^t \frac{\mathrm{d}t'}{a(t')}} \tag{4.28}$$

La région de l'Univers à laquelle on peut (en principe) accéder aujourd'hui est donc limitée par $d_H(t_0)$. Si l'on prend par exemple le cas simple d'un Univers plat et dominé par la matière, on montre immédiatement de (4.28) que

$$t_0 = \frac{2}{3}\, t_H \qquad d_H(t_0) = 3\, t_0$$

soit

$$d_H(t_0) = 2\, t_H = 2.7 \times 10^{10}\ \text{a.l}$$

Avec les données du modèle standard actuel (ΛCDM, section 4.5) : $\Omega_m = 30\%$, $\Omega_r \simeq 0$ et $\Omega_v = 70\%$ on trouve numériquement

$$d_H(t_0) \simeq 4.5 \times 10^{10}\ \text{a.l}$$

En fait l'Univers est opaque jusqu'au découplage des photons, qui s'effectue au bout de 4×10^5 ans environ après le Big Bang, et l'horizon effectif est plus petit (figure 4.7).

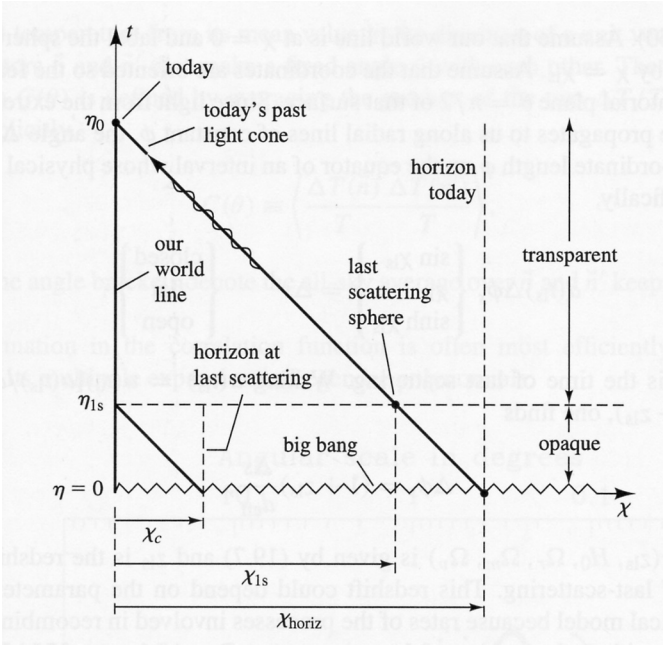

FIG. 4.7 – Univers visible aujourd'hui en tenant compte de l'opacité initiale. D'après Hartle [2003].

4.3.3 Relation entre la luminosité et le décalage vers le rouge.

Nous avons vu que pour des galaxies proches le rapport f/L du flux terrestre f à la luminosité L est donné en fonction du décalage z par (4.4), $c = 1$

$$\frac{f}{L} = \frac{H_0^2}{4\pi z^2}$$

Nous allons établir la généralisation de cette relation avec une première application au cas de l'espace plat, la discussion pour un courbure non nulle, légèrement plus complexe, étant renvoyée à la section 4.4.2. On définit une distance effective d_{eff} entre l'étoile et le point d'observation de la façon suivante : l'aire de la sphère lieu des points ayant la même distance à la source que le point d'observation est $4\pi d_{\text{eff}}^2$, et $d_{\text{eff}} \neq d(t)$ si la courbure spatiale est $\neq 0$. On observe une réduction du flux d'énergie pour deux raisons.

1. La fréquence (et donc l'énergie en raison de la loi de Planck-Einstein) des photons est plus faible en raison du décalage vers le rouge cosmologique (4.11)

$$\omega_0 = \frac{\omega_e}{(1+z)}$$

2. Comme l'intervalle δt_0 entre la réception de deux photons successifs est plus grand que l'intervalle δt_e entre leur émission, on compte moins de photons par unité de temps

$$\delta t_0 = \delta t_e (1+z)$$

Le rapport f/L vaut dans ces conditions

$$\frac{f}{L} = \frac{1}{4\pi d_{\text{eff}}^2} \frac{1}{(1+z)^2} \tag{4.29}$$

Prenons un modèle simple pour fixer les idées, celui d'Einstein et de Sitter : $\Omega_m = 1$, $\Omega_r = \Omega_v = 0$ et $k = 0$, un espace plat. Il est instructif de conduire le calcul en utilisant la variable z, qui est la "bonne" variable cosmologique, plutôt que t. Nous appellerons z' la variable d'intégration et suivant (4.11)

$$1 + z' = \frac{a_0}{a(t)} \qquad \frac{\mathrm{d}z'}{\mathrm{d}t} = -\frac{a_0\,\dot{a}(t)}{a^2(t)}$$

On écrit d'abord en suivant (4.7) que $d_{\rm eff} = a_0 d_{\rm com}$

$$d_{\rm eff} = a_0 \int_{t_e}^{t_0} \frac{{\rm d}t}{a(t)} = \int_0^z {\rm d}z' \, \frac{a(t)}{\dot{a}(t)}$$

et on exprime $a(t)/\dot{a}(t)$ en utilisant l'équation de Friedmann (4.20) pour $k = 0$ et $\rho = \rho_c(1+z')^3$, ce qui donne $\dot{a}/a = H_0(1+z')^{3/2}$, soit

$$d_{\rm eff} = \frac{1}{H_0} \int_0^z \frac{{\rm d}z'}{(1+z')^{3/2}} = \frac{2}{H_0} \left(1 - \frac{1}{\sqrt{1+z}}\right)$$

Cette expression se réduit à z/H_0 pour $z \ll 1$ et on retrouve (4.4). Écrivons f/L sous sa forme finale

$$\frac{f}{L} = \frac{H_0^2}{16\pi} \frac{1}{(1+z)[(1+z)^{1/2} - 1]^2} \tag{4.30}$$

Le cas général est examiné au § 4.4.2. La figure 4.8 illustre le calcul précédent dans le cas à deux dimensions d'espace.

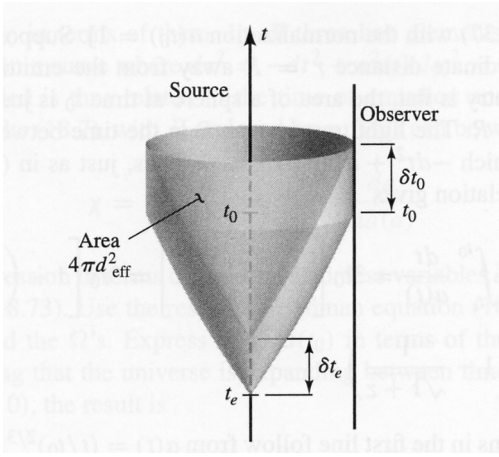

FIG. 4.8 – Relation flux/luminosité. D'après Hartle [2003].

4.4 Cas de la courbure spatiale non nulle

4.4.1 Métrique de Friedmann-Robertson-Walker

Traitons brièvement le cas de courbure spatiale non nulle, puisque la nature semble avoir choisi la courbure nulle ! Nous avons déjà mentionné qu'il n'existe que trois types de variétés à trois dimensions homogènes et isotropes

1. L'espace plat.
2. La sphère S^3 à courbure constante > 0.
3. L'hyperboloïde à courbure constante < 0.

L'espace plat vient d'être étudié ; passons au cas de la sphère S^3, qui est sans doute le plus intuitif. La sphère unité S^3 peut être paramétrée par trois angles (χ, θ, φ)

$$0 \le \chi \le \pi \qquad 0 \le \theta \le \pi \qquad 0 \le \varphi \le 2\pi$$

Les coordonnés cartésiennes (X, Y, Z, W) sont données par

$$\begin{aligned} X &= \sin\chi \sin\theta \cos\varphi \qquad & Z &= \sin\chi\cos\theta \\ Y &= \sin\chi \sin\theta \sin\varphi \qquad & W &= \cos\chi \end{aligned} \tag{4.31}$$

La métrique spatiale se calcule immédiatement à partir de (4.31)

$$\mathrm{d}S^2 = \mathrm{d}\chi^2 + \sin^2\chi(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2) \tag{4.32}$$

Le volume de l'espace est fini au temps t et est proportionnel à $2\pi^2 a^3(t)$. Le cas de l'hyperboloïde s'obtient en remplaçant $\sin\chi$ par $\sinh\chi$

$$\mathrm{d}S^2 = \mathrm{d}\chi^2 + \sinh^2\chi(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2) \tag{4.33}$$

Pour regrouper les trois cas de figure possibles,on écrit

$$\mathrm{d}S^2 = \mathrm{d}\chi^2 + S_k^2(\chi)(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2) \tag{4.34}$$

le facteur $S_k(\chi)$ étant donné par :

1. sphère : $S_1(\chi) = \sin\chi$;
2. espace plat : $S_0(\chi) = \chi$;
3. hyperboloïde : $S_{-1}(\chi) = \sinh\chi$.

Contrairement aux apparences, on voit que r est en fait une coordonnée angulaire ! L'équation de propagation d'un photon est (voir (4.6))

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)\mathrm{d}S^2 = 0$$

Si un photon reçu au temps t_0 a été émis par une galaxie dont la coordonnée est χ_{com}, le temps d'émission t_e est donné par

$$\chi_{\mathrm{com}} = \int_{t_e}^{t_0} \frac{\mathrm{d}t}{a(t)}$$

et on en déduit que (4.11) reste valable pour un espace courbe. L'aire d'une sphère à deux dimensions dont les points sont à une coordonnée comobile χ_{com} de l'origine est

$$4\pi a_0^2\, S_k^2(\chi_{\mathrm{com}})$$

On trouve souvent la métrique de l'espace-temps écrite sous la forme de Friedmann-Robertson-Walker (FRW), obtenue grâce à un changement de variables élémentaire

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)\right]$$

mais nous ne nous servirons pas de cette forme.

Dans un traitement rigoureux, l'évolution du facteur de dilatation $a(t)$ est fixée par l'équation d'Einstein (3.20)

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = -8\pi G T_{\mu\nu}$$

et par l'expression (3.19) de $T_{\mu\nu}$ du fluide parfait de galaxies

$$T_{\mu\nu} = (\mathcal{P} + \rho)u_\mu u_\nu - \mathcal{P}g_{\mu\nu}$$

La combinaison de ces deux équations permet de démontrer rigoureusement l'équation de Friedmann (4.20). Ensuite, étant donné une équation d'état, on peut en déduire la loi d'évolution de $a(t)$. Pour $k = +1$, l'expansion s'arrête au bout d'un certain temps et l'Univers se met à se contracter : c'est le Big Crunch. Pour $k = 0$ et $k = -1$, l'expansion se poursuit indéfiniment. Toutefois, si une énergie du vide est présente, on observe toujours une expansion indéfinie. Une prédiction remarquable de la relativité générale est que le type d'Univers homogène et isotrope, caractérisé par sa courbure, est lié à la densité d'énergie qu'il contient.

4.4.2 Distance de luminosité dans le cas général

Revenons sur la distance de luminosité lorsque la courbure spatiale est non nulle, en écrivant l'équation de Friedmann sous la forme

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2\left[\frac{\rho}{\rho_c} + (1-\Omega_t)\frac{a_0^2}{a^2}\right] \tag{4.35}$$

où nous avons utilisé l'expression (4.22) de la densité critique ρ_c et défini $\Omega_t = \rho_0/\rho_c$; il est instructif de vérifier que cette équation est bien correcte à $t = 0$ en retrouvant (4.22). On décompose ρ en une partie matière (non relativiste) ρ_m, une partie rayonnement ρ_r et une partie vide ρ_v suivant (4.23). Dans le cas de la matière par exemple on écrit

$$\frac{\rho_m}{\rho_c} = \frac{\rho_m^0\, a_0^3}{\rho_c\, a^3} = \Omega_m(1+z')^3$$

et en procédant de la même manière avec ρ_r et ρ_v on aboutit à

$$\frac{\dot{a}}{a} = H_0\left[\Omega_m(1+z')^3 + \Omega_r(1+z')^4 + \Omega_v + (1-\Omega_t)(1+z')^2\right]^{1/2} \tag{4.36}$$

On en déduit $\chi_{\rm com}$

$$\chi_{\rm com} = \frac{1}{a_0 H_0}\int_0^z \frac{{\rm d}z'}{\left[\Omega_m(1+z')^3 + \Omega_r(1+z')^4 + \Omega_v + (1-\Omega_t)(1+z')^2\right]^{1/2}} \tag{4.37}$$

Ainsi que nous l'avons déjà mentionné, la surface de la sphère à utiliser dans le calcul est

$$4\pi d_{\rm eff}^2 = 4\pi a_0^2 S_k^2(\chi_{\rm com})$$

et la distance effective est $d_{\rm eff} = a_0 S_k(\chi_{\rm com})$ où l'on rappelle que

$$S_0(\chi) = \chi \qquad S_1(\chi) = \sin\chi \qquad S_{-1}(\chi) = \sinh\chi$$

Ceci donne le rapport f/L

$$\frac{f}{L} = \frac{1}{4\pi a_0^2 S_k^2(\chi_{\rm com})(1+z)^2} = \frac{1}{4\pi d_{\rm eff}^2(1+z)^2} = \frac{1}{4\pi d_L^2} \tag{4.38}$$

Cette expression permet d'identifier la *distance de luminosité*

$$d_L = d_{\rm eff}(1+z) = a_0 S_k(\chi_{\rm com})(1+z)$$

Une autre distance comunément utilisée est la *distance angulaire* d_A : considérons un objet que nous voyons sous un angle $\Delta\theta$. Les photons se sont propagés depuis cet objet jusqu'à nous à θ et φ constants. La taille de l'objet au moment de l'émission était $\Delta L = a(t)S_k(\chi_{\rm com})\Delta\theta$, et comme l'angle n'a pas varié, on a aujourd'hui

$$\Delta\theta = \frac{\Delta L}{a(t)S_k(\chi_{\rm com})} = \frac{\Delta L(1+z)}{a_0 S_k(\chi_{\rm com})} = \frac{\Delta L}{d_A}$$

par définition de la distance angulaire d_A. On a donc la relation suivante entre d_A et d_L

$$d_A = a_0 S_k(\chi)(1+z)^{-1} = d_L(1+z)^{-2}$$

4.5 Le modèle ΛCDM

Le rapport f/L dépend des paramètres $\Omega_m, \ldots \Omega_t$ du modèle d'Univers, et sa mesure en fonction de z, qui est une donné observationnelle indépendante, nous donne donc accès à ces paramètres. Il est commode de se placer à $t = t_0$ est de définir le *paramètre de décélération* q_0 par

$$q_0 = -\frac{a(t)\,\ddot{a}(t)}{\dot{a}^2(t)}\Big|_{t_0} \tag{4.39}$$

Ce paramètre a été initialement défini avec un signe moins, car tout le monde s'attendait d'après l'équation (4.21) à ce que q_0 soit positif, c'est-à-dire que l'expansion de l'Univers ralentisse.

