Leçon 1 : Principe d'équivalence et tenseurs

Andy : *Alors si je suis dans un ascenseur sans fenêtre et que je me sens vraiment lourd, je ne peux pas savoir si l'ascenseur est en accélération ou si par espièglerie tu m'as déposé sur Jupiter.*

Lenny : *C'est exact, tu ne peux pas savoir.*

Andy : *Mais au moins sur Jupiter, si je ne bouge pas, les rayons lumineux iront tout droit.*

Lenny : *Oh non, ils seront courbés comme dans un ascenseur en accélération.*

Andy : *Hum, je vois.*

Lenny : *Et si tu tombes dans un trou noir, prends garde, les choses vont devenir réellement étranges. Mais, ne t'en fais pas, je vais éclairer tout ça.*

Andy : *Euh, avec des rayons droits ou courbes ?*

Introduction

Ce livre d'introduction à la relativité générale (RG) est le quatrième de la série Le Minimum Théorique (LMT). Les trois premiers étaient consacrés respectivement à la mécanique classique, à la mécanique quantique, et à la relativité restreinte et la théorie classique des champs. Le premier volume expose la description lagrangienne et hamiltonienne des phénomènes physiques et le principe de moindre action, qui est l'un des principes fondamentaux sous-jacents à toute la physique (voir volume 3, leçon 7 sur les principes fondamentaux et l'invariance de jauge). Ils ont été utilisés dans les trois premiers volumes et le seront encore dans celui-ci et les suivants.

La physique fait grand usage des mathématiques comme boîte à outils pour construire des théories formelles et quantifiables modélisant les phénomènes naturels. Les principaux outils que nous avons utilisés jusqu'ici sont la trigonométrie, l'algèbre linéaire – c'est-à-dire les espaces vectoriels – et le calcul différentiel et intégral. Ils ont été expliqués dans le volume 1 ainsi que dans de brefs rappels dans les autres volumes. Nous supposons que le lecteur et la lectrice sont familiers avec ces outils mathématiques et avec les notions physiques présentées dans les volumes 1 et 3. Le présent volume 4, comme les volumes 1 et 3, mais contrairement au volume 2, traite de la physique classique dans le sens où aucun phénomène quantique aléatoire n'intervient.

Nous avons également commencé à utiliser de manière légère les tenseurs dans le volume 3 sur la relativité restreinte et la théorie classique des champs. À présent, avec la relativité générale, nous allons les utiliser intensivement. Nous allons commencer par les présenter en détail. On se souvient que les tenseurs généralisent les vecteurs. De même qu'un vecteur est représenté par différents jeux de composantes selon la base utilisée pour repérer l'espace vectoriel dans lequel il se trouve, un tenseur aura des composantes différentes dans différents systèmes de coordonnées. Les règles pour passer d'une collection de composantes à une autre joueront un rôle fondamental. En outre, nous travaillerons généralement avec des *champs tensoriels*, qui sont des ensembles de tenseurs – un tenseur différent attaché à chaque point de l'espace, ou de l'espace-temps.

Les tenseurs ont été inventés par Ricci-Curbastro et Levi-Civita [1] pour développer les travaux de Gauss [2] sur la courbure des surfaces, et les travaux de Riemann [3] sur la géométrie non euclidienne. Einstein [4] a largement utilisé les tenseurs pour construire sa théorie de la relativité générale. Il a également fait d'importantes contributions à leur emploi : la notation des indices en haut ou en bas et la convention de sommation d'Einstein.

Dans *Savants et écrivains* (1910), Poincaré [5] écrit qu' « en sciences mathématiques, une bonne notation a la même importance philosophique qu'une bonne classification en sciences naturelles ». Dans ce livre nous nous efforçons de toujours utiliser la notation la plus simple et la plus claire possible.

Principe d'équivalence

Les articles révolutionnaires publiés par Einstein en 1905 sur la relativité restreinte ont profondément clarifié et étendu les idées sur lesquelles plusieurs autres physiciens et mathématiciens – Lorentz [6], Poincaré, et d'autres - travaillaient depuis quelques années. Einstein a étudié les conséquences du fait que les lois de la physique, en particulier le comportement de la lumière, sont les mêmes dans différents référentiels inertiels. Cela s'appelle le *principe de relativité galiléenne*. Il en a déduit une nouvelle explication des transformations de Lorentz, de la relativité du temps, de l'équivalence de la masse et de l'énergie, etc.

Après 1905, Einstein a commencé à réfléchir à une généralisation du principe de relativité galiléenne à tout type de référentiel, pas seulement des référentiels inertiels, mais aussi des référentiels qui

1. Gregorio Ricci-Curbastro (1853–1925) et son élève Tullio Levi-Civita (1873–1941) étaient des mathématiciens italiens. Leur article en commun le plus important est "Méthodes de calcul différentiel absolu et leurs applications", dans *Mathematische Annalen* 54 (1900), pp. 125–201. Ils n'utilisaient pas le mot *tenseur*, qui a été introduit plus tard par d'autres auteurs.
2. Carl Friedrich Gauss (1777–1855), mathématicien allemand.
3. Bernhard Riemann (1826–1866), mathématicien allemand.
4. Albert Einstein (1879–1955), physicien allemand, puis suisse, puis à nouveau allemand, enfin plus tard naturalisé américain.
5. Henri Poincaré (1854–1912), mathématicien français.
6. Hendrik Antoon Lorentz (1853–1928), physicien néerlandais.

peuvent être en accélération. Un *référentiel inertiel* – dit encore *référentiel galiléen* – est un référentiel dans lequel les lois de Newton, reliant les forces et les mouvements, ont des expressions simples. Ou, si vous préférez une image plus concrète, et que vous savez jongler, c'est un référentiel dans lequel vous pouvez jongler sans problème – par exemple dans un wagon roulant à vitesse constante, sans à-coups ni accélérations d'aucune sorte. Après dix ans d'efforts pour construire une théorie étendant le principe de relativité aux référentiels en accélération, et présentant la gravitation sous un jour nouveau, Einstein publia ses travaux en novembre 1915. Contrairement à la relativité restreinte qui couronne les travaux de beaucoup de savants, la relativité générale est essentiellement l'œuvre d'un seul homme.

Nous allons commencer notre étude de la relativité générale à peu près où Einstein a commencé. C'est une constante chez Einstein de démarrer avec un fait élémentaire, tellement simple que presque un enfant peut le comprendre, et d'en déduire des conséquences incroyablement profondes. Nous pensons que c'est aussi la meilleure façon de l'enseigner – de commencer par les choses les plus simples et d'en déduire les conséquences. Nous allons donc partir du *principe d'équivalence.*

Qu'est-ce que le principe d'équivalence ? C'est le principe qui dit que *la gravité est en quelque sorte la même chose que l'accélération.* Nous expliquerons précisément ce que l'on entend par là et montrerons comment Einstein l'a utilisé. À partir de là, nous nous demanderons quelle sorte de structure mathématique doit avoir une théorie pour que le principe d'équivalence soit vrai. Quel genre de mathématiques devons-nous utiliser pour le décrire ?

La plupart des lecteurs et des lectrices ont probablement entendu dire que la relativité générale était une théorie non seulement sur la gravité, mais aussi sur la géométrie de l'espace-temps. Il est donc intéressant de commencer par le début et de se demander ce qui a amené Einstein à dire que la gravité a quelque chose à voir avec la géométrie. Qu'est-ce que cela veut dire que « la gravité est égale à l'accélération » ? Réponse : si nous sommes dans un référentiel accéléré, par exemple un ascenseur accélérant vers le haut, nous ressentons un champ gravitationnel supplémentaire dit fictif ou « effectif », par opposition à « dû réellement à la présence d'un corps massif ». Les enfants le savent parce qu'ils le ressentent.

Ce qui suit est peut-être un marteau pour écraser une mouche, cependant mettre en équations le mouvement d'un ascenseur éclaire avec un exemple très simple comment les physiciens « mathématisent » un phénomène naturel, puis comment les mathématiques en retour permettent de faire des prédictions sur le phénomène.

Avant de continuer, soulignons que l'étude d'un ascenseur, et des lois de la physique telles qu'elles sont observées à l'intérieur, est très simple. C'est néanmoins une première présentation de concepts très importants. C'est pourquoi il est fondamental de bien la comprendre. Nous nous y référerons souvent. Dans les leçons 4 à 9, cela nous aidera à comprendre l'accélération, la gravitation et comment la gravitation « déforme » l'espace-temps.

Imaginons donc l'expérience de pensée conçue par Einstein où quelqu'un est dans un ascenseur, figure 1. Plus tard dans les manuels, l'ascenseur a été promu en fusée. Mais je ne suis jamais monté dans une fusée, alors que j'ai déjà pris l'ascenseur. Je connais la sensation que l'on éprouve quand il accélère ou décélère. Mettons que l'ascenseur se déplace vers le haut avec une vélocité v.

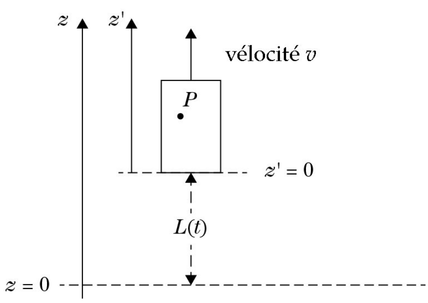


Figure 1 : Ascenseur et deux référentiels.

