

Une introduction aux aspects géométriques
de la Relativité Générale

François Béguin

Ce texte, issu d'exposés effectués à El Oued en mars 2005, est une introduction très partielle à la Relativité Générale. En fait, il serait plus juste de parler d'une introduction à quelques problèmes de géométrie lorentzienne issus de la Relativité Générale. En particulier, on parlera de physique que pour justifier brièvement certaines définitions. De même, les problèmes d'analyse (et de géométrie) soulevés par l'étude de l'équation d'Einstein ne seront qu'à peine évoqués. J'espère que les lecteurs voudront bien excuser l'étroitesse de ce point de vue (hélas, due à l'étroitesse de mes compétences).

Prérequis. Ce texte s'adresse à un lecteur familier des concepts de base de géométrie différentielle (variété, espace tangent, différentielle, champs de vecteurs, etc.). Par ailleurs, si on a rappelé dans un appendice les définitions des concepts de géométrie pseudo-riemannienne utilisés dans le texte (métrique, dérivée covariante, géodésiques, tenseur de courbure, champs de Jacobi), un si court appendice ne saurait avoir d'autre ambition que de rafraîchir la mémoire d'un lecteur ayant déjà un peu cotoyé ces objets.

Indications bibliographiques. Il existe au moins deux bons livres classiques de géométrie lorentzienne : celui de J.K. Beem, P.E. Ehrlich et K.L. Easley ([5]), et celui de B. O'Neill ([14]) ; tout deux abordent un certain nombre de notions et problèmes issues de la Relativité Générale. Les livres de Relativité Générale à destination des physiciens sont naturellement plus nombreux (mais parfois un peu obscurs pour un mathématicien) ; parmi ceux-ci, les livres de S. Carroll ([7]) et de R. Wald ([16]) sont tout deux très bien fait (le premier étant nettement plus élémentaire mais moins complet que le second) ; le livre de S.W. Hawking et R.F. Ellis ([12]) est une véritable mine d'or, hélas parfois un peu difficile à explorer. Des références concernant certains points plus spécifiques (modèles d'espace-temps à symétries exactes, problème de Cauchy pour l'équation d'Einstein, théorèmes d'existence de singularités, etc.) seront indiquées dans le corps du texte.

0.1 L'univers de la Relativité Générale

La modélisation de l'espace et du temps en Relativité Générale est issue de deux "sauts conceptuels" fondamentaux, tout deux dus à A. Einstein.

L'expérience de Michelson-Morley est à l'origine du premier de ces deux sauts conceptuels. Cette expérience montre que la lumière a une vitesse finie, et indépendante du référentiel galiléen dans laquelle on la mesure. Einstein a

compris que ceci implique que, s'il n'existe pas de notion absolue d'*immobilité*, alors il n'existe pas non plus de notion absolue de *simultanéité*. Autrement dit, deux évènements qui paraissent simultanés à un observateur A peuvent ne pas paraître simultanés à un observateur B (en mouvement par rapport à A). La seule notion qui garde un sens absolu est celle d'évènements ayant lieu en même temps *et* au même endroit. On doit donc remplacer la conception newtonienne d'un espace (tridimensionnel) et d'un temps (unidimensionnel) indépendants par celle d'un espace-temps (quadridimensionnel).

Le second saut conceptuel est né du constat qu'on ne peut pas distinguer, par une expérience physique, un repère non-accélééré plongé dans un champ gravitationnel \vec{G} d'un repère en apesanteur soumis à une accélération $-\vec{G}$. Autrement dit, les effets d'un champ gravitationnel peuvent toujours être interprétés comme les effets d'une accélération. Einstein explique que ceci doit nous conduire à modéliser l'univers, non plus par un espace-temps plat muni des champ gravitationnel, mais par un espace-temps "courbe" sans champ gravitationnel.

Mathématiquement, l'*espace-temps* de la relativité générale est une variété connexe lisse M de dimension 4, munie d'une métrique lorentzienne g (par habitude, on considère des métriques lorentziennes de signature $(-, +, +, +)$, plutôt que $(-, -, -, +)$). La courbure de la métrique g est relié à la distribution de matière¹ dans l'espace-temps par une équation (appelée *équation d'Einstein*) que nous verrons plus tard.

Ceci appelle un certain nombre de commentaires.

— Un point important est qu'on ne connaît pas *a priori* la variété M et la métrique lorentzienne g qui modélisent l'univers. C'est ce point qui rend la relativité générale très différente des autres théories physiques : l'espace de la mécanique newtonienne est l'espace euclidien \mathbb{R}^3 , l'espace-temps de la relativité restreinte est l'espace de Minkovski, l'espace de la mécanique quantique est l'espace de Hilbert $L^2(\mathbb{R}^3)$, etc. ; en relativité générale, l'espace-temps est une variété lorentzienne qu'on ne connaît pas *a priori*. Il y a plusieurs raisons à cela. Tout d'abord, nous ignorons la distribution exacte de matière dans l'espace-temps. Comme c'est la matière qui "courbe" l'espace-temps, l'ignorance de la distribution de matière implique l'ignorance de la métrique lorentzienne g . Par ailleurs, la théorie de la Relativité Générale ne nous indique pas *a priori* la topologie de la variété M . Enfin, l'*équation d'Einstein* qui relie

¹et aux propriétés (mécaniques, électromagnétiques, etc.) de celle-ci.

la métrique lorentzienne g à la distribution d'énergie matière est telle que, même si on connaissait la topologie de M et la distribution exacte de matière dans M , ceci ne déterminerait pas² pour autant la métrique g .

— Ceci étant, toutes les variétés lorentziennes de dimension 4 ne constituent pas des espaces-temps intéressants d'un point de vue physique. La principale raison en est très simple : nous avons déjà dit que la courbure de l'espace-temps est reliée à la distribution de matière ; un espace-temps censé modéliser l'espace-temps physique doit donc avoir une distribution de courbure qui correspond à une distribution de matière "raisonnable". Par ailleurs, il existe également d'autres postulats physiques qui limite les topologies possibles pour l'espace-temps ; en particulier, on verra que l'espace-temps physique ne peut être une variété compacte (un espace-temps compact violerait le *principe de causalité*).

— L'espace-temps "physique" est quadridimensionnel³. Néanmoins, rien n'interdit au mathématiciens de considérer des espace-temps de dimension différente de 4, si ceux-ci présentent un intérêt. Il n'est d'ailleurs pas inhabituel de voir des physiciens théoriciens s'intéresser à des espaces-temps de dimension 3 ; la raison est en général assez triviale : certains problèmes étant inextricables en dimension 4, on commence par essayer de comprendre la situation en dimension 3.

Dans toute la suite, on suppose donné un espace-temps (M, g) . Sauf mention du contraire, tous les objets considérés sont lisses.

0.2 Vecteurs, courbes, surfaces et hypersurfaces de type temps, lumière ou espace

Le rôle de la métrique lorentzienne g est double. D'une part, une telle métrique fournit une notion de *courbure*, ce qui permet de donner un sens au principe selon lequel "on doit remplacer le champ gravitationnel par de la courbure" (voir la partie 0.6). D'autre part, une telle métrique permet de traduire les résultats de l'expérience de Michelson-Morley. Rappelons que cette expérience a montré que la lumière a une vitesse finie et indépendante du référentiel galiléen dans lequel on la mesure. Ceci implique qu'aucun objet

²Plus précisément, la distribution d'énergie matière détermine la courbure de Ricci de la métrique g ; mais, en dimension 4, une métrique n'est pas uniquement caractérisée par sa courbure de Ricci.

³si tant est que cette phrase a un sens.

matériel ne peut avoir une vitesse supérieure à celle de la lumière. Par suite, certaines courbes dans l'espace-temps ne peuvent correspondre à des trajectoires d'objets matériels, d'autres ne peuvent correspondre à des trajectoires de rayons lumineux, *etc.*. La métrique lorentzienne g permet de distinguer ces différents types de trajectoires, comme nous allons le voir maintenant.

Définition. On dit qu'un vecteur v tangent à M est de type *temps*, *lumière* ou *espace* suivant que sa pseudo-norme $g(v, v)$ est strictement négative, nulle, ou strictement positive. On dit que v est un vecteur *causal* si v est de type temps ou lumière (c'est-à-dire si sa pseudo-norme $g(v, v)$ est négative ou nulle).

Pour chaque point p de M , on peut trouver des coordonnées linéaires (x, y, z, t) sur T_pM dans lesquelles la forme quadratique g_p s'écrit $x^2 + y^2 + z^2 - t^2$ (les formes quadratiques sur \mathbb{R}^n sont classifiées par leur signature). Par conséquent, l'ensemble des vecteurs de type temps de T_pM est un cône ouvert, union de deux demi-cônes convexes. La frontière de ce cône ouvert est l'ensemble des vecteurs de type lumière de T_pM .

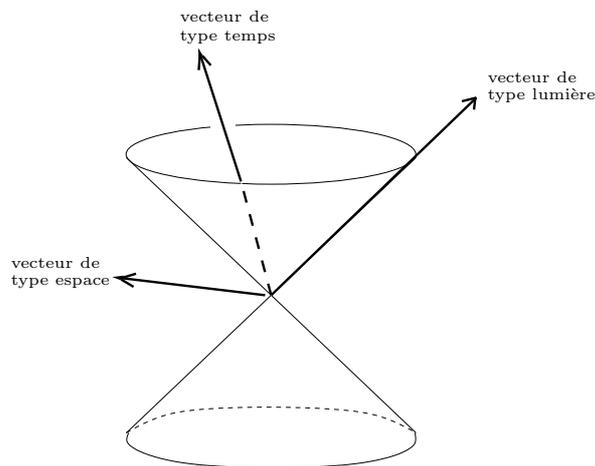


Figure 0.1: Vecteurs de types temps, lumière, espace.

Une courbe (lisse) $\gamma : I \rightarrow M$ est dite *de type temps* si son vecteur tangent $\frac{d\gamma}{ds}(s)$ est de type temps quel que soit $s \in I$. On définit de même les notions de courbe *de type lumière*, de courbe *causale*, et de courbe *de type espace*.

Remarques. En général, le type (temps, lumière, espace) du vecteur tangent à une courbe paramétrée change lorsqu'on se déplace le long de la courbe.

Une courbe paramétrée n'est donc, en général, ni de type temps, ni de type lumière, ni causale, ni de type espace.

Par contre, le type du vecteur tangent à une géodésique ne change pas lorsqu'on se déplace le long de cette géodésique⁴. Une géodésique a donc toujours un type (temps, lumière ou espace) bien défini.

D'un point de vue physique, les trajectoires (dans l'espace-temps) les objets matériels sont toujours des courbes de type temps, et les trajectoires des photons sont toujours des courbes de type lumière⁵. Réciproquement, toute courbe de type temps (resp. lumière) est une trajectoire potentielle d'objet matériel (resp. de photon).

Définition. Une surface (resp. une hypersurface) Σ dans M est dite de type *espace* si tous les vecteurs tangents à Σ sont de type espace. En d'autres termes, une surface (resp. une hypersurface) Σ dans M est de type espace si toutes les courbes tracées sur Σ sont de type espace.

Remarque. N'importe quel surface (et *a fortiori* n'importe quelle hypersurface) dans M a des vecteurs tangents de type espace (ceci provient du fait que, en tout point p de M , le cône des vecteurs temps ou lumière de T_pM est saillant, *i.e.* ne contient aucun sous-espace vectoriel de dimension 2). Autrement dit, il n'existe pas *a strictement parler* de surface ou d'hypersurface de type temps ou de type lumière⁶.

Orientation temporelle

Définitions. Une *orientation temporelle* de l'espace-temps (M, g) est un choix continu, pour tout $p \in M$, d'une des deux composantes connexes du cône des vecteurs de type temps de T_pM . On dit alors que les vecteurs situés dans (l'adhérence de) la composante connexe choisie en p sont *dirigés vers le futur*, et que les vecteurs situés dans (l'adhérence de) l'autre composante connexe sont *dirigés vers le passé*.

⁴En effet, si $\gamma : I \rightarrow M$ est une géodésique, alors on a

$$\frac{d}{ds}g\left(\frac{d\gamma}{ds}, \frac{d\gamma}{ds}\right) = 2g\left(\frac{D}{ds}\frac{d\gamma}{ds}, \frac{d\gamma}{ds}\right) = 2g\left(0, \frac{d\gamma}{ds}\right) = 0$$

⁵Ce qui traduit l'invariance de la vitesse de la lumière par changement de repère, et le fait que toute particule matérielle "va moins vite que la lumière".

⁶Cependant, par abus de langage, on appelle parfois *surface ou hypersurface lumière* une surface ou hypersurface en restriction à laquelle la métrique lorentzienne est dégénérée en tout point.

Remarque et convention. Certaines variétés lorentziennes n’admettent pas d’orientation temporelle. Cependant, d’un point de vue physique, il semble “évident” que le temps est (au moins localement) orienté. C’est pourquoi, les espaces-temps considérés en Relativité Générale sont presque toujours munis d’une orientation temporelle. En particulier, ce sera le cas dans la suite de ce texte.

Définitions. Une courbe de type temps $\gamma : I \rightarrow M$ est dite *dirigée vers le futur* si son vecteur tangent en chaque point $\frac{d\gamma}{ds}(s)$ est dans le demi-cône futur de $T_{\gamma(s)}M$ pour tout $s \in I$. Une courbe causale $\gamma : I \rightarrow M$ est dite *dirigée vers le futur* si son vecteur tangent $\frac{d\gamma}{ds}(s)$ est dans l’adhérence du demi-cône futur de $T_{\gamma(s)}M$ pour tout $s \in I$. On définit de même la notion de courbe de type temps (resp. de courbe causale) *dirigée vers le passé*.

Pour des raisons techniques, il est utile d’étendre les notions ci-dessus à des courbes qui ne sont pas lisses en certains points. On dira qu’une courbe $\gamma : I \rightarrow M$ est *lisse par morceaux* si γ est continue, si γ est lisse sauf en un nombre fini de points $x_0, \dots, x_p \in I$, et si γ admet des vecteurs tangents à gauche et à droite en chacun des points x_0, \dots, x_p . On dira qu’une telle courbe γ est *de type temps dirigée vers le futur* si le vecteur tangent à γ aux points de $I \setminus \{x_0, \dots, x_p\}$, ainsi que les vecteurs tangents à gauche et à droite aux points x_0, \dots, x_p sont de type temps dirigés vers le futur. On définit de même les notions de courbe lisse par morceaux *causale dirigée vers le futur*, *de type temps dirigée vers le passé*, *causale dirigée vers le futur* et *causale dirigée vers le passé*.

Remarques. La définition ci-dessus est justifiée par le fait que toute courbe lisse par morceaux de type temps (resp. causale) dirigée vers le futur peut être *lissée*, c’est-à-dire déformée en une courbe lisse de type temps (resp. causale) dirigée vers le futur. Ce fait repose sur la convexité de l’ensemble des vecteurs de types temps dirigés vers le futur en chaque point de M .

Par contre, il n’existe pas de bonne notion de “courbe lisse par morceaux de type temps” (sans préciser si les vecteurs tangents pointent vers le passé ou le futur). En effet, si les vecteurs tangents à gauche et à droite en un point où la courbe n’est pas lisse ne sont pas dans le même demi-cône, alors la courbe ne peut pas être déformée par une courbe lisse de type temps. Pour le même genre de raison, il n’existe pas de bonne notion de “courbe lisse par morceaux de type espace”.

Invariance conforme

Si g' est une deuxième métrique lorentzienne sur M , on dit que g et g' sont dans la même *classe conforme* si, en tout point p de M , les formes quadratiques g_p et g'_p sont proportionnelles. Pour des raisons de signature, le coefficient de proportionnalité est alors automatiquement strictement positif. On fait alors la remarque triviale mais importante suivante :

Remarque. Si g et g' sont dans la même classe conforme, alors un vecteur est de type temps (resp. lumière, espace) pour g si et seulement s'il est de type temps (resp. lumière, espace) pour g' .

Il en résulte les notions de courbe, surface, hypersurface de type temps, lumière et espace ne dépendent que de la classe conforme de la métrique lorentzienne g . Il en sera de même pour les différentes notions de passé et de futur d'un ensemble qui seront définies dans la prochaine partie.

0.3 Passé(s) et futur(s) d'un sous-ensemble

Puisqu'il ne peut emprunter que des courbes de type temps (resp. lumière) dirigées vers le futur, un objet matériel (resp. un signal lumineux) ne peut pas *a priori* aller de n'importe quel point de l'espace-temps à n'importe quel autre. Le but essentiel de cette partie est d'étudier l'ensemble des points de l'espace-temps que peut atteindre un objet matériel (resp. un signal lumineux) en partant d'un point p donné. Cet ensemble s'appelle le *futur chronologique* (resp. le *futur causal*) du point p .

Définition. Soit p un point de M et U un sous-ensemble de M contenant p . Le *futur chronologique* de p dans U , noté $I^+(p, U)$, est l'ensemble des points $q \in M$ tels qu'il existe une courbe $\gamma : I \rightarrow U$ non-triviale⁷, lisse par morceaux, de type temps, dirigée vers le futur, allant de p à q .

On définit de façon similaire le *passé chronologique* de p dans U , que l'on note $I^-(p, U)$. Les notations $I^+(p)$ et $I^-(p)$ désignent le futur et le passé chronologiques de p dans M . Le passé chronologique du point p (dans M) est l'ensemble des points $q \in M$ tels qu'un objet matériel peut aller en p en partant de q (autrement dit tel que p est dans le futur chronologique de q).

Remarques. On utilisera sans cesse les faits suivants :

⁷c'est-à-dire non-réduite à un seul point

— En général, un point p n'appartient en général pas à son futur chronologique $I^+(p, U)$. En effet, dans la définition ci-dessus, on ne considère que les courbes *non-triviales*.

— Un point q est dans le futur chronologique d'un point p si et seulement si le point p est dans le passé chronologique du point q .

— La relation “être dans le futur chronologique de” est transitive. Autrement dit : pour tous points $p, q, r \in M$ et tout ensemble U contenant ces trois points, si $q \in I^+(p, U)$ et si $r \in I^+(q, U)$ alors $r \in I^+(p, U)$. En effet, quand on met bout-à-bout une courbe lisse par morceaux, de type temps, dirigée vers le futur, allant d'un point p à un point q , et une courbe lisse par morceaux, de type temps, dirigée vers le futur, allant de q à un point r , on obtient une courbe lisse par morceaux, de type temps, dirigée vers le futur, allant de p à r . C'est d'ailleurs le principal intérêt de considérer des courbes lisses par morceaux.