En 1998, deux groupes indépendants réussirent à mesurer la décélération de l'Univers en utilisant comme bougies standard des supernovae de type Ia et la grande surprise fut l'observation d'une valeur négative de ce paramètre q_0, soit une expansion de l'Univers qui s'accélère! En fait les supernovae à grand z sont moins lumineuses que dans un Univers où l'expansion ralentirait. Comme nous l'avons vu, un tel comportement ne peut s'expliquer que par la présence d'une énergie du vide. Les résultats sont résumés dans la figure 4.9.

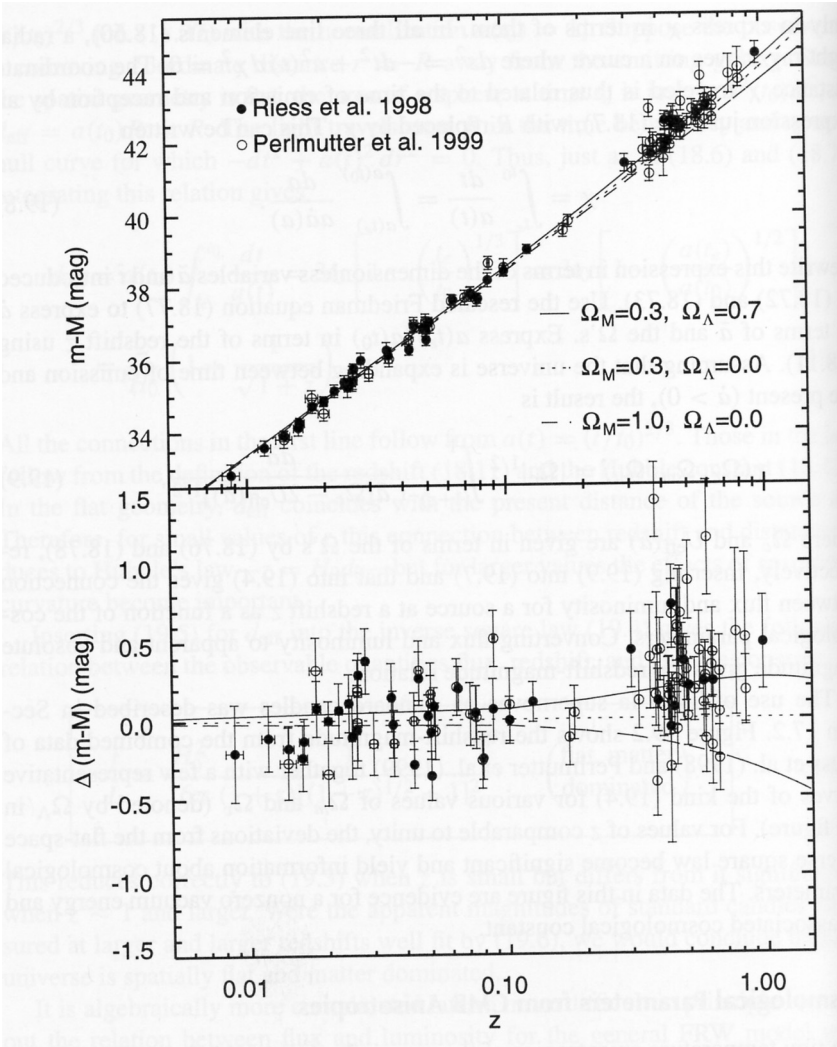

FIG. 4.9 – Accélération de l'expansion de l'Univers. Ω_M est noté Ω_m dans cet exposé, Ω_Λ est noté Ω_v et Ω_T est noté Ω_t ; z est défini en (4.11).

Il est intéressant de calculer le rapport $\ddot{a}/a$. En différentiant (4.36), ou mieux en utilisant directement (4.21), on trouve

$$\frac{\ddot{a}}{a} = -H_0^2 \left[\frac{1}{2}\,\Omega_m (1+z)^3 + \Omega_r (1+z)^4 - \Omega_v \right] \tag{4.40}$$

soit aujourd'hui, en négligeant le rayonnnement

$$\frac{\ddot{a}}{a} = -H_0^2 \left[\frac{1}{2}\,\Omega_m - \Omega_v \right] \qquad q_0 = \frac{1}{2}\,\Omega_m - \Omega_v \tag{4.41}$$

En l'absence d'énergie du vide on aurait bien $\ddot{a} < 0$.

Nous avons déjà mentionné que la matière visible n'est qu'une faible faction de la matière contenue dans l'Univers. Mais il y a pire : on pourrait imaginer que la matière non visible soit ordinaire, c'est-à-dire constituée de protons et de neutrons, une matière sombre baryonique. Malheureusement (?) les observations sont incompatibles avec cette hypothèse. La matière sombre doit être pour l'essentiel non baryonique, car l'étude des grandes structures de l'Univers montre qu'elle doit interagir très faiblement avec la matière ordinaire. De plus cette matière sombre doit être froide, ou non relativiste[6] : l'énergie cinétique des particules qui composent la matière sombre doit être petite par rapport à leur énergie de masse. Les deux données précédentes conduisent au modèle standard actuel de la cosmologie, modèle appelé ΛCDM : CDM = Cold Dark Matter, et Λ fait référence à la constante cosmologique introduite par Einstein. En effet une modification possible de (3.20) est

$$R_{\mu\nu} - \frac{1}{2}\, g_{\mu\nu} = -8\pi G T_{\mu\nu} - \Lambda g_{\mu\nu} \tag{4.42}$$

Cette modification est équivalente à l'introduction d'une énergie du vide constante (voir cependant la note 5) avec la correspondance

$$\rho_v = \frac{c^4 \Lambda}{8\pi G} \tag{4.43}$$

Paramètres	2002	WMAP
Constante de Hubble H_0 (km/s/Mpc	72 ± 7	71 ± 4
Décélération $q_0 = -a\ddot{a}/\dot{a}^2\|_{t_0}$	-0.67 ± 0.25	-0.66 ± 0.10
Âge de l'Univers (Gan)	13 ± 1.5	13.7 ± 0.2
Ω_t	1.03 ± 0.03	1.02 ± 0.02
Ω_b	0.039 ± 0.008	0.044 ± 0.004
$\Omega_{\rm cdm}$	0.29 ± 0.04	0.23 ± 0.04
Ω_v	0.67 ± 0.06	0.73 ± 0.06

TAB. 4.1 – Paramètres du modèle ΛCDM (a) Paramètres obtenus avant 2003 (b) Résultats du satellite WMAP (Wilkinson Microwave Anisotropy Probe). Ω_t est le rapport de la densité mesurée à la densité critique ρ_c (4.35), Ω_b est la fraction de matière baryonique, $\Omega_{\rm cdm}$ est la fraction de matière sombre froide ($\Omega_m = \Omega_b + \Omega_{\rm cdm}$) et Ω_v la fraction d'énergie du vide. D'après W. Freeman et M. Turner, *Rev. Mod.Phys*, **75**, 1433 (2003).

La table 4.1 compare les résultats disponibles en 2002 à ceux du satellite WMAP, qui a mesuré avec une grande précision les anisotropies du rayonnement cosmologique. Ces anisotropies sont liées aux fluctuations qui ont donné naissance aux grandes structures par instabilité gravitationnelle et leur mesure permet de remonter à un grand nombre de paramètres du modèle de Big Bang.

Comment comprendre l'énergie du vide ? Il n'y a pas pour le moment d'autre solution que de faire appel aux fluctuations quantiques. Prenons comme exemple le cas familier du champ électromagnétique quantifié dans une cavité de volume V. Chaque mode normal de la cavité est un oscillateur harmonique de fréquence $\omega_k = c|\vec{k}|$, où $\vec{k}$ est le vecteur d'onde. Il lui correspond en physique quantique un oscillateur harmonique quantifié dont l'énergie de point zéro, ou énergie de l'état fondamental, est $\hbar\omega_k/2$. La somme des énergies de point zéro de ces modes, ou énergie du vide, est infinie et vaut

$$V\rho_v = \sum_{\vec{k}} \hbar\omega_k = \frac{\hbar c V}{2\pi^2} \int_0^\infty k^3 \mathrm{d}k \tag{4.44}$$

[6] Le scenario CDM prédit que les petites structures se forment avant les grandes (scenario bottom up), alors que le scenario HDM (Hot Dark Matter) prédit au contraire que les petites structures s'obtiennent par fractionnement des grandes (scenario top-down). C'est le scenario bottom up qui est favorisé par l'observation.

Par une curieuse revanche de l'histoire, la quantification du rayonnement, postulée par Planck et Einstein pour se débarrasser des infinis dans le rayonnement du corps noir, introduit une autre source d'infinis, cette fois à température nulle ! En théorie quantique des champs on "renormalise" en décidant de prendre comme zéro d'énergie l'énergie du vide, qui est inobservable en valeur absolue, sauf en relativité générale, où l'on doit prendre en compte *toutes* les formes d'énergie. En revanche on observe des *différences* d'énergie du vide en disposant par exemple dans le vide deux plaques conductrices qui se font face. La modification des modes normaux due à la présence des plaques change l'énergie du vide par une quantité *finie*, et se traduit physiquement par une attraction entre les plaques. C'est *l'effet Casimir*, qui a été vérifié récemment avec une précision $\sim 10^{-3}$. Si l'on essaie de deviner un cut-off pour limiter l'intégrale dans (4.44), le seul cut-off naturel est fixé par l'échelle de Planck $l_P = \sqrt{G\hbar/c^3} \sim 10^{-35}$ m, et on trouve que la densité d'énergie du vide ainsi estimée est d'un facteur 10^{120} supérieure à celle mesurée en cosmologie ! Cependant ce calcul pose le problème du référentiel : il n'est pas invariant de Lorentz.

En résumé le modèle ΛCDM est en excellent accord avec les observations. En particulier la concordance des résultats pré-WMAP avec ceux de WMAP est impressionnante. Cependant il reste deux questions fondamentales, auxquelles les physiciens des particules et de la théorie quantique des champs n'ont pour le moment aucune réponse. Quelle est la particule (ou les particules) qui entrent dans la composition de la matière sombre froide ? Quelle est l'origine de l'énergie du vide ? En l'absence d'une réponse satisfaisante à ces deux questions, un doute continuera à planer sur la pertinence du modèle ΛCDM, et on ne pourra pas empêcher les astrophysiciens de spéculer sur des solutions encore plus radicales, comme la variabilité dans le temps des constantes fondamentales ou bien la faillite de la relativité générale aux très grandes distances.

Bibliographie.

Hartle [2003], chapitres 17 à 19 ; F. Bouchet, *Relativité, cosmologie et évolution de l'Univers*, contribution à l'ouvrage collectif *Actualité d'Einstein*, EDPSciences/Editions du CNRS, à paraître en 2005 ; J. Rich *Principes de la cosmologie*, Éditions de l'École Polytechnique, (2002) ; P. Coles et F. Lucchini *Cosmology : the origin and evolution of cosmic structure*, John Wiley, New-York (2002) ; W. Freeman et M. Turner, *Rev. Mod.Phys*, **75**, 1433 (2003) ; P. Peebles et B. Ratra, *Rev. Mod.Phys*, **75**, 559 (2003) ; S. Perlemutter, *Supernovae, Dark Energy and the Accelerating Universe*, Physics Today, avril 2003, p. 53 ; L. Koopmans et R. Blanford, *Gravitational Lenses*, Physics Today, juin 2004, p. 45.