Pour l'instant le problème est à une dimension. Nous ne nous intéressons qu'à la direction verticale. Il y a deux repères : l'un est fixe par rapport à la Terre. Il utilise la coordonnée z. L'autre est fixe par rapport à l'ascenseur. Il utilise la coordonnée z'. Un point P n'importe où le long de l'axe vertical a deux coordonnées : la coordonnée z dans le repère fixe (appelé aussi le repère *stationnaire*), et la coordonnée z' dans le repère de l'ascenseur. Par exemple, le plancher de l'ascenseur a la coordonnée $z' = 0$. Sa coordonnée z, en revanche, est la distance L, qui est manifestement une fonction du temps. On peut donc écrire pour n'importe quel point P

$$z' = z - L(t) \tag{1}$$

On va s'intéresser à la question suivante : si nous connaissons les lois de la physique dans le repère z, que deviennent-elles dans le repère z' ?

Avertissement : au moins au début, nous allons ignorer la relativité restreinte. Cela revient à prétendre que la vitesse de la lumière est infinie, ou que nous parlons de mouvements si lents qu'elle peut être considérée comme infinie. Vous vous demandez peut-être, si la relativité générale est la généralisation de la relativité restreinte, comment Einstein a pu commencer à y penser sans inclure la relativité restreinte ?

La réponse est que la relativité restreinte concerne les vitesses élevées, tandis que la gravité concerne les masses. Il y a des situations où la gravité est importante, mais les vitesses élevées ne le sont pas. Einstein a donc commencé à penser à la gravité pour des vitesses lentes. Ce n'est que plus tard qu'il l'a combinée avec la relativité restreinte pour considérer la conjonction des vitesses élevées et de la gravité. C'est devenu la relativité générale.

Voyons ce que nous savons pour les vitesses lentes. Supposons que les repères z' et z soient tous deux des référentiels inertiels. Cela signifie, entre autres, qu'ils sont liés par une vitesse uniforme :

$$L(t) = vt \tag{2}$$

Nous avons choisi les coordonnées de sorte que lorsque $t = 0$, elles ont la même valeur. Au temps $t = 0$, pour n'importe quel point, z et z' sont égaux. Par exemple, quand $t = 0$ le plancher de l'ascenseur a la coordonnée 0 dans les deux référentiels. Puis, le plancher commence à s'élever, sa position z étant égale à vt. Pour n'importe quel point – fixe ou mobile, peu importe –, nous pouvons donc écrire l'équation (1). Compte tenu de l'équation (2), quand les deux référentiels sont galiléens l'équation (1) se réécrit

$$z' = z - vt \tag{3}$$

Observez qu'il s'agit d'un *changement de coordonnées* mêlant l'espace et le temps. Pour les lectrices et les lecteurs familiers avec le volume 3 de LMT sur la relativité restreinte, cela pose naturellement la question : qu'en est-il du temps dans le référentiel de l'ascenseur ? Si nous laissons de côté la relativité restreinte, nous pouvons simplement dire que t' et t sont les mêmes. Nous n'avons pas à prendre en compte les transformations de Lorentz. Ainsi, l'autre moitié du changement de coordonnées serait $t' = t$.

Nous pourrions également ajouter au repère stationnaire une coordonnée x horizontale et une coordonnée y sortant de la page. De manière analogue, des coordonnées x' et y' pourraient être attachées à l'ascenseur, figure 2. La coordonnée x jouera un rôle quand nous étudierons la trajectoire d'un rayon lumineux. Toutefois tant que l'ascenseur ne glisse pas horizontalement, x' et x peuvent être considérés comme égaux. De même pour y' et y.

Dans un souci de clarté sur la figure 2, nous avons un peu décalé l'ascenseur à droite de l'axe z. Mais il faut imaginer que les deux axes verticaux sont les mêmes, glissant le long l'un de l'autre, et qu'à $t = 0$ les deux origines O et O' coïncident. Encore une fois, l'ascenseur ne se déplace que verticalement.

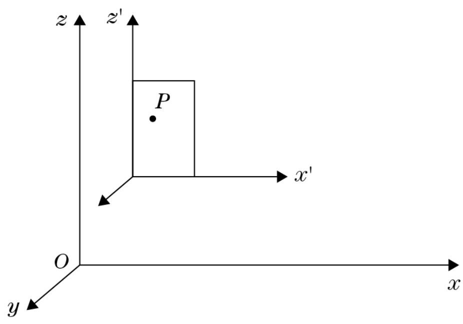


Figure 2 : Ascenseur et deux référentiels à trois dimensions.

Finalement le changement complet de coordonnées est

$$\begin{aligned} z' &= z - vt \\ t' &= t \\ x' &= x \\ y' &= y \end{aligned} \tag{4}$$

C'est un *changement de coordonnées spatio-temporelles.* Pour tout point P *de l'espace-temps*, il exprime ses coordonnées dans le référentiel mobile de l'ascenseur en fonction de ses coordonnées dans le référentiel fixe. Ce changement est trivial. Une seule coordonnée, à savoir z, est transformée de manière intéressante. Nous pouvons aussi écrire, si l'on veut, $z = z' + vt'$, puisque t et t' sont les mêmes.

Regardons une loi de la physique exprimée dans le repère stationnaire. Prenons la seconde loi de Newton, $F = ma$, appliquée à un objet ou une particule. L'accélération a est $\ddot{z}$, où z est la coordonnée verticale de la particule. Alors on peut écrire

$$F = m\ddot{z} \tag{5}$$

Comme nous le savons, $\ddot{z}$ est la dérivée seconde de z par rapport au temps – c'est ce qu'on appelle l'accélération verticale – et F est bien sûr la composante verticale de la force. Les autres composantes sont prises égales à zéro ; quelle que soit la force exercée, elle s'exerce verticalement. À quoi cette force pourrait-elle être due ? Elle pourrait être liée à l'ascenseur – mais ce n'est pas nécessaire. Il pourrait y avoir quelque chose dans l'ascenseur poussant sur la particule. Ou cela pourrait être une corde attachée au plafond et à la particule et qui la tire. Il pourrait y avoir un champ de force le long de l'axe vertical. N'importe quelle sorte de force *liée à l'ascenseur ou pas* pourrait agir sur la particule. Quelle que soit la cause, nous savons, d'après la loi de Newton, que l'équation du mouvement de la particule, exprimée dans le référentiel stationnaire, est donnée par l'équation (5).

Quelle est l'équation du mouvement exprimée dans le référentiel de l'ascenseur ? Il est facile d'y répondre. Il suffit de trouver l'accélération dans le « repère prime » (c'est le nom que nous allons donner au repère de l'ascenseur) exprimée à l'aide de l'accélération initiale, puis d'aller dans l'autre sens. Quelle est l'accélération « prime » ? C'est la dérivée seconde par rapport au temps de z'. Si l'on prend la première des quatre équations dans le système (4)

$$z' = z - vt$$

une première différentiation donne

$$\dot{z}' = \dot{z} - v$$

et une seconde différentiation donne

$$\ddot{z}' = \ddot{z}$$

On parvient au fait que dans les deux référentiels l'accélération est la même.

Tout cela devrait être familier. Mais je tenais à le formaliser pour faire ressortir certains points. En particulier, je voulais souligner que *nous effectuons un changement de coordonnées*. Et nous nous demandons comment les lois de la physique changent quand on passe d'un repère à l'autre. Que peut-on dire maintenant de la loi de Newton dans le référentiel prime ? Dans l'équation (5), nous exprimons $\ddot{z}$ en fonction des coordonnées primes. Étant donné que les deux quantités $\ddot{z}$ et $\ddot{z}'$ sont égales, on obtient simplement

$$F = m\ddot{z}' \tag{6}$$

Nous avons trouvé que la loi de Newton dans le référentiel de l'ascenseur est exactement la même que dans le référentiel stationnaire. Ce n'est pas surprenant. Les deux référentiels se déplacent avec une vitesse uniforme l'un par rapport à l'autre. Si l'un d'eux est un référentiel inertiel, l'autre est aussi un référentiel inertiel. Newton nous a appris que les lois de la physique sont les mêmes dans tous les référentiels inertiels. On l'appelle parfois le *principe de relativité galiléenne*. Nous venons de l'établir formellement.

Tournons-nous maintenant vers un référentiel mobile *accéléré*.

Référentiels accélérés

Supposons que la distance $L(t)$, sur la figure 1, s'accroisse avec une accélération constante. La position du plancher de l'ascenseur n'est plus une fonction linéaire de t, mais est donnée par

$$L(t) = \frac{1}{2}gt^2 \tag{7}$$

Nous utilisons la lettre g pour l'accélération parce que nous allons découvrir que l'accélération imite un champ gravitationnel – ainsi que nous le ressentons lorsque nous montons dans un ascenseur et qu'il accélère. On sait depuis le volume 1 de LMT sur la mécanique classique, ou depuis le lycée, que l'équation (2) correspond à une accélération uniforme. En effet, si on dérive $L(t)$ par rapport au temps, après une différentiation on obtient

$$\dot{L} = gt$$

ce qui signifie que la vitesse de l'ascenseur augmente linéairement avec le temps. Et après une seconde différentiation par rapport au temps, on obtient

$$\ddot{L} = g$$

Cela signifie que l'accélération de l'ascenseur est constante. L'ascenseur est uniformément accéléré vers le haut. Les équations reliant les coordonnées primes et non primes sont différentes des équations (4). Pour les coordonnées verticales, le changement est maintenant

$$z' = z - \frac{1}{2}gt^2 \tag{8}$$

Les autres équations du système (4) restent inchangées :

$$\begin{aligned} t' &= t \\ x' &= x \\ y' &= y \end{aligned}$$

Ces quatre équations sont notre nouveau changement de coordonnées. Elles expriment la relation entre des coordonnées accélérées les unes par rapport aux autres.