Définition. Soit p un point de M , et U un sous-ensemble de M contenant p . Le *futur causal* de p dans U , noté $J^+(p, U)$, est l'ensemble des points q de M tels qu'il existe une courbe (éventuellement réduite à un point) lisse par morceaux, causale, dirigée vers le futur allant de p à q .

On définit de même le *passé causal* de p dans U , noté $J^-(p, U)$. On note $J^+(p)$ et $J^-(p)$ le futur causal et le passé causal de p dans M . D'un point de vue physique, le futur causal de p est la région de l'espace-temps où peuvent se rendre des objets et des signaux lumineux en partant de p . En d'autres termes, c'est la région de l'espace-temps qui peut être influencée par ce qui se passe en p .

Remarques. Les faits suivants découlent immédiatement des définitions du futur chronologique et du futur causal d'un point :

— On a bien sûr $I^+(p, U) \subset J^+(p, U)$. De plus, $J^+(p, U)$ est inclus dans l'adhérence de $I^+(p, U)$ (en effet, il est facile d'approcher une courbe causale issue de p par une suite de courbes de type temps issues de p).

— Un point p appartient toujours à son futur causal $J^+(p, U)$. En effet, dans la définition de $J^+(p, U)$, on prend en compte les courbes réduites à un point.

— Un point q est dans le futur causal d'un point p si et seulement si le point p est dans le passé causal de q .

— La relation “être dans le futur causal de” est transitive : pour tous p, q, r et tout U contenant ces trois points, si $q \in J^+(p, U)$ et $r \in J^+(q, U)$, alors $r \in J^+(p, U)$.

Le lemme ci-dessous, intuitivement évident mais non-trivial, est à la clef de nombreux résultats portant sur le futur des points de M :

Lemme 1. *Soit p un point de M et U un ouvert convexe contenant p . Alors, :*

- *l'ensemble $I^+(p, U)$ est constitué des points $q \in U$ tels que l'unique arc géodésique qui joint p à q est de type temps, dirigé vers le futur.*
- *l'ensemble $J^+(p, U)$ est constitué des points $q \in U$ tels que l'unique arc géodésique qui joint p à q est de type temps ou lumière, dirigé vers le futur.*

Preuve. Puisque U est un voisinage ouvert convexe de p , l'ensemble $W = \exp_p^{-1}(U)$ est un voisinage de 0 dans T_pM et l'application \exp_p réalise un difféomorphisme de W sur U . On note C le demi-cône de T_pM constitué des vecteurs de type temps dirigés vers le futur. On considère un point $q \in I^+(p, U)$. Pour démontrer la première assertion du lemme, on doit démontrer que $q \in \exp_p(C \cap W)$.

Puisque $q \in I^+(p, U)$, il existe une courbe $\gamma : [0, 1] \rightarrow U$ de type temps dirigée vers le futur qui joint p à q . On note alors $\hat{\gamma} : [0, 1] \rightarrow W \subset T_pM$ la courbe définie par $\hat{\gamma} = (\exp_p)^{-1} \circ \gamma$. On va montrer que $\hat{\gamma}(t) \in C$ pour tout $t \in [0, 1]$. Ceci entraînera en particulier que $q = \gamma(1) \in \exp_p(C \cap W)$.

Affirmation 1 : pour $t > 0$ assez petit, $\hat{\gamma}(t)$ est dans le demi-cône C .

Pour prouver cette affirmation, on identifie (via l'identification canonique) l'espace vectoriel T_pM à son espace tangent en 0. Avec cette identification, la différentielle en 0 de l'application \exp_p est l'identité de T_pM (puisque la courbe $t \mapsto \exp_p(tv)$ a pour vecteur tangent v en 0). Par conséquent, le vecteur $\frac{d\hat{\gamma}}{dt}|_{t=0}$ est identifié au vecteur $\frac{d\gamma}{dt}|_{t=0}$. En particulier, le vecteur $\frac{d\hat{\gamma}}{dt}|_{t=0}$ est de type temps dirigé vers le futur. Par conséquent, le point $\hat{\gamma}(t)$ est dans le demi-cône C pour tout $t > 0$ assez petit, ce qui termine la preuve de l'affirmation 1.

Affirmation 2 : la courbe $\hat{\gamma}$ ne ressort pas du demi-cône C .

Pour prouver cette affirmation, on considère les hypersurfaces de niveaux de la fonction $v \mapsto g_p(v, v)$. Ces hypersurfaces définissent un feuilletage trivial du demi-cône $C \subset T_pM$. L'image de ce feuilletage par l'application \exp_p est un feuilletage trivial, qu'on notera \mathcal{F} , de l'ouvert $\exp_p(C) \cap U$. Un calcul classique montre que, pour tout $v \in C \cap W$, la feuille de \mathcal{F} passant par le point $\exp_p(v)$ est orthogonale en ce point à la géodésique $t \mapsto \exp_p(t.v)$ (voir, par exemple, [12, lemme 4.5.2]). Comme $t \mapsto \exp_p(t.v)$ est une géodésique de type temps pour tout $v \in C$, on en déduit que les feuilles de \mathcal{F} sont des hypersurfaces de type espace. Et, puisque γ est de type temps, on en déduit

que γ est transverse aux feuilles de \mathcal{F} . De manière équivalente, la courbe $\hat{\gamma}$ est transverse aux niveaux de $v \mapsto g_p(v, v)$ dans T_pM . On voit facilement sur la figure 0.2 que ceci empêche la courbe $\hat{\gamma}$ de sortir du demi-cône C , et termine la preuve de l'affirmation 2.

On a donc montré en particulier que $\hat{\gamma}(1)$ appartient à C . Autrement dit, le point $q = \gamma(1)$ est de la forme $\exp_p(v)$ avec $v \in C$, c'est-à-dire v de type temps dirigé vers le futur. Ceci termine la preuve de la première assertion du lemme. La seconde assertion se déduit facilement de la première en utilisant le fait que tout point de $J^+(p, U)$ est limite de point de $I^+(p, U)$. \square

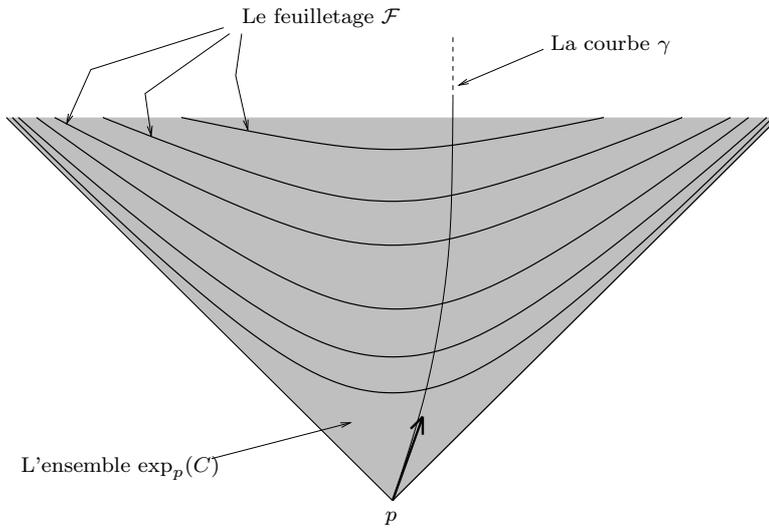


Figure 0.2: Le feuilletage \mathcal{F} et la courbe γ dans la preuve du lemme 1

Proposition 2. *Pour tout point $p \in M$ et tout voisinage ouvert U de p , l'ensemble $I^+(p, U)$ est ouvert.*

Preuve de la proposition 2. Pour éviter les détails techniques, on fait la preuve dans le cas où $U = M$. Soit $r \in I^+(p)$; on doit montrer que $I^+(p)$ est un voisinage de r . Pour cela, on considère un voisinage convexe V de r , et un point $q \in V \cap I^+(p) \cap I^-(r, V)$ (il suffit de prendre q sur $\gamma \cap V$ où γ est une courbe de type temps, dirigée vers le futur, joignant p à r). L'application \exp_q réalise un difféomorphisme de $W = (\exp_q)^{-1}(V) \subset T_qM$ sur V . Par ailleurs, puisque $q \in I^-(r, V)$ (c'est-à-dire $r \in I^+(q, V)$), le lemme 1 montre qu'il existe un

vecteur $v \in W$ de type temps dirigé vers le futur tel que $r = \exp_q(v)$. Notons C le demi-cône des vecteurs de type temps dirigés vers le futur dans T_qM . Alors, $\exp_q(C \cap W)$ est inclus dans $I^+(q, V)$ (d'après le lemme 1), donc aussi dans $I^+(p)$ (par transitivité de la relation "être dans le futur chronologique de"). Par ailleurs, $C \cap W$ est un voisinage ouvert du vecteur v dans T_qM ; donc $\exp_q(C \cap W)$ est un voisinage de $r = \exp_q(v)$ dans M . On a ainsi trouvé un voisinage de r inclus dans $I^+(p)$. \square

Remarque. Si U est un voisinage convexe de p , le lemme 1 implique que l'ensemble $J^+(p, U)$ est fermé dans U . Cependant, ceci est faux en général lorsque U n'est pas un voisinage convexe. En particulier, $J^+(p) = J^+(p, M)$ n'est en général pas fermé (voir l'exemple de la figure 0.3).

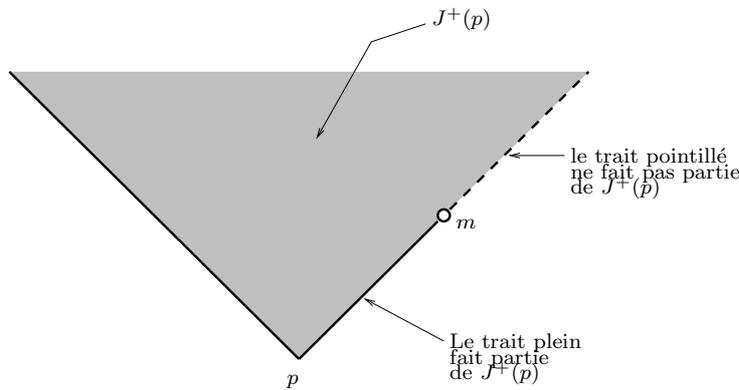


Figure 0.3: La figure représente le cas où l'espace-temps M est l'espace de Minkowski (c'est-à-dire \mathbb{R}^4 muni de la métrique lorentzienne standard $dx^2 + dy^2 + dz^2 - dt^2$) privé d'un point quelconque m . On considère alors un point p tel qu'une demi-géodésique lumière issue de p et dirigée vers le futur va "buter" sur m . On voit facilement que le futur causal $J^+(p)$ de p dans M n'est pas fermé.

Proposition 3. *Pour tous points p, q, r de M , et tout ensemble U contenant ces trois points,*

- si $q \in I^+(p, U)$ et $r \in J^+(q, U)$ alors $r \in I^+(p, U)$;
- si $q \in J^+(p, U)$ et $r \in I^+(q, U)$ alors $r \in I^+(p, U)$.

Schéma de preuve de la première assertion (la seconde se montre de même).

On remarque tout d'abord que le cas particulier où U est un ouvert convexe

découle des arguments utilisés dans la preuve du lemme 1. On s'attaque alors au cas général (U quelconque). Soit γ une courbe causale dirigée vers le futur joignant q à r . On choisit alors des points $q = m_0, m_1, \dots, m_k, m_{k+1} = r$ sur γ , placés dans cet ordre sur γ , et tels que deux points successifs m_i et m_{i+1} sont dans un même ouvert convexe U_i . En utilisant le fait que la proposition est vraie dans les ouverts convexes simples U_0, \dots, U_{k-1} , on obtient alors que $m_1 \in I^+(p)$, puis que $m_2 \in I^+(p), \dots$, puis que $m_{k+1} = r \in I^+(p)$. \square

Du lemme 1, on déduit facilement que, pour tous points $p, q \in M$, si $q \in J^+(p) \setminus I^+(p)$, alors il existe un arc géodésique de type lumière joignant p à q . En fait, la proposition suivante montre un résultat beaucoup plus fort : toute courbe causale joignant p à q est un arc géodésique de type lumière.

Proposition 4. *Soient p et q deux points de M . Si $q \in J^+(p) \setminus I^+(p)$, alors, pour toute courbe $\gamma : I \rightarrow M$ causale, dirigée vers le futur, lisse par morceaux, allant de p à q ,*

- à reparamétrage près, γ est un arc géodésique de type lumière,
- il n'existe pas de point conjugué à p le long de γ .

Preuve de la première assertion. Soit $\gamma : [0, 1] \rightarrow M$ une courbe lisse par morceaux, causale dirigée vers le futur, joignant p à q . On suppose que γ ne peut pas être reparamétrée en un arc géodésique de type lumière ; on va montrer que ceci implique $q \in I^+(p)$.

On considère un sous-intervalle compact $[t, t']$ de $[0, 1]$ tel que $\gamma|_{[t, t']}$ ne peut pas être reparamétrée en un arc géodésique, et tel que $\gamma|_{[t, t']}$ est contenue dans un voisinage convexe $U \subset M$. On note $r = \gamma(t)$. Puisque V est un voisinage ouvert convexe de r , l'ensemble $W = \exp_r^{-1}(U)$ est un voisinage de 0 dans $T_r M$ et l'application \exp_r réalise un difféomorphisme de W sur U . On note C (resp. $\partial C, \overline{C}$) l'ensemble des vecteurs de type temps (resp. de type lumière, de type temps ou lumière) dirigés vers le futur de $T_r M$. La seconde assertion du lemme 1 implique que la courbe $\gamma|_{[t_1, t_2]}$ est contenue dans $\exp_r(\overline{C} \cap W)$.

Affirmation : la courbe $\gamma|_{[t, t']}$ ne peut pas être rester sur l'hypersurface $\exp_r(\partial C \cap W)$.

En effet, l'hypersurface $\exp_r(\partial C \cap W)$ est réglée par les demi-géodésiques lumières issues de r , et tout vecteur tangent à l'hypersurface et transverse à ce réglage est de type espace. Puisque les vecteurs tangents à γ ne sont jamais de type espace, si γ était contenue dans l'hypersurface $\exp_r(\partial C \cap W)$, alors γ serait tangente en tout point à une géodésique lumière issue de r .

Autrement dit, γ pourrait être reparamétrée en une géodésique lumière, ce qui est contraire à notre hypothèse.

Il va donc nécessairement exister un $t'' \in [t, t']$ tel que $\gamma(t'') \in \exp_r(C \cap W) \subset I^+(r)$. En notant $s = \gamma(t'')$, on a alors la situation suivante : $r \in J^+(p)$, $s \in I^+(r)$ et $q \in J^+(s)$. D'après la proposition 3, ceci implique $q \in I^+(p)$. \square

Esquisse de preuve de la seconde assertion. Soit $\gamma : [0, 1] \rightarrow M$ un arc géodésique de type lumière joignant p à q . On suppose qu'il existe un point $r = \gamma(s_0)$ conjugué à p le long de γ .

Puisque $r = \gamma(s_0)$ est conjugué à $p = \gamma(0)$, il existe une déformation $(\gamma_u)_{u \in [-\varepsilon, \varepsilon]}$ de γ , telle que γ_u est un arc géodésique de type lumière vérifiant $\gamma_u(0) = p$ pour tout u , et telle que $\frac{\partial}{\partial u} \gamma_u(s_0) = 0$. Cette dernière condition signifie que $\gamma_u(s_0)$ est très proche du point r pour u petit. Si $\gamma_u(s_0)$ était exactement égal à r , en concaténant l'arc $\gamma_u|_{[0, s_0]}$ et l'arc $\gamma|_{[s_0, 1]}$, on obtiendrait une courbe causale qui joint p à q , mais qui n'est pas reparamétrable en un arc géodésique. D'après la première assertion de la proposition 4, ceci impliquerait que $q \in I^+(p)$.

Pour faire une transformer cette idée en une vraie preuve, on procède comme suit. Étant donné une fonction $f : [0, 1] \rightarrow \mathbb{R}$, on considère, pour tout $\varepsilon > 0$, la courbe $\gamma_{f, \varepsilon} : [0, 1] \rightarrow M$ définie par

$$\gamma_{f, \varepsilon}(t) = \exp_{\gamma(t)}(\varepsilon \cdot f(t) \cdot \widehat{J}(t)) \quad \text{où} \quad \widehat{J} = \frac{J}{\|J\|}$$

(le champ \widehat{J} n'est *a priori* pas défini en $t = t_0$, mais on vérifie facilement qu'il se prolonge par continuité en ce point). On cherche alors f tel que la courbe $\gamma_{f, \varepsilon}$ soit de type temps (pour ε assez petit), passe par p et un point $\gamma(t_1)$. Ces conditions se traduisent par l'annulation de f en 0 et en t_1 , et par une équation différentielle du second ordre sur f . On montre alors facilement que le problème admet une solution (voir, par exemple, [12, proposition 4.5.10]). \square

Jusqu'ici nous n'avons parlé que du futur (resp. du passé) d'un point. Cependant, on peut également définir le futur de n'importe quelle partie de M : pour toutes parties A, U de M telles que $A \subset U$, on appelle *futur chronologique* et *futur causal* de A dans U les ensembles

$$I^+(A, U) = \bigcup_{p \in A} I^+(p, U) \quad \text{et} \quad J^+(A, U) = \bigcup_{p \in A} J^+(p, U)$$

La plupart des propriétés énoncées ci-dessus dans le cas où A est un point se généralisent immédiatement au cas où A est quelconque. Le cas particulier où

A est une surface de type espace jouera un rôle important à la fin de ce texte ; dans ce cas particulier, on a l'analogie suivant de la proposition 4 :

Proposition 5. *Soit Σ est une surface de type espace dans M et q un point de $J^+(\Sigma) \setminus I^+(\Sigma)$. Alors, pour toute courbe causale $\gamma : I \rightarrow M$ allant d'un point de Σ à q ,*

- γ est un arc géodésique de type lumière,
- γ est orthogonal à Σ ,
- il n'existe pas de point focal à Σ le long de γ .