Chapitre 5

Boîte à outils de géométrie différentielle

N.B. Ce chapitre n'est évidemment pas une introduction, même succincte, à la géométrie différentielle. Il se borne à rassembler les notions qui sont strictement indispensables à la compréhension des équations fondamentales de la relativité générale.

5.1 Espace tangent à une variété

Une *variété* M de dimension N est un espace topologique qui est *localement* isomorphe à $\mathbb{R}^N$. Cela ne dit rien sur les propriétés topologiques globales de la variété. Ainsi le plan réel à deux dimensions $\mathbb{R}^2$, le tore à deux dimensions T^2 et la sphère S^2 sont tous trois des variétés de dimension 2, bien que leurs propriétés topologiques globales soient très différentes. Nous n'aurons besoin dans ce cours que des propriété locales. On peut introduire localement sur la variété des coordonnées x^i qui repèrent un point de la variété. Toutefois, on ne peut pas en général repérer tout point P de la variété par un système de coordonnées unique. Il faut en général un système de cartes dont chacune recouvre un partie de la variété et qui se raccordent entre elles, l'ensemble des ces cartes formant un *atlas*.

Soit une courbe C tracée sur M paramétrée par t : les coordonnées d'un point de C sont $x^i(t)$. On appelle *dérivée directionnelle* ξ par rapport à la courbe C au point $P \in C$

$$\xi = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{\text{le long de } C \text{ en } P} \equiv \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{C,P}$$

l'application qui fait correspondre à toute fonction $f(t)$ sa dérivée $\mathrm{d}f/\mathrm{d}t$. L'ensemble de ces dérivées directionnelles forme *l'espace tangent T_P en P à la variété*; T_P est donc l'espace des dérivées directionnelles, et les dérivées directionnelles sont les vecteurs de T_P. Nous verrons ci-dessous que T_P est un espace vectoriel de dimension N. Montrons la linéarité : si Γ est une autre courbe paramétrée par t et coupant C en P (figure 5.1) avec $\eta = \mathrm{d}/\mathrm{d}t|_{\Gamma,P}$, on peut former la combinaison linéaire avec des coefficients a et b[1]

$$a\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{C,P} + b\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{\Gamma,P}$$

et cette combinaison est bien une *dérivation*, c'est-à-dire qu'elle obéit bien à la règle de Leibniz

$$(a\xi + b\eta)(fg) = g(a\xi + b\eta)f + f(a\xi + b\eta)g$$

On peut trouver aisément une base de T_P en utilisant un système de coordonnées $x^i(t)$ valable dans le

[1] Pour le lecteur qui souhaite des notations plus explicites : soit $x^i(t) = c^i(t)$ la paramétrisation de la courbe C par N fonctions $c^i(t)$ et $\gamma^i(t)$ celle de Γ : $x^i(t) = \gamma^i(t)$. Alors $\xi = \dot{c}^i\partial_i$ et $\eta = \dot{\gamma}^i\partial_i$. La combinaison linéaire des deux vecteurs est

$$a\xi + b\eta = (a\dot{c}^i + b\dot{\gamma}^i)\partial_i$$

.

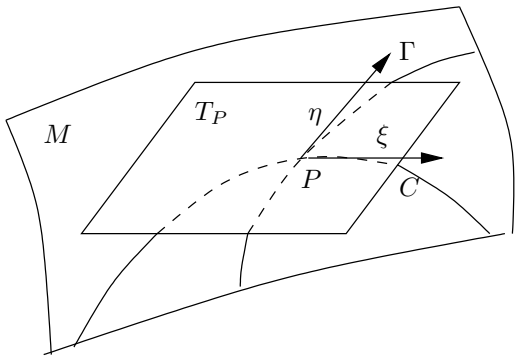

FIG. 5.1 – (a) Courbes $C(t)$ et $\Gamma(\tau)$, vecteur vitesse en P et interprétation géométrique du plan tangent en P.

voisinage du point P. Un point sur C étant alors repéré par $x^i(t)$, on peut écrire

$$\begin{aligned} \frac{\mathrm{d}f}{\mathrm{d}t}\Big|_{C,P} &= \frac{\mathrm{d}x^i}{\mathrm{d}t}\,\frac{\partial f}{\partial x^i} = \frac{\mathrm{d}x^i}{\mathrm{d}t}\,\partial_i f \\ \xi = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{C,P} &= \frac{\mathrm{d}x^i}{\mathrm{d}t}\,\partial_i = \xi^i\partial_i \end{aligned} \tag{5.1}$$

Si l'on interprète le paramètre t comme un temps, $x^i(t)$ décrit le parcours d'un point sur C en fonction du temps, et $\mathrm{d}x^i/\mathrm{d}t$ n'est autre que le vecteur vitesse de ce point en P : figure 5.1. Ceci donne une interprétation géométrique utile de la dérivée directionnelle, mais cette interprétation a l'inconvénient de masquer le caractère intrinsèque de cette dérivée : il n'est pas nécessaire de plonger M dans un espace de dimension $> N$, même si la figure 5.1 peut être utile pour l'intuition.

L'équation (5.1) montre que le vecteur $\xi \in T_P$, qui, soulignons le une fois de plus, *existe indépendamment de tout système de coordonnées*, a pour composantes $\xi^i = \mathrm{d}x^i/\mathrm{d}t$ dans la base $e_i = \partial_i$ de T_P. L'ensemble des e_i forme une base de T_P, appelée *base de coordonnées*, et comme il y a N dérivées partielles ∂_i indépendantes, T_P est manifestement de dimension N. Il est crucial de comprendre que l'espace tangent à M est associé au point P, et on ne saura pas *a priori* comparer les vecteurs de deux plans tangents T_P et $T_{P'}$ associés à deux points P et P' différents. En résumé, dans une base de coordonnées, un vecteur ξ s'écrit comme

$$\xi = \xi^i e_i = \xi^i\partial_i \qquad e_i^k = \delta_i^k \tag{5.2}$$

Le vecteur ξ est appelé *vecteur contravariant*, ses composantes sont en exposant. En ce qui concerne les vecteurs de base contravariants e_k, k étiquette les vecteurs de base, de composantes $e_k^i = \delta_k^i$. Il est très important de comprendre que, contrairement à ce que pourrait laisser supposer la notation x^i, la coordonnée x^i *n'est pas* la composante i d'un vecteur contravariant. Dans le cas de l'espace-temps plat du chapitre 3, x^μ était *à la fois* une coordonnée et la composante μ d'un vecteur qui relie l'origine au point de coordonnées x^μ. Ce n'est plus le cas dans l'espace-temps de la relativité générale. En fait on a noté x^i par commodité, mais on aurait aussi bien pu utiliser x_i ; la notation x^i a l'avantage de la cohérence, car $u^i = \mathrm{d}x^i/\mathrm{d}t|_{C,P}$ est une composante de vecteur, contrairement à x^i, alors que dans le cas du chapitre 3 on distingue x^μ et x_μ (*cf.* (3.7)) : voir l'exemple de la section 5.3.3 où $x^1 = \theta$ et $x^2 = \varphi$. On n'insistera jamais assez sur le fait que les vecteurs sur une variété sont *attachés à l'espace tangent en un point de ladite variété.*

La *dérivée directionnelle* ∂_ξ d'une fonction $f(x^i)$ est par définition

$$\partial_\xi f = \xi^i\partial_i f \tag{5.3}$$

Si l'on choisit pour ξ un vecteur de base e_k, $\xi = e_k$

$$e_k^i\partial_i f = \partial_{e_k} f = \delta_k^i\partial_i f = \partial_k f$$

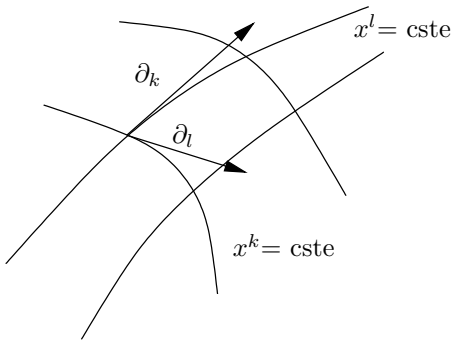

FIG. 5.2 – Dérivées directionnelles ∂_k et ∂_l.

∂_k est la dérivée dans la direction k, les autres coordonnées restant constantes : figure 5.2. Dans un changement de système de coordonnées $x^i(x'^j)$, la loi de transformation des dérivées directionnelles est

$$\frac{\mathrm{d}x^i}{\mathrm{d}t}\Big|_P = \frac{\partial x^i}{\partial x'^j}\Big|_P \frac{\mathrm{d}x'^j}{\mathrm{d}t}\Big|_{C,P} \quad \text{ou} \quad \xi^i = \frac{\partial x^i}{\partial x'^j}\Big|_P \xi'^j \tag{5.4}$$

On peut donc écrire une forme matricielle de cette loi de transformation

$$\xi^i = A^i_{\ j}(P)\xi'^j \qquad A^i_{\ j}(P) = \frac{\partial x^i}{\partial x'^j}\Big|_P \tag{5.5}$$

C'est bien sûr une transformation linéraire sur des vecteurs, mais cette transformation est spécifique du point P : la matrice $A^i_{\ j}(P)$ dépend du point P.

Passons ensuite à la notion de *covecteur = vecteur covariant* ou encore *1-forme*. Un covecteur est simplement une forme linéaire sur les vecteurs, c'est-à-dire une application $T_P \to \mathbb{R}$ qui à tout vecteur fait correspondre un nombre réel en respectant la propriété de linéarité. Nous noterons Λ un covecteur et $< \Lambda, \xi >$ l'application linéaire. La linéarité implique l'action suivante sur la combinaison $a\xi + b\eta$ des vecteurs ξ et η

$$< \Lambda, a\xi + b\eta >= a < \Lambda, \xi > +b < \Lambda, \eta >$$

L'application $\mathrm{d}f$ qui au vecteur $\xi = \mathrm{d}/\mathrm{d}t|_{C,P}$ fait correspondre le nombre $\mathrm{d}f/\mathrm{d}t|_{C,P}$

$$< \mathrm{d}f, \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{C,P} >\equiv \frac{\mathrm{d}f}{\mathrm{d}t}\Big|_{C,P} \tag{5.6}$$

est bien un covecteur. En effet

$$< \mathrm{d}f, a\xi + b\eta >= a\frac{\mathrm{d}f}{\mathrm{d}t}\Big|_{C,P} + b\frac{\mathrm{d}f}{\mathrm{d}t}\Big|_{\Gamma,P}$$

Soit e^i la base duale de e_i, c'est-à-dire la base telle que $< e^i, e_k >= \delta^i_k$. Les composantes de $\mathrm{d}f$ ne sont autres que $\partial_i f$. En effet

$$< \frac{\partial f}{\partial x^i}e^i, \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{C,P} >=< \partial_i f e^i, \xi^k e_k >= \frac{\mathrm{d}x^i}{\mathrm{d}t}\,\partial_i f = \frac{\mathrm{d}f}{\mathrm{d}t}\Big|_{C,P}$$

La base e^i est souvent notée $e^i \equiv \mathrm{d}x^i$. La loi de transformation des composantes $\partial_i f$ d'un covecteur sont *différentes* de celles d'un vecteur. Écrivons les explicitement

$$\frac{\partial f}{\partial x^i} = \frac{\partial f}{\partial x'^j}\,\frac{\partial x'^j}{\partial x^i}\Big|_P \quad \text{soit} \quad \Lambda_i = \frac{\partial x'^j}{\partial x^i}\,\Lambda'_j\Big|_P \tag{5.7}$$

5.2 Champs de tenseurs

Un champ de vecteurs est un ensemble de vecteurs dépendant du point P, de composantes $T^i(P) \equiv \xi^i(P) = \xi^i(x^j)$ dans la base de coordonnées e_i, et se transformant en chaque point selon (5.4). On définit de même des champs de covecteurs $T_i(P) \equiv \Lambda_i(P) = \Lambda_i(x^j)$ se transformant suivant (5.7). Un tenseur de type (m, n) aura m composantes contravariantes et n composantes covariantes

$$T^{i_1 \ldots i_m}{}_{j_1 \ldots j_n} \tag{5.8}$$

se transformant par (5.4) pour les composantes contravariantes et suivant (5.7) pour les composantes covariantes. Par exemple le tenseur de type (0,2) T_{ij} se transformera comme

$$T_{ij} = \frac{\partial x'^k}{\partial x^i} \frac{\partial x'^l}{\partial x^j} T'_{kl} \tag{5.9}$$

Tout comme le vecteur ξ, le tenseur T existe indépendamment de ses composantes

$$T = T^{i_1 \ldots i_m}{}_{j_1 \ldots j_n} \, e_{i_1} \otimes \cdots \otimes e_{i_m} \, e^{j_1} \otimes \cdots \otimes e^{j_n} \tag{5.10}$$

où $\otimes$ indique le produit tensoriel. Le type de tenseur est défini par la donnée du couple (m, n). Il existe deux opérations sur les champs de tenseurs qui sont intrinsèques à la variété, c'est-à-dire qui ne dépendent d'aucune structure additionnelle

(1) *La différentiation extérieure* : nous nous limiterons à en donner la définition dans le cas d'un covecteur T_i, dont la dérivée extérieure est[2]

$$(\mathrm{d}T)_{ij} = \frac{\partial T_i}{\partial x^j} - \frac{\partial T_j}{\partial x^i} \tag{5.11}$$

C'est un bon exercice que de vérifier que $(\mathrm{d}T)_{ij}$ se transforme comme un tenseur (0,2).