Nous continuerons à supposer que dans le système de coordonnées z, ou « non prime », les lois de la physique sont exactement celles que Newton nous a apprises. En d'autres termes, le référentiel stationnaire est inertiel et nous avons $F = m\ddot{z}$. Mais le référentiel prime n'est plus inertiel. Il est en accélération uniforme par rapport au référentiel stationnaire. Demandons-nous quelles sont maintenant les lois de la physique dans ce référentiel accéléré. Nous devons à nouveau effectuer l'opération de différentiation deux fois sur l'équation (8). Nous connaissons le résultat :

$$\ddot{z}' = \ddot{z} - g \tag{9}$$

Ah ah ! Maintenant, l'accélération prime et l'accélération non prime diffèrent par une quantité g. Pour écrire l'équation de Newton dans le référentiel prime, on multiplie les deux membres de l'équation (9) par m, la masse de la particule, et on remplace $m\ddot{z}$ par F. On obtient

$$m\ddot{z}' = F - mg \tag{10}$$

Nous avons atteint notre but : l'équation (10) ressemble à la seconde loi de Newton, c'est-à-dire que la masse multipliée par l'accélération est égale à un certain terme [7].

Ce terme, $F - mg$, on l'appelle la *force dans le référentiel prime.* On remarque, comme on s'y attendait, que la force dans le référentiel prime a un terme supplémentaire : la masse de la particule multipliée par l'accélération de l'ascenseur, avec un signe moins.

Ce qui est intéressant à propos de ce terme supplémentaire $-mg$ dans l'équation (10) – que l'on appelle la « force fictive » – c'est qu'il ressemble exactement à la force exercée sur la particule par la gravité à la surface de la Terre ou à la surface de n'importe quelle sorte de gros corps massif [8]. C'est pourquoi nous avons dénoté l'accélération par la lettre g. Cette lettre g représentait une sorte de gravité. Autrement dit l'accélération ressemble à un champ gravitationnel uniforme.

Précisons en quel sens l'accélération ressemble à la gravité. Il faut d'abord souligner une propriété importante de la gravité :

Une particularité de la gravité est que les forces gravitationnelles sont proportionnelles à la masse – la même masse *que celle qui apparaît dans l'équation du mouvement de Newton.*

On dit parfois que *la masse gravitationnelle est la même que la masse inertielle.* Cela a des conséquences profondes. Si l'équation du mouvement est

$$F = ma \tag{11}$$

et la force elle-même est proportionnelle à la masse, alors la masse s'élimine dans l'équation (11). C'est une caractéristique des forces gravitationnelles : pour un petit objet se déplaçant dans un champ gravitationnel, son mouvement ne dépend pas de sa masse. Un

7. Noter que si, sur la figure 1, la force F que l'on considère est simplement la force gravitationnelle terrestre, elle pointe vers le bas. Donc le terme $F - mg$ pointe aussi vers le bas avec une magnitude encore plus grande que F. Quoi qu'il en soit l'équation (10) est vraie pour n'importe quelle force F.

8. Y compris – on le verra – dans le voisinage d'un trou noir.

exemple est le mouvement de la Terre autour du Soleil. Il est indépendant de la masse de la Terre. Si vous savez où se trouve la Terre au temps t et que vous connaissez aussi sa vélocité à ce moment-là, vous pouvez prédire sa trajectoire. Vous n'avez pas besoin de connaître la masse de la Terre.

L'équation (10) contient un exemple de force fictive imitant l'effet de la gravité. La plupart des gens avant Einstein considéraient cela en grande partie comme une coïncidence accidentelle. Ils savaient bien sûr que l'effet de l'accélération imitait celui de la gravité, mais ils n'y avaient pas prêté beaucoup d'attention. C'est Einstein qui a dit : regardez, c'est un principe profond de la nature que les forces gravitationnelles ne puissent pas être distinguées de l'effet d'un référentiel accéléré.

Si vous êtes dans un ascenseur sans fenêtre et que vous sentez que votre corps a un certain poids – ce que vous pouvez vérifier avec un pèse-personne –, vous ne pouvez pas dire si l'ascenseur, avec vous à l'intérieur, repose immobile sur la surface d'une planète ou bien si, loin de tout corps massif dans l'univers, quelque démon espiègle accélère votre ascenseur. C'est le *principe d'équivalence*. Il étend le principe de relativité galiléenne qui énonçait qu'on peut jongler de la même manière au repos ou dans un wagon en mouvement uniforme. Ainsi, avec un exemple simple, nous avons montré la similitude entre le mouvement accéléré et la gravité. Nous avons commencé à donner un sens à la phrase « la gravité est en quelque sorte la même chose que l'accélération ».

Nous devons cependant discuter un peu ce résultat. Y croyons-nous tout à fait ou faut-il le nuancer ? Mais, avant cela, faisons quelques dessins montrant à quoi ressemblent les deux changements de coordonnées qu'on a effectués.

Changement de coordonnées curvilinéaire

Considérons pour commencer le cas où $L(t)$ est proportionnel à t. On a alors

$$z' = z - vt$$

Sur la figure 3, chaque point – appelé aussi un *événement* – dans l'espace-temps a une paire de coordonnées z et t dans le repère stationnaire et une paire de coordonnées z' et t' dans le repère de l'ascenseur. Bien sûr $t' = t$, et nous avons omis les deux autres coordonnées spatiales x et y, qui ne changent pas entre le repère fixe et celui de l'ascenseur. Nous avons représenté les trajectoires temporelles des points pour lesquels z est fixe en pointillés et ceux pour lesquels z' est fixe en traits pleins [9].

Une idée fondamentale qu'il est important de comprendre est que les événements dans l'espace-temps existent indépendamment de leurs coordonnées, tout comme les points dans l'espace ne dépendent pas de la carte que nous utilisons. Les coordonnées ne sont qu'une sorte de *nomenclature* ou *étiquetage* pratique. Nous pouvons utiliser le repérage que l'on préfère. Nous y reviendrons après avoir examiné les figures 3 et 4.

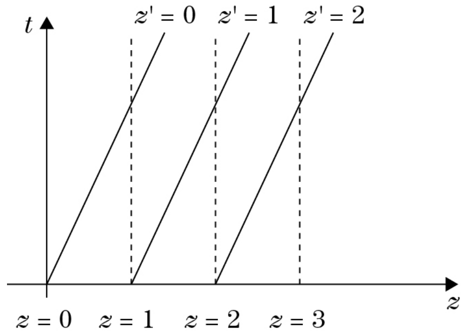


Figure 3 : Transformation de coordonnées linéaire. On peut dire, si l'on veut, que les coordonnées $(z',\ t')$ sont représentées *dans le système de coordonnées de base* $(z,\ t)$.

Un événement est *un point* sur la page. Il a une seule position sur la page, mais il a deux paires de coordonnées : une paire dans le repère $(z,\ t)$ et une autre paire dans le repère $(z',\ t')$.

C'est ce que l'on appelle une *transformation de coordonnées linéaire* entre les deux référentiels ou repères. Les trajectoires droites dans $(z,\ t)$ sont envoyées vers des trajectoires droites dans $(z',\ t')$. Noter qu'une trajectoire droite dans le repère de Minkowski à une

9. Le lecteur et la lectrice sont invités à se reporter au volume 3 de LMT s'ils ont besoin de rafraîchir leurs connaissances sur le diagramme de Minkowski représentant un espace-temps à une dimension spatiale et une dimension temporelle. Il est important de comprendre qu'un point *fixe* dans l'espace a une *trajectoire* dans l'espace-temps.

seule dimension spatiale est simplement une progression le long de l'unique axe spatial à vitesse constante.

Cela prend plus de signification si l'on ajoute les deux autres axes spatiaux x et y. Le diagramme de Minkowski devient un schéma dans un espace-temps à quatre dimensions, trois spatiales et une temporelle. On ne le représente plus graphiquement, mais on peut l'imaginer dans sa tête. Là encore ce qui est une trajectoire droite dans le repère spatio-temporel $(x,\ y,\ z,\ t)$ – c'est-à-dire une trajectoire spatiale à vitesse constante le long d'une droite quelconque dans l'espace – reste une trajectoire droite dans le repère spatio-temporel $(x',\ y',\ z',\ t')$.

Faisons la même chose pour le système de coordonnées accéléré. Pour rester simples, nous sommes à nouveau dans un espace-temps à deux dimensions : une seule dimension spatiale et la dimension temporelle. Nous pourrions si nous voulions imaginer aussi des coordonnées x et y, mais elles ne joueraient aucun rôle. Nous regardons donc les évènements dans l'espace-temps repérés dans le repère fixe $(z,\ t)$. Et nous regardons aussi leurs coordonnées dans le repère mobile $(z',\ t')$.

L'équation de transformation est maintenant l'équation (8) reliant z' et z. À nouveau, sur la figure 4, chaque point de l'espace-temps a deux paires de coordonnées : $(z,\ t)$ et $(z',\ t')$. Les trajectoires temporelles des évènements ayant leur coordonnée z fixe (c'est-à-dire qui « ne bougent pas » dans l'espace repéré par le repère stationnaire) sont toujours représentées par les lignes pointillées verticales.

Mais maintenant, sur la figure 4 qui utilise les coordonnées z et t comme axes perpendiculaires de représentation, les trajectoires temporelles des évènements ayant leur coordonnée z' fixe (c'est-à-dire les points qui ne bougent pas dans l'ascenseur en accélération) sont des paraboles couchées sur le côté. Nous pouvons même représenter des temps négatifs dans le passé. Imaginez en effet l'ascenseur qui se déplaçait initialement vers le bas avec une vitesse négative mais une accélération positive g (en d'autres termes, qui ralentissait). À un moment l'ascenseur s'arrête, puis repart vers le haut avec toujours la même accélération constante g.

Chaque parabole est décalée par rapport à la précédente d'une unité de longueur spatiale vers la droite [10].