Dans un certains nombre de problèmes, la stratégie est la suivante. On a un ensemble⁸ A , et on voudrait un maximum d'informations sur le futur chronologique ou causal de A . Pour cela, on cherche à étudier la frontière de $I^+(A)$ (qui coïncide avec celle de $J^+(A)$). Mais cela n'est en général guère facile. Par contre, les propositions 4 et 5 nous fournissent des informations sur l'ensemble $J^+(A) \setminus I^+(A)$. Il reste donc à montrer que, sous certaines hypothèses raisonnables, la frontière de $J^+(A)$ n'est pas trop différente de $J^+(A) \setminus I^+(A)$. On verra un résultat de ce type dans la partie 0.5.

0.4 Hypothèses de causalité

Puisque, d'un point de vue physique, toute courbe de types temps dans M représente une trajectoire putative de particule matérielle, l'existence de courbes de type temps fermées impliquerait qu'un objet matériel pourrait "revenir dans son passé". On considère généralement qu'un tel évènement contreviendrait aux principes philosophiques sur lesquels est bâtie la physique (on parle de "violation du principe de causalité"). C'est pourquoi, on postule presque toujours que l'espace-temps (M, g) satisfait l'hypothèse suivante :

Hypothèse chronologique faible : il n'existe aucune courbe de type temps fermée dans l'espace-temps M .

En fait, l'existence d'une particule qui reviendrait, non pas à strictement parler dans son passé, mais arbitrairement proche d'un point de son passé peut sembler également paradoxale ; il est donc fréquent d'utiliser une hypothèse plus forte que celle ci-dessus :

⁸Par exemple, dans la dernière partie de ce texte, cet ensemble sera une surface entourant un astre dont on veut montrer qu'il va, en temps fini, s'effondrer pour créer un trou noir, et donc forcer l'existence d'une "singularité".

Hypothèse chronologique forte : tout point $p \in M$ admet un voisinage U arbitrairement petit, tel que toute courbe de type temps qui sort de U n'y revient plus.

Il existe de nombreuses variantes des hypothèses ci-dessus. Par exemple, on peut renforcer chacune des deux hypothèses ci-dessus en prenant en compte toutes les courbes causales, au lieu des seules courbes de type temps (*hypothèse de causalité faible et forte*). On peut aussi supposer qu'il n'existe pas de courbe causale fermée non seulement pour la métrique considérée, mais aussi pour toute perturbation assez petite de cette métrique (*hypothèse de causalité stable*). Toutes ces hypothèses sont regroupées sous l'appellation générique d'*hypothèses de causalité*.

Formellement, les hypothèses de causalité évoquées ci-dessus sont toutes différentes les unes des autres (voir, par exemple, [12, partie 6.4]) . Cependant, les exemples de variétés lorentziennes qui vérifient une des hypothèses mais pas les autres semblent assez artificiels. On peut d'ailleurs montrer des résultats qui affirment que, pour des variétés lorentziennes satisfaisant certaines propriétés géométriques raisonnables, toutes les hypothèses évoquées ci-dessus sont équivalentes. Voici un exemple de tel résultat (voir [12, page 192] pour une démonstration) :

Proposition 6. *Supposons que (M, g) satisfait la hypothèse chronologique faible, et supposons que toute géodésique de type temps ou lumière de (M, g) admet une paire de points conjugués. Alors, (M, g) satisfait l'hypothèse chronologique forte.*

Question. *Est-il vrai que, si l'espace-temps (M, g) satisfait l'hypothèse chronologique faible, alors il existe une métrique lorentzienne g' arbitrairement proche⁹ de g tel que l'espace (M, g') vérifie toutes les hypothèses de causalité évoquées ci-dessus ?*

Les hypothèses de causalité imposent des restrictions assez fortes sur la topologie de l'espace-temps. Par exemple, elles excluent la possibilité que l'espace-temps puisse être compact :

Proposition 7. *Si l'espace-temps (M, g) satisfait l'hypothèse chronologique faible, alors M n'est pas compact.*

⁹pour la topologie raisonnable à déterminer

Preuve. Supposons M compacte. Pour tout point q dans M , l'ensemble $I^+(q)$ est ouvert. Par compacité de M , on peut donc trouver un nombre fini de points q_1, \dots, q_n tels que les ensembles $I^+(q_1), \dots, I^+(q_n)$ recouvrent M .

Notons $i_1 = 1$. Puisque les ensembles $I^+(q_1), \dots, I^+(q_n)$ recouvrent M , il existe $i_2 \in \{1, \dots, n\}$ tel que q_{i_1} est dans le futur de q_{i_2} . Pour la même raison, il existe $i_3 \in \{1, \dots, n\}$ tel que q_{i_2} est dans le futur de q_{i_3} . De proche en proche, on construit une suite (i_k) d'éléments de $\{1, \dots, n\}$ telle que q_{i_k} est dans le futur de $q_{i_{k+1}}$ pour tout k . Notons que, par transitivité de la relation "être dans le futur de", ceci implique que q_{i_k} est dans le futur de q_{i_l} pour tous k, l tels que $k < l$. Puisque la suite (i_k) prend ses valeurs dans un ensemble fini, il doit exister deux entiers r et s tel que $i_r = i_s$. Mais alors le point q_{i_r} est dans son propre futur. Donc M ne satisfait pas la hypothèse chronologique faible. \square

En fait, les hypothèses de causalité ont tendance à forcer l'espace-temps à être peu différent (au niveau topologique) du produit d'une hypersurface de type espace par une droite de type temps. Voici, un résultat qui va dans ce sens :

Proposition 8. *Supposons que (M, g) satisfait l'hypothèse de causalité forte, et supposons qu'il existe deux hypersurfaces de type espace disjointes S et S' dans M qui bordent un compact M_0 de M . Alors S est homéomorphe à S' et M_0 est homéomorphe à $S \times [0, 1]$.*

Schéma de preuve. On peut construire un champ de vecteurs X sur M , qui ne s'annule pas, n'est tangent à S et S' en aucun point, et pointe du côté de M_0 le long de S (par exemple, on peut choisir X de type temps). Soit γ une courbe intégrale de X partant d'un point p de S . L'hypothèse de causalité forte interdit à γ de rester à tout jamais dans une région compacte (on utilise ici le fait que X ne s'annule pas). Donc γ doit sortir de M_0 , et la seule possibilité pour sortir de M_0 est de couper S' au bout d'un temps $\tau(p)$ (car X pointe vers l'intérieur de M_0 le long de S). On vérifie alors que l'application $\Phi : (p, t) \mapsto X^{\frac{t}{\tau(p)}}(p)$ réalise un homéomorphisme de $S \times [0, 1]$ sur M_0 . \square

0.5 Horloges globales, hypersurfaces de Cauchy et hyperbolicité globale

Le principe fondateur de la Relativité Restreinte est qu'il n'existe pas de notion absolue de simultanéité. En d'autres termes : il n'y a pas d'*horloge* privilégiée

(une “horloge” est une fonction qui assigne une “heure” à chaque évènement, c’est-à-dire à chaque point de l’espace-temps ; voir la définition précise ci-dessous). De plus, en Relativité Générale, il arrive qu’il n’existe aucune horloge *globale* (c’est-à-dire définie sur l’espace-temps en entier). Le but de cette partie est de relier l’existence d’horloges globales à d’autres propriétés fondamentales des espaces-temps.

Définition. Une *horloge* définie sur un sous-ensemble U de l’espace-temps M est une fonction continue $t : U \rightarrow \mathbb{R}$ qui satisfait la propriété suivante : si q est dans le futur chronologique de p , alors $t(q) > t(p)$. Une *horloge globale* est une horloge définie sur M en entier.

Notons que si (M, g) admet une horloge globale, alors (M, g) satisfait toutes les hypothèses de causalité énoncée dans la partie précédente. Ceci étant, la plupart des modèles d’espace-temps qu’utilisent les physiciens admettent une horloge globale. Par ailleurs, nous verrons dans la partie 0.6 que les espace-temps construits numériquement par résolution d’un problème de Cauchy pour l’équation d’Einstein admettent automatiquement une horloge globale.

Définition. On dit que l’espace-temps (M, g) est *globalement hyperbolique* s’il satisfait l’hypothèse chronologique forte, et si $J^+(p) \cap J^-(q)$ est compact pour tous $p, q \in M$.

Proposition 9. *Si M est un espace temps globalement hyperbolique, alors le futur causal $J^+(K)$ de n’importe quel compact K de M est fermé.*

Preuve. Montrons tout d’abord que le futur causal de tout point de M est fermé. Considérons un point p de M . Supposons que $J^+(p)$ n’est pas fermé, et considérons alors $q \in \text{adh}(J^+(p)) \setminus J^+(p)$. Soit alors $r \in I^+(q)$. Alors $q \in \text{adh}(J^+(p)) \cap I^-(r) = \text{adh}(J^+(p) \cap I^-(r))$ (l’égalité provient du fait que $I^-(r)$ est ouvert). Et puisque $J^+(p) \cap J^-(r)$ est compact, ceci implique que $q \in J^+(p) \cap J^-(r)$. D’où une contradiction. Donc $J^+(p)$ est fermé.

Soit maintenant K un compact de M . Supposons que $J^+(K)$ n’est pas fermé, et considérons un point q dans $\text{adh}(J^+(K)) \setminus J^+(K)$. Soit q_n une suite de points de $I^+(q)$ convergeant vers q , telle que $q_{n+1} \in I^-(q_n)$ pour tout n . Alors, pour chaque n , l’ensemble $J^-(q_n)$ est fermé, donc $J^-(q_n) \cap K$ est compact. Donc $\bigcap_n J^-(q_n) \cap K$ est un compact non-vide ; soit p est un point de cet ensemble. Alors le point q_n est dans $J^+(p)$ pour tout n . Et puisque $J^+(p)$ est fermé d’après la première partie de la preuve, le point q est dans $J^+(p)$. Comme $p \in K$, ceci contredit le fait que $q \notin J^+(K)$. \square

Remarque. Si l'espace-temps (M, g) est globalement hyperbolique, alors, pour tout point $p \in M$, les propositions 4 et 9 fournissent de précieuses informations sur la frontière du futur de p : cette frontière est constituée de points q reliés à p par un arc géodésique de type lumière ne contenant pas de point conjugué à p .

Définition. Une *hypersurface de Cauchy* de M est une hypersurface de type espace Σ dans M telle que toute courbe causale inextensible¹⁰ intersecte Σ en un point et un seul.

Proposition 10 (R. Geroch). *Si l'espace-temps (M, g) est globalement hyperbolique, alors il admet une horloge globale et une hypersurface de Cauchy. De plus, M est difféomorphe à un produit $S \times \mathbb{R}$ où S est une hypersurface de Cauchy.*

Preuve partielle. Supposons que (M, g) est globalement hyperbolique. On choisit un volume riemannien μ sur M tel que le volume total de M soit égal à 1. Remarquons tout d'abord que pour tout point $p \in M$, on a $\mu(J^+(p)) = \mu(I^+(p))$; en effet, $J^+(p) \setminus I^+(p)$ est inclus dans l'image par une application différentiable (l'application \exp_p) d'une hypersurface (le cône des vecteurs de type lumière de $T_p M$). Bien sûr, on a de même $\mu(J^-(p)) = \mu(I^-(p))$ pour tout $p \in M$. On considère alors les fonctions $f^+ : M \rightarrow \mathbb{R}$ et $f^- : M \rightarrow \mathbb{R}$ définies par $f^+(p) = \mu(J^+(p)) = \mu(I^+(p))$ et $f^-(p) = \mu(J^-(p)) = \mu(I^-(p))$. Clairement, f^+ et f^- sont respectivement strictement décroissante et strictement croissante le long de n'importe quelle courbe causale dirigée vers le futur.

Montrons que f^+ tend vers 0 le long de toute courbe causale, dirigée vers le futur, inextensible de M . Soit $\gamma : \mathbb{R} \rightarrow M$ une telle courbe, et supposons que $f^+(\gamma(t))$ ne tend pas vers 0 quand t tend vers $+\infty$. Alors, l'intersection décroissante $\bigcap_{t \in \mathbb{R}} J^+(\gamma(t))$ est non-vide ; soit q un point de cette ensemble, et t_0 un réel. Comme M est globalement hyperbolique, $J^-(q) \cap J^+(\gamma(t_0))$ est compact. Et comme $\gamma(t)$ sort de tout compact lorsque t tend vers $+\infty$, il existe $t > t_0$ tel que $\gamma(t) \notin J^-(q) \cap J^+(\gamma(t_0))$. Mais $\gamma(t) \in J^+(\gamma(t_0))$ (puisque $t > t_0$). Donc $\gamma(t) \notin J^-(q)$, c'est-à-dire $q \notin J^+(\gamma(t))$, ce qui contredit la définition de q . Donc $f^+(\gamma(t))$ tend vers 0 quand t tend vers $+\infty$. On montre de même que $f^-(\gamma(t))$ tend vers 0 quand t tend vers $-\infty$.

Montrons que f^+ et f^- sont continues. Remarquons tout d'abord qu'il suffit pour cela de montrer qu'elles sont continues le long de toute courbe causale : en effet, n'importe quelle courbe lisse dans M peut-être approchée (en

¹⁰Une courbe causale $\gamma : I \rightarrow M$ est dite *inextensible* s'il n'existe aucune courbe causale $\hat{\gamma} : J \rightarrow M$ telle que J contient strictement I et $\hat{\gamma}|_I = \gamma$.

topologie C^0) par une courbe continue, C^1 par morceaux, dont chaque morceau est un arc causal. Considérons donc une courbe causale γ , et $r \in \gamma$. Supposons que f^+ n'est pas semi-continue supérieurement en r le long de γ . Alors, il existe une suite $(r_n)_{n \in \mathbb{N}}$ de points de γ , qui tend vers r , telle que les points r_n sont tous strictement dans le passé de r , et telle que $\mu(\bigcap_n J^+(r_n)) > \mu(J^+(r))$. En particulier, il existe $q \in (\bigcap_n J^+(r_n)) \setminus J^+(r)$. Par définition, r n'est pas dans $J^-(q)$, mais r_n est dans $J^-(q)$ pour tout n , et donc r est dans l'adhérence de $J^-(q)$. Ceci constitue une contradiction, puisque $J^-(q)$ est fermé d'après la proposition 9 ; donc f^+ est semi-continue supérieurement en r le long de γ . On montre de même que f^+ est semi-continue inférieurement le long de γ (cette fois-ci, les points r_n sont dans le futur de r , on a $\mu(\bigcup_n I^+(r_n)) > \mu(I^+(r))$, et on obtient une contradiction en utilisant le fait que $I^+(r)$ est ouvert). On procède bien sûr de même avec f^- .

Considérons maintenant la fonction $f : M \rightarrow \mathbb{R}$ définie par $f(p) = f^+(p)/f^-(p)$. C'est une horloge globale sur M . Ceci implique immédiatement que toute courbe causale coupe n'importe quel niveau (non-vide) de f en au plus un point. Par ailleurs, toute courbe causale coupe n'importe quel niveau non-vide de f en au moins un point : ceci provient simplement du fait que les valeurs de f varient continument de $-\infty$ à $+\infty$ le long de n'importe quelle courbe causale inextensible.

La fonction f est *a priori* seulement continue. Si la fonction $f : M \rightarrow \mathbb{R}$ est lisse, et si la dérivée de f le long de toute courbe causale dirigée vers le futur est strictement positive, alors n'importe quel niveau non-vide de f est une hypersurface lisse. Les propriétés de f impliquent que toute courbe causale coupe cette hypersurface en un en un point et un seul. On a donc trouvé l'hypersurface de Cauchy recherchée. Si f n'est pas lisse, il faut montrer qu'on peut lisser f et obtenir une fonction dont la dérivée le long de toute courbe causale est strictement positive ; pour ce point, je renvoie à [6].

Le dernier point à montrer est la décomposition $M \approx S \times \mathbb{R}$, où S est une hypersurface de Cauchy - par exemple, un niveau de f . Soit ∇f le gradient de f . C'est un champ de vecteurs de type temps. On peut le renormaliser par une fonction réelle a de sorte que $X = a\nabla f$ soit dirigé vers le futur, et de sorte que le flot φ^t engendré par X soit complet, c'est-à-dire défini pour tous les temps. On peut donc définir une application $F : S \times \mathbb{R} \rightarrow M$ par $F(x, t) = \varphi^t(x)$. C'est clairement un difféomorphisme local, injectif, et le fait que S soit une hypersurface de Cauchy implique la surjectivité de F puisque chaque trajectoire de φ est une courbe causale (et même de type

temps) inextensible. Ceci achève la preuve de la proposition. \square

Réciproquement, l'existence d'une surface de Cauchy implique l'hyperbolicité globale :

Proposition 11. *Si l'espace-temps (M, g) admet une hypersurface de Cauchy, alors il est globalement hyperbolique.*

Schéma de preuve. La preuve peut se décomposer en trois étapes :

1) On commence par montrer que toute suite $(\gamma_n)_{n \in \mathbb{N}}$ de courbes causales admet une courbe limite γ (au sens où tout voisinage de tout point de γ rencontre une infinité de γ_n). On obtient de plus que, si p est un point de M tel que tout voisinage de p rencontre une infinité de γ_n , alors on peut choisir γ passant par p . Enfin, on obtient que, si les γ_n sont toutes inextensibles dans un ouvert U , alors on peut choisir γ inextensible dans U .

2) On considère alors l'hypersurface de Cauchy $S \subset M$, deux points $p, q \in M$, et une suite (γ_n) de courbes causales joignant p à q . En distinguant deux cas (suivant que S sépare p et q , ou pas), en utilisant le fait que toute courbe causale inextensible coupe S , et le lemme de l'étape 1), on trouve une courbe limite γ de la suite (γ_n) , qui est causale et joint p à q .

3) L'étape 2) montre que, pour tous $p, q \in M$, l'ensemble $C(p, q)$ des arcs causaux joignant p à q est compact (la topologie étant celle pour laquelle un voisinage de $\gamma_0 \in C(p, q)$ est constitué des arcs $\gamma \in C(p, q)$ tels que l'image de γ dans M est contenue dans un voisinage de l'image de γ_0 dans M). On montre enfin que la compacité de $C(p, q)$ pour tous $p, q \in M$ est équivalente à l'hyperbolicité globale de l'espace-temps (M, g) .