(2) *La dérivée de Lie* L_ξ d'un champ de tenseurs, qui est définie de la façon suivante. Pour une fonction scalaire $f(x^i)$, $L_\xi f$ est simplement la dérivée directionnelle $\partial_\xi f$

$$L_\xi f = \partial_\xi f = \xi^i \partial_i f \tag{5.12}$$

Pour un champ de vecteurs η^i

$$L_\xi \eta^i = \xi^j \partial_j \eta^i - \eta^j \partial_j \xi^i = \partial_\xi \eta^i - \partial_\eta \xi^i = -L_\eta \xi^i \tag{5.13}$$

Un résultat important est que $[\partial_\xi, \partial_\eta]$ définit un champ de vecteurs. En effet

$$[\partial_\xi, \partial_\eta] f = \partial_\xi \partial_\eta f - \partial_\eta \partial_\xi f = \partial_{L_\xi \eta} f \tag{5.14}$$

La démonstration de ce résultat est immédiate

$$\begin{aligned} [\partial_\xi, \partial_\eta] f &= \xi^j \partial_j (\eta^i \partial_i f) - \eta^j \partial_j (\xi^i \partial_i f) \\ &= (\xi^j \partial_j \eta^i)(\partial_i f) - (\eta^j \partial_j \xi^i)(\partial_i f) = \partial_{L_\xi \eta} f \end{aligned}$$

Autrement dit, $[\partial_\xi, \partial_\eta]$ est un opérateur différentiel du premier ordre. Le champ de vecteurs $L_\xi \eta = [\xi, \eta]$ est le *crochet de Lie* des champs de vecteurs ξ et η.

5.3 Connexions

5.3.1 Dérivée covariante

On pourrait penser qu'un objet comme $\partial_i T_j$ se comporte comme un tenseur (0,2) dans un changement de coordonnées, mais tel n'est pas le cas. En revanche, les termes indésirables, ceux qui ne correspondent

[2] Nous n'aurons pas l'occasion de nous servir de la différentielle extérieure, ni de la dérivée de Lie. Cependant ces deux notions donnent une bonne occasion de se familiariser avec les champs de vecteurs.

pas à une loi de transformation de type (0,2) disparaissent dans le cas de la différentielle extérieure $(\mathrm{d}T)_{ij}$, qui est bien un tenseur de type (0,2). De même $\partial_i T^j$ n'est pas un tenseur de type (1,1). On va donc chercher au lieu de la dérivée partielle ∂_i une opération ∇_i, qui, appliquée sur un vecteur T^j, donne bien un tenseur (1,1). Cette opération sera appelée la *dérivée covariante*. Nous allons exiger de cette opération les propriétés suivantes

1. C'est un opération linéaire.
2. Elle se réduit à la dérivée ordinaire pour une fonction scalaire f

$$\nabla_i f = \partial_i f \tag{5.15}$$

3. C'est une dérivation : elle obéit à la règle de Leibniz.

Soit T^j un vecteur (contravariant). Des propriétés (2) et (3) on tire

$$\partial_i(fT^j) - \nabla_i(fT^j) = f[\partial_i T^j - \nabla_i T^j]$$

Cette équation montre que la quantité $[\partial_i T^j - \nabla_i T^j]$ ne dépend que de la valeur de T^j au point P, $T^j(P)$. En effet, si les champs de vecteurs T'^j et T^j coïncident au point P : $T'^j(P) = T^j(P)$, on peut écrire un développement du type

$$T'^j - T^j = \sum_\alpha f^\alpha \omega^j_\alpha \quad \text{avec} \quad f^\alpha(P) = 0$$

On a donc

$$\partial_i(T'^j - T^j) - \nabla_i(T'^j - T^j) = \partial_i \left(\sum_\alpha f^\alpha \omega^j_\alpha \right) - \nabla_i \left(\sum_\alpha f^\alpha \omega^j_\alpha \right) \propto f^\alpha(P) = 0$$

Enfin, d'après la propriété (1), $[\partial_i T^j - \nabla_i T^j]$ est une combinaison linéaire des T^l, soit

$$\partial_i T^j - \nabla_i T^j = \Gamma^j_{li} T^l$$

La quantité Γ^j_{li} est un *coefficient de connexion*, ou simplement *connexion*, ou encore un symbole de Christoffel. Une même variété peut être munie de connexions différentes. La seule condition est que la connexion doit obéir aux propriétés (1−3). En résumé, l'action de la dérivée covariante sur un vecteur T^j est

$$\boxed{\nabla_i T^j = \partial_i T^j + \Gamma^j_{li} T^l} \tag{5.16}$$

L'action de la dérivée covariante sur un covecteur s'obtient aisément en remarquant que la quantité $T_i S^i$, construite à partir du covecteur T_i et du vecteur S^i est une fonction scalaire, comme on le vérifie à partir de (5.5) et (5.7)

$$T_i S^i = T'_i S'^i$$

et par conséquent

$$\nabla_i(T_j S^j) = \partial_j(T_j S^j)$$

On en déduit

$$\boxed{\nabla_i T_j = \partial_i T_j - \Gamma^l_{ij} T_l} \tag{5.17}$$

Attention cependant : contrairement à ce que la notation pourrait laisser croire, Γ^l_{ij} *n'est pas* un tenseur de type (1,2) ! La *torsion* Θ^l_{ij} est définie par

$$\Theta^l_{ij} = \Gamma^l_{ij} - \Gamma^l_{ji} = \Gamma^l_{[ij]} \tag{5.18}$$

où nous avons introduit une notation standard

$$A_{[ij]} = A_{ij} - A_{ji} \qquad A_{(ij)} = A_{ij} + A_{ji} \tag{5.19}$$

Contrairement au cas de la connexion, on peut montrer que la torsion est bien un tenseur de type (1,2) : comme dans le cas de différentielle extérieure, les termes indésirables disparaissent à cause de l'antisymétrisation dans (5.18).

L'action sur les vecteurs d'une base de coordonnées donne directement la connexion. En effet, d'après (5.16)

$$(\nabla_i e_m)^j = \partial_i e_m^j + \Gamma_{li}^j e_m^l = \Gamma_{mi}^j$$

car $e_m^j = \delta_m^j$. On peut donc écrire

$$\boxed{\nabla_i\, e_m = \Gamma_{mi}^l\, e_l} \tag{5.20}$$

5.3.2 Transport parallèle et géodésiques

La notion qui suit logiquement celle de connexion est celle de *transport parallèle*. Examinons d'abord cette notion dans le cas d'une fonction. Soit une courbe C paramétrée par $x^i(t)$. Si une fonction $f(x^i)$ est constante le long de la courbe C, on dira que la fonction f est transportée parallèlement le long de la courbe C

$$f(x^i(t+\mathrm{d}t)) \simeq f\left(x^i + \frac{\mathrm{d}x^i}{\mathrm{d}t}\,\mathrm{d}t\right) \simeq f(x^i) + \partial_i f \frac{\mathrm{d}x^i}{\mathrm{d}t}\,\mathrm{d}t$$

et $f(x^i(t)) = f(x^i(t+\mathrm{d}t))$ implique

$$\frac{\mathrm{d}x^i}{\mathrm{d}t}\,\partial_i f = 0 \tag{5.21}$$

En termes imagés : "le gradient $\partial_i f$ est perpendiculaire au vecteur vitesse". Un vecteur T^i sera transporté parallèlement le long de C si sa dérivée covariante est "perpendiculaire" au vecteur vitesse

$$0 = \frac{\mathrm{d}x^k}{\mathrm{d}t}\,\nabla_k T^i = \frac{\mathrm{d}x^k}{\mathrm{d}t}\,\left(\partial_k T^i + \Gamma_{jk}^i T^j\right)$$

soit

$$\boxed{\frac{\mathrm{d}T^i}{\mathrm{d}t} + u^k \Gamma_{jk}^i T^j = 0 \qquad u^k = \frac{\mathrm{d}x^k}{\mathrm{d}t}} \tag{5.22}$$

Il est essentiel de remarquer que l'équation (5.22) est une équation tensorielle admissible, valable dans tout système de coordonnées si elle est valable dans un système particulier, car

$$\frac{\mathrm{d}x^k}{\mathrm{d}t}\,\nabla_k T^i$$

est un vecteur, alors que

$$\frac{\mathrm{d}x^k}{\mathrm{d}t}\,\partial_k T^i = \frac{\mathrm{d}T^i}{\mathrm{d}t}$$

n'est pas un vecteur. Si $\mathrm{d}T^i/\mathrm{d}t = 0$ dans un système de coordonnées particulier, rien ne garantit qu'il en sera de même dans un autre système ; cette équation n'a pas une structure tensorielle admissible. Le transport parallèle dépend de la courbe (figure 5.3), mais non du système de coordonnées. Si l'on se donne une courbe C entre deux points A et B paramétrés par exemple par $t = 0$ et $t = 1$ et une condition initiale $T^i(t = 0)$, alors $T^i(t = 1)$ est déterminé de façon unique car (5.22) est un système d'équations différentielles du premier ordre.

La notion de *géodésique* est étroitement associée à celle de transport parallèle : une courbe G est une géodésique si son vecteur vitesse est transporté parallèlement à lui-même le long de G, soit (rappelons que $u^i = \mathrm{d}x^i/\mathrm{d}t$)

$$u^i \nabla_i u^j = u^i(\partial_i u^j + \Gamma_{ki}^j\, u^k) = 0$$

ce qui donne l'équation des géodésiques

$$\boxed{\frac{\mathrm{d}^2 x^j}{\mathrm{d}t^2} + \Gamma_{ki}^j\,\frac{\mathrm{d}x^i}{\mathrm{d}t}\,\frac{\mathrm{d}x^k}{\mathrm{d}t} = 0} \tag{5.23}$$

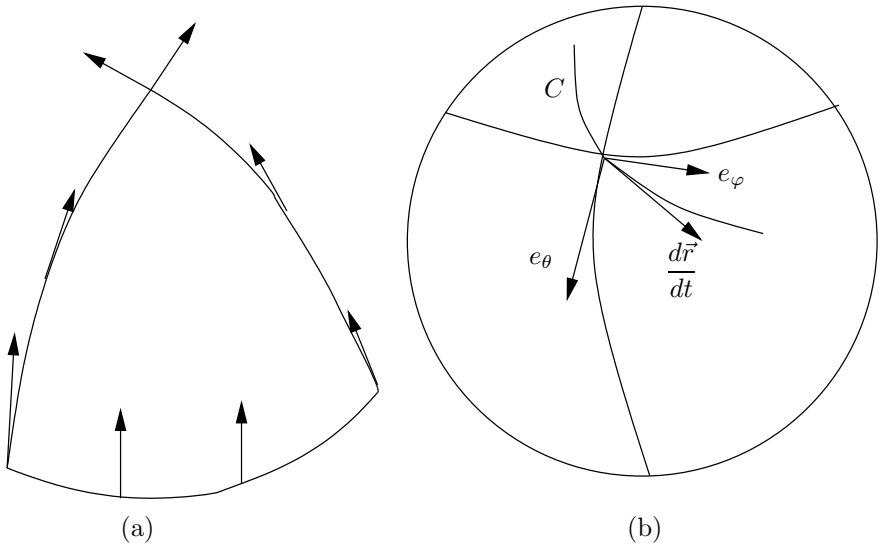

FIG. 5.3 – (a) Transport parallèle sur une sphère suivant deux courbes (des arcs de cercle) différentes. (b) Vecteurs e_θ et e_φ du plan tangent.

où nous avons utilisé

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\mathrm{d}x^i}{\mathrm{d}t}\,\partial_i \quad \text{soit} \quad \frac{\mathrm{d}x^i}{\mathrm{d}t}\,\partial_i\left(\frac{\mathrm{d}x^j}{\mathrm{d}t}\right) = \frac{\mathrm{d}^2x^j}{\mathrm{d}t^2}$$

Deux remarques sur les géodésiques.

1. Seule la partie symétrique dans les deux indices inférieurs de la connexion Γ intervient dans l'équation des géodésiques ; en d'autres termes, cette équation ne dépend pas de la torsion.
2. Une géodésique $x^i(t)$ est déterminée de façon unique par la donnée d'une position initiale $x^i(t=0)$ et d'une vitesse initiale $u^i(t=0)$. En effet (5.23) est un système d'équations différentielles du second ordre.