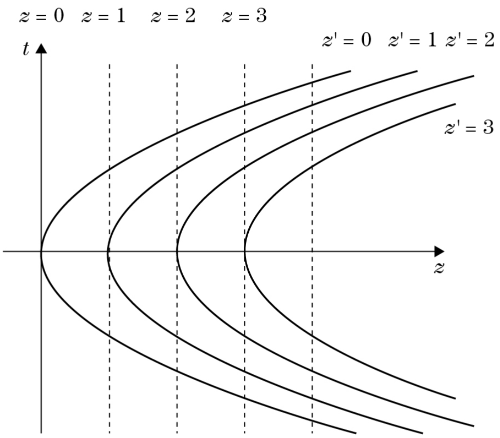


Figure 4 : Changement de coordonnées curvilinéaire.

Sans surprise, la figure 4 montre que les trajectoires droites dans un repère, par exemple celles qui auraient été droites dans le repère de l'ascenseur, ne sont plus des trajectoires droites dans l'autre, ici celui stationnaire. Elles deviennent des courbes. C'est évidemment vrai aussi quand on inverse le rôle des deux repères.

En ce qui concerne le temps, les droites correspondant à t fixe ou à t' fixe restent les mêmes droites horizontales dans les deux repères. Nous ne les avons pas représentées.

Il faut voir la figure 4 comme deux ensembles de coordonnées pour localiser chaque point de l'espace-temps. Un ensemble de coordonnées a des axes droits, tandis que le second – représenté dans le premier repère – est curviligne. Ses lignes $z' =$ constante sont en fait des courbes, tandis que ses lignes $t' =$ constante sont des droites horizontales. Il s'agit donc d'une *transformation de coordonnées curvilinéaire* dite encore *curviligne.*

10. Comme on n'incorpore pas encore les effets de la relativité restreinte, il n'y a pas de contraction des longueurs, et l'unité de longueur est la même dans les deux repères. On verra un traitement plus correct des repères accélérés à partir de la leçon 4, section Relativité restreinte, et dans les leçons ultérieures.

Insistons sur la manière d'interpréter et d'utiliser la figure 4, car il est fondamental de bien la comprendre si l'on veut comprendre la théorie de la relativité – la relativité restreinte et encore plus la relativité générale.

La page représente l'espace-temps – ici, une dimension spatiale et une dimension temporelle. Les points (= événements) dans l'espace-temps sont des points sur la page. Un événement *n'a pas deux positions* sur la page, c'est-à-dire dans l'espace-temps. *Il n'a qu'une seule position* sur la page. Mais cette position peut être localisée à l'aide de plusieurs systèmes de repérage c'est-à-dire plusieurs référentiels différents. Un référentiel n'est rien de plus qu'un ensemble complet d' « étiquettes », avec une étiquette (composée de deux nombres, car notre espace-temps ici est bidimensionnel) attachée à chaque point de l'espace-temps, c'est-à-dire à chaque événement.

Dans un espace à deux dimensions, le système de coordonnées peut être géométriquement simple, comme les axes orthogonaux d'un repère cartésien dans le plan. Mais ce n'est pas une nécessité. D'une part, sur Terre, qui n'est pas un plan, les axes ne sont pas des droites. Les axes utilisés par les cartographes et les navigateurs sont les méridiens et les parallèles. D'autre part, sur une surface 2D quelconque, plane ou pas, nous pouvons imaginer des lignes évoquant une assiette de nouilles pour servir de repère – tant que le système attache de manière bijective une paire de coordonnées à chaque point. C'est ce que fait la figure 4 dans l'espace-temps constitué d'une dimension temporelle et d'une dimension spatiale représentées sur la page. Nous en verrons d'autres dans la leçon 2.

Une chose qu'Einstein a comprise très tôt est celle-ci :

Il y a un lien entre la gravité et les changements de coordonnées curvilinéaires dans l'espace-temps.

La relativité restreinte portait seulement sur les transformations linéaires – des transformations qui envoient une vélocité uniforme dans un repère vers une vélocité uniforme dans un autre repère. Les transformations de Lorentz sont de cette nature. Elles transforment des lignes droites dans l'espace-temps en lignes droites dans l'espace-temps [11]. Cependant, si nous voulons imiter un champ gra-

11. Il faut bien comprendre que le concept de « ligne droite » dans l'espace-temps est en réalité relatif au *système de repérage* utilisé. En relativité restreinte on s'intéresse, essentiellement, seulement aux repères galiléens. Dans

vitationnel à l'aide d'une accélération, nous sommes alors en train de considérer un changement de coordonnées de l'espace-temps *curviligne*. Cette observation semble triviale. Et quand Einstein l'a faite, tous les physiciens le savaient et ont probablement pensé : « Oh oui, pas de quoi fouetter un chat. » Mais Einstein était très intelligent et très persistant. Il s'est rendu compte que s'il en poursuivait très loin les conséquences, il pourrait alors répondre à des questions auxquelles personne ne savait répondre.

Regardons un exemple simple de question à laquelle Einstein a répondu en utilisant les coordonnées curvilignes de l'espace-temps représentant l'accélération, et par conséquent, si les deux sont la même chose, la gravité. La question est : quel est l'effet de la gravité sur la lumière ?

Effet de la gravité sur la lumière

Quand Einstein s'est posé pour la première fois, vers 1907, la question de savoir « quel est l'effet de la gravité sur la lumière ? », la plupart des physiciens auraient répondu : « Il n'y a pas d'effet de la gravité sur la lumière. La lumière est lumière. La gravité est la gravité. Un rayon lumineux passant à proximité d'un objet massif se déplace en ligne droite. C'est une loi de la nature que la lumière se déplace de manière rectiligne. Et il n'y a aucune raison de penser que la gravité a un effet sur elle. »

Mais Einstein a dit : « Non, si ce principe d'équivalence entre l'accélération et la gravité est vrai, alors la gravité doit affecter la lumière. Pourquoi ? *Parce que l'accélération affecte la lumière.* » C'était encore un de ces arguments que vous pourriez expliquer à un enfant éveillé.

Imaginons qu'au temps $t = 0$, une lampe torche, ou si nous voulons être plus modernes un crayon laser, émette une impulsion de lumière dans une direction horizontale depuis le côté gauche de l'ascenseur, figure 5.

Noter que la question ne se pose pas de savoir si la lampe est fixe dans l'ascenseur ou fixe dans le repère stationnaire. On est en train

ces repères, une progression à vitesse constante est toujours vue comme une ligne droite dans l'espace-temps. Et ça reste vrai quand on change de repère, passant d'un repère galiléen à un autre repère galiléen.

de considérer un *évènement* – l'émission d'une impulsion (pensez par exemple à *un photon*) à un moment donné à un endroit donné dans l'espace-temps. Nous avons vu plus haut que les évènements ont une et une seule position dans l'espace-temps. Cette position peut, en revanche, être repérée de différentes façons.

La lumière se déplace ensuite vers la droite à la vitesse habituelle de la lumière, dénotée c. Étant donné que nous avons supposé que le référentiel stationnaire était inertiel, la lumière se déplace en ligne droite dans ce référentiel stationnaire.

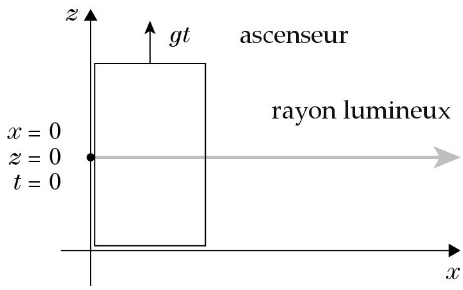


Figure 5 : Trajectoire d'un rayon lumineux représentée dans le *référentiel stationnaire*. Noter que cette fois on utilise *deux coordonnées spatiales*. Et on n'a pas représenté la coordonnée temporelle.

Les équations de la trajectoire du rayon lumineux sont

$$\begin{aligned} x &= ct \\ z &= 0 \end{aligned} \tag{12}$$

La première équation dit simplement que la lumière progresse à travers l'ascenseur à la vitesse de la lumière, et la deuxième équation dit que, dans le repère stationnaire, c'est une trajectoire rectiligne horizontale – pas de surprise.

Regardons la forme que prennent ces équations dans le repère prime. Étant donné que dans le changement de coordonnées, $x' = x$ et $t' = t$ (pour simplifier, on va même continuer à utiliser le temps universel t), la première équation devient

$$x' = ct$$

Et la deuxième équation du système (12), $z = 0$, en vertu de l'équation (8) page 10, prend la forme plus intéressante

$$z' = -\frac{g}{2}t^2$$

Elle indique qu'en même temps qu'il traverse l'ascenseur, le rayon lumineux accélère vers le bas – vers le plancher – comme attiré par la gravité.

On peut même éliminer la variable t dans les deux équations écrites dans le repère prime. Le temps t joue en effet simplement le rôle de paramètre dans une représentation paramétrique d'une courbe. On obtient une fonction liant la variable indépendante x' et la variable dépendante z'. Son graphe est la trajectoire du rayon lumineux dans le repère $(x',\ z')$. L'expression de la fonction est

$$z' = -\frac{g}{2c^2}x'^2 \tag{13}$$

On voit, comme on s'y attendait, que la trajectoire du rayon lumineux vue dans le repère de l'ascenseur n'est plus une droite, mais une parabole [12], figure 6.

Mais, dit Einstein, si l'effet de l'accélération est de courber la trajectoire d'un rayon lumineux, il doit en être de même de l'effet de la gravité.