On trouve une preuve complète, par exemple, dans [12, partie 6.6]. \square

Remarquons que les propositions 11 et 10 mises bout-à-bout montrent que l'existence d'une surface de Cauchy implique l'existence d'une horloge globale. Par contre, la réciproque est fautive : en considérant l'ensemble

$$\{(x, y, z, t) \in \mathbb{R}^4 \mid x^2 + y^2 + z^2 < 1\}$$

muni de la métrique induite par la forme quadratique $x^2 + y^2 + z^2 - t^2$, on obtient un espace-temps qui admet une horloge globale, mais pas d'hypersurface de Cauchy. Cependant, on a des réciproques partielles, comme par exemple :

Proposition 12. *Si l'espace-temps (M, g) admet une horloge globale dont les hypersurfaces de niveaux sont compactes, alors (M, g) admet une hypersurface de Cauchy.*

Preuve. Soit $t : M \rightarrow \mathbb{R}$ une horloge globale et S une hypersurface de niveau de t . Puisque t est strictement croissante le long de chaque courbe causale orientée vers le futur, S est nécessairement une hypersurface de type espace que chaque courbe causale coupe au plus une fois. Par ailleurs, l'ensemble des courbes causales qui coupent S est clairement ouvert (pour la topologie compacte-ouverte), et la compacité de S implique facilement qu'il est fermé. Comme l'ensemble des courbes causales (paramétrées et dirigées vers le futur) est connexe, on en déduit que toute courbe causale coupe S , ce qui conclut la preuve. \square

Concluons par une remarque. Nous avons affirmé au début de cette partie qu'un des principes fondateurs de la Relativité est qu'il n'existe pas d'horloge privilégiée. En fait, en Relativité Générale, cette affirmation n'est vraie que *localement*. En effet, sur certains types d'espace-temps particuliers, il arrive qu'on réussisse à définir une horloge globale privilégiée. La valeur en un point p d'une telle horloge dépend alors nécessairement de la géométrie globale de l'espace-temps, et non seulement d'un voisinage de p .

Exemple. Pour tout point p de l'espace-temps (M, g) , on note $\tau(p)$ le supremum des longueurs de toutes les courbes de type temps, partant de p , et dirigées vers le passé. En général, $\tau(p) = +\infty$ au moins pour certains points $p \in M$. Mais, si $\tau(p)$ est fini pour tout $p \in M$, et si $p \mapsto \tau(p)$ est continue, alors τ est une horloge globale sur M . Cette horloge est canonique au sens où sa définition ne dépend pas du choix de coordonnées particulières. On l'appelle l'*horloge cosmologique*. Il existe des exemples importants d'espace-temps (le modèle d'univers de Robertson-Walker, par exemple) pour lesquels l'horloge cosmologique est bien définie. Sur ce sujet, voir [3].

Problème. *Trouver des hypothèses les plus générales possibles qui permettent de définir une horloge globale privilégiée.*

0.6 L'équation d'Einstein

L'idée de base de la relativité générale consiste à remplacer un espace-temps plat muni d'un champ gravitationnel par un "espace-temps courbe" sans champ gravitationnel. Autrement dit, en Relativité Générale, la matière ne crée plus de champ gravitationnel, mais "courbe l'espace". Reste à savoir de quelle manière une distribution de matière donnée courbe l'espace ? La réponse nous est donnée par l'*équation d'Einstein*, qui relie la métrique

lorentzienne g et un tenseur T qui décrit la distribution de matière dans l'espace-temps :

$$\text{Ric}_g - \frac{1}{2}\text{Scal}_g.g + \Lambda.g = 8\pi.T$$

où :

- Ric_g désigne le tenseur de Ricci de la métrique lorentzienne g ,
- Scal_g désigne la courbure scalaire de g ,
- Λ est une constante, appelée *constante cosmologique*,
- T est un champ de formes bilinéaires symétriques, de divergence nulle, qui décrit la distribution de matière et d'énergie dans l'espace-temps, appelé *tenseur énergie-impulsion*.

Essayons de justifier (partiellement) cette équation. Il nous faut tout d'abord admettre que la matière et les divers champs présents dans l'espace-temps peuvent être décrits par un champ de formes bilinéaires symétriques, de divergence nulle, qu'on note T (le fait que T soit de divergence nulle traduit la conservation de l'énergie). Pour exprimer le fait que la matière courbe l'espace, on doit alors chercher une équation du type $G(g) = T$, où $G(g)$ est un champ de formes bilinéaires symétriques, de divergence nulle, qui dépend de la métrique g . De plus, on veut que cette équation ne dépendent pas du choix d'un système de coordonnées, c'est-à-dire que T soit équivariant sous l'action du groupe des difféomorphismes, c'est-à-dire vérifie $G(h_*g) = h_*G(g)$ pour tout difféomorphisme $h : M \rightarrow M$. Or, il résulte des travaux de D. Hilbert qu'un tel champ de formes bilinéaires est nécessairement de la forme $G(g) = k(\text{Ric}_g - \frac{1}{2}\text{Scal}_g.g + \Lambda.g)$ où k et Λ sont des constantes. Enfin, on trouve la valeur $k = \frac{1}{8\pi}$ en comparant l'équation d'Einstein pour des métriques proches de la métrique de Minkowski et les lois de la mécanique newtonnienne¹¹. Bien sûr, d'un point de vue physique, la principale justification *a posteriori* de cette équation est le fait qu'elle semble rendre compte avec une très grande précision des résultats des expériences.

Il est important de comprendre que l'équation d'Einstein ne permet pas, à elle seule, de déterminer l'espace-temps (M, g) . Tout d'abord, cette équation suppose connue la variété M puisqu'elle relie des objets (la courbure de la métrique lorentzienne g et le tenseur énergie-impulsion T) qui vivent sur cette variété. De plus, à moins de supposer qu'on connaît la distribution exacte de matière dans l'espace-temps en entier, on ne connaît ni la valeur exacte du tenseur énergie-impulsion T en chaque point de M , ni la valeur exacte de la

¹¹En fait, la valeur $k = \frac{1}{8\pi}$ dépend d'un choix d'unités que nous ne précisons pas ici

constante cosmologique¹² Λ . Enfin, quand bien même on connaîtrait M , T et Λ , il faut remarquer que la connaissance du tenseur $\text{Ric}_g - \frac{1}{2}\text{Scal}_g \cdot g + \Lambda \cdot g$ (qui est équivalente à la connaissance du tenseur de Ricci Ric_g) ne permet pas de déterminer entièrement¹³ la métrique lorentzienne g .

L'équation d'Einstein peut néanmoins être utilisée de plusieurs façons :

— On peut tout d'abord chercher des solutions particulières de cette équation. En général, on cherche des solutions qui possèdent beaucoup de symétries (on parle de *solutions à symétrie exactes*). Nous évoquerons ci-dessous deux types de solutions à symétries exactes qui jouent un rôle particulièrement important en astrophysique : les *espace-temps de Robertson-Walker* et l'*espace-temps de Schwarzschild-Kruskal*.

— On peut également formuler un problème de Cauchy pour l'équation d'Einstein, ce qui permet de calculer numériquement des certaines solutions de cette équation.

— Enfin, on peut utiliser l'équation d'Einstein comme un dictionnaire qui permet de traduire des propriétés physiques du tenseur énergie-impulsion T en propriétés géométriques de la métrique g . Nous montrerons un exemple de cette démarche ci-après (voir le paragraphe sur *l'hypothèse d'énergie positive*).

Solutions à symétries exactes I : les espaces-temps de Robertson-Walker

Les *espaces-temps de Robertson-Walker* sont des solutions de l'équation d'Einstein qui modélisent la forme globale de l'univers, en négligeant les irrégularités locales de la distribution de matière. Les premières solutions de ce type ont été proposées dans les années 20, indépendamment par G. Lemaître et A. Friedman. Ces solutions ont une importance considérable en Astrophysique : ils sont à l'origine de la théorie prédisant l'existence du *Big Bang*, c'est grâce à eux qu'on estime l'âge de l'univers, etc.

Un *espace-temps de Robertson-Walker* est un espace-temps (M, g) qui satisfait les trois hypothèses suivantes :

(1) (M, g) admet une horloge globale. On rappelle que ceci implique que $M = S \times I$, où S est une variété de dimension 3 et d'un intervalle $I =]t_*, t^*[$

¹²La valeur de cette constante (en particulier son signe) est reliée à la quantité de "matière noire" présente dans l'univers. L'estimation de cette valeur fait l'objet d'intenses recherches et d'interminables débats chez les physiciens.

¹³Si j'ai bien compris, c'est d'ailleurs ceci qui est à l'origine des fameuses *ondes gravitationnelles* dont plusieurs expériences essaient en ce moment de détecter l'existence.

de \mathbb{R} (*a priori* t_* et t^* peuvent être finis ou infinis), où la surface $S \times \{t\}$ est de type espace pour tout $t \in I$, et la courbe $\{x\} \times I$ est de type temps pour tout $x \in S$.

(2) (M, g) est *spatialement isotrope* : pour tout $(x, t) \in S \times I = M$, et tous vecteurs unitaires v, v' tangents à $S \times \{t\}$ en (x, t) , il existe une isométrie Φ de (M, g) qui préserve l'hypersurface $S \times \{t\}$, qui fixe le point (x, t) , et telle que $d\Phi_{(x,t)} \cdot v = v'$.

(3) (M, g) est *maximal* : si h est un plongement isométrique de (M, g) dans un espace-temps (M', g') satisfaisant la première et la deuxième hypothèse ci-dessus, alors h est surjectif.

Des trois hypothèses de la définition ci-dessus, la plus importante est certainement l'hypothèse (2). De manière informelle, cette hypothèse affirme que, quelque soit le point de l'espace-temps où on se trouve, ce que l'on voit ne dépend pas de la direction dans laquelle on regarde. Bien entendu, cette hypothèse n'est pas vérifiée dans l'univers physique aux échelles usuelles d'observation. Néanmoins, les mesures des astrophysiciens¹⁴ indiquent qu'à très grande échelle, la distribution d'énergie-matière dans l'univers observée depuis la Terre est quasiment isotrope. On fait, par ailleurs, le *postulat Copernicien* que la Terre n'a pas une position spéciale dans l'univers, et que la distribution de matière observée depuis un autre point de l'univers serait donc également isotrope. L'hypothèse ci-dessus semble donc justifiée si on ne s'intéresse qu'à la forme globale de l'univers, sans tenir compte des irrégularités locales.

On montre facilement que l'hypothèse (2) implique d'une part que la métrique riemannienne induite par g sur l'hypersurface $S \times \{t\}$ est à courbure constante pour tout t , et d'autre part qu'on passe de la métrique sur l'hypersurface $S \times \{t_1\}$ à la métrique sur l'hypersurface $S \times \{t_2\}$ par une simple homothétie. Autrement dit, la métrique lorentzienne g peut s'écrire :

$$g = -dt^2 + f(t)g_S$$

où $f : I \rightarrow]0, +\infty[$ est une fonction lisse et g_S est une métrique riemannienne à courbure constante sur la 3-variété S .

Soit maintenant $T = \text{Ric}_g - \frac{1}{2}\text{Scal}_g \cdot g + \Lambda \cdot g$. Lorsqu'on écrit T en fonction de f et g_S , on voit que T a la forme d'un tenseur énergie-impulsion bien connu des physiciens : le tenseur énergie-impulsion d'un *fluide parfait*. Un tel fluide

¹⁴le célèbre *rayonnement cosmique de fond*

est caractérisé par sa *densité d'énergie* ρ et sa *pression* p . Ici, on a

$$\rho = \frac{3}{8\pi} \left(\left(\frac{f'}{f} \right)^2 + \frac{k}{f^2} \right) \quad \text{et} \quad p = -\frac{1}{8\pi} \left(2\frac{f''}{f} + \left(\frac{f'}{f} \right)^2 + \frac{k}{f^2} \right)$$

On remarque alors le lemme suivant :

Lemme 13. *Supposons que $\rho + 3p$ est strictement positif en tout point de l'espace-temps M , et qu'il existe $t_0 \in I$ tel que $f'(t_0) > 0$. Alors, t_* est fini, et $f(t)$ tend vers 0 lorsque t tend vers t_* .*

Preuve. Des expressions de ρ et p , on tire :

$$\frac{f''}{f} = -\frac{4\pi}{3}(\rho + 3p)$$

Comme f et $\rho + 3p$ sont strictement positifs, on en déduit que f'' est strictement positif, et donc que $f(t) < f(t_0) + f'(t_0)(t - t_0)$. Comme $f'(t_0) > 0$, et comme f doit être strictement positive sur $]t_*, t^*[$, on doit donc avoir $t_* \geq t_0 - f(t_0)/f'(t_0) > -\infty$. Par ailleurs, si f ne tendait pas vers 0 en t_* alors on pourrait prolonger la métrique g sur $S \times]t_* - \varepsilon, t^*[$, ce qui contredirait la maximalité de l'espace-temps (M, g) . \square

Dans les espaces-temps de Robertson Walker, le signe de $f'(t)$ indique si on passe de la métrique sur l'hypersurface $S(t)$ à la métrique sur l'hypersurface $S(t + h)$ via une contraction ou une dilatation. Dans l'univers physique, les observations des astrophysiciens (le célèbre *décalage vers le rouge*) semblent indiquer que les galaxies lointaines s'éloignent nous. Par conséquent, si on modélise l'univers physique par un espace-temps de Robertson-Walker, alors f' devrait être positif au point de l'espace-temps où nous nous trouvons actuellement. Par ailleurs, dans tous les fluides connus, la densité d'énergie est très supérieure à la pression. Les physiciens pensent donc que, si on modélise l'univers physique par un espace-temps de Robertson-Walker, alors on doit supposer que la quantité $\rho + 3p$ est partout positive. Le lemme 13 implique alors que, sous ces conditions, la borne inférieure t_* de l'intervalle I est finie. Dans ce modèle, "l'âge de notre univers" est donc fini. Contrairement aux apparences, cette affirmation ne dépend pas du choix d'une horloge ; en effet, si t_* est fini, alors toute géodésiques de type temps ou lumière dirigée vers le passé est incomplète (c'est-à-dire "sort de l'espace-temps en temps fini"). Physiquement, ceci traduit le fait que, si on suit à rebours la trajectoire d'un

objet matériel ou d'un photon, cette trajectoire "arrive au bord de l'espace-temps" ("à la naissance de l'univers") au bout d'un temps propre fini. De plus, f tend vers 0 lorsque t tend vers t_* . Ceci implique que, la densité d'énergie ρ , la pression p , et les coefficients du tenseur de courbure de g (donc l'intensité des forces gravitationnelles subies par n'importe quel objet matériel) tendent vers l'infini lorsqu'on s'approche de la "naissance de l'univers". C'est pourquoi on parle de *Big Bang*.

Solutions à symétrie exactes II : l'espace-temps de Schwarzschild-Kruskal

En 1916, moins d'un an après la publication de l'article dans lequel Einstein formule la théorie de la Relativité Générale, K. Schwarzschild a proposé une solution de l'équation d'Einstein destinée à modéliser l'univers au voisinage d'un astre sphérique isolé. C'est cette solution qui a été utilisée dans les différents tests de la Relativité Générale (mesure de l'avance du périhélie de Mercure, de la déviation des rayons lumineux lors du passage près d'un astre, etc). C'est également cette solution qui est à l'origine de la découverte du concept de *trou noir*.

Une *espace-temps de Schwarzschild* est un espace-temps (M, g) qui satisfait les conditions suivantes :

- (1) M est difféomorphe à $S \times \mathbb{R}$, où S est le complémentaire d'une boule centrée en 0 dans \mathbb{R}^3 (ceci traduit le fait qu'on se place à l'extérieur d'un astre sphérique). On considèrera des coordonnées sphériques (r, θ, φ) sur \mathbb{R}^3 , et une coordonnée t sur \mathbb{R} .
- (2) La coordonnée t est une horloge globale et $\frac{\partial}{\partial t}$ est un *champ de Killing* (c'est-à-dire que le flot du champ $\frac{\partial}{\partial t}$ est un groupe à un paramètre d'isométries de g).
- (3) La métrique g est à *symétrie sphérique*, c'est-à-dire que, pour tout $\Psi \in O(3)$, l'application $(x, t) \mapsto (\Psi(x), t)$ est une isométrie de g .
- (4) (M, g) est *vide*, c'est-à-dire solution de l'équation d'Einstein avec tenseur énergie-impulsion nul.
- (5) (M, g) est *asymptotiquement isométrique à l'espace de Minkowski*, c'est-à-dire que la métrique g est asymptote à la métrique de Minkowski $-dt^2 + dr^2 + r^2 \sin \theta d\theta d\varphi$ quand r tend vers l'infini.

Ces hypothèses impliquent que la métrique g peut se mettre sous la forme

$$g = - \left(1 - \frac{2m}{r}\right) dt^2 + \left(1 - \frac{2m}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2),$$

où m est un paramètre qui s'interprète comme la masse de l'astre considéré.

La formule ci-dessus dégénère en $r = 2m$. Il est donc naturel de considérer que l'espace-temps de Schwarzschild doit se limiter à la région $\{(r, \theta, \varphi, t) \mid r > 2m\}$ (ce qui n'est guère choquant puisqu'on a supposé dès le début de la construction qu'on se plaçait à l'extérieur d'un astre sphérique). Cependant, si on s'intéresse de plus près à la métrique g définie ci-dessus, on remarque que la métrique g ne dégénère absolument pas lorsqu'on s'approche de $r = 2m$. En fait, la divergence de la formule ci-dessus est un simple artefact du choix du système de coordonnées. On le vérifie facilement en considérant les *coordonnées de Kruskal* (r, θ, φ, v) où $v = t + 2m \log(r - 2m)$; dans ces coordonnées, la métrique g s'écrit

$$g = - \left(1 - \frac{2m}{r}\right) dv^2 + 2dvdr + r^2(d\theta^2 + \sin^2\theta d\varphi^2)$$

Cette formule est parfaitement régulière en $r = 2m$ (elle ne dégénère qu'en $r = 0$). On peut donc prolonger la métrique g sur $M' = (\mathbb{R}^3 \setminus \{0\}) \times \mathbb{R}$ par la formule ci-dessus (et, puisque toutes les formules sont analytiques, le prolongement est unique).