5.3.3 Exemple de la sphère

Nous allons illustrer les notions de connexion et de transport parallèle sur le cas familier de la sphère S^2, choisie de rayon unité. Dans un premier temps il sera commode de considérer S^2 comme une variété plongée dans l'espace $\mathbb{R}^3$. Le mouvement $\vec{r}(t) = (x(t), y(t), z(t))$ d'un point sur S^2 est repéré par exemple par ses coordonnées polaires $\theta(t)$ et $\varphi(t)$

$$\begin{aligned} x(t) &= \sin\theta(t)\cos\varphi(t) \\ y(t) &= \sin\theta(t)\sin\varphi(t) \\ z(t) &= \cos\theta(t) \end{aligned} \tag{5.24}$$

En différentiant ces équations par rapport à t on obtient le vecteur vitesse u

$$u = \dot{\theta}\, e_\theta + \dot{\varphi}\, e_\varphi \tag{5.25}$$

où les vecteurs e_θ et e_φ sont des vecteurs du plan tangent T_P à la sphère

$$e_\theta = (\cos\theta\cos\varphi, \cos\theta\sin\varphi, -\sin\theta) = \frac{\partial}{\partial\theta} \tag{5.26}$$

$$e_\varphi = (-\sin\theta\sin\varphi, \sin\theta\cos\varphi, 0) = \frac{\partial}{\partial\varphi} \tag{5.27}$$

Par convention, nous noterons les vecteurs sans flèche lorsqu'ils sont considérés comme des vecteurs du plan tangent : (e_θ, e_φ), et avec une flèche lorsqu'ils sont considérés comme vecteurs de $\mathbb{R}^3$: $(\vec{e}_\theta, \vec{e}_\varphi)$. Tout champ de vecteurs sur la sphère $X(\theta,\varphi)$ peut se décomposer sur les vecteurs de base e_θ et e_φ

$$X(\theta,\varphi) = X^\theta(\theta,\varphi)\, e_\theta + X^\varphi(\theta,\varphi)\, e_\varphi \tag{5.28}$$

Si l'on considère $\vec{e}_\theta$ et $\vec{e}_\varphi$ comme des vecteurs de $\mathbb{R}^3$, on peut les différentier par rapport à θ et φ ($\hat{r} = \vec{r}/r$)

$$\begin{aligned}
\partial_\theta\, \vec{e}_\theta &= (-\sin\theta\cos\varphi, -\sin\theta\sin\varphi, -\cos\theta) = -\hat{r}\\
\partial_\varphi \vec{e}_\theta &= (-\cos\theta\sin\varphi, \cos\theta\cos\varphi, 0) = \cot\theta\, \vec{e}_\varphi\\
\partial_\theta \vec{e}_\varphi &= (-\cos\theta\sin\varphi, \cos\theta\cos\varphi, 0) = \cot\theta\, \vec{e}_\varphi\\
\partial_\varphi \vec{e}_\varphi &= (-\sin\theta\cos\varphi, -\sin\theta\sin\varphi, 0) = -\sin^2\theta\, \hat{r} - \sin\theta\cos\theta\, \vec{e}_\theta
\end{aligned} \tag{5.29}$$

On constate que $\partial_\theta \vec{e}_\theta$ est orthogonal à T_P, que $\partial_\varphi \vec{e}_\theta$ et $\partial_\theta \vec{e}_\varphi$ sont contenus dans T_P et que $\partial_\varphi \vec{e}_\varphi$ possède une composante dans T_P et une composante orthogonale à T_P. Soit

$$\partial_X = X^\theta \partial_\theta + X^\varphi \partial_\varphi$$

la dérivé directionnelle suivant X. Si $Y(\theta,\varphi)$ est un autre champ de vecteurs, l'expression $(\vec{X}\cdot\vec{\partial})\vec{Y}$ ne définit pas un vecteur de T_P : c'est bien un vecteur de $\mathbb{R}^3$, mais ce n'est pas un vecteur de T_P. Pour obtenir un vecteur de T_P, il faut projeter sur ce plan à l'aide du projecteur $\mathcal{P}$

$$\mathcal{P}\left[(\vec{X}\cdot\vec{\partial})\vec{Y}\right] = (\vec{X}\cdot\vec{\nabla})\vec{Y}$$

et cette projection définit une dérivée covariante. En effet

- c'est une opération linéaire
- elle se réduit à la dérivée ordinaire pour une fonction
- c'est une dérivation

Pour faire le lien avec les notations utilisées jusqu'ici, on se donne la correspondance $\theta \to x^1$ et $\varphi \to x^2$: $e_\theta \to e_1$, $e_\varphi \to e_2$ De l'équation (5.20) $\nabla_i\, e_m = \Gamma^l_{mi}\, e_l$ on tire par identification avec (5.29)

$$\begin{aligned}
\nabla_\theta\, e_\theta &= 0 & \nabla_\varphi\, e_\theta &= \cot\theta\, e_\varphi\\
\nabla_\theta\, e_\varphi &= \cot\theta\, e_\varphi & \nabla_\varphi\, e_\varphi &= -\sin\theta\cos\theta\, e_\theta
\end{aligned} \tag{5.30}$$

d'où la connexion

$$\Gamma^\varphi_{\theta\varphi} = \Gamma^\varphi_{\varphi\theta} = \cot\theta \qquad \Gamma^\theta_{\varphi\varphi} = -\sin\theta\cos\theta \tag{5.31}$$

les autres composantes étant nulles. À partir de ce point on peut oublier que S^2 est plongée dans $\mathbb{R}^3$: les équations (5.31) définissent de façon *intrinsèque* une connexion sur S^2. Ceci donne l'équation des géodésiques

$$\begin{aligned}
\ddot{\theta} - \sin\theta\cos\theta\ \dot{\varphi}^2 &= 0\\
\ddot{\varphi} + 2\cot\theta\ \dot{\theta}\dot{\varphi} &= 0
\end{aligned} \tag{5.32}$$

C'est un exercice instructif, mais pas entièrement trivial, de montrer à partir de ces équations, que les géodésiques de la sphère sont des grands cercles parcourus à vitesse constante.

5.4 Métrique et courbure

5.4.1 Connexion associée à une métrique

Dans l'exemple de la sphère S^2 donné ci-dessus, on sait que la métrique euclidienne sur $\mathbb{R}^3$

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \tag{5.33}$$

induit sur la sphère unité une métrique

$$\mathrm{d}s^2 = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2 = g_{ij}(\theta,\varphi)\,\mathrm{d}x^i\mathrm{d}x^j \tag{5.34}$$

avec comme précédemment $x^1 = \theta$ et $x^2 = \varphi$; g_{ij} est appelé le *tenseur métrique*, et sous forme matricielle

$$g(\theta,\varphi) = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix} \tag{5.35}$$

La longueur d'un arc de courbe $C : (\theta(t), \varphi(t))$ sur la sphère est

$$\ell = \int_{t_A}^{t_B} \left[\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2 + \sin^2\theta \left(\frac{\mathrm{d}\varphi}{\mathrm{d}t}\right)^2 \right]^{1/2} \mathrm{d}t \tag{5.36}$$

Une métrique $g_{ij}(x)$ sur une variété M est un tenseur symétrique ($g_{ij} = g_{ji}$) (0,2) qui définit une forme bilinéaire $(\bullet,\bullet)$ sur les vecteurs : étant donné deux vecteurs ξ et η, cette forme bilinéaire est

$$(\xi,\eta) = g_{ij}(x)\,\xi^i\eta^j \tag{5.37}$$

De plus la matrice g_{ij} doit être une matrice positive. Si l'on sait munir la variété M d'une métrique, le tenseur métrique $g_{ij}(x)$ permet de calculer la longueur d'un arc de courbe sur la variété[3] par une généralisation immédiate de (5.36)

$$\ell = \int_{t_A}^{t_B} \left[g_{ij}(x) \frac{\mathrm{d}x^i}{\mathrm{d}t}\,\frac{\mathrm{d}x^j}{\mathrm{d}t} \right]^{1/2} \mathrm{d}t \tag{5.38}$$

Définissons le covecteur $\overline{\xi}$ associé au vecteur ξ par

$$\overline{\xi}_i = g_{ij}(x)\,\xi^j \tag{5.39}$$

Alors la forme bilinéaire (ξ,η) peut s'écrire

$$(\xi,\eta) =< \overline{\xi},\eta >= \overline{\xi}_i\eta^i$$

et peut être interprétée comme le produit scalaire des vecteurs ξ et η. On voit que le tenseur métrique permet de "descendre les indices" suivant (5.39), en associant par exemple un covecteur à un vecteur. On note $g^{ij}(x)$ la matrice inverse de $g_{ij}(x)$

$$g_{ij}(x)\,g^{jk}(x) = \delta_i^k$$

qui existe car g_{ij} est supposée définie positive. Le tenseur g^{ij} permet de "monter les indices", par exemple

$$\xi^i = g^{ij}(x)\,\overline{\xi}_j$$

Cette équation, tout comme (5.39), se généralisent trivialement à un tenseur quelconque, par exemple on passe d'un tenseur (2,0) à un tenseur (1,1) par

$$T^i_{\ j} = g_{jk}\,T^{ik}$$

La *trace* d'un tenseur (1,1) est définie par

$$\mathrm{Tr}\,T = T^i\,i = g_{ik}\,T^{ik} = g^{ik}\,T_{ik} \tag{5.40}$$

Les notions de connexion et de métrique sont *a priori* des notions indépendantes. Cependant, ainsi que nous allons le voir, on peut déduire d'une métrique une connexion *unique*, *la connexion associée à la métrique*, pourvu que l'on exige les deux conditions suivantes

[3] Nous nous plaçons pour l'instant dans le cas d'une métrique euclidienne, cas où la matrice g_{ij} est définie positive.

1. La torsion est nulle est dans une base de coordonnées : $\Gamma^i_{jk} = \Gamma^i_{kj}$
2. La dérivée covariante "tue la métrique" : $\nabla_k g_{ij} = 0$

En manipulant les indices dans l'équation $\nabla_k g_{ij} = 0$ et en utilisant la symétrie $\Gamma^i_{jk} = \Gamma^i_{kj}$, on trouve l'expression explicite de la connexion associée à la métrique

$$\Gamma^k_{ij} = \Gamma^k_{ji} = \frac{1}{2}\, g^{kl}\left(\partial_j\, g_{li} + \partial_i\, g_{lj} - \partial_l g_{ij}\right) \tag{5.41}$$

Le fait que l'on obtienne cette formule explicite montre que la connexion associée à la métrique est bien unique.

Pour illustrer ces concepts, revenons à l'exemple de la sphère où

$$\begin{aligned} \mathrm{d}s^2 &= \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2 \\ g_{\theta\theta} &= 1 \qquad g_{\theta\varphi} = g_{\varphi\theta} = 0 \qquad g_{\varphi\varphi} = \sin^2\theta \end{aligned}$$

On vérifie imédiatement que la connexion trouvée en (5.31) est symétrique et tue la métrique (5.34)

$$\nabla_i g_{jk} = 0$$

La connexion Γ (5.31) est bien la connexion associée à la métrique, comme on peut le vérifier explicitement en évaluant (5.41) dans le cas de la métrique de la sphère.

Terminons cette brève revue de la connexion associée à la métrique par l'énoncé de quelques propriétés utiles. Tout d'abord on remarque que si l'on utilise dans la dérivée covariante (5.16) la connexion associée à la métrique, alors cette dérivée commute avec l'opération de montée ou de descente des indices

$$\nabla_k T_i = \nabla_k(g_{ij}\, T^j) = g_{ij}(\nabla_k\, T^j) \tag{5.42}$$

où nous avons utilisé $\nabla_k g_{ij} = 0$. Ensuite, si $T^i(t)$ et $S^i(t)$ sont des vecteurs transportés parallèlement le long d'une courbe C, leur produit scalaire est invariant le long de cette courbe

$$u^k \nabla_k(g_{ij} T^i S^j) = u^k\left[g_{ij}(\nabla_k T^i) S^j + g_{ij} T^i (\nabla_k S^j)\right] = 0 \tag{5.43}$$

car par définition du transport parallèle

$$u^k \nabla_k T^i = u^k \nabla_k S^j = 0$$

La divergence (ordinaire) d'un champ de vecteurs $\partial_i T^i$ n'est pas un scalaire, mais $\nabla_i T^i$ *est* un scalaire

$$\begin{aligned} \nabla_i T^i &= \partial_i T^i + \Gamma^i_{ki} T^k \\ &= \frac{1}{\sqrt{g}}\,\partial_i(\sqrt{g}\; T^i) \end{aligned} \tag{5.44}$$

avec $g = \det g_{ij}$. Enfin une remarque importante concerne les géodésiques. Nous avons donné en (5.38) une expression de la longueur d'une courbe entre deux points de paramètres t_A et t_B. On peut montrer par des techniques standard de calcul variationnel que l'on retrouve l'équation (5.23) des géodésiques avec la connexion associée à la métrique (5.41) en cherchant la courbe qui minimise (plus généralement extrémise) la longueur ℓ entre deux points A et B fixés : $\delta\ell = 0$. Mais en fait l'équation (5.23) $u^k \nabla_k u^i = 0$ donne en plus la façon dont est parcourue la géodésique. Ceci s'explique par le fait que t est un paramètre affine (*cf.* (3.11). Si l'on choisit un paramètre non affine λ, alors $u^k \nabla_k u^i = \alpha\, u^i$, où α est une constante. On le voit dans le cas d'un mouvement rectiligne uniforme dans l'espace euclidien ordinaire, où les géodésiques sont des droites.