Andy : *Fichtre, Lenny, c'est vraiment simple ! C'est tout ?*

Lenny : *Oui, Andy, c'est tout. Et tu peux être sûr qu'il y a beaucoup de physiciens qui se sont mordus les lèvres de ne pas l'avoir vu.*

Pour résumer, dans le repère stationnaire (figure 5), la trajectoire du photon est une droite, alors que dans le repère de l'ascenseur (figure 6) c'est une parabole.

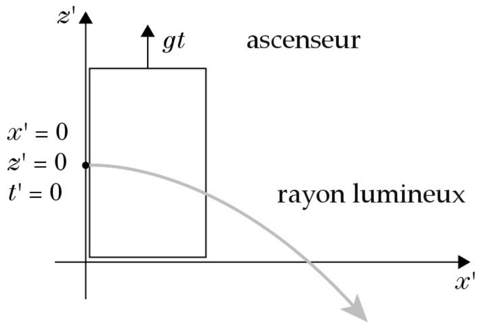


Figure 6 : Trajectoire du rayon dans le *référentiel de l'ascenseur.*

12. Noter qu'il est important que l'ascenseur soit *en accélération.*

Imaginons trois personnes en train de discuter. Je suis dans l'ascenseur, et je dis : « La gravité attire le rayon lumineux vers le bas. » Vous êtes dans le repère stationnaire, et vous dites : « Non, c'est juste que l'ascenseur accélère vers le haut ; cela donne l'impression que le rayon lumineux suit une trajectoire courbe. » Et Einstein dit : « C'est la même chose ! »

Ce raisonnement le convainquit qu'un champ gravitationnel devait courber un rayon lumineux. Pour autant que je sache, aucun autre physicien à l'époque n'avait déjà compris cela.

En conclusion, nous voyons qu'il est utile de penser aux changements de coordonnées curvilignes dans l'espace-temps. Quand on examine les changements de coordonnées curvilignes, *la forme des lois de Newton change*. L'une des choses qui se produisent est que des champs gravitationnels *apparents* – appelés aussi *fictifs* ou parfois *effectifs* – se matérialisent, qui sont physiquement impossibles à distinguer des champs gravitationnels concrets.

Mais sont-ils vraiment physiquement indiscernables ? Dans certains contextes oui, mais pas dans tous. Passons maintenant aux champs gravitationnels concrets, à savoir les champs gravitationnels créés par des objets massifs comme le Soleil ou la Terre.

Forces de marée

La figure 7 représente la Terre, ou le Soleil, ou n'importe quel corps massif. L'accélération gravitationnelle ne pointe pas verticalement vers le bas de la page. Elle pointe vers le centre du corps.

Il est évident qu'il n'y a aucun moyen d'effectuer un changement de coordonnées, comme nous en avons fait dans les sections précédentes, qui éliminerait l'effet du champ gravitationnel. Néanmoins, si vous êtes dans un petit laboratoire dans l'espace en chute libre vers la Terre, ou vers n'importe quel objet massif que vous êtes en train d'étudier, vous penserez que dans ce laboratoire il n'y a pas de champ gravitationnel.

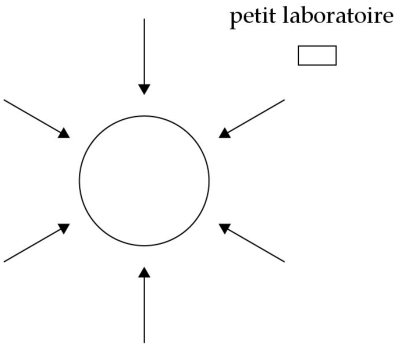


Figure 7 : Champ gravitationnel centripète d'un corps massif et petit laboratoire tombant en chute libre vers le corps. On ne ressent aucune gravité dans le laboratoire.

Exercice 1 : Si nous tombons en chute libre dans un champ gravitationnel uniforme, montrez que nous ne ressentons aucune gravité et que les objets flottent autour de nous comme dans la Station spatiale internationale.

Mais, encore une fois, il n'y a aucun moyen d'effectuer un changement de coordonnées qui éliminerait *globalement*, dans tout l'espace, le fait qu'il existe un champ gravitationnel pointant vers le centre du corps. Par exemple, une transformation très simple similaire aux équations (12) pourrait éliminer la gravité dans une région de l'espace d'un côté de la Terre, mais la même transformation augmenterait le champ gravitationnel de l'autre côté. Aucune transformation même plus complexe ne résoudrait le problème.

Une façon de comprendre pourquoi nous ne pouvons pas nous débarrasser de la gravité est de penser à un objet qui n'est pas petit par rapport au champ gravitationnel. Mon exemple favori est un homme d'une taille de 1000 kilomètres qui tombe en position verticale dans le champ gravitationnel terrestre, figure 8. Comme il est très grand, les différentes parties de son corps ressentent une attraction différente. N'oubliez pas que plus vous êtes loin, plus la gravité est faible. Dans sa position, sa tête ressent une gravité plus faible que ses pieds. Ceux-ci sont attirés plus fortement. L'homme a l'impression d'être étiré, et cette sensation lui signale qu'il y a un objet massif à proximité. Le sentiment d'inconfort qu'il ressent, en raison du champ gravitationnel non uniforme, ne peut pas être supprimé en passant à un repère en chute libre. En effet, aucun changement dans la description mathématique de l'espace-temps ou de la nature en général ne peut modifier un phénomène physique.

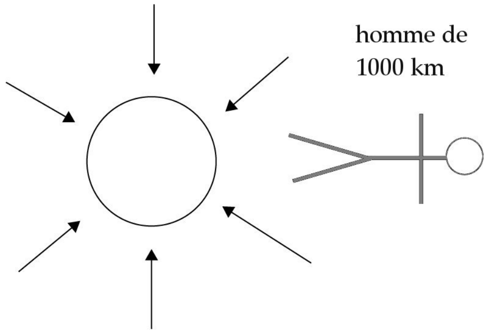


Figure 8 : Un homme de 1000 km tombant vers la Terre.

Les forces qu'il ressent sont appelées *forces de marée*, car elles jouent également un rôle important dans le phénomène des marées. Elles ne peuvent pas être uniformément supprimées par un changement de coordonnées.

Voyons aussi ce qui se passe s'il ne tombe pas verticalement mais latéralement, perpendiculaire à un rayon. Dans ce cas, sa tête et ses pieds seront à la même distance de la Terre. Les deux seront soumis à la même force en magnitude pointant vers la Terre. Mais comme les directions des forces sont radiales, elles ne sont pas parallèles. La force sur sa tête et la force sur ses pieds auront toutes deux une composante le long de son corps. On voit tout de suite que les forces de marée le comprimeront un peu. Cette sensation de compression n'est pas non plus quelque chose que nous pouvons supprimer par un changement de coordonnées. Être étiré ou comprimé, ou les deux, par le champ gravitationnel terrestre – si vous êtes assez grand – est un phénomène concret qu'on ne peut pas éliminer.

En résumé, il n'est pas tout à fait exact que la gravité soit équivalente à se déplacer dans un référentiel accéléré.

Andy : *Ah ah ! En définitive Einstein avait tort.*

Lenny : *Il est vrai, Andy, qu'Einstein s'est parfois trompé. Mais ce n'est pas le cas ici. Il lui suffisait de préciser que son principe d'équivalence s'appliquait à un champ gravitationnel* uniforme.

Ce qu'Einstein voulait réellement dire, c'est que des petits objets, pendant un court laps de temps, ne peuvent pas faire la différence entre un champ gravitationnel et un référentiel accéléré.

Cela soulève la question suivante : si je vous présente un champ de force, existe-t-il une transformation des coordonnées qui l'éliminerait ?

Par exemple, le champ de force à l'intérieur de l'ascenseur, créé par son accélération uniforme par rapport à un référentiel inertiel[13], est un champ de force vertical pointant vers le bas et uniforme partout. Il y a une transformation qui l'annule : prendre simplement les coordonnées z au lieu des coordonnées z'. C'est un changement de coordonnées non linéaire qui élimine le champ de force.

Avec d'autres types de changement de coordonnées, vous pouvez rendre le champ gravitationnel plus compliqué, par exemple des changements qui affectent également la coordonnée x. Ils peuvent déformer ou « tordre » le champ gravitationnel dans la direction de l'axe des x. Vous pouvez simultanément accélérer le long de l'axe des z tout en oscillant le long de l'axe des x. Quel type de champ gravitationnel percevrez-vous ? Un champ compliqué : il aura une composante verticale et une composante oscillante fonction du temps le long de l'axe horizontal.

Si au lieu de l'ascenseur vous montez dans un manège, et au lieu des coordonnées primes précédentes vous utilisez des coordonnées polaires $(r,\ \theta,\ t)$ fixes par rapport au manège, un objet fixe dans le repère stationnaire initial, ou avec un mouvement simple comme le rayon lumineux, va avoir un mouvement étrange dans le repère du manège. Vous penserez peut-être avoir découvert un champ gravitationnel répulsif. Quoi qu'il en soit, le changement de coordonnées inverse révélera que votre champ apparemment compliqué n'est que

13. Pour bien comprendre cet exemple, on peut oublier le champ gravitationnel terrestre, qui ne joue aucun rôle important. Ou si l'on préfère on peut imaginer l'ascenseur loin de tout corps massif dans l'univers et accéléré par exemple par un réacteur – ou par un physicien espiègle.

la conséquence d'un changement de repère. En choisissant des coordonnées bizarres, vous pouvez créer des champs gravitationnels fictifs, autrement dit apparents, très compliqués. Néanmoins, ils ne seront pas réels, au sens où ils ne seront pas dus à la présence de corps massifs.

Si je vous donne le champ de force partout dans l'espace, ou dans une région de l'espace, comment déterminez-vous s'il est fictif ou réel, c'est-à-dire s'il s'agit de la sorte de faux champ gravitationnel résultant d'un changement de coordonnées vers un référentiel avec des accélérations dans tous les sens par rapport à un repère inertiel, ou bien si c'est un authentique champ gravitationnel ?