Remarque. Pour trouver la formule définissant la métrique g dans le système de coordonnées (r, θ, φ, t) , nous avons explicitement supposé que la coordonnée t définissait une horloge globale, et donc, en particulier, que le vecteur $\frac{\partial}{\partial t}$ était de type temps. Or le vecteur $\frac{\partial}{\partial t}$ n'est de type temps que dans la zone $r > 2m$ (il est de type espace dans la zone $r < 2m$). Il n'est donc guère étonnant que la formule définissant g dans les coordonnées (r, θ, φ, t) dégénère sur l'hypersurface $r = 2m$.

Notons, par contre, que la dégénérescence des formules ci-dessus en $r = 0$ n'est pas due à des choix artificiels de coordonnées. On peut en effet vérifier que la courbure scalaire de g (qui est un objet intrinsèque) explose lorsque r tend vers 0.

En étudiant l'espace-temps (M', g) , on remarque que celui-ci possède un certain nombre de propriétés dont les interprétations physiques sont pour le moins surprenantes :

— Tout d’abord, on constate que les courbes causales dirigées vers le futur peuvent entrer dans la région $r < 2m$, mais ne peuvent jamais sortir de cette région. D’un point de vue physique, ceci signifie qu’un objet matériel ou un rayon lumineux peut entrer dans la région $r < 2m$ (qui est un voisinage de l’astre considéré), mais ne peut jamais sortir de cette région. C’est pourquoi, on qualifie la région $r < 2m$ de *trou noir*. L’hypersurface $r = 2m$ est l’*horizon* du trou noir.

— Une fois entré dans la région $r < 2m$, une courbe causale dirigée vers le futur n’a d’autre avenir que d’aller buter en temps fini vers la “ligne de singularités” $r = 0$. De plus, lorsqu’on s’approche de $r = 0$, le tenseur de courbure de g explose. D’un point de vue physique, ceci signifie que, dans le modèle de Schwarzschild-Kruskal, tout objet matériel ou rayon lumineux qui pénètre dans le “trou noir” va cesser d’exister au bout d’un temps fini, et, vers la fin de sa courte vie, subir des forces dont l’intensité tend vers l’infini !

Remarque. La validité des différentes interprétations physiques ci-dessus repose sur des hypothèses plus ou moins fortes suivant les cas :

— l’existence d’un trou noir (*i.e.* d’une zone d’où aucun objet matériel ni rayon lumineux ne peut sortir) suppose que la métrique g modélise fidèlement l’univers sur $(\mathbb{R}^3 \setminus B) \times \mathbb{R}$ où B est une boule centrée en 0 de rayon strictement plus petit que $2m$;

— l’existence de trajectoires d’objets matériels qui cessent d’exister au bout d’un temps fini et qui subissent des forces d’intensité infinie suppose que la métrique g modélise fidèlement l’univers sur $(\mathbb{R}^3 \setminus \{0\}) \times \mathbb{R}$ en entier (y compris à l’intérieur de l’astre sphérique qu’on a considéré au départ), et donc que le tenseur énergie-impulsion est nul sur $(\mathbb{R}^3 \setminus \{0\}) \times \mathbb{R}$ en entier (y compris à l’intérieur de l’astre considéré).

La majorité des astrophysiciens semble aujourd’hui croire en l’existence de trous noirs dans notre univers : tout astre suffisamment massif terminerait sa vie entouré d’un horizon d’où ne peut ressortir aucun objet matériel ou rayon lumineux. Pour une présentation un peu moins sommaire des espaces-temps de Robertson-Walker et de Schwarzschild-Kruskal, le lecteur est invité à consulter par exemple [7], [12],[14] ou [16]. On trouvera par ailleurs dans ces ouvrages de nombreux autres modèles d’espace-temps à symétries exactes.

Le problème de Cauchy

Une autre approche de l’équation d’Einstein, plus proche du traitement usuel des EDP, consiste à se restreindre aux espace-temps (M, g) , tels que M est

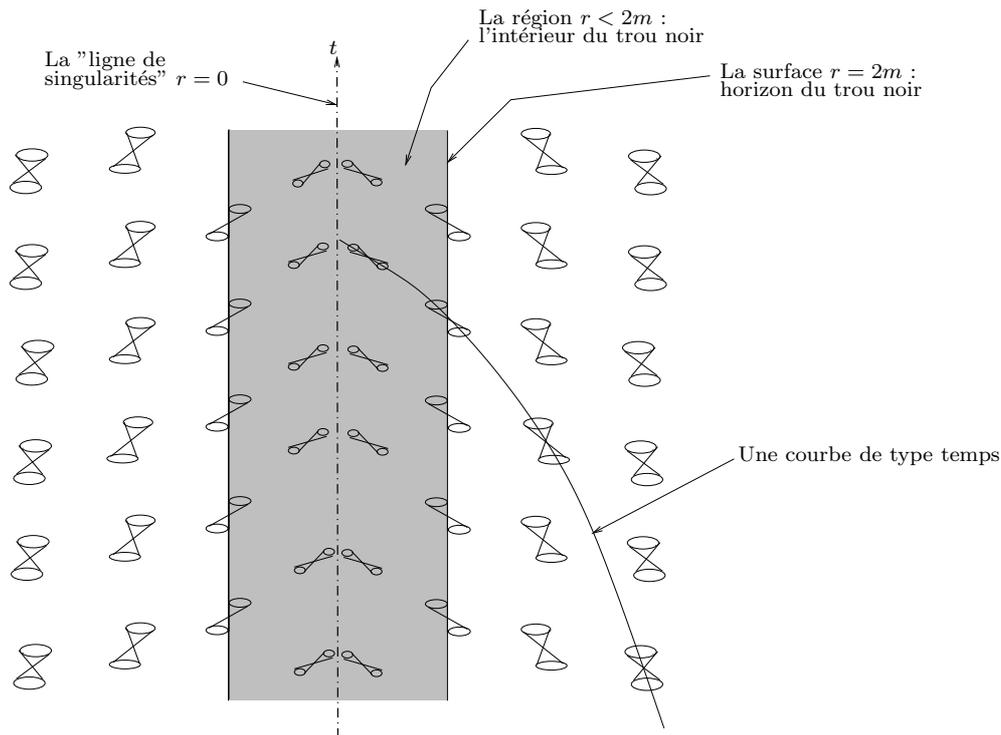


Figure 0.4: La figure représente l'espace-temps de Schwarzschild-Kruskal lorsqu'on "oublie" les variables θ et φ . On a dessiné une courbe de type temps typique qui traverse la surface $r = 2m$, se trouve dès lors emprisonnée dans la région $r < 2m$, et va, au bout d'un temps fini, s'échouer sur la "ligne de singularités" $r = 0$.

difféomorphe à un produit $\Sigma \times \mathbb{R}$ (où Σ est une variété de dimension 3), et tels que chaque hypersurface $\Sigma \times \{t\}$ est de type espace pour la métrique g . Ainsi, on peut voir la métrique lorentzienne g comme une famille de métriques riemanniennes (les métriques induites par g sur les "tranches" $\Sigma \times \{t\}$), et l'équation d'Einstein comme une équation dévolution dans l'espace des métriques riemanniennes sur la 3-variété Σ . Dans ce cadre, il est alors possible de formuler un *problème de Cauchy* (pour simplifier, on ne considère ici que l'équation d'Einstein dans le vide, c'est-à-dire avec tenseur énergie-impulsion identiquement nul) :

Formulation du problème de Cauchy. Soient Σ une variété de dimension 3, γ une métrique riemannienne sur Σ , et B un champ de formes bilinéaires symétriques sur Σ . Existe-t-il une métrique lorentzienne g , définie sur un ouvert M de $\Sigma \times \mathbb{R}$ contenant $\Sigma \times \{0\}$, qui satisfait l'équation d'Einstein telle que γ soit la métrique induite par g sur $\Sigma \times \{0\}$, et telle que B soit la seconde forme fondamentale de $\Sigma \times \{0\}$ pour la métrique g ?

Pour que le problème de Cauchy formulé ci-dessus admette une solution, la métrique γ et le champ de forme bilinéaires B ne peuvent être choisis de manière arbitraire. En effet, la métrique induite et la seconde forme fondamentale d'une hypersurface dans une variété pseudo-riemannienne satisfont toujours certaines relations : les équations de Gauss-Codazzi. Ces relations donnent lieu à des *équations de contraintes* (quatre équations scalaires) qui doivent être satisfaites par γ et B .

Même lorsque les équations de contraintes sont satisfaites, l'existence et l'unicité des solutions de l'équation d'Einstein sont loin d'être des problèmes triviaux. Le problème majeur est l'invariance de l'équation d'Einstein sous l'action du groupe des difféomorphismes : pour tout difféomorphisme h de M , si g est une métrique lorentzienne solution de l'équation d'Einstein, alors h^*g est aussi solution de cette équation¹⁵. L'espace des solutions, s'il est non-vide, est donc automatiquement de dimension infinie. Pour obtenir des résultats d'existence et unicité, il faut donc "fixer une jauge", c'est-à-dire choisir un système de coordonnées. Bien sûr, on veut que, dans ce système de coordonnées, l'équation d'Einstein est une forme "la plus sympathique possible". Le choix le plus classique consiste à travailler dans un système de *coordonnées harmoniques* (c'est-à-dire, dans un système de coordonnées que l'on suppose isothermes pour le Laplacien de la métrique lorentzienne g recherchée). Dans de telles coordonnées, l'équation d'Einstein est de type *quasi-linéaire hyperbolique*¹⁶, ce qui a permis à Y. Choquet-Bruhat et R. Geroch de montrer le résultat fondamental suivant :

Théorème 14 (Y. Choquet-Bruhat, [9]). Soient Σ , γ et B une variété de dimension 3, une métrique riemannienne et un champ de formes bilinéaire sur Σ qui satisfont les équations de contraintes. Alors, le problème de Cauchy formulé ci-dessus admet une solution (M, g) . De plus, (M, g) est globalement hyperbolique et $\Sigma \times \{0\}$ est une surface de Cauchy dans M .

¹⁵Les physiciens disent que "l'équation d'Einstein a un groupe de jauge de dimension infinie"

¹⁶J'avoue ne pas bien comprendre ce que cela signifie.

Théorème 15 (Y. Choquet-Bruhat et R. Geroch, [8]). *Soient Σ , γ et B qui satisfont les équations de contraintes. Alors le problème de Cauchy admet une unique solution maximale parmi les solutions globalement hyperboliques pour lesquelles $\Sigma \times \{0\}$ est une surface de Cauchy.*

L'étude (théorique et numérique) du problème de Cauchy pour l'équation d'Einstein est un problème très riche, dont je ne dirai rien de plus¹⁷. Pour plus de détails, je renvoie par exemple aux chapitres traitant du problème de Cauchy dans [12] ou [16], ou bien aux articles de survol [4] et [1].

L' hypothèse de positivité de l'énergie

L'équation d'Einstein peut également être utilisée de manière indirecte dans des situations où on ne veut pas supposer que l'on connaît la forme exacte du tenseur énergie-impulsion T . Cette équation permet en effet de traduire n'importe quelle hypothèse "physique" portant sur le tenseur énergie-impulsion en une hypothèse géométrique portant sur la métrique lorentzienne g . Réciproquement, elle permet de vérifier qu'une hypothèse géométrique sur la métrique g correspond à une condition "physiquement raisonnable" sur le tenseur énergie-impulsion.

Par exemple, l'hypothèse géométrique suivante :

$$\text{pour tout vecteur causal } v, \text{ on a } \text{Ric}_g(v, v) \geq 0$$

appelée *hypothèse de positivité de l'énergie*, joue un rôle crucial dans la preuve de nombreux résultats (comme, par exemple, les théorèmes d'existence de singularités qu'on évoquera dans la partie 0.8). En utilisant l'équation d'Einstein, on voit que cette hypothèse géométrique est équivalente à l'hypothèse "physique" suivante :

$$\text{pour tout vecteur causal } v, \text{ on a } T(v, v) \geq \left(\frac{1}{2}C(T) - \frac{\Lambda}{8\pi} \right) \cdot g(v, v)$$

où $C(T)$ désigne la *contraction* du tenseur T (c'est-à-dire $C(T) = \sum_{i=1}^4 T(e_i, e_i)$) où $(e_i)_{i=1\dots 4}$ est une base de l'espace tangent à M au point considéré). Les physiciens affirment que cette hypothèse est satisfaite par tous les tenseurs énergie-impulsion classiques. C'est pourquoi, il est généralement admis que les espaces-temps physiques satisfont l'hypothèse de positivité de l'énergie¹⁸.

¹⁷principalement, parce que j'en serais incapable

¹⁸Ceci étant, Hawking a expliqué que la dilatation de la durée de vie des anti-particules au voisinage d'un trou noir, pourrait entraîner la violation de cette hypothèse

Remarque. En fait, l’hypothèse de positivité de l’énergie semble traduire le fait que “la gravitation est attractive”. En effet, si v est le vecteur tangent d’une géodésique γ , alors la positivité de $\text{Ric}_g(v, v)$ implique que les géodésiques “infiniment proches de γ ” ont tendance, en moyenne, à se rapprocher de γ .

0.7 Existence de points conjugués et focaux

Dans la partie 0.3, nous avons vu rôle crucial joué par les points conjugués et des points focaux dans la délimitation du futur d’un point ou d’une surface de type espace (propositions 4 et 5). Par ailleurs, dans la partie 0.6, nous avons dégagé une hypothèse portant sur la courbure de Ricci, qui semble correspondre à une hypothèse physique très raisonnable : l’hypothèse de positivité de l’énergie. Nous allons maintenant relier tout cela : nous allons montrer comment l’hypothèse de positivité de l’énergie force, dans de nombreux cas, l’existence de points conjugués et de points focaux le long des géodésiques de type temps et lumière.

Points conjugués le long d’une géodésique de type temps

On commence par l’étude la plus simple : celle de l’existence de points conjugués le long d’une géodésique de type temps. Soit donc $\gamma : I \rightarrow M$ une géodésique de type temps et $p = \gamma(s_0)$ un point de γ . On veut savoir s’il existe un point $q = \gamma(s_1)$ conjugué à p le long de γ . Pour cela, il faut tenir compte de tous les champs de Jacobi le long de γ qui s’annulent en p et sont orthogonaux à γ (il y en a une infinité). Heureusement, en utilisant la linéarité de l’équation différentielle satisfaite par les champs de Jacobi, nous allons nous ramener à l’étude d’un champ d’applications linéaires le long de γ .

Pour tout champ de vecteurs V défini le long de γ , on note $V(s)$ le vecteur $V(\gamma(s))$. Pour tout $s \in I$, on note $E(s)$ l’orthogonal dans $T_{\gamma(s)}M$ du vecteur $\frac{d\gamma}{ds}(s)$. On considère une base $(e_1(s_0), e_2(s_0), e_3(s_0))$ de $E(s_0)$, que transporte parallèlement en une base $(e_1(s), e_2(s), e_3(s))$ de $E(s)$ pour tout $s \in I$. Remarquons que, pour tout $s \in I$, la métrique lorentzienne g induit sur $E(s)$ une forme bilinéaire *définie positive* (ceci jouera un rôle crucial ci-dessous).

On notera \mathcal{J} l’ensemble des champs de Jacobi le long de γ qui s’annulent en p et sont orthogonaux à γ . Les champs de Jacobi satisfont une équation différentielle linéaire du second ordre. L’application qui à un champ de Jacobi $J \in \mathcal{J}$ associe sa dérivée $\frac{DJ}{ds}(s_0)$ en s_0 réalise donc un isomorphisme entre l’espace vectoriel \mathcal{J} et l’espace vectoriel $E(s_0)$. En particulier, \mathcal{J} est un espace

vectorel de dimension 3.

On choisit une base $(s \mapsto J_1(s), s \mapsto J_2(s), s \mapsto J_3(s))$ de \mathcal{J} . Pour tout $s \in I$, on considère alors l'endomorphisme $A(s)$ de l'espace vectoriel $E(s)$ qui envoie les vecteurs $e_1(s), e_2(s), e_3(s)$ sur les vecteurs $J_1(s), J_2(s), J_3(s)$. Pour tout champ parallèle $s \mapsto V(s)$ le long de γ à valeur dans $E(s)$, le champ $s \mapsto A(s).V(s)$ est un champ de Jacobi appartenant à \mathcal{J} . Réciproquement, tout champ de Jacobi $J \in \mathcal{J}$ s'écrit alors $s \mapsto A(s).V(s)$ où $s \mapsto V(s)$ est un champ de vecteurs parallèle le long de γ . Par conséquent, il existe un point conjugué à p le long de γ si et seulement si il existe $s \neq s_0$ tel que $A(s)$ n'est pas inversible. C'est pourquoi on va s'intéresser à l'annulation du déterminant de $A(s)$.

Rappelons que, pour tout $s \in I$, la métrique g induit un produit scalaire (*défini positif*) sur l'espace vectoriel $E(s)$. Les traces, adjoints, etc. d'endomorphismes de $E(s)$ considérés ci-dessous sont relatifs à ce produit scalaire. Pour tout s tel que $A(s)$ est inversible, on note

$$\theta(s) := \operatorname{tr} \left(\frac{DA}{ds} \cdot A^{-1} \right) (s) = \frac{1}{\det A} \cdot \frac{d}{ds} \det(A)(s)$$

(l'égalité entre les deux expressions est un exercice d'algèbre linéaire élémentaire laissé au lecteur). On dit que $\theta(s)$ est la *divergence*¹⁹ au point $\gamma(s)$ des géodésiques de passant par p . On note par ailleurs

$$\Sigma(s) := \left(\frac{DA}{ds} \cdot A^{-1} - \frac{1}{3}\theta \right) (s)$$

la partie sans trace de $\frac{DA}{ds} \cdot A^{-1}(s)$, et

$$\sigma(s) := \sqrt{\operatorname{tr}(\Sigma\Sigma^*)}(s)$$

la norme euclidienne de l'endomorphisme Σ . On dit que $\sigma(s)$ est la *distortion* au point $\gamma(s)$ des géodésiques passant par p .

Afin de trouver des conditions suffisantes à l'annulation de $\theta(s)$ (et donc à l'existence de points conjugués à p), nous allons établir une équation différentielle satisfaite par $\theta(s)$.