1. Géodésique avec paramètre affine : $\dfrac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2} = 0$
2. Géodésiques avec paramètre non affine : $\dfrac{\mathrm{d}^2\vec{r}}{\mathrm{d}\lambda^2} = \alpha\dfrac{\mathrm{d}\vec{r}}{\mathrm{d}\lambda}$

5.4.2 Tenseur de courbure

Il nous reste à définir la notion de *courbure*. À partir d'une connexion quelconque (non nécesairement associée à une métrique) et d'une base de coordonnées e_i on définit le tenseur de torsion Θ^i_{kl}

$$\nabla_k e_l - \nabla_l e_k = \Theta^i_{kl} e_i$$

et le *tenseur de courbure* ${R^i}_{qkl}$[4]

$$\boxed{[\nabla_k, \nabla_l] e_q = -{R^i}_{qkl}\, e_i} \tag{5.45}$$

La définition de la torsion suit de (**??**), compte tenu de (5.20). En utilisant une approche analogue à celle qui suit (5.15), on peut montrer que $[\nabla_k, \nabla_l] e_q$ est une fonction linéaire des e_i, ce qui justifie (5.45). Si la connexion est symétrique, une manipulation d'indices permet de calculer explicitement ${R^i}_{qkl}$

$${R^i}_{qkl} = -\partial_k \Gamma^i_{ql} + \partial_l \Gamma^i_{qk} - \Gamma^i_{pk}\Gamma^p_{ql} + \Gamma^i_{pl}\Gamma^p_{qk} \tag{5.46}$$

L'interprétation géométrique la plus parlante du tenseur de courbure est la suivante : considérons un covecteur T_i transporté parallèlement le long d'une courbe fermée C. Après avoir parcouru une fois la courbe fermée, le covecteur T_i est différent du covecteur initial par ΔT_i, qui vaut

$$\Delta T_i = \frac{1}{2}\, {R^j}_{ikl} T_j \oint x^k \mathrm{d}x^l \tag{5.47}$$

Si un vecteur est transporté parallèlement à lui-même le long d'une courbe fermée et qu'il ne coïncide pas avec le vecteur initial, la courbure est nécessairement $\neq 0$

En relativité générale, la situation est très souvent la suivante. On dispose d'une métrique g, à partir de laquelle on calcule la connexion associée à la métrique via (5.41), puis le tenseur de courbure via (5.46). Le schéma est donc le suivant[5]

métrique → connexion → courbure

Les expressions (5.41) et (5.46) montrent que le tenseur de courbure dépend *non linéairement* de la métrique. Le tenseur ${R^i}_{qkl}$ vérifie plusieurs relations de symétrie, par exemple

$${R^i}_{qkl} = -{R^i}_{qlk}$$

évidente d'après (5.45). En raison de ces propriétés de symétrie, on montre que dans une variété de dimension N le tenseur de courbure possède

$$d_N = \frac{1}{12}\, N^2 (N^2 - 1)$$

[4]Il existe une correspondance intéressante avec les théories de jauge, abéliennes et non abéliennes. Pour simplifier la discussion, je me limiterai aux théories abéliennes dans un situation indépendante du temps. Dans le formalisme de l'intégrale de chemin, le poids statistique d'une trajectoire d'une particule chargée de charge q est donnée par

$$\exp\left(-\mathrm{i}\frac{q}{\hbar}\int_{1(C)}^{2} \mathrm{d}\vec{r}\cdot\vec{A}(\vec{r})\right)$$

où $\vec{A}$ est le potentiel vecteur. Cette expression peut être interprétée comme le transport parallèle de la fonction d'onde du point 1 au point 2 en suivant la courbe C. L'analogue de la courbure est le champ magnétique

$$\vec{B} = \vec{\nabla} \times \vec{A}$$

et si la courbure est non nulle, le transport parallèle le long d'une courbe fermée est non trivial

$$\oint \vec{A}\cdot \mathrm{d}\vec{r} = \int\int \vec{B}\cdot \mathrm{d}\vec{S}$$

d'après le théorème de Stokes. L'analogue de la dérivé covariante est

$$\vec{D} = \vec{\nabla} - \mathrm{i}\frac{q}{\hbar}\,\vec{A}$$

.

[5]Les calculs explicites de (5.41) et (5.45) étant donné une métrique sont en général longs et pénibles et ils ne sont pas instructifs. Ils servent juste à se convaincre que l'on serait encore capable de passer le concours de l'Ecole Polytechnique. Il existe des programme *Mathematica* pour effectuer ces calculs de façon automatique (voir le site WEB du livre de Hartle).

composantes indépendantes, soit

$$\begin{aligned} N &= 2 \qquad d_2 = 1 \\ N &= 3 \qquad d_3 = 6 \\ N &= 4 \qquad d_4 = 20 \end{aligned}$$

Le tenseur de courbure obéit aussi à *l'identité de Bianchi*, qui est en fait reliée à l'identité de Jacobi

$$\nabla_{[i} R_{jk]lq} = 0 \tag{5.48}$$

où [ ] indique l'antisymétrisation comme dans (5.19). À partir du tenseur de courbure on construit le *tenseur de Ricci* R_{ij}, qui est un tenseur symétrique

$$R_{ij} = R^q{}_{iqj} = R_{ji} = \partial_j \Gamma^k_{ik} - \partial_k \Gamma^k_{ij} + \Gamma^k_{lj}\Gamma^l_{ik} - \Gamma^k_{lk}\Gamma^l_{ij} \tag{5.49}$$

et la courbure $R = R^i{}_i$, qui est la trace du tenseur de Ricci, et est donc un scalaire. Le tenseur symétrique G_{ij}

$$G_{ij} = R_{ij} - \frac{1}{2} R\, g_{ij} \tag{5.50}$$

vérifie, en raison de l'identité de Bianchi, l'identité importante

$$\boxed{\nabla^i G_{ij} = 0} \tag{5.51}$$

Pour $N = 2$ et $N = 3$, le tenseur de courbure peut être construit à partir du tenseur de Ricci et de la métrique, et si $R_{ij} = 0$, le tenseur de courbure s'annule. Nous verrons que les équations d'Einstein impliquent que le tenseur de Ricci s'annule en l'absence de matière, et donc la courbure doit s'annuler dans le vide en dimension 2 et en dimension 3, ce qui entraîne l'absence de forces de gravitation : selon Einstein, la gravitation ne peut exister que pour des dimensions d'espace-temps ≥ 4 !

5.5 Adaptation à la relativité générale

La dimension de l'espace-temps étant $N = 4$, la relativité générale utilisera une variété à quatre dimensions. Un point de cette variété sera repéré par quatre coordonnées x^μ, $\mu = (0, 1, 2, 3)$. Comme dans le chapitre 3, $x^0 = 0$ est une coordonnée de type temps, et x^i une coordonnée de type espace. Cependant on doit tenir compte de ce que la métrique $g_{\mu\nu}$ n'est pas définie positive, mais que la matrice $g_{\mu\nu}$, tout comme $\eta_{\mu\nu}$, a une valeur propre positive et trois valeurs propres négatives. Il est facile de se convaincre que le nombre de valeurs propres négatives est inchangé dans un changement de coordonnées (5.5), et la seule conséquence pratique est que, tout comme dans l'espace de Minkowski (*cf.* (3.7)) , la quantité $\mathrm{d}s^2$ ("l'élément de longueur")

$$\mathrm{d}s^2 = g_{\mu\nu}\, x^\mu\, x^\nu \tag{5.52}$$

peut être positive, négative ou nulle[6]. Comme la matrice $g_{\mu\nu}$ est symétrique, on peut la diagonaliser par une transformation orthogonale en un point donné x^μ_0 ; évidemment la diagonalisation ne sera en général pas valable en un autre point. On peut ensuite effectuer une dilatation sur chacune des quatre coordonnés de façon à se ramener à la situation de l'espace- temps plat : $g_{\mu\nu}(x_0) \to \eta_{\mu\nu}$. Un nouveau changement de coordonnées permet d'éliminer les termes linéaires $\mathcal{O}(x^\mu - x^\mu_0)$ avec pour résultat

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + \mathcal{O}(x^\mu - x^\mu_0)^2 \tag{5.53}$$

D'après (5.41) la connexion Γ s'annule en x_0

$$\Gamma^\mu_{\rho\sigma}(x_0) = 0 \tag{5.54}$$

[6] On doit remplacer dans (5.44) $\sqrt{g}$ par $\sqrt{|g|}$.

L'équation (5.53) (ou (5.54)) définit un *référentiel d'inertie local* (RIL) : c'est le référentiel en chute libre du chapitre 2. On peut associer à ce RIL une base pseudo-orthogonale comme celle de l'espace-temps plat. Cette base ne doit pas être confondue avec une base de coordonnées.

Dans le RIL, l'équation d'une géodésique est *localement* celle d'un mouvement rectiligne uniforme (*cf.* (3.11))

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = 0 \quad \text{ou} \quad \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = 0 \tag{5.55}$$

où u^μ est la quadrivitesse et τ le temps propre, mais (5.55) ne peut pas être valable partout : ce n'est pas une équation tensoriellement admissible. L'équation tensoriellement admissible qui se réduit à (5.55) est l'équation d'une géodésique

$$u^\nu \nabla_\nu u^\mu = \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu_{\rho\sigma} \frac{\mathrm{d}u^\rho}{\mathrm{d}\tau} \frac{\mathrm{d}u^\sigma}{\mathrm{d}\tau} = 0 \tag{5.56}$$

Les particules, massives ou non, suivent des géodésiques de l'espace-temps : (5.56) se réduit à (5.55) dans un RIL.

Une dernière remarque concerne la loi de conservation du tenseur énergie-impulsion. L'équation de conservation (3.18) $\partial_\mu T^{\mu\nu} = 0$ n'est pas une équation tensoriellement admissible. L'équation tensoriellement admissible qui se réduit à celle-ci dans un RIL est

$$\boxed{\nabla_\mu T^{\mu\nu} = \partial_\mu T^{\mu\nu} + \Gamma^\mu_{\sigma\mu} T^{\sigma\nu} + \Gamma^\nu_{\sigma\mu} T^{\mu\sigma} = 0} \tag{5.57}$$

Bibliographie.

Tous les livres de relativité générale contiennent une introduction à la géométrie différentielle, par exemple Hartle [2003], chapitres 7, 8 et 20. À mon avis aucune de ces introductions ne vaut l'exposé de Doubrovine *et al.* [1983], chapitres 3 et 4.

Chapitre 6

Solutions à symétrie sphérique

6.1 Équation d'Einstein

Le principe qui sous-tend l'équation d'Einstein consiste à écrire une relation entre la géométrie et l'énergie-impulsion. On ne peut pas démontrer l'équation d'Einstein, tout comme on ne peut pas démontrer les équations de Maxwell ou de Newton. Toutefois on peut argumenter de la façon suivante.

1. C'est l'équation la plus simple possible satisfaisant au principe précédent.
2. Elle est mathématiquement cohérente et définit un problème de valeurs initiales.
3. Elle redonne l'équation de Newton dans une limite appropriée.

L'équation la plus simple possible est la suivante

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}\, R g_{\mu\nu} = -\kappa T_{\mu\nu} \tag{6.1}$$

où κ est une constante de proportionnalité à déterminer. En effet, si l'on veut avoir dans le membre de droite le tenseur énergie-impulsion, le membre de gauche doit aussi être un tenseur symétrique d'ordre deux construit avec la courbure. Les deux tenseurs symétriques les plus simples à notre disposition sont le tenseur de Ricci $R_{\mu\nu}$ et le tenseur métrique $g_{\mu\nu}$. Comme le tenseur énergie-impulsion obéit à l'équation de conservation

$$\nabla^\mu T_{\mu\nu} = 0$$

il doit en être de même du membre de gauche de (6.1), et compte tenu de l'identité de Bianchi (5.51), la seule combinaison possible de $R_{\mu\nu}$ et de $g_{\mu\nu}$ convenable est précisément $G_{\mu\nu}$. Prenant la trace de (6.1) on déduit la courbure scalaire sous la forme

$$R = -\kappa\, T^\mu{}_\mu = -\kappa T$$

ce qui permet de récrire (6.1) sous la forme souvent utile

$$R_{\mu\nu} = -\kappa \left(T_{\mu\nu} - \frac{1}{2}\, T\, g_{\mu\nu} \right) \tag{6.2}$$

Sous cette forme, on voit immédiatement qu'en l'absence de matière ($T_{\mu\nu} = 0$), le tenseur de Ricci et la courbure sont nuls. Dans un espace-temps de dimension ≤ 3, ceci implique que le tenseur de courbure est nul, et il ne peut y avoir de gravitation dans un sens usuel que pour une dimension ≥ 4 !