Quand on parle de gravité newtonienne, il existe un moyen simple. Vous calculez si le champ cause des forces de marée, c'est-à-dire s'il comprimera ou étirera un objet comme l'homme de 1000 km. Si les calculs sont trop gros car le champ est très compliqué, et si vous êtes dans des conditions expérimentales, vous pouvez toujours prendre un objet, une masse, un cristal. Vous le laissez tomber librement et vous voyez quelles contraintes et tensions s'exercent dessus. Si le cristal est assez grand, les phénomènes seront détectables. Et si vous détectez des contraintes et déformations, vous en déduisez qu'il s'agit d'un champ gravitationnel réel.

Nous y reviendrons dans le dernier chapitre sur les ondes gravitationnelles. Il s'agira de champ gravitationnel réel, i.e. qu'on ne peut pas éliminer par un changement de repère, mais il ne s'agira pas de champ dû à un corps massif à proximité.

Einstein s'est posé la question suivante : *quel type de mathématiques permettent de répondre à la question de savoir si un champ de force est un authentique champ gravitationnel ou pas ?*

Géométrie non euclidienne

Après son travail sur la relativité restreinte, et après avoir pris connaissance de la structure mathématique que Minkowski[14] avait conçue pour la décrire, Einstein savait qu'à la relativité restreinte était associée une géométrie. Quittons un instant la gravité, pour

14. Hermann Minkowski (1864–1909), mathématicien et physicien théorique polonais puis allemand.

nous rafraîchir la mémoire sur cette idée importante. La relativité restreinte et la théorie classique des champs était le sujet principal du troisième volume de la série LMT. Ici la seule chose que nous allons utiliser est que l'espace-temps a une géométrie.

Dans la géométrie de Minkowski de la relativité restreinte, il existe une sorte de distance entre deux points, c'est-à-dire entre deux *événements* dans l'espace-temps, figure 9.

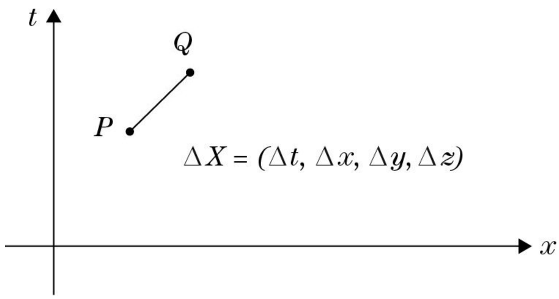


Figure 9 : Géométrie de Minkowski : 4-vecteur joignant P et Q.

La distance entre P et Q n'est pas la distance euclidienne ordinaire à laquelle on pourrait être tenté de penser. Elle est définie comme suit. Appelons ΔX le 4-vecteur allant de P à Q. Au couple de points P et Q on associe une quantité notée $\Delta\tau$ définie par

$$\Delta\tau^2 = \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$$

Observez que $\Delta\tau$ ne satisfait pas les propriétés habituelles d'une distance[15]. En particulier, $\Delta\tau^2$ peut être positif ou négatif ; et il peut être nul pour deux événements qui ne coïncident pas. Le lecteur et la lectrice sont renvoyés au volume 3 de LMT pour plus de détails. Nous ne faisons ici qu'un bref rappel.

La quantité $\Delta\tau$ est appelée le *temps propre* entre P et Q. C'est un invariant dans les transformations de Lorentz. C'est pourquoi il est considéré comme une sorte de distance, tout comme dans l'espace euclidien ordinaire à trois dimensions (« l'espace 3D »), la distance entre deux points, dont le carré est défini par $\Delta x^2 + \Delta y^2 + \Delta z^2$, est invariante dans les isométries (translations, rotations, etc.).

15. Pour cette raison, on l'appelle parfois une *pseudo distance* ou une *pseudo métrique*.

Nous définissons aussi la quantité Δs par

$$\Delta s^2 = -\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$$

On appelle Δs la *distance propre* entre P et Q. Bien sûr, $\Delta\tau$ et Δs ne sont pas deux concepts réellement différents. Ils diffèrent juste par le facteur imaginaire i. Ce ne sont que deux façons de parler de la « distance » de Minkowski entre P et Q. Certains physiciens préfèrent utiliser la distance propre et d'autres le temps propre entre deux évènements.

Einstein connaissait la géométrie non euclidienne de la relativité restreinte. Dans son travail pour inclure la gravité et pour étudier les conséquences du principe d'équivalence, il s'est également rendu compte que la question que nous avons posée à la fin de la section précédente – y a-t-il un changement de coordonnées qui supprime les forces ? – était très similaire à un certain problème de mathématiques qui avait été beaucoup étudié par Riemann. Il s'agit de déterminer si une géométrie est plane ou pas.

Géométrie riemannienne

Quand peut-on dire d'une géométrie qu'elle est *plane* ? Intuitivement, l'idée est la suivante : la géométrie d'une page est plane. La géométrie de la surface d'une sphère ou d'une section de la surface d'une sphère n'est pas plane. La *géométrie intrinsèque* de la page reste plane même si on enroule la page comme sur la figure 10. Nous développerons mathématiquement l'idée plus loin.

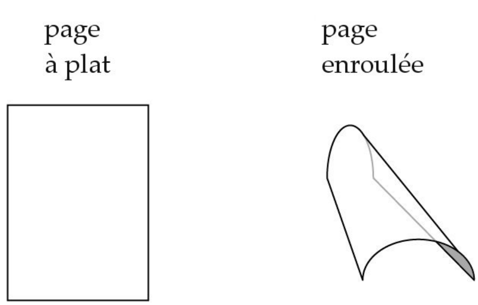


Figure 10 : La géométrie intrinsèque de la page reste plane.

Pour l'instant, disons simplement que la géométrie intrinsèque d'une surface est la géométrie qu'une coccinelle se déplaçant librement dessus, équipée de minuscules outils d'arpentage, verrait si elle essayait d'établir un carte d'état-major de la surface.

Si la coccinelle travaille soigneusement, elle pourra identifier des collines et des vallées s'il y en a, mais elle n'aura aucun moyen de déterminer si la page est enroulée ou pas. Nous, nous le voyons parce que *pour nous* la page est plongée dans l'espace euclidien à trois dimensions dans lequel nous vivons. En déployant la page, nous pouvons rendre sa planéité évidente.

Einstein s'est rendu compte qu'il y avait une grande similitude entre les deux questions :

1. Savoir si une géométrie est intrinsèquement plane ou pas.
2. Savoir si un espace-temps contient un champ gravitationnel réel, ou bien si le champ éventuel que l'on ressent peut être éliminé par un changement de référentiel.

Riemann avait étudié la première question. Mais il n'avait jamais pensé même en rêve à des géométries avec un signe moins dans la définition du carré de la distance. Il pensait à des géométries non euclidiennes comparables à beaucoup d'égards à la géométrie euclidienne – pas à la géométrie de Minkowski.

Commençons par les mathématiques de la géométrie riemannienne, c'est-à-dire la géométrie d'espaces où la distance entre deux points peut ne pas être la distance euclidienne, mais dans lesquels le carré de la distance est toujours positif[16].

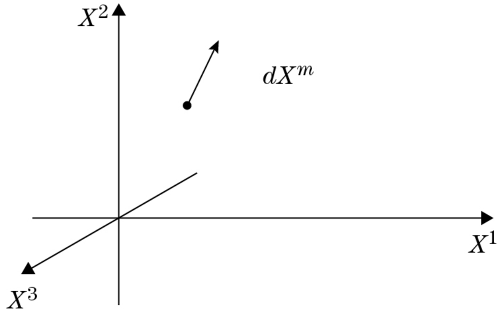


Figure 11 : Petit déplacement entre deux points proches dans l'espace.

16. En mathématiques, on les appelle des *distances définies positives*.

Nous regardons deux points dans un espace, figure 11. Dans l'exemple, il y a trois dimensions, donc trois axes, X^1, X^2 et X^3. Il pourrait y en avoir davantage. Ainsi, un point a trois coordonnées, que nous pouvons noter sous la forme X^m, où l'index m va 1 à 3, ou au nombre d'axes quel qu'il soit. Le petit déplacement entre un point et un autre à proximité a trois composantes, qui peuvent être notées ΔX^m ou, si le déplacement est infinitésimal, dX^m.

Si cet espace a la géométrie euclidienne ordinaire, le carré de la longueur de dX^m est donné par le théorème de Pythagore

$$dS^2 = (dX^1)^2 + (dX^2)^2 + (dX^3)^2 + \dots \tag{14}$$

Il est d'usage d'utiliser S majuscule. En trois dimensions, il y a trois termes dans la somme. Si nous sommes en deux dimensions, il y a deux termes. Si l'espace est à 26 dimensions, il y en a 26, et ainsi de suite. La formule (14) est la formule de la distance euclidienne entre deux points dans un espace euclidien.

Par souci de simplicité et de facilité de visualisation, concentrons-nous sur un espace à deux dimensions. Il peut s'agir du plan ordinaire ou d'une surface bidimensionnelle quelconque que nous pouvons commodément visualiser plongée dans l'espace euclidien à trois dimensions, figure 12.