¹⁹Rappelons que, pour toute famille $(\gamma_u)_{u \in [-\varepsilon, \varepsilon]}$ de géodésiques passant par p telle que $\gamma_0 = \gamma$, le champ $s \mapsto \frac{\partial}{\partial u} \gamma_u$ est un champ de Jacobi appartenant à \mathcal{J} (quitte à reparamétriser les géodésiques γ_u). On en déduit que le signe de $\theta(s)$ indique si, en moyenne, les géodésiques passant par p (et proches de γ) ont tendance à s'éloigner ou se rapprocher de γ . Ce qui justifie le terme "divergence".

Proposition 16. *Au voisinage de toute valeur de s telle que $A(s)$ soit inversible, la divergence $s \mapsto \theta(s)$ est solution de l'équation de Raychudhuri*

$$\frac{d\theta}{ds} = -\text{Ric}_g \left(\frac{d\gamma}{ds}, \frac{d\gamma}{ds} \right) - \sigma^2 - \frac{1}{3}\theta^2$$

Au cours de la preuve de la proposition 16, nous utiliserons le lemme et les deux remarques ci-dessous :

Lemme 17. *Pour tout $s \in I$, l'endomorphisme $\frac{DA}{ds} \cdot A^{-1}(s)$ est auto-adjoint.*

Remarque. Le lemme 17 implique que l'endomorphisme $\Sigma(s)$ est également auto-adjoint pour tout s , et qu'on a donc

$$\sigma(s) = \sqrt{\text{tr}(\Sigma^2)}(s).$$

Remarque. Pour tout $s \in I$, par définition de la courbure de Ricci, on a

$$\text{Ric}_g \left(\frac{d\gamma}{ds}, \frac{d\gamma}{ds} \right) = \text{tr} \left(\mathbf{R}_g \left(\frac{d\gamma}{ds}, \bullet \right) \cdot \frac{d\gamma}{ds} \right) \quad (\star)$$

où la trace est prise sur l'espace vectoriel $T_{\gamma(s)}M$ en entier. Cependant, par antisymétrie du tenseur de courbure, on a $\mathbf{R} \left(\frac{d\gamma}{ds}, \frac{d\gamma}{ds} \right) \cdot \frac{d\gamma}{ds} = 0$. Dans l'égalité (\star) ci-dessus, il suffit donc de prendre la trace en restriction à un supplémentaire de $\mathbb{R} \cdot \frac{d\gamma}{ds}$ dans $T_{\gamma(s)}M$ (par exemple, en restriction à $E(s)$).

Preuve de la proposition 16. Pour tout champ de vecteurs parallèle $s \mapsto V(s)$ le long de γ , le champ $s \mapsto A.V(s)$ est un champ de Jacobi, et satisfait donc l'équation différentielle

$$\frac{D^2}{ds^2} A.V = -\mathbf{R}_g \left(\frac{d\gamma}{ds}, A.V \right) \cdot \frac{d\gamma}{ds}$$

Par suite, au voisinage d'une valeur de s où $A(s)$ est inversible, V étant parallèle, on a

$$\frac{D^2 A}{ds^2} \cdot A^{-1} = -\mathbf{R}_g \left(\frac{d\gamma}{ds}, \bullet \right) \cdot \frac{d\gamma}{ds}$$

et, par suite,

$$\begin{aligned} \frac{D}{ds} \left(\frac{DA}{ds} \cdot A^{-1} \right) &= \frac{D^2 A}{ds^2} \cdot A^{-1} - \frac{DA}{ds} \cdot A^{-1} \cdot \frac{DA}{ds} \cdot A^{-1} \\ &= -\mathbf{R}_g \left(\frac{d\gamma}{ds}, \bullet \right) \cdot \frac{d\gamma}{ds} - \left(\frac{DA}{ds} \cdot A^{-1} \right)^2 \end{aligned}$$

En prenant la trace (par rapport, à la métrique induite par g sur $E(s)$) des deux membres de cette équation, et en utilisant les deux remarques ci-dessus, on obtient

$$\begin{aligned} \frac{D}{ds}\theta &= -\operatorname{tr}\left(\operatorname{R}_g\left(\frac{d\gamma}{ds}, \bullet\right) \cdot \frac{d\gamma}{ds}\right) - \operatorname{tr}\left(\left(\frac{DA}{ds} \cdot A^{-1}\right)^2\right) \\ &= -\operatorname{Ric}_g\left(\frac{d\gamma}{ds}, \frac{d\gamma}{ds}\right) - \operatorname{tr}\left(\left(\Sigma + \frac{1}{3}\theta\operatorname{Id}\right)^2\right) \\ &= -\operatorname{Ric}_g\left(\frac{d\gamma}{ds}, \frac{d\gamma}{ds}\right) - \sigma^2 - \frac{1}{3}\theta^2 \end{aligned}$$

□

Preuve du lemme 17. L'ingrédient fondamental de la preuve est la “symétrie par paire” du tenseur de courbure :

$$g(\operatorname{R}_g(x, y) \cdot z, t) = g(\operatorname{R}_g(x, t) \cdot z, y)$$

On commence par montrer que la matrice $B = A^* \frac{DA}{ds} - \frac{DA^*}{ds} A$ est nulle pour toute valeur du paramètre s . On a

$$\frac{DB}{ds} = A^* \frac{D^2 A}{ds^2} - \frac{D^2 A^*}{ds^2} A$$

Or, pour tous vecteurs v, w , on a

$$\begin{aligned} &g\left(\left(A^* \frac{D^2 A}{ds^2} - \frac{D^2 A^*}{ds^2} \cdot A\right) \cdot v, w\right) \\ &= g\left(A^* \frac{D^2 A}{ds^2} \cdot v, w\right) - g\left(A^* \cdot \frac{D^2 A}{ds^2} \cdot w, v\right) \\ &= g\left(A^* \cdot R\left(\frac{d\gamma}{ds}, v\right) \cdot \frac{d\gamma}{ds}, w\right) - g\left(A^* \cdot R\left(\frac{d\gamma}{ds}, w\right) \cdot \frac{d\gamma}{ds}, v\right) \\ &= 0 \end{aligned}$$

(la dernière égalité provient de la “symétrie par paire” du tenseur de courbure rappelée ci-dessus). La dérivée de $B(s)$ est donc nulle pour tout s . Comme $B(s_0)$ est nulle (car $A(s_0)$ l'est), on en déduit que $B(s)$ est nulle pour tout s . Enfin, pour tout $s \in I$, on a

$$\begin{aligned} \frac{DA}{ds} \cdot A^{-1} - \left(\frac{DA}{ds} \cdot A^{-1}\right)^* &= \frac{DA}{ds} \cdot A^{-1} - (A^{-1})^* \cdot \frac{DA^*}{ds} \\ &= (A^{-1})^* \cdot B \cdot A^{-1} \\ &= 0 \end{aligned}$$

ce qui prouve que $\frac{DA}{ds}.A^{-1}(s)$ est auto-adjoint pour tout s . \square

L'équation de Raychaudhuri et l'hypothèse de positivité de l'énergie vont maintenant nous permettre de trouver un critère d'existence de point conjugué.

Proposition 18. *Supposons que l'espace-temps (M, g) satisfait l'hypothèse de positivité de l'énergie : $\text{Ric}_g(v, v) \geq 0$ pour tout v de type temps ou lumière. Supposons que la géodésique γ est complète (c'est-à-dire que $I = \mathbb{R}$). Enfin, supposons qu'il existe une valeur $s_1 > s_0$ du paramètre²⁰ tel que la divergence $\theta(s_1)$ est strictement négative. Alors, il existe un point $q = \gamma(s_2)$ conjugué à p le long de γ . De plus, on a $s_2 < s_1 + \frac{3}{-\theta(s_1)}$.*

Preuve. La divergence $s \mapsto \theta(s)$ satisfait à l'équation de Raychaudhuri

$$\frac{d\theta}{ds}(s) = -\text{Ric}_g\left(\frac{d\gamma}{ds}(s), \frac{d\gamma}{ds}(s)\right) - \sigma(s)^2 - \frac{1}{3}\theta(s)^2$$

Puisque (M, g) satisfait l'hypothèse de positivité de l'énergie, le terme $-\text{Ric}_g\left(\frac{d\gamma}{ds}(s), \frac{d\gamma}{ds}(s)\right)$ est négatif. Donc $s \mapsto \theta(s)$ satisfait l'inéquation différentielle

$$\frac{d\theta}{ds}(s) < \frac{1}{3}\theta(s)^2$$

Ceci implique que, si θ garde une valeur finie sur $[s_0, s]$, on a

$$\theta(s) \leq \frac{3}{s - \left(s_1 + \frac{3}{-\theta(s_1)}\right)}$$

Par conséquent, $\theta(s)$ "explose" nécessairement pour une s_2 du paramètre telle que $s_1 \leq s_1 + \frac{3}{-\theta(s_1)}$. On a vu que ceci implique que le point $q = \gamma(s_2)$ est un point conjugué à p le long de γ . \square

Points conjugués le long d'une géodésique de type lumière

On considère maintenant une géodésique de type *lumière* $\gamma : I \rightarrow M$ et un point $p = \gamma(s_0)$ sur γ . Comme précédemment, on cherche des conditions suffisantes à l'existence d'un point conjugué à p le long de γ . Pour cela, on reprend le raisonnement effectué précédemment dans le cas où γ était de type temps. Il apparaît cependant une difficulté supplémentaire : la restriction de $g_{\gamma(s)}$ à $E(s)$ est maintenant une forme bilinéaire *dégénérée*. Pour contourner cette

²⁰Rappelons que s_0 est la valeur du paramètre telle que $p = \gamma(s_0)$.

difficulté, on va devoir remplacer l'espace vectoriel $E(s)$ par l'espace vectoriel $\widehat{E}(s)$ obtenu comme quotient abstrait de $E(s)$ par la direction $\mathbb{R} \cdot \frac{d\gamma}{ds}(s)$.

On appellera *champ de vecteurs le long de γ à valeurs dans \widehat{E}* toute application de la forme $s \mapsto \widehat{J}(s)$ avec $\widehat{J}(s) \in \widehat{E}(s)$ pour tout s . On peut alors vérifier (et j'invite le lecteur à le faire) les points suivants :

- pour tout $s \in I$, la restriction de $g_{\gamma(s)}$ au sous-espace $E(s)$ passe au quotient en une forme bilinéaire définie positive sur $\widehat{E}(s)$ (qu'on notera $\widehat{g}_{\gamma(s)}$) ;
- la dérivation covariante $\frac{D}{ds}$ passe au quotient en une dérivation covariante $\frac{\widehat{D}}{ds}$ opérant sur les champs le long de γ à valeurs dans \widehat{E} ; ceci définit, pour tous s_1, s_2 , un transport parallèle le long de γ de $\widehat{E}(s_1)$ dans $\widehat{E}(s_2)$;
- l'endomorphisme

$$R \left(\frac{d\gamma}{ds}(s), \bullet \right) \cdot \frac{d\gamma}{ds}(s) : E(s) \rightarrow E(s)$$

passe au quotient en un endomorphisme

$$\widehat{R} \left(\frac{d\gamma}{ds}(s), \bullet \right) \cdot \frac{d\gamma}{ds}(s) : \widehat{E}(s) \rightarrow \widehat{E}(s) ;$$

— la trace de ce dernier endomorphisme (calculée à l'aide de la forme bilinéaire $\widehat{g}_{\gamma(s)}$) de cet endomorphisme de $\widehat{E}(s)$ vaut encore

$$\text{Ric} \left(\frac{d\gamma}{ds}, \frac{d\gamma}{ds} \right) (s).$$

— un champ \widehat{J} le long de γ à valeurs dans \widehat{E} est la projection d'un champ de Jacobi (à valeurs dans E) si et seulement si il satisfait l'équation différentielle

$$\frac{\widehat{D}^2 \widehat{J}}{ds^2}(s) = -\widehat{R} \left(\frac{d\gamma}{ds}, \widehat{J} \right) \cdot \frac{d\gamma}{ds}(s)$$

(on dira alors que \widehat{J} est un champ de Jacobi le long de γ à valeurs dans \widehat{E}) ;

— si \widehat{J} est un champ de Jacobi le long de γ à valeurs dans \widehat{E} , et si \widehat{J} s'annule pour deux valeurs s_1, s_2 du paramètre, alors il existe un champ de Jacobi J à valeurs dans E qui relève \widehat{J} et qui s'annule en s_1 et s_2 .

On note maintenant $\widehat{\mathcal{J}}$ l'ensemble des champs de Jacobi le long de γ à valeurs dans \widehat{E} , qui s'annulent en p . Cet ensemble est naturellement un espace vectoriel de dimension 2 ; on en choisit une base $s \mapsto \widehat{J}_1(s), s \mapsto \widehat{J}_2(s)$.

Par ailleurs, on choisit une base $(\widehat{e}_1(s_0), \widehat{e}_2(s_0))$ de $\widehat{E}(s_0)$, que l'on transporte parallèlement en une base $(\widehat{e}_1(s), \widehat{e}_2(s))$ pour tout s (c'est-à-dire que $s \mapsto \widehat{e}_i(s)$ sont solutions de l'équation différentielle $\frac{D\widehat{e}_i}{ds}(s) = 0$). Puis, pour tout s , on note $\widehat{A}(s)$ l'endomorphisme de $\widehat{E}(s)$ qui envoie les vecteurs $\widehat{e}_1(s), \widehat{e}_2(s)$ sur les vecteurs $\widehat{J}_1(s), \widehat{J}_2(s)$.

On définit la divergence $\widehat{\theta}(s)$ au point $\gamma(s)$ des géodésiques de type lumière passant par p :

$$\widehat{\theta}(s) := \operatorname{tr} \left(\frac{D\widehat{A}}{ds} \cdot \widehat{A}^{-1} \right) (s) = \frac{1}{\det \widehat{A}} \cdot \frac{d}{ds} \det(\widehat{A})(s)$$

Puis, on considère la partie sans trace de $\frac{D\widehat{A}}{ds} \cdot \widehat{A}^{-1}(s)$:

$$\widehat{\Sigma}(s) := \left(\frac{D\widehat{A}}{ds} \cdot \widehat{A}^{-1} - \frac{1}{2} \widehat{\theta} \right) (s)$$

et enfin la norme euclidienne de $\widehat{\Sigma}(s)$

$$\widehat{\sigma}(s) := \sqrt{\operatorname{tr} \left(\widehat{\Sigma} \widehat{\Sigma}^* \right) (s)}.$$

En raisonnant comme dans la preuve de la proposition 16, on montre que $\widehat{\theta}$ satisfait l'équation de Raychaudhuri

$$\frac{d\widehat{\theta}}{ds} = -\operatorname{Ric}_g \left(\frac{d\gamma}{ds}, \frac{d\gamma}{ds} \right) - \widehat{\sigma}^2 - \frac{1}{2} \widehat{\theta}^2$$

On en déduit alors un analogue de la proposition 18 :

Proposition 19. *Supposons que l'espace-temps (M, g) satisfait l'hypothèse de positivité de l'énergie : $\operatorname{Ric}_g(v, v) \geq 0$ pour tout v de type temps ou lumière. Supposons que la géodésique γ est complète (c'est-à-dire que $I = \mathbb{R}$). Enfin, supposons qu'il existe une valeur $s_1 > s_0$ du paramètre tel que la divergence $\widehat{\theta}(s_1)$ est strictement négative. Alors, il existe un point $q = \gamma(s_2)$ conjugué à p le long de γ . De plus, on a $s_2 < s_1 + \frac{2}{-\widehat{\theta}(s_1)}$.*

0.7.1 Points focaux à une surface de type espace

On considère maintenant une surface Σ de type espace dans M , et une géodésique lumière $\gamma : I \rightarrow M$ orthogonale à la surface Σ en un point

$p = \gamma(s_0)$. On cherche des conditions suffisantes à l'existence de points focaux à Σ le long de γ .

On rappelle qu'en chaque point q de Σ , il existe exactement deux vecteurs unitaires de type lumière, dirigés vers le futur, orthogonaux à Σ , qu'on note $n(q)$ et $n'(q)$ (de manière à ce que $n(q)$ et $n'(q)$ dépendent continûment de q). Pour tout $q \in \Sigma$, on définit des opérateurs de Weingarten $W_q, W'_q : T_q\Sigma \rightarrow T_q\Sigma$ par $W_q(v) = D_v n$ et $W'_q(v) = D_v n'$.

Quitte à échanger n et n' , on peut supposer que $\frac{d\gamma}{ds}(s_0)$ est colinéaire à $n(q)$. On rappelle maintenant que le point $q = \gamma(s_1)$ est un point focal à Σ (le long de γ) s'il existe un champ de Jacobi J le long de γ tel que $J(s_0) \in T_p\Sigma$, tel que $\frac{DJ}{ds}(s_0) = W_p(J(s_0))$, et tel que $J(s_1) = 0$.

On note maintenant $\widehat{\mathcal{J}}$ l'ensemble des champs de Jacobi \widehat{J} à valeurs dans \widehat{E} qui satisfont

$$\frac{\widehat{D}\widehat{J}}{ds}(s_0) = \widehat{W}_p(\widehat{J}(s_0))$$

Notons que le plan $T_p\Sigma$ est un supplémentaire à la droite $\mathbb{R} \cdot \frac{d\gamma}{ds}(s_0)$ dans $E(s_0)$; par conséquent, $\widehat{E}(s_0)$ s'identifie à $T_p\Sigma$. L'opérateur $W_p : T_p\Sigma \rightarrow T_p\Sigma$ induit ainsi un opérateur $\widehat{W}_p : \widehat{E}(s_0) \rightarrow \widehat{E}(s_0)$. On note alors $\widehat{\mathcal{J}}$ l'ensemble des champs de Jacobi \widehat{J} à valeurs dans \widehat{E} qui satisfont

$$\frac{\widehat{D}\widehat{J}}{ds}(s_0) = \widehat{W}_p(\widehat{J}(s_0))$$

C'est clairement un espace vectoriel de dimension 2, et on vérifie que $q = \gamma(s_1)$ est un point focal à γ si et seulement si il existe un élément de $\widehat{\mathcal{J}}$ qui s'annule en s_1 .