Il reste à fixer la constante κ. Pour ce faire, nous allons nous placer en champ gravitationnel faible indépendant du temps et considérer une situation où les vitesses des particules contribuant à $T_{\mu\nu}$ sont petites par rapport à c. Dans cette limite $T_{00} \simeq T \simeq \rho$, où ρ est la densité de matière, et toutes les autres composantes du tenseur sont négligeables par rapport à T_{00}. L'hypothèse du champ faible permet d'écrire $g_{00} = 1 + h_{00}$, avec $|h_{00}| \ll 1$, et la composante (00) de (6.2) devient

$$R_{00} = -\frac{1}{2}\, \kappa T_{00}$$

Nous devons évaluer la composante R_{00} du tenseur de Ricci à partir de la métrique. D'après (5.49)

$$R_{00} = \partial_0 \Gamma^i_{0i} - \partial_i \Gamma^i_{00} + \mathcal{O}(\Gamma^2)$$

et les termes $\mathcal{O}(\Gamma^2)$ peuvent être négligés car ils sont d'ordre $(h_{00})^2$. La contribution de ∂_0 s'annule dans une situation stationnaire et il ne reste que $-\partial_i \Gamma^i_{00}$. En utilisant (5.41) on évalue Γ^i_{00}

$$\Gamma^i_{00} = \frac{1}{2}\, g^{il}(\partial_0 g_{l0} + \partial_0 g_{0l} - \partial_l g_{00}) \simeq -\frac{1}{2}\, \eta^{il} \partial_l h_{00}$$

où η^{il} représente les composantes spatiales du tenseur de Minkowski (rappelons que les lettres latines vont de 1 à 3 et étiquettent les composantes spatiales). On en déduit

$$R_{00} = -\frac{1}{2}\, \eta^{ij} \partial_i \partial_j\, h_{00} = \frac{1}{2}\, \nabla^2\, h_{00}$$

soit

$$\nabla^2 h_{00} = -\kappa T_{00} = -\kappa \rho$$

Mais nous avons vu au chapitre 1 que $h_{00} = 2\Phi(\vec{r})$, où Φ est le potentiel gravitationnel, et donc

$$\nabla^2 \Phi(\vec{r}) = \frac{\kappa}{2}\, \rho = 4\pi G \rho$$

d'après l'équation de Poisson pour Φ. Nous pouvons donc faire l'identification $\kappa = 8\pi G$, ce qui donne pour (6.1)

$$\boxed{G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}\, R g_{\mu\nu} = -8\pi G\, T_{\mu\nu}} \tag{6.3}$$

L'équation d'Einstein peut être vue comme un ensemble d'équations aux dérivées partielles non linéaires et du second ordre pour la métrique $g_{\mu\nu}$. Les dix équations sont en fait réduites à 6 équations indépendantes en raison de l'identité de Bianchi $\nabla^\mu G_{\mu\nu} = 0$. En principe, si l'on se donne $g_{\mu\nu}$ et ses dérivées premières par rapport au temps sur une surface du genre espace (par exemple sur une surface $t = t_0 =$cste), on doit pouvoir calculer $g_{\mu\nu}(x)$ pour tout temps $t > t_0$. En fait la situation est plus complexe en raison de la possibilité de transformations de jauge, et en pratique on essaiera plutôt de construire une métrique compatible avec les symétries du problème.

Afin de discuter commodément des généralisations possibles de l'équation d'Einstein (6.3), il est utile d'introduire *l'action d'Einstein-Hilbert* S_H : en effet il est possible de déduire (6.1) d'un principe de moindre action, en définissant

$$S_H = \int \mathrm{d}^4 x\, \sqrt{|g|}\, R \tag{6.4}$$

Un calcul un peu long montre alors que

$$\frac{1}{\sqrt{|g|}}\, \frac{\delta S_H}{\delta g_{\mu\nu}} = -G_{\mu\nu} \tag{6.5}$$

Pour obtenir le second membre de (6.1), il faut se donner une action S_M pour la matière, et *par définition* le tenseur énergie-impulsion est

$$T^{\mu\nu} = \frac{1}{\sqrt{|g|}}\, \frac{\delta S_M}{\delta g_{\mu\nu}} \tag{6.6}$$

Par exemple S_M pourrait être l'action d'un champ scalaire. Dans ces conditions l'équation d'Einstein se déduit du principe de moindre action

$$\frac{\delta}{\delta g_{\mu\nu}} \left(\frac{1}{8\pi G}\, S_H + S_M \right) = 0 \tag{6.7}$$

On peut géneraliser (6.1) en ajoutant un terme de constante cosmologique

$$G_{\mu\nu} = -8\pi G T_{\mu\nu} - \Lambda g_{\mu\nu} \tag{6.8}$$

Ceci est bien compatible avec la structure tensorielle, mais ne redonne pas la gravité newtonienne dans les limites convenables. On peut réinterpréter cette constante cosmologique comme une contribution à $T_{\mu\nu}$ provenant de l'énergie du vide[1]. Une autre généralisation possible consiste à ajouter à S_H des dérivées de $g_{\mu\nu}$ d'ordre plus élevé que deux, par exemple en introduisant des termes non linéaires en $R_{\mu\nu}$

$$S_H = \int \mathrm{d}^4x \sqrt{|g|} \left(R + \alpha_1 R^2 + \alpha_2 R_{\mu\nu}R^{\mu\nu} + \ldots\right)$$

ou à ajouter un mélange avec un champ scalaire λ couplé à la courbure, comme dans la théorie de Brans-Dicke

$$S_{\mathrm{BD}} = \int \mathrm{d}^4x \sqrt{|g|} \left[f(\lambda R) - \frac{1}{2} g^{\mu\nu}\partial_\mu\lambda\partial_\nu\lambda - V(\lambda)\right]$$

.

À ce point, il vaut la peine de bien caractériser la relativité générale par rapport à la relativité restreinte. Cette caractérisation est parfaitement résumée par T. Damour (Damour [2005]). "Le principe de relativité générale a un statut physique différent du principe de relativité restreinte. Le principe de relativité restreinte est un principe de symétrie de la structure de l'espace-temps qui affirme que la physique est *la même* dans une classe particulière de référentiels, et que donc certains phénomènes "correspondants" se déroulent de la même façon dans des référentiels différents ("transformations actives"). En revanche, le principe de relativité générale est un *principe d'indifférence* : les phénomènes ne se déroulent (en général) pas de la même façon dans des systèmes de coordonnées différents, mais aucun des systèmes de coordonnées (étendu) n'a de statut privilégié par rapport aux autres."

Nous n'insisterons pas sur les vérifications de la relativité générale et renvoyons à la revue récente de T. Damour (Damour [2005]) : en résumé, la relativité générale a été testée aujourd'hui dans des situations très diverses, depuis notre environnement immédiat (système GPS) jusqu'aux confins de l'Univers (lentilles gravitationnelles). Les tests les plus précis atteignent une précision relative de 10^{-5}.

6.2 Métrique de Schwarzschild

Lorsqu'une métrique est invariante dans un changement de coordonnées particulier, on dit que l'on a une symétrie de la métrique. Dans le cas simple où la métrique est invariante par une translation, par exemple $x^1 \to x^1 + a$, on associe à cette invariance un *vecteur de Killing*

$$\xi = (0, 1, 0, 0) \tag{6.9}$$

Les conséquences d'une symétrie de la métrique se déduisent commodément de l'étude des vecteurs de Killing. En particulier on montre que $\xi \cdot u$ est invariant le long d'une géodésique, $u^\mu = \mathrm{d}x^\mu/\mathrm{d}\tau$ étant la quadrivitesse le long de cette géodésique.

Nous allons nous intéresser au cas de la *symétrie sphérique*. Il est intuitif, mais long à montrer rigoureusement, que la partie spatiale de la métrique $g_{ij} \propto \delta_{ij}$. En effet ceci assure que la partie spatiale du $\mathrm{d}s^2$ a la forme

$$\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2$$

les termes tels que $\mathrm{d}x\mathrm{d}y$ étant incompatibles avec la symétrie sphérique. Écrivant

$$\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 = r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2) = r^2\mathrm{d}\Omega^2$$

la forme la plus générale de la métrique compatible avec la symétrie sphérique est, en fonction de deux coordonnées a et r

$$\mathrm{d}s^2 = g_{aa}(a,r)\mathrm{d}a^2 + 2g_{ar}\mathrm{d}a\mathrm{d}r - g_{rr}(a,r)\mathrm{d}r^2 - r^2\mathrm{d}\Omega^2 \tag{6.10}$$

Des termes croisés (a,θ), (a,φ), (r,θ) et (r,φ) sont exclus par la symétrie sphérique. On peut donc se limiter à l'étude des coordonnées (a,r). On effectue un changement de coordonnées $a \to t(a,r)$

$$\mathrm{d}t = \frac{\partial t}{\partial a}\,\mathrm{d}a + \frac{\partial t}{\partial r}\,\mathrm{d}r$$

[1] Voir cependant la note 4 du chapitre 4.

et on souhaiterait avoir

$$\mathrm{d}s^2 = A(t,r)\mathrm{d}t^2 - B(t,r)\mathrm{d}r^2 - r^2\mathrm{d}\Omega^2$$

ce qui donne 3 équations pour 3 fonctions inconnues, t, A et B. On peut donc toujours mettre la métrique sous la forme

$$\mathrm{d}s^2 = \mathrm{e}^{2\alpha(r,t)}\,\mathrm{d}t^2 - \mathrm{e}^{2\beta(r,t)}\,\mathrm{d}r^2 - r^2\,\mathrm{d}\Omega^2 \tag{6.11}$$

Il faut évidemment prendre garde au fait que la relation entre t et le temps effectivement mesuré par un observateur est indirecte (*cf.* (6.15)), et de même r n'est pas la distance au centre de symétrie. En revanche la surface d'une sphère de rayon r est bien $4\pi r^2$. On va s'intéresser pour l'instant uniquement à la région extérieure à la source du champ de gravitation, où le tenseur de Ricci $R_{\mu\nu} = 0$ d'après (6.3). Notre programme consiste maintenant à calculer le tenseur de Ricci à partir de la métrique (6.11) et à l'annuler. On trouve à partir de (5.46)

$$\begin{aligned} R_{tr} &= \frac{2}{r}\,\partial_t\beta = 0 \\ R_{\theta\theta} &= \mathrm{e}^{-2\beta}\Big[r(\partial_r\beta - \partial_r\alpha) - 1\Big] + 1 = 0 \end{aligned}$$

On en déduit que β est seulement une fonction de r, $\beta(r)$; de plus $\partial_t R_{\theta\theta} = 0 \Longrightarrow \partial_t\partial_r\alpha = 0$ soit

$$\alpha(r,t) = f(r) + g(t)$$

ce qui donne pour la métrique

$$\mathrm{d}s^2 = \mathrm{e}^{2f(r)}\,\mathrm{e}^{2g(t)}\,\mathrm{d}t^2 - \mathrm{e}^{2\beta(r)}\mathrm{d}r^2 - r^2\mathrm{d}\Omega^2 \tag{6.12}$$

On effectue un changement de variables $t \to t'$ tel que

$$\mathrm{d}t' = \mathrm{e}^{g(t)}\mathrm{d}t$$

et en réétiquetant $t' \to t$ on obtient le *théorème de Birkhoff*

$$\mathrm{d}s^2 = \mathrm{e}^{2\alpha(r)}\,\mathrm{d}t^2 - \mathrm{e}^{2\beta(r)}\mathrm{d}r^2 - r^2\mathrm{d}\Omega^2 \tag{6.13}$$

En mots, le théorème de Birkhoff nous dit que dans le cas d'une symétrie sphérique, *la métrique est stationnaire*, ce qui implique l'existence d'un vecteur de Killing $\xi = (1,0,0,0)$. Une distribution de matière à symétrie sphérique ne rayonne pas d'ondes gravitationnelles, de même qu'en électromagnétisme une distribution de charges à symétrie sphérique ne rayonne pas : il n'existe pas de rayonnement monopolaire !

Pour finir de déterminer la métrique, examinons la combinaison suivante de R_{tt} et de R_{rr}

$$\mathrm{e}^{2(\beta-\alpha)}R_{tt} - R_{rr} = \frac{2}{r}(\partial_r\alpha + \partial_r\beta) = 0$$

d'où $\alpha = -\beta + \text{cste}$. Enfin $R_{\theta\theta} = 0$ devient, compte tenu des résultats précédents

$$\mathrm{e}^{2\alpha}(2r\partial_r\alpha + 1) = 1 \Longrightarrow \partial_r\left(r\mathrm{e}^{2\alpha}\right) = 1 \Longrightarrow \mathrm{e}^{2\alpha} = 1 + \frac{C}{r}$$

Ceci conduit à la forme suivante de la métrique

$$\mathrm{d}s^2 = \left(1 + \frac{C}{r}\right)\mathrm{d}t^2 - \left(1 + \frac{C}{r}\right)^{-1}\mathrm{d}r^2 - r^2\mathrm{d}\Omega^2$$

Pour déterminer la constante C, il reste à examiner la limite de champ faible où

$$g_{tt} \simeq 1 + 2\Phi(r) = 1 - \frac{2GM}{r}$$

ce qui donne la forme finale de la métrique se Schwarzschild

$$\boxed{\mathrm{d}s^2 = \left(1 - \frac{2GM}{r}\right)\mathrm{d}t^2 - \left(1 - \frac{2GM}{r}\right)^{-1}\mathrm{d}r^2 - r^2\mathrm{d}\Omega^2} \tag{6.14}$$

La métrique de Schwarzschild donne accès à tous les résultats classiques de la relativité générale

- déviation de la lumière par une masse ;
- précession du périhélie de Mercure ;
- décalage vers le rouge gravitationnel ;
- retard de l'écho radar
- *etc.*

Il faut calculer les géodésiques et utiliser le fait que $\xi \cdot u$ est une constante le long de ces géodésiques, si ξ est un vecteur de Killing. Les calculs exacts n'étant pas possibles en général, on doit effectuer un développement en puissances de v/c appelé approximation post-newtonienne.