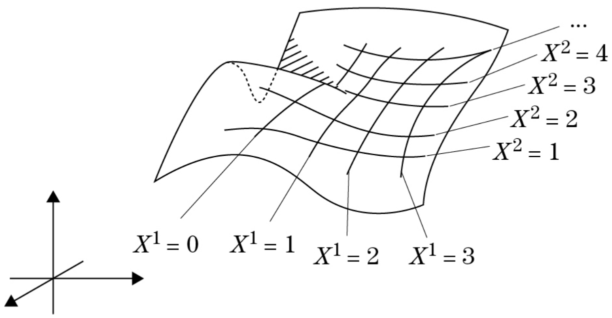


Figure 12 : Variété bidimensionnelle (c'est-à-dire surface 2D) et ses coordonnées curvilignes, vues plongées dans l'espace euclidien à trois dimensions ordinaire.

Il n'y a rien de spécial au nombre 2 des dimensions de l'espace considéré, sauf qu'il s'agit alors d'une surface dans le sens ordinaire du terme et qu'elle est facile à visualiser, y compris sa géométrie avec éventuellement des bosses et des creux intrinsèques, dans l'espace 3D ordinaire. Les mathématiciens parlent parfois de « surfaces » même lorsqu'elles ont plus que deux dimensions [17]. Ou alors ils emploient le terme de *variété*, que Lenny n'aime pas qu'Andy utilise car ça fait jargon de spécialiste de géométrie différentielle :-)

Gauss avait déjà compris que sur les surfaces gondolées, ou pour parler plus solennellement ayant en chaque point une certaine courbure, la formule de la distance entre deux points proches n'était plus aussi simple que l'équation (14). En effet, il ne faut pas se méprendre quand on regarde la figure 12. La surface est représentée *plongée dans l'espace euclidien 3D ordinaire*. Mais c'est seulement pour la commodité de la perception.

Nous devons penser à cette surface comme à un espace en soi, muni d'un système de coordonnées pour localiser n'importe quel point, avec des lignes courbes correspondant à une coordonnée constante, etc., et où une distance a été définie. Il faut oublier l'espace euclidien 3D dans lequel nous avons plongé notre surface. La distance entre deux points sur la surface n'est certainement pas leur distance dans l'espace 3D de plongement, et n'est même pas nécessairement définie sur la surface avec l'équivalent de l'équation (14). Nous verrons bientôt exactement comment ces distances sont exprimées mathématiquement.

Riemann a généralisé ces surfaces et leur métrique (= la formule pour calculer les distances, voir infra) à n'importe quelle dimension. Mais continuons à utiliser notre image à deux dimensions pour soutenir l'intuition. Et comme toujours progressons pas à pas pour ne rater aucun détail important.

La première chose que nous faisons avec une surface est de la munir d'un système de coordonnées, ce qui nous permettra de travailler aisément avec les points. Imaginons que nous tracions à la craie des lignes de coordonnées. On ne s'occupe pas d'essayer de les faire « droites ». Du reste quand la surface est vraiment gondolée, il n'y a probablement pas de lignes que nous pourrions qualifier clairement de droites. Mais nous les dénotons toujours par X.

17. Leur logique est que dans un espace de dimension n, un sous-espace de dimension $n-1$ est une « surface » dans l'espace total, comme un drap est une surface dans l'espace 3D.

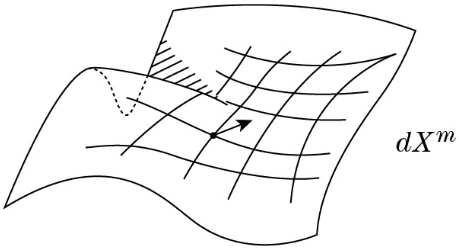


Figure 13 : Deux points proches et le petit déplacement dX^m allant de l'un à l'autre.

Les valeurs des coordonnées X^m d'un point ne sont pas directement liées à des distances. Ce ne sont que des étiquettes numériques. Les points $(X^1 = 0,\ X^2 = 0)$ et $(X^1 = 1,\ X^2 = 0)$ ne sont pas nécessairement séparés par une distance d'une unité. Considérons maintenant deux points voisins, figure 13. Les coordonnées du deuxième sont seulement un peu différentes de celles du premier. Mais, contrairement à la figure 11, qui était encore un espace euclidien, nous sommes maintenant sur une surface à la courbure partout arbitraire et avec un système de repérage arbitraire.

Maintenant, nous *définissons* une distance sur la surface pour des points séparés par un petit décalage comme dX^m. Ce n'est plus aussi simple que l'équation (14), bien qu'il y ait des similitudes avec elle. Voici la nouvelle définition de dS^2 :

$$\boxed{dS^2 = \sum_{m,\ n} g_{mn}(X)\ dX^m dX^n \qquad (15)}$$

Les fonctions $g_{mn}(X)$'s, prises dans leur ensemble, forment ce que l'on appelle la *métrique* de l'espace considéré. C'est un ensemble de fonctions *qui dépendent de la position* X. En chaque point il y a une collection différente. Mais les fonctions g_{mn} ont les propriétés habituelles de continuité et de dérivabilité, par rapport aux déplacements, de fonctions « qui se comportent bien ».

La formule (15) est très générale et s'applique que la variété soit plane ou qu'elle ait une certaine courbure en chaque point. Elle est très importante en géométrie riemannienne – c'est pourquoi nous l'avons encadrée. Nous verrons bientôt qu'elle est aussi très importante dans la géométrie de Minkowski de la relativité.

La convention de sommation d'Einstein nous permettra de réécrire la formule (15) sous une forme plus légère. Elle est expliquée plus loin dans la section « Interlude mathématique : Convention de sommation d'Einstein ».

Soit dit en passant, la formule (15) s'applique même aux géométries *planes* munies de coordonnées curvilignes. Supposons que vous preniez une géométrie plane, comme la surface de la page, mais que pour une raison ou pour une autre vous utilisiez des coordonnées curvilignes pour localiser les points et que vous vous demandiez comment calculer la distance entre deux points proches. Le carré de la distance entre deux points proches l'un de l'autre sera alors une *forme quadratique* des dX^m. Une forme quadratique veut dire une somme de termes, dont chacun est le produit de deux décalages infinitésimaux de coordonnées, multiplié par un coefficient comme g_{mn} qui dépend de X.

La surface de la Terre offre un bon exemple de calcul de distance sur une surface bombée. Regardons la distance entre deux points proches repérés par leur longitude et leur latitude, figure 14. Soit R le rayon de la Terre. Les deux points ont pour coordonnées $(\phi,\ \theta)$ et $(\phi + d\phi,\ \theta + d\theta)$, où θ est la latitude et ϕ la longitude.

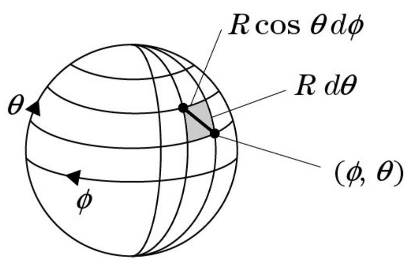


Figure 14 : Calcul de distance à la surface de la Terre. On a représenté les points $(\phi,\ \theta)$ et $(\phi + d\phi,\ \theta + d\theta)$, et le segment qui les joint.

On applique le théorème de Pythagore dans la petite région grisée, approximativement rectangulaire et plate, pour calculer le carré de la longueur de sa diagonale. Les côtés le long d'un méridien ont pour longueur $Rd\theta$. Et les côtés le long d'un parallèle ont pour longueur $Rd\phi$ corrigée du cosinus de la latitude. À l'équateur c'est $Rd\phi$ sans correction, tandis qu'au pôle c'est zéro.

La formule générale pour le carré de la distance est

$$dS^2 = R^2 \left[d\theta^2 + (\cos\theta)^2 d\phi^2\right] \tag{16}$$

C'est un exemple de distance au carré qui n'est pas simplement égale à $d\theta^2 + d\phi^2$, mais a des coefficients devant les produits de différentielles. Ici, le coefficient intéressant est $(\cos\theta)^2$ devant l'un des termes $(dX^m)^2$. Ce coefficient $(\cos\theta)^2$ s'écrit plus généralement $\cos^2\theta$. Notez aussi que dans la métrique (16), il n'y a pas de terme croisé de la forme $d\theta\, d\phi$, parce que les coordonnées curvilignes naturelles que nous avons utilisées sur la sphère sont toujours localement orthogonales quel que soit le point considéré [18].

Dans d'autres exemples – sur la sphère avec des coordonnées plus compliquées, ou sur une surface courbe plus générale comme sur la figure 13 – où les coordonnées ne sont pas nécessairement localement perpendiculaires, la formule pour dS^2 comprendra des termes en $dX^m dX^n$. Mais ce sera toujours une forme quadratique. Il n'y aura jamais de terme comme $d\theta^3$. Il n'y aura jamais non plus de terme du premier degré. Chaque terme sera quadratique. En outre, en géométrie riemannienne, en chaque point X, la forme quadratique définissant la métrique à cet endroit-là sera définie positive.

Vous vous demandez peut-être pourquoi nous définissons la distance dS seulement pour de petits déplacements (en fait infinitésimaux). C'est que pour définir la distance entre deux points A et B éloignés, il faut d'abord savoir de quoi on parle. Il peut y avoir des bosses et des creux sur la surface entre les deux. Nous pourrions définir la distance comme ceci : placer une cheville en A, une autre en B, et tendre une ficelle entre les deux. Cela serait une notion de distance. Mais une ficelle peut passer par plusieurs chemins tout en étant tendue. On peut faire le tour de la colline d'une façon et quelqu'un d'autre le ferait d'une autre. Pensez simplement sur Terre à aller du pôle Nord au pôle Sud.

18. Les coordonnées sphériques de la figure 14, bien qu'elles soient un peu plus sophistiquées que les coordonnées cartésiennes, étaient déjà utilisées au XVIe siècle, tandis que les coordonnées cartésiennes n'ont commencé à être utilisées pour faire de la géométrie analytique qu'au XVIIe siècle. Descartes s'est inspiré des travaux des cartographes avant lui.