En considérant une base de $\widehat{\mathcal{J}}$, on définit alors, pour tout $s \in I$, les endomorphismes $\widehat{A}(s)$ et $\widehat{\Sigma}(s)$, ainsi que les nombres $\widehat{\theta}(s)$, et $\widehat{\sigma}(s)$. On dit que $\widehat{\theta}(s)$ est la *divergence au point $\gamma(s)$ des géodésiques de type lumière orthogonales à Σ* . On remarque que, par construction de $\widehat{\mathcal{J}}$ et de \widehat{A} , on a

$$\frac{\widehat{D}\widehat{A}}{ds} \cdot \widehat{A}^{-1}(s_0) = W_p$$

En particulier, $\widehat{\theta}(s_0)$ n'est autre que la trace de "l'opérateur de Weingarten" W_p (c'est donc une sorte de "courbure moyenne" de la surface Σ associé à la normale de type lumière n).

En raisonnant exactement comme dans les parties précédentes, on montre que $\widehat{\theta}$ satisfait l'équation de Raychaudhuri

$$\frac{d\widehat{\theta}}{ds} = -\text{Ric}_g \left(\frac{d\gamma}{ds}, \frac{d\gamma}{ds} \right) - \widehat{\sigma}^2 - \frac{1}{2}\widehat{\theta}^2$$

puis on montre :

Proposition 20. *Supposons que l'espace-temps (M, g) satisfait l'hypothèse de positivité de l'énergie : $\text{Ric}_g(v, v) \geq 0$ pour tout v de type temps ou lumière. Supposons que la géodésique γ est complète (c'est-à-dire que $I = \mathbb{R}$). Enfin, supposons qu'il existe une valeur $s_1 > s_0$ du paramètre tel que la divergence $\widehat{\theta}(s_1)$ est strictement négative. Alors, il existe un point $q = \gamma(s_2)$ focal à Σ le long de γ . De plus, on a $s_2 < s_1 + \frac{2}{-\widehat{\theta}(s_1)}$.*

En particulier, si la trace de l'opérateur de Weingarten W_p est strictement négative, alors il existe un point $q = \gamma(s_2)$ focal à Σ le long de γ avec $s_2 < s_0 + \frac{2}{-\widehat{\theta}(s_0)}$.

0.8 Théorèmes d'existence de singularités

Comme nous l'avons vu dans la partie 0.6, certains modèles d'espaces-temps (espaces temps de Robertson-Walker, de Schwarzschild-Kruskal) présentent des propriétés assez étonnantes (du moins lorsqu'on interprète ces propriétés en termes physiques) : existence de géodésiques causales incomplètes (certaines trajectoires potentielles d'objets matériels et de photons cessent brutalement d'exister au bout d'un temps fini), explosion de la courbure le long de ces géodésiques incomplètes (certains objets matériels et photons sont soumis à des forces gravitationnelles qui tendent vers l'infini en temps fini). Cependant, certains physiciens ont suggéré dans les années 50-60 que ces propriétés pourraient être de simples artefacts des symétries exactes présentes dans les modèles de Robertson-Walker et de Schwarzschild-Kruskal. C'est ce qui a conduit R. Penrose, puis S. Hawking, à démontrer l'existence de géodésiques causales incomplètes, sous des hypothèses très générales. En particulier, les théorèmes de Penrose et Hawking ne suppose aucunement que l'espace-temps admet des symétries exactes. Ces théorèmes sont généralement regroupés sous l'expression "théorèmes d'existence de singularités".

Remarque. Les physiciens ont pris l'habitude de parler d'existence de *singularités* en cas d'existence de géodésiques causales incomplètes. Il faut prendre garde à ne pas se laisser abuser par ce vocabulaire : l'espace-temps est une

variété parfaitement lisse, et la métrique lorentzienne n'est singulière en aucun point de l'espace-temps.

Par ailleurs, les physiciens associent souvent l'existence de géodésiques causales incomplètes, et la divergence du tenseur de courbure le long de ces géodésiques. Si ces deux phénomènes coexistent effectivement dans de nombreux modèles d'espaces-temps (par exemples, les modèles de Robertson-Walker et de Schwarzschild-Kruskal), ils n'ont aucune raison d'être associés en général. Il existe d'ailleurs des modèles d'espace-temps en dimension 3 (appelés *modèles BTZ*) qui possèdent les caractéristiques usuelles d'un modèle de "trou noir" (géodésiques causales incomplètes, existence d'une zone de laquelle aucun objet ou rayon lumineux ne peut s'échapper, etc.), à ceci près que la courbure reste parfaitement bornée le long des géodésiques causales incomplètes. Les théorèmes d'existence de singularités de Hawking et Penrose ne montrent que l'existence de géodésiques incomplètes, sans rien affirmer sur le comportement du tenseur de courbure²¹.

Afin de pouvoir énoncer le plus simple des théorèmes d'existence de singularités, il nous faut tout d'abord expliquer ce qu'est une *surface piégée*. Cette notion sert à modéliser la situation lors de la formation d'un trou noir. Plus précisément, on prévoit que certains astres, s'effondrant sur eux-mêmes sous l'effet de la gravitation "après avoir épuisé leur carburant nucléaire", deviennent suffisamment denses pour qu'aucune particule matérielle, ni même aucun photon, ne puisse plus s'échapper de leur voisinage. La notion de surface piégée modélise le bord d'une région de l'espace²² dont aucune particule matérielle ou photon ne peut s'échapper. Pour obtenir une définition mathématique indépendante de tout choix de coordonnées, on doit néanmoins s'éloigner un peu de cette intuition physique :

Définition. Une *surface piégée* Σ est une surface compacte de type espace dans M telle que les deux familles de géodésiques de type lumière orthogonales à Σ ont une divergence négative en chaque point de Σ . Autrement dit, une *surface piégée* Σ est une surface compacte de type espace dans M telle que, pour tout $q \in \Sigma$, les opérateurs de Weingarten W_q et W_q' définis ont une trace strictement négative (voir la fin de la partie 0.7 et la fin de l'appendice).

²¹Ceci est certainement dû aux preuves de ces théorèmes qui, bien que très élégantes, sont un peu frustrantes : on suppose que toutes les géodésiques de type temps et lumière sont complètes, et on montre que ceci conduit à une contradiction, mais on n'a rien appris de la localisation des géodésiques incomplètes, ni de la géométrie au voisinage de ces géodésiques.

²²Ici, il s'agit bien de *l'espace*, et non de *l'espace-temps*. Lorsqu'on s'exprime ainsi, on sous-entend donc un choix d'une horloge (locale).

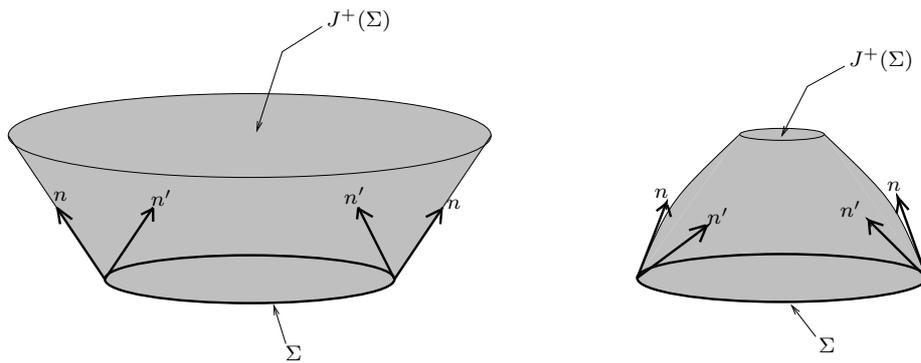


Figure 0.5: On a représenté ci-dessus (dans deux situations différentes) une surface de type espace Σ , les vecteurs n, n' unitaires de type lumière dirigés vers le futur orthogonaux à Σ , et le futur causal $J^+(\Sigma)$. Pour les besoins de la figure, on a “oublié” une dimension ; la surface Σ est donc représentée par une ellipse. À gauche, on a schématisé la situation “normale” ; à droite, on a schématisé la situation lorsque Σ est une surface piégée.

On comprend mieux le lien entre l'idée physique de “surface piégée” et la définition mathématique ci-dessus à la lecture de la preuve du théorème 21 ci-dessous. En effet, on montre dans cette preuve que (sous certaines hypothèses raisonnables) la frontière du futur causal d'une surface piégée est une hypersurface compacte. Les exemples typiques de surfaces piégées sont les surfaces $\{(r, \theta, \varphi, t) \mid r = r_0 \text{ et } t = t_0\}$ avec $r_0 < 2m$ dans le modèle de Schwarzschild-Kruskal. Il semble que les observations astronomiques indiquent maintenant que des surfaces piégées ont de très bonnes chances d'exister dans notre univers (ceci, bien sûr, si l'on admet que notre univers peut être modélisé par la Relativité Générale, c'est-à-dire que cela a un sens d'interpréter les observations expérimentales en termes de sous-variété d'une variété riemannienne, de seconde forme fondamentale, etc.)

Voici le premier théorème d'existence de singularités, que l'on doit à Penrose :

Théorème 21 (Penrose, 1965). *On suppose que l'espace-temps (M, g) satisfait les hypothèses suivantes :*

- *M satisfait l'hypothèse de positivité de l'énergie, i.e. $\text{Ric}_g(v, v) \geq 0$ pour tout v causal ;*
- *il existe une surface piégée Σ dans M ;*
- *M admet une hypersurface de Cauchy S non-compacte.*

Alors, M a une géodésique de type lumière incomplète.

Schéma de la preuve. ²³ On raisonne par l'absurde : on suppose que toutes les géodésiques sont complètes. On étudie alors la frontière du futur de la surface piégée Σ . Cette frontière est un ensemble *achronal* : deux quelconques de ses points ne peuvent jamais être joints par une courbe causale. Comme tout ensemble achronal "sans bord", c'est une hypersurface topologique. Par ailleurs, puisque M admet une hypersurface de Cauchy, M est globalement hyperbolique, et la frontière de $J^+(\Sigma)$ coïncide donc avec $J^+(\Sigma) \setminus I^+(\Sigma)$ (voir propositions 11 et 9). On a vu que tout point de $J^+(\Sigma) \setminus I^+(\Sigma)$ est relié à Σ par un segment de géodésique lumière, orthogonal à Σ , qui ne contient pas de point focal à Σ (proposition 5). Mais, puisque toute géodésique lumière est complète, le fait que Σ soit une surface piégée et l'hypothèse de positivité de l'énergie fournissent une borne uniforme sur la longueur des segments géodésiques lumières partant de Σ , orthogonaux à Σ , et qui ne contiennent pas de points conjugués à Σ (voir proposition 20). On en déduit que l'hypersurface topologique $\text{fr}(J^+(\Sigma))$ est compacte.

On choisit alors un champ de vecteurs de type temps X sur M . Chaque orbite de X coupe la surface de Cauchy S en un point et un seul, la projection sur S le long des orbites de X est donc bien définie. Cette projection induit un plongement (topologique) de $\text{fr}(J^+(\Sigma))$ dans S (l'injectivité provient du fait que $\text{fr}(J^+(\Sigma))$ est un ensemble achronal). Ceci est absurde puisqu'on ne peut pas plonger une variété compacte dans une variété non-compacte connexe de même dimension. \square

Quelques années après que Penrose a obtenu le résultat ci-dessus, Hawking et Penrose ont prouvé des énoncés du même type, qui permettent de se débarrasser de l'hypothèse d'existence d'une hypersurface de Cauchy²⁴. De plus, contrairement au théorème 21 qui ne traite que des situations du type "trou noir", les résultats ultérieurs de Hawking et Penrose montrent également qu'une hypothèse du type "univers en expansion" implique l'existence de géodésiques de type temps incomplète vers le passé (ce qui semble indiquer l'existence d'une singularité initiale du type "Big Bang").

²³On trouvera une preuve complète dans [15] ou dans [12, page 263]. On trouvera également une description informelle, mais lumineuse, de la preuve dans [13]

²⁴L'hypothèse d'hyperbolicité globale de l'espace-temps semblait particulièrement gênante à certains physiciens (comme Hawking), parce qu'elle n'a pas d'interprétation physique claire et parce qu'elle concerne la géométrie globale de l'espace-temps. Hawking et Penrose l'ont remplacée par une hypothèse de causalité tout aussi globale, mais qui est plus nettement plus faible, et dont l'interprétation physique est plus claire.

Ces résultats soulèvent un certain nombre de questions.

Tout d'abord, comme on l'a déjà dit, les physiciens s'attendent à ce que la relativité générale conduisent non seulement à l'existence de géodésiques de type temps ou lumière incomplète, mais aussi (et surtout) à une divergence de la densité d'énergie (et donc du tenseur de courbure) le long de telles géodésiques. D'où le problème suivant :

Problème. *Montrer que, sous des hypothèses similaires à celles des théorèmes de Hawking et Penrose (et éventuellement d'autres hypothèses techniques appropriées), certaines composantes du tenseur de courbure explosent (i.e tendent vers l'infini) le long des géodésiques de type temps et lumière incomplètes²⁵ ? (Voir [12, parties 8.3, 8.4 et 8.5] à ce sujet)*

Par ailleurs, d'un point de vue plus mathématique, on peut se demander si l'absence de géodésique incomplète, combinée avec une hypothèse de positivité du tenseur de Ricci dans certaines directions (type hypothèse de positivité de l'énergie) est une contrainte forte pour une variété lorentzienne (par exemple : est-ce que cela interdit certaines topologies). Dans cette direction, D. Gannon ([11]), et plus récemment L. Andersson et G. Galloway ([2]), ont montré qu'un espace-temps globalement hyperbolique, "pas trop compliqué au voisinage de l'infini", qui satisfait une hypothèse de positivité de l'énergie, et dont toutes les géodésiques de type lumière sont complètes, a nécessairement une topologie "triviale". Je renvoie à [10] pour une présentation plus large de résultats de ce type.

Appendice. Rappels de géométrie pseudo-riemannienne

Dans cet appendice, on suppose fixée une variété différentielle lisse M . On note n la dimension de cette variété.

Métriques pseudo-riemanniennes

Une *métrique pseudo-riemannienne* g de signature (k, l) sur M est un champ de formes bilinéaires symétriques de signature $(k, n - k)$ sur M (c'est-à-dire la donnée, pour chaque $p \in M$, d'une forme bilinéaire symétrique g_p sur l'espace

²⁵Pour que la notion d'explosion d'une composante du tenseur de courbure le long d'une courbe γ ait un sens intrinsèque, il faut considérer les composantes par rapport à une base que l'on transporte parallèlement le long de γ .

tangent $T_p M$, qui dépend de manière lisse du point p). Une *métrique riemannienne* est une métrique pseudo-riemannienne de signature $(n, 0)$. Une *métrique lorentzienne* est une métrique pseudo-riemannienne de signature $(1, n - 1)$ ou de signature $(n - 1, 1)$.

On supposera maintenant fixée une métrique pseudo-riemannienne g sur M .

Dérivation covariante et transport parallèle

Si p et q sont deux points d'un espace affine E , on a une identification canonique entre les espaces tangents $T_p E$ et $T_q E$. De même, la métrique pseudo-riemannienne g sur la variété M permet de définir une identification entre les espaces tangents $T_p M$ et $T_q M$, lorsque p et q sont deux points de M (cependant, cette identification n'est pas complètement canonique, elle dépend du choix d'un arc joignant p à q). Pour construire cette identification, on doit d'abord définir la *dérivation covariante* et le *transport parallèle* associés à g .

Considérons tout d'abord un système de coordonnées locales $(x_i)_{i=1\dots n}$ défini sur un voisinage V d'un point p de M , et notons $\partial_1, \dots, \partial_n$ les champs de vecteurs élémentaires associés à ce système de coordonnées. On appelle *dérivée covariante* du champ de vecteurs ∂_i dans la direction du champ de vecteurs ∂_j le champ de vecteurs

$$D_{\partial_j} \partial_i := \frac{1}{2} \sum_{k=1}^n g(\partial_k, \partial_l) \left(\frac{\partial}{\partial x_i} g(\partial_j, \partial_m) + \frac{\partial}{\partial x_j} g(\partial_i, \partial_m) - \frac{\partial}{\partial x_m} g(\partial_i, \partial_j) \right) \partial_k$$

Il suffit alors d'étendre cette formule par linéarité pour définir la dérivée covariante $D_Y X$ d'un champ de vecteurs quelconque X (défini sur U) dans la direction d'un autre champ de vecteurs quelconque Y . Le point important est que $D_Y X$ ne dépend en fait pas du choix du système de coordonnées $(x_i)_{i=1\dots n}$ (voir remarque ci-dessous).

On vérifie facilement que la valeur du champ de vecteurs $D_Y X$ en un point p ne dépend que de la valeur de Y en p , et de la restriction de X au germe de courbe intégrale de Y passant par p . Ainsi, si $\gamma : I \rightarrow M$ est une courbe lisse, et si X est un champ de vecteurs le long de M , alors $D_{\frac{d\gamma}{ds}} X$ est un champ de vecteurs bien défini le long de γ . Lorsque X n'est défini que le long de γ , il n'y a pas d'ambiguïté quand à la direction dans laquelle on dérive ; on utilise alors la notation $\frac{DX}{ds}$ pour désigner le champ $D_{\frac{d\gamma}{ds}} X$.

Étant donné un arc paramétré lisse $\gamma : I \rightarrow M$ (sans point double) et un vecteur $v \in T_{\gamma(s_0)} M$ pour un certain $s_0 \in I$, le théorème de Cauchy-Lipschitz

implique qu'il existe un unique champ de vecteurs V le long de γ , tel que $V(\gamma(s_0)) = v$ et tel que $D_{\frac{d\gamma}{ds}}V = 0$. On dit que le champ de vecteurs V est obtenu par *transport parallèle* de v le long de γ . Quels que soient $s_0, s_1 \in I$, ceci définit une application linéaire

$$\tau_{\gamma, s_0, s_1} : T_{\gamma(s_0)} \rightarrow T_{\gamma(s_1)}$$

qui envoie un vecteur $v \in T_{\gamma(s_0)}$ sur la valeur en $\gamma(s_1)$ du champ V obtenu par transport parallèle de v le long de γ . On peut alors montrer :

Proposition 22. *L'application linéaire τ_{γ, s_0, s_1} est une isométrie : pour tous $v, w \in T_{\gamma(s_0)}M$, on a*

$$g_{\gamma(s_1)}(\tau_{\gamma, s_0, s_1}(v), \tau_{\gamma, s_0, s_1}(w)) = g_{\gamma(s_0)}(v, w)$$

.