6.3 Trous noirs

La métrique de Schwarzschild semble singulière lorsque la coordonnée $r = r_S = 2GM$ ($r_S = 2GM/c^2$ si l'on rétablit c) ; r_S est apelé le *rayon de Schwarzschild*. Dans le cas du Soleil $r_S \simeq 3\,\mathrm{km}$, ce qui est évidemment négligeable par rapport au rayon du Soleil, que l'on peut donc assimiler en pratique à une masse ponctuelle. En revanche le rayon d'une étoile à neutrons est du même ordre de grandeur que son rayon de Schwarzschild[2], et la question de savoir ce qui se passe lorsque $r \to r_S$ est intéressante. Le calcul de la courbure scalaire R pour $r = r_S$ donne un résultat fini, ce qui suggère que la singularité de la métrique de Schwarzschild à $r = r_S$ est un artefact du système de coordonnées, ce que nous allons montrer en exhibant un système de coordonnées manifestement non singulier à $r = r_S$. Nous allons toujours nous placer à (θ, φ) fixés, et prendre en compte uniquement les coordonnées t et r. Examinons d'abord le décalage vers le rouge gravitationnel (figure 6.1). L'observateur O_1 en r_1 émet des signaux à des intervalles Δt réguliers dans une direction fixée ($\mathrm{d}\theta = \mathrm{d}\varphi = 0$), mais pour lui les intervalles de temps propre entre l'émission de deux signaux consécutifs sont

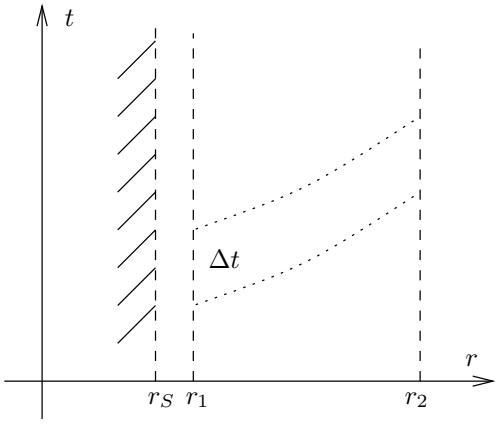

FIG. 6.1 – Décalage vers le rouge gravitationnel. La trajectoire des photons est en pointillés.

$$\Delta\tau_1 = \left(1 - \frac{2GM}{r_1}\right)\Delta t \tag{6.15}$$

Compte tenu d'une relation similaire entre Δt et les intervalles de temps propre de l'observateur O_2 en r_2 entre la réception de deux signaux consécutifs, on obtient pour le rapport entre fréquences reçues et émises

$$\frac{\omega_2}{\omega_1} = \frac{\Delta\tau_1}{\Delta\tau_2} = \left(\frac{1 - 2GM/r_1}{1 - 2GM/r_2}\right)^{1/2} \xrightarrow[r_2 \gg r_S]{} \left(1 - \frac{2GM}{r_1}\right)^{1/2}$$

[2] $r_S \simeq 0.4r$ si $r \simeq 10\,\mathrm{km}$ est le rayon d'une étoile à neutrons dont la masse est égale à $\simeq 1.5$ fois celle du Soleil.

et pour un observateur O_2 à l'infini, avec $r_1 = r_{\rm em}$

$$\boxed{\omega_\infty = \omega_{\rm em}\left(1 - \frac{2GM}{r_{\rm em}}\right)^{1/2}} \tag{6.16}$$

Lorsque $r_{\rm em} \to r_S$, le rythme d'émission des signaux et la fréquence des photons tendent vers zéro pour un observateur à l'infini, et bientôt il ne verra plus rien : au fur et à mesure que $r_{\rm em} \to r_S$, le rythme des signaux est de plus en plus ralenti et l'énergie des photons devient de plus en plus petite. Il est aussi intéressant d'examiner le cône de lumière

$$\frac{{\rm d}t}{{\rm d}r} = \pm\left(1 - \frac{2GM}{r}\right)^{-1} \tag{6.17}$$

ce qui montre que l'angle au sommet du cône devient de plus en plus aigu quand $r \to r_S$ (figure 6.2).

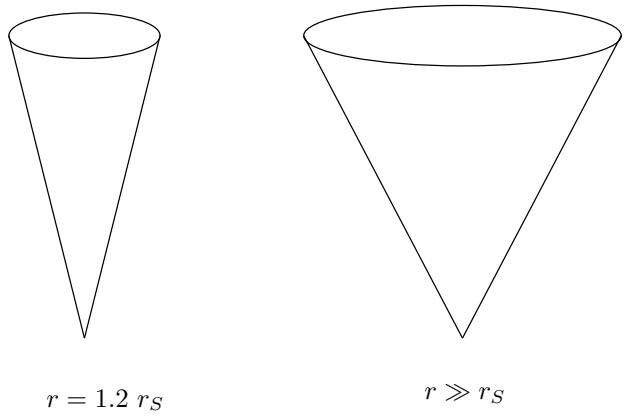

FIG. 6.2 – Évolution des cônes de lumière dans les coordonnées (t, r).

Pour comprendre ce qui se passe lorsque $r < r_S$, il faut faire appel à un autre système de coordonnées. Le plus simple est sans doute celui d'Eddington-Finkelstein (EF) où $(t, r) \to (v, r)$, avec

$$\boxed{t = v - r - 2GM \ln\left|\frac{r}{2GM} - 1\right| = v - r - r_S \ln\left|\frac{r}{r_S} - 1\right|} \tag{6.18}$$

À partir de (6.18) on obtient immédiatement

$${\rm d}t = {\rm d}v - \frac{1}{1 - r_S/r}$$

d'où le ${\rm d}s^2$

$$\begin{aligned} {\rm d}s^2 &= \left(1 - \frac{r_S}{r}\right)\left({\rm d}v - \frac{1}{1 - r_S/r}\right)^2 - \left(1 - \frac{r_S}{r}\right)^{-1} {\rm d}r^2 \\ &= \left(1 - \frac{r_S}{r}\right){\rm d}v^2 + 2{\rm d}v{\rm d}r - r^2{\rm d}\Omega^2 \end{aligned} \tag{6.19}$$

On constate que la métrique n'est plus singulière pour $r = r_S$! De plus le déterminant g de la métrique vaut

$$g|_{r=r_S} = -r_S^4 \sin^2\theta$$

et g^{-1} existe pour $r = r_S$. Étudions le cône de lumière ; il faut résoudre l'équation ${\rm d}s^2 = 0$

$${\rm d}s^2 = \left(1 - \frac{r_S}{r}\right){\rm d}v^2 + 2{\rm d}v{\rm d}r = 0 \tag{6.20}$$

Il y a deux possibilités

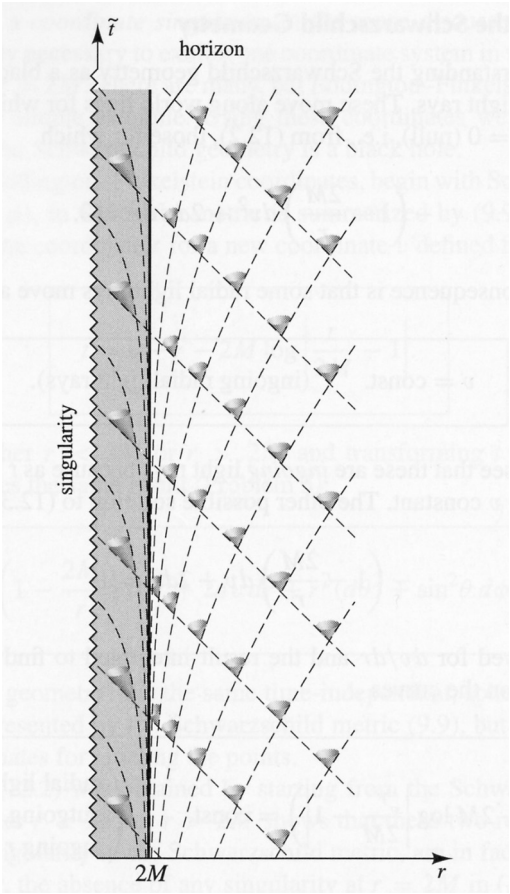

FIG. 6.3 – Les rayons lumineux entrant et sortants dans les variables $(\tilde{t}, r)$. D'après Hartle [2003].

1. $\mathrm{d}v = 0$, $v = \text{cste}$. Pour $r \gg r_S$, cela correspond à $t + r = \text{cste}$, c'est-à-dire à des rayons lumineux entrants : r décroît si t croît.
2. $$\left(1 - \frac{r_S}{r}\right) \mathrm{d}v + 2\mathrm{d}r = 0$$

   soit

   $$v - 2\left(r + r_S \ln\left|\frac{r}{r_S} - 1\right|\right) = \text{cste}$$

   Si $r \gg r_S$

   $$v - 2r \simeq t + r - 2r = t - r = \text{cste}$$

   ce qui correspond à des rayons lumineux sortants : r croît si t croît.

Pour la discussion générale, il est commode d'introduire la variable $\tilde{t} = v - r$. Lorsque l'on se trouve dans le cas (1), on aura $\tilde{t} + r = \text{cste}$, ce qui correspond à des droites entrantes pour $r > r_S$ aussi bien que pour $r < r_S$. Dans le cas (2)

$$\tilde{t} = \text{cste} - r + 2\left(r + r_S \ln\left|\frac{r}{r_S} - 1\right|\right)$$

soit

$$\frac{\mathrm{d}\tilde{t}}{\mathrm{d}r} = 1 + \frac{2}{r/r_S - 1} \tag{6.21}$$

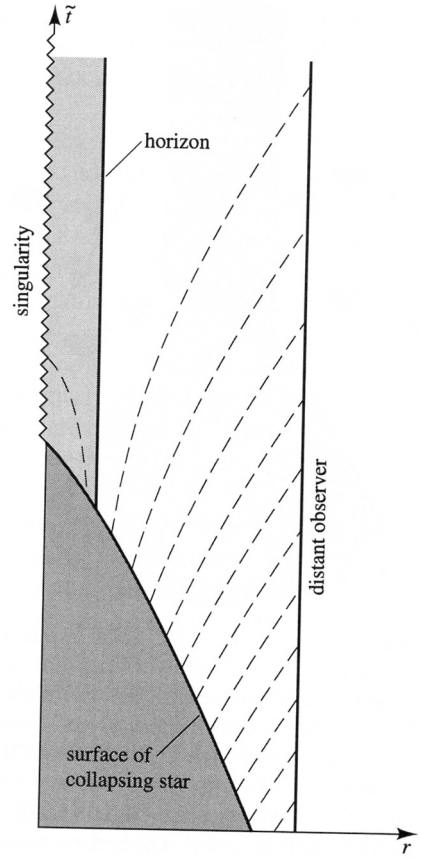

FIG. 6.4 – Schéma de l'effondrement gravitationnel d'une étoile. D'après Hartle [2003].

ce qui veut dire que

$$\frac{\mathrm{d}\tilde{t}}{\mathrm{d}r} \xrightarrow[r \to r_S^+]{} +\infty \qquad\qquad \frac{\mathrm{d}\tilde{t}}{\mathrm{d}r} \xrightarrow[r \to r_S^-]{} -\infty$$

d'où le schéma de la figure 6.3. La surface $r = r_S$ est une surface à trois dimensions du genre lumière ; les rayons lumineux se propagent le long de cette surface, qui est appelée *l'horizon du trou noir*. Aucun rayon lumineux émis depuis $r \leq r_S$ ne peut se propager à l'infini. La figure 6.4 donne le schéma de l'effondrement gravitationnel d'une étoile qui termine sa vie comme un trou noir. Les rayons émis depuis la surface de l'étoile sont d'abord reçus par un observateur à l'infini, puis leur fréquence diminue progressivement au fur et à mesure de l'effondrement, jusqu'au moment où l'observateur ne reçoit plus rien.

Comme par définition on ne peut pas voir un trou noir, on ne peut avoir que des présomptions sur leur existence. La première possibilité de trou noir vient de l'effondrement gravitationnel d'une étoile dont la masse est supérieure à 3 fois la masse solaire environ. L'existence d'un tel trou noir peut être "mise en évidence" s'il a un compagnon dont il perturbe l'orbite. On a ainsi "détecté" dans la Galaxie une dizaine de trous noirs plausibles, dont celui du Cygne. Une autre possibilité est que l'on ait des trous noirs au centre de certaines galaxies, dont la masse pourrait atteindre un million voire un milliard de masses solaires. Il est à peu près certain qu'il existe un trou noir d'environ un million de fois la masse solaire au centre de notre Galaxie. Enfin les quasars sont probablement des trous noirs géants, dont la masse est de l'ordre de 10^9 masses solaires, et qui "avalent" une quantité de matière énorme autour d'eux. Cette matière rayonne des rayons X en abondance, ce qui permet de conclure à l'existence plausible d'un trou noir. Toutes les galaxies ont probablement en leur centre un tel trou noir. Mais comme il n'y a plus de gaz à aspirer, le trou noir devient "invisible" faute d'être alimenté.

Bibliographie

Hartle [2003], chapitres 9, 12 et 13 ; Carroll [2004], chapitres 4 et 7, Wald [1984], chapitre 6 ; Damour [2005] ; R. Blanford et N. Gehrels, *Revisiting the black hole*, Physics Today, juin 1999, p. 40.

6.4 Remerciements.

Je suis très reconnaissant à Pierre Coullet et Yves Pomeau qui m'ont incité à donner ce cours aux rencontres non linéaires de Peyresq. Je remercie également Thierry Grandou et Mathieu Le Bellac pour leur lecture attentive et critique du manuscrit.