Un champ de vecteurs W le long de γ est dit *parallèle* si il satisfait $D_{\frac{d\gamma}{ds}}W = 0$, ou de manière équivalente, si on a $W(\gamma(s_1)) = \tau_{\gamma, s_0, s_1}(W(\gamma(s_0)))$ pour tous s_0, s_1 .

Remarque. On peut montrer que la dérivé covariante $(X, Y) \mapsto D_X Y$ est entièrement caractérisée par le fait que l'application τ_{γ, s_0, s_1} est une isométrie (pour tout arc γ et tous réels s_0, s_1), et par l'égalité $D_X Y - D_Y X = [X, Y]$. On en déduit alors immédiatement que la définition de $D_X Y$ donnée ci-dessus est en fait indépendante du choix du système de coordonnées $(x_i)_{i=1\dots n}$.

Géodésiques, application exponentielle, longueur des arcs

Une *géodésique* de (M, g) est une courbe paramétrée $\gamma : I \rightarrow M$ qui est une solution maximale de l'équation différentielle

$$\frac{D}{ds} \frac{d\gamma}{dt} = 0.$$

Un *arc géodésique* est un arc paramétré que l'on peut prolonger en une géodésique.

Il découle alors du théorème de Cauchy-Lipschitz que, pour tout point p et tout vecteur $v \in T_p M$, il existe une unique géodésique $\gamma : I \rightarrow M$ telle que $\gamma(0) = p$ et $\frac{d\gamma}{ds}(0) = v$. Ceci permet de définir, en chaque point p de M , une *application exponentielle*

$$\exp_p : T_p M \rightarrow M$$

qui à un vecteur $v \in T_p M$ associe le point $\gamma(1) \in M$ où γ est l'unique géodésique telle que $\gamma(0) = p$ et $\frac{d\gamma}{ds}(0) = v$.

On appelle *voisinage normal* d'un point p de M tout voisinage V tel que l'application \exp_p réalise un difféomorphisme d'un voisinage de l'origine dans $T_p M$ sur V . Un ouvert de M est dit *convexe* si c'est un voisinage normal de chacun de ces points. Cette définition implique en particulier que deux points quelconques p et q d'un ouvert convexe V sont joints par un unique arc géodésique (à reparamétrisation près) dans V . En appliquant les théorèmes généraux sur les équations différentielles à l'équation qui définit les géodésiques, on obtient facilement :

Proposition 23. *Tout point de M admet un voisinage convexe.*

L'importance des géodésiques en géométrie pseudo-riemannienne provient essentiellement du lien entre les géodésiques et les points critiques de la fonctionnelle d'énergie. Plus précisément, considérons la fonctionnelle E qui à tout arc paramétré $\gamma : [a, b] \rightarrow M$ associe son *énergie* $E(\gamma)$ définie par

$$E(\gamma) = \int_a^b g \left(\frac{d\gamma}{ds}, \frac{d\gamma}{ds} \right) ds$$

Étant deux points distincts $p, q \in M$, on note $\Gamma(p, q)$ l'ensemble des arcs paramétrés joignant p à q . Étant donné une sous-variété Σ de M , et un point $q \in M \setminus \Sigma$, on note $\Gamma(\Sigma, q)$ l'ensemble des arcs paramétrés qui joignent Σ à q . On a alors la proposition suivante :

Proposition 24. *Quels que soient $p, q \in M$ distincts, les points critiques de la restriction de la fonctionnelle E à $\Gamma(p, q)$ sont exactement les arcs géodésiques joignant p à q .*

Quel que soit la sous-variété Σ de M , et un point $q \in M \setminus \Sigma$, les points critiques de la restriction de la fonctionnelle E à $\Gamma(\Sigma, q)$ sont exactement les arcs géodésiques joignant Σ à q , et orthogonaux à Σ .

Une géodésique $\gamma : I \rightarrow M$ est dite *complète* si l'intervalle I est égal à \mathbb{R} en entier ; sinon elle est dite *incomplète*. Une variété est dite *géodésiquement complète* si toutes les géodésiques sont complètes.

Notons que les variétés pseudo-riemanniennes géodésiquement complètes sont rares. Par exemple, contrairement au cas Riemannien, il n'est pas vrai en général qu'une variété pseudo-riemannienne compacte est géodésiquement complète.

Tenseur de courbure

On s'intéresse à la question suivante : deux métriques pseudo-riemanniennes de même signature sont-elles toujours isométriques ? Autrement dit : au voisinage d'un point p de M , peut-on trouver un système de coordonnées locales (x_1, \dots, x_n) tel que, dans ce système de coordonnées, la métrique g s'écrive $g_x = -dx_1^2 - \dots - dx_k^2 + dx_{k+1}^2 + \dots + dx_n^2$?

La réponse à la question ci-dessus est connue depuis Riemann. Tout d'abord, on peut vérifier que, pour tout point $p \in M$, l'application exponentielle permet de construire un système de coordonnées locales (x_1, \dots, x_n) d'origine p tel que, dans ces coordonnées, g s'écrit $g_x = -dx_1^2 - \dots - dx_k^2 + dx_{k+1}^2 + \dots + dx_n^2 + o(x)$. Par contre, il n'existe en général aucun système de coordonnées locales (x_1, \dots, x_n) dans lesquelles le terme d'ordre 2 en x s'annule. Le terme d'ordre 2 qui apparaît dans l'écriture de g est le *tenseur de courbure* de g en p .

Par définition, le *tenseur de courbure* R_g de g est l'application trilineaire qui, à trois champs de vecteurs X, Y, Z associe le champ de vecteurs $R_g(X, Y).Z$ défini par

$$R_g(X, Y).Z = -D_X D_Y Z + D_Y D_X Z + D_{[X, Y]} Z$$

Remarque importante. Les valeurs des champs de vecteurs $D_X D_Y Z$, $D_Y D_X Z$ et $D_{[X, Y]} Z$ en un point p ne dépendent pas que des valeurs de X, Y et Z au seul point p . Cependant, on peut montrer que la valeur du champ $R_g(X, Y).Z$ en p ne dépend que des valeurs de X, Y et Z au point p . L'expression $R_g(X, Y).Z$ est donc définie dès que X, Y, Z sont des vecteurs tangents à M en un même point p .

Le tenseur de courbure possède de nombreuses symétries : pour tout point $p \in M$ et tout quadruplet (X, Y, Z, W) de vecteurs de $(T_p M)^4$, on a

1. $R_g(X, Y).Z = -R_g(Y, X).Z$
2. $g(R_g(X, Y).Z, W) = -g(R_g(X, Y).W, Z)$
3. $R_g(X, Y).Z + R_g(Y, Z).X + R_g(Z, X).Y = 0$
4. $g(R_g(X, Y).Z, W) = g(R_g(Z, W).X, Y)$

Le tenseur de courbure est un objet complexe, fort difficile à manipuler et à interpréter ; c'est pourquoi on cherche à en extraire des objets qui contiennent moins d'informations, mais sont plus simples. Les plus importants de ces objets sont le *tenseur de Ricci* et la *courbure scalaire* de g . Le *tenseur de Ricci*

de g en p est la forme bilinéaire symétrique qui à deux vecteurs $X, Y \in T_pM$ associe le nombre réel $\text{Ric}_g(X, Y)$ défini par

$$\text{Ric}_g(X, Z) := \text{tr}(V \mapsto R_g(X, V).Y) = \sum_{i=1}^n g(e_i, e_i)g(R_g(X, e_i).Y, e_i)$$

où $(e_i)_{i=1\dots n}$ est une base pseudo-orthonormée de T_pM . La *courbure scalaire* de g au point p est le nombre réel

$$\text{Scal}_g := \sum_{i=1}^n \text{Ric}_g(e_i, e_i)$$

où $(e_i)_{i=1\dots n}$ est une base pseudo-orthonormée de T_pM .

Champs de Jacobi et points conjugués

On a vu ci-dessus que les points critiques de la fonctionnelle d'énergie E sont des arcs géodésiques. Nous allons maintenant étudier les familles continues d'arcs géodésiques et décrire partiellement le comportement de E au voisinage d'un de ses points critiques.

Soit $\gamma : I \rightarrow M$ une géodésique. Étant donné un champ de vecteurs V le long de γ , une *déformation de γ de direction V* est une famille $(\gamma_u)_{u \in [-\varepsilon, \varepsilon]}$ de courbes paramétrée par I , telle que $\gamma_0 = \gamma$, et telle que $\frac{\partial}{\partial u}|_{u=0} \gamma_u(s) = V(\gamma(s))$. On dit que $(\gamma_u)_{u \in [-\varepsilon, \varepsilon]}$ est une *déformation géodésique* si γ_u est une géodésique pour tout u .

Un *champ de Jacobi le long de γ* est un champ de vecteurs V le long de γ qui satisfait l'équation différentielle

$$\frac{D^2V}{ds^2} + R_g\left(\frac{d\gamma}{ds}, V\right) \cdot \frac{d\gamma}{ds} = 0$$

Le lien entre les champs de Jacobi le long de γ et les déformation géodésiques de γ est donné par la proposition suivante :

Proposition 25. *Si $(\gamma_u)_{u \in [-\varepsilon, \varepsilon]}$ est une déformation géodésique de γ de direction V , alors V est un champ de Jacobi. Réciproquement, si V est un champ de Jacobi le long de γ , alors il existe une déformation géodésique de γ de direction V .*

On dit que deux points $p = \gamma(a)$ et $q = \gamma(b)$ sont *conjugués* le long de γ s'il existe un champ de Jacobi le long de γ qui s'annule en p et en q .

Remarque. Si V est un champ de Jacobi le long de γ qui s'annule en deux points distincts, alors V est nécessairement orthogonal à γ . Pour déterminer l'existence de points conjugués le long de γ , il suffit donc de considérer les champs de Jacobi orthogonaux à γ .

On a alors les propositions suivantes :

Proposition 26. *Supposons que γ n'est pas de type lumière.*

S'il existe une déformation géodésique $(\gamma_u)_{u \in [-\varepsilon, \varepsilon]}$ de γ telle que $\gamma_u(a) = p$ et $\gamma_u(b) = q$ pour tout u , alors les points p et q sont conjugués le long de γ .

Réciproquement, si les points p et q sont conjugués le long de γ , alors il existe une déformation géodésique $(\gamma_u)_{u \in [\varepsilon, \varepsilon]}$ de γ , telle que $\gamma_u(a) = p$ pour tout u , et telle que $\frac{\partial \gamma_u(b)}{\partial u}|_{u=0} = 0$ (c'est-à-dire que la courbe γ_u passe presque par q pour u petit). De plus, $\gamma_u|_{[a, b]}$ a la même énergie que $\gamma|_{[a, b]}$ à des termes d'ordre 3 en u près.

Proposition 27. *Supposons maintenant que γ est de type lumière.*

S'il existe une déformation géodésique $(\gamma_u)_{u \in [-\varepsilon, \varepsilon]}$ de γ , telle que γ_u est de type lumière pour tout u , et telle que $\gamma_u(a) = p$ et $\gamma_u(b) = q$ pour tout u , alors les points p et q sont conjugués le long de γ .

Réciproquement, si les points p et q sont conjugués le long de γ , alors il existe une déformation géodésique $(\gamma_u)_{u \in [\varepsilon, \varepsilon]}$ de γ , telle que γ_u est de type lumière pour tout u , telle que $\gamma_u(a) = p$ pour tout u , et telle que $\frac{\partial \gamma_u(b)}{\partial u}|_{u=0} = 0$ (c'est-à-dire que la courbe γ_u passe presque par q pour u petit).

Points focaux

On suppose maintenant que la métrique g est lorentzienne (de signature $(-, +, \dots, +)$). On considère une hypersurface Σ de type espace dans M . On note n le champ des normales à Σ , dirigées vers le futur, de pseudo-norme -1 . En tout point p de Σ , on considère l'opérateur de Weingarten $W_p : T_p \Sigma \rightarrow T_p \Sigma$ défini par $W_p(v) = D_v n$.

On suppose maintenant que la géodésique γ est orthogonale à Σ en un point $p = \gamma(a)$ (ce qui entraîne que γ est de type temps), et on considère un point $q = \gamma(b)$ sur γ . Pour fixer les idées, on suppose que $b > a$ et que γ est dirigée vers le futur. On dit que q est un *point focal* à Σ (le long de γ) s'il existe un champ de Jacobi V le long de γ , orthogonal à γ (et donc tangent à Σ en p), qui s'annule en q , et tel que $\frac{DV}{ds}(p) = W_p(V(p))$. On a alors la proposition suivante :

Proposition 28. *S'il existe une déformation géodésique $(\gamma_u)_{u \in [-\varepsilon, \varepsilon]}$ de γ telle que γ_u est orthogonale à Σ en $\gamma_u(a)$, et telle que $\gamma_u(b) = q$ pour tout u , alors q est un point focal à Σ le long de γ .*

Réciproquement, si q est un point focal à Σ le long de γ , alors il existe une déformation géodésique $(\gamma_u)_{u \in [\varepsilon, \varepsilon]}$ de γ , telle que γ_u est orthogonale à Σ pour tout u , et $\frac{\partial \gamma_u(b)}{\partial u}|_{u=0} = 0$. De plus, $\gamma_u|_{[a, b]}$ a la même énergie que $\gamma|_{[a, b]}$ à des termes d'ordre 3 en u près.

On s'intéresse maintenant au cas spécifique où M est de dimension 4. On considère alors une surface Σ de type espace dans M . En chaque point de Σ , il existe deux directions de type lumière orthogonales à Σ . Il existe donc deux champs n et n' de normales à Σ , de type lumière, dirigées vers le futur, et de norme 1 pour une métrique riemannienne auxiliaire arbitraire. En chaque point p de Σ , on a donc deux opérateurs de Weingarten $W_p : T_p \Sigma \rightarrow T_p \Sigma$ et $W'_p : T_p \Sigma \rightarrow T_p \Sigma$ définis par $W_p(v) = D_v n$ et $W'_p(v) = D_v n'$.

On suppose maintenant que la géodésique γ est de type lumière et orthogonale à Σ en un point $p = \gamma(a)$; par exemple, on supposera que $\frac{d\gamma}{ds}|_{s=a}$ est colinéaire à la normale n . On considère un point $q = \gamma(b)$ sur γ . Pour fixer les idées, on suppose que $b > a$ et que γ est dirigée vers le futur. On dit que q est un *point focal* à Σ (le long de γ) s'il existe un champ de Jacobi V le long de γ , orthogonal à γ et tangent à Σ en p , qui s'annule en q , et tel que $\frac{DV}{ds}(p) = W_p(V(p))$. On a alors la proposition suivante :

Proposition 29. *S'il existe une déformation géodésique $(\gamma_u)_{u \in [-\varepsilon, \varepsilon]}$ de γ telle que γ_u est de type lumière, telle que γ_u est orthogonale à Σ en $\gamma_u(a)$ pour tout u et telle que $\gamma_u(b) = q$ pour tout u , alors q est un point focal à Σ le long de γ .*

Réciproquement, si q est un point focal à Σ le long de γ , alors il existe une déformation géodésique $(\gamma_u)_{u \in [\varepsilon, \varepsilon]}$ de γ telle que γ_u est de type lumière et orthogonale à Σ pour tout u , et telle que $\frac{\partial \gamma_u(b)}{\partial u}|_{u=0} = 0$.

*Bibliography

- [1] L. Andersson. The Global Existence Problem in General Relativity. In *The Einstein Equations and the Large Scale Behaviour of Gravitational Fields*, P.T. Chrusciel and H. Friedrich (eds), Birkhäuser Verlag, 2004.
- [2] L. Andersson and G. Galloway. dS/CFT and spacetime topology. *Adv. Theor. Math. Phys.*, **6** (2003), 307-327.

- [3] L. Andersson, G. Galloway and R. Howard. The cosmological time function. *Class. Quantum Grav.*, **15** (1998) 309-322.
- [4] R. Bartnik and J. Isenberg. The Constraint Equations. In *The Einstein Equations and the Large Scale Behaviour of Gravitational Fields*, P.T. Chrusciel and H. Friedrich (eds), Birkhäuser Verlag, 2004.
- [5] J.K. Beem, P.E. Ehrlich and K.L. Easley. *Global Lorentzian Geometry*. Marcel Dekker, 1996.
- [6] A.N. Bernal and M. Sánchez. Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes. *Commun. Math. Phys.* **257** (2005) 43-50.
- [7] S.M. Carroll. *Spacetime and geometry : an introduction to Relativity*. Addison Wesley, 2003.
- [8] Y. Choquet-Bruhat and R. Geroch. Global aspects of the Cauchy problem in general relativity. *Comm. Math. Phys.* **14** (1969), 329–335.
- [9] Y. Fourès-Bruhat. Théorème d'existence pour certains systèmes d'équations aux dérivées partielles non linéaires. *Acta Mathematica* **88** (1952), 141-225.
- [10] G.J. Galloway. Null Geometry and the Einstein Equations. In *The Einstein Equations and the Large Scale Behaviour of Gravitational Fields*, P.T. Chrusciel and H. Friedrich (eds), Birkhäuser Verlag, 2004.
- [11] D. Gannon. Singularities in nonsimply connected space-times. *J. Math. Phys.*, **16** (1975), 2364–2367.
- [12] G. Ellis, S. W. Hawking. *The large scale structure of spacetime*. Cambridge Monographs on Mathematical Physics, Cambridge university Press, 1973.
- [13] S.W. Hawking and R. Penrose. *The Nature of Space and Time*. Princeton University Press, 1996. Traduction en Français : *La Nature de l'Espace et du Temps*, Folio Essais, Gallimard, 2003.
- [14] B. O'Neill *Semi-riemannian Geometry (with Applications to Relativity)*. Academic Press, 1983.

- [15] R. Penrose. *Techniques of Differential Topology in Relativity*. CBMS-NSF Regional Conference Series, SIAM, 1972.
- [16] R.M. Wald. *General Relativity*. University of Chicago Press, 1984.

François Béguin
Laboratoire de Mathématiques (UMR CNRS 8628)
Université Paris Sud - Bâtiment 425
91405 ORSAY