En outre, même s'il n'y a qu'une seule réponse, nous devons connaître la géométrie de la surface en chaque point de la région où se trouvent A et B, non seulement pour calculer la distance mais pour savoir où placer la corde. C'est pourquoi, la notion de distance entre deux points quelconques est plus compliquée qu'en géométrie euclidienne. En revanche entre deux *points voisins* le problème ne se pose pas. En effet, localement, une surface lisse peut être approximée par son plan tangent et les lignes de coordonnées curvilignes par des lignes droites – pas nécessairement perpendiculaires mais droites.

Tenseur métrique

Approfondissons la géométrie d'une surface courbe et ses liens avec l'équation (15) qui définit la distance entre deux points voisins sur celle-ci. Rappel de l'équation

$$dS^2 = \sum_{m,\ n} g_{mn}(X)\ dX^m dX^n$$

Afin de bien comprendre la géométrie de la surface et son comportement, imaginons que nous disposions des petits éléments du jeu de construction appelé Tinkertoys le long de la surface courbe. Par exemple, ils pourraient suivre approximativement les lignes de coordonnées sur la surface. Nous ajouterions également des éléments en diagonale. Cela créerait une sorte de treillis comme le montre la figure 15. Mais tout treillis raisonnablement dense, disposé comme on veut, ferait aussi l'affaire.

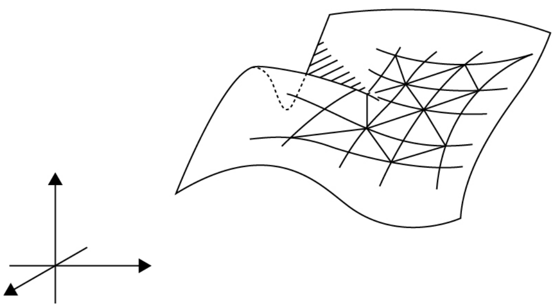


Figure 15 : Treillis d'éléments Tinkertoys disposés sur la surface.

Imaginons maintenant que nous soulevions notre treillis de la surface. Parfois, il gardera sa forme de manière rigide, parfois non. Il ne conservera pas sa forme s'il est possible de passer de la forme initiale à une nouvelle forme sans forcer aucun élément Tinkertoys à être étiré, comprimé ou plié.

Dans certains cas, il sera même possible *de le mettre à plat.* C'est le cas, par exemple, sur la figure 10 allant de la forme de droite à la forme de gauche – qui est juste une page à plat.

Exercice 2 : Existe-t-il une surface courbe avec un treillis de petites tiges disposées dessus qui ne puisse pas être mise à plat, mais qui puisse néanmoins changer de forme ?

Réponse : Oui. D'après le *Theorema egregium* de Gauss, une surface peut être modifiée sans l'étirer ni la comprimer nulle part à condition que nous préservions partout sa courbure gaussienne. Par exemple, il est possible de modifier ainsi une section de paraboloïde hyperbolique.

Nous verrons que pour que la surface initiale puisse prendre d'autres formes il faut que la collection des coefficients g_{mn} de l'équation (15) satisfasse certaines propriétés.

La collection des g_{mn} porte un nom. Elle s'appelle le *tenseur métrique.* C'est l'objet mathématique qui nous permet de calculer la distance entre deux points voisins sur notre surface de Riemann.

Prêtez attention au fait que les g_{mn} sont des fonctions de la position X (le point sur la variété). Donc, à proprement parler, nous parlons d'un *champ de tenseurs*, appelé aussi *champ tensoriel.*

Lorsque le treillis d'éléments Tinkertoys – dont la fonction est simplement d'imposer que la distance entre n'importe quelle paire de points proches ne change pas – peut être mis à plat, la géométrie de la surface est dite *intrinsèquement plane* ou *intrinsèquement plate.* Nous le définirons plus rigoureusement plus loin.

Parfois, au contraire, le treillis de petites tiges ne peut pas être mis à plat. Par exemple sur la sphère, si nous disposons initialement un treillis triangulant une grande partie de la sphère, nous ne pourrons pas le mettre à plat [19].

19. C'est un problème bien connu des cartographes, qui a conduit à l'invention de divers types de cartes du monde, la plus célèbre étant la projection Mercator inventée par le cartographe flamand Gerardus Mercator (1512–1594).

La question que nous nous posons – en poursuivant notre exemple imagé avec le jeu Tinkertoys – est la suivante : si je construisais un treillis de petits bâtonnets couvrant une surface, et que je vous donnais la longueur de chaque bâtonnet, comment pourriez-vous me dire, sans construire vous-même le treillis pour essayer, s'il s'agit d'une surface intrinsèquement plate ou d'une surface intrinsèquement courbe, c'est-à-dire ne pouvant pas être mise à plat ?

Formulons le problème plus précisément et mathématiquement. Nous partons du tenseur métrique $g_{mn}(X)$, qui est une fonction de la position, dans un système de coordonnées.

Gardez à l'esprit qu'il existe de nombreux systèmes de coordonnées curvilignes différents possibles sur la surface, et que dans chaque système de coordonnées, le tenseur métrique sera différent. Il aura des composantes différentes, tout comme un même 3-vecteur dans l'espace euclidien 3D ordinaire a des composantes différentes selon la base utilisée pour le représenter, mais en plus les composantes dans un système varieront avec la position.

Je sélectionne un système de coordonnées et je vous donne le tenseur métrique de ma surface. C'est-à-dire, je vous donne la distance entre chaque paire de points voisins. Puis je vous pose la question : ma surface est-elle intrinsèquement plane ou non ?

Pour répondre, vous pourriez penser à « vérifier π ». Voici comment vous procèderiez. Travaillons avec une surface 2D pour rester simple. Et visualisons-la plongée dans l'espace euclidien 3D habituel, comme montré sur la figure 12. Vous choisissez un point et tracez un cercle à une distance r autour de lui. Puis vous mesurez sa circonférence et la divisez par $2r$. Si vous obtenez 3,14159… vous dites que la surface est plane. Sinon vous dites qu'elle est intrinsèquement courbée. Notez que cette procédure est adaptée pour une surface bidimensionnelle, sous certaines conditions, mais qu'elle est plus compliquée pour des variétés à plus de dimensions.

Mathématiquement parlant, en quoi consiste prendre le tenseur métrique et répondre à la question mon espace est-il plat ? Et d'abord que veut dire être plat ? Par définition, cela signifie ceci :

L'espace est plat[20] *si nous pouvons trouver un changement de coordonnées, c'est-à-dire un nouveau système de coordonnées, dans lequel,* en tout point de l'espace, *la formule du carré de la distance* dS^2 *devient simplement* $(dX^1)^2 + (dX^2)^2 + \ldots + (dX^n)^2$, *comme c'est le cas en géométrie euclidienne.*

Il n'est pas nécessaire que les $g_{mn}(X)$ *initiaux* soient partout la matrice unitaire, avec des uns sur la diagonale et des zéros ailleurs – comme si l'équation (15) n'était que le théorème de Pythagore. Mais nous devons trouver un nouveau système de coordonnées dans lequel le tenseur métrique soit partout la matrice unitaire.

Vu comme ça, il y a une vague similitude avec la question de savoir si on peut trouver un changement de référentiel qui supprime le champ gravitationnel. En fait, il s'avère qu'il ne s'agit pas d'une vague similitude, mais d'un parallèle étroit. La question est : pouvons-nous trouver un nouveau système de coordonnées dans notre espace de Riemann (et plus tard de Minkowski) qui supprime le caractère courbe du tenseur métrique initial g_{mn} ?

Pour répondre à cette question de nature géométrique, nous devons faire des mathématiques sans lesquelles il est impossible de comprendre la relativité générale. Ces mathématiques sont du calcul tensoriel et de la géométrie différentielle. Au début, cela semblera rébarbatif car nous travaillerons avec différents systèmes de coordonnées, des dérivées partielles de composantes, des flopées d'indices, etc. Mais nous verrons bientôt que c'est simple. Ces mathématiques, on l'a vu, ont été inventées par Ricci-Curbastro et Levi-Civita à la fin du XIXe siècle pour développer des travaux de Gauss et Riemann. Elles ont été un peu simplifiées par Einstein, qui a édicté des règles pour la position des indices et s'est astucieusement débarrassé de la plupart des symboles de sommation.

Avant d'expliquer ce qu'est la convention de sommation d'Einstein éliminant la plupart des symboles de sommation, consacrons une section au concept simple de variable muette.

20. On ne le répètera plus, mais quand on dit d'un espace qu'il est plat, on veut dire qu'il est intrinsèquement plat ou, s'il s'agit d'une surface 2D, intrinsèquement plan. Cependant le terme plat est d'emploi plus général. Par exemple on dit de l'espace euclidien 3D ordinaire qu'il est plat.

Interlude mathématique : Variables muettes

Nous sommes habitués aux équations où toutes les variables ont une signification mathématique ou physique. Un exemple physique est l'équation (7), reproduite ici :

$$L(t) = \frac{1}{2}gt^2$$

Cette célèbre équation a été trouvée par Galilée dans la première moitié du XVIIe siècle [21], avant l'invention du calcul différentiel et intégral. C'est d'ailleurs l'une des équations qui a conduit à l'invention du calcul différentiel et intégral par Newton et Leibniz. Elle décrit mathématiquement la chute d'un objet : L représente la distance de chute en fonction du temps, g représente l'accélération à la surface terrestre et t représente le temps.

Un autre exemple encore plus simple et purement mathématique est

$$A = ab$$

où a est la longueur d'un rectangle, b sa largeur et A son aire.

Nous connaissons aussi des équations où l'une des variables n'est qu'une notation mathématique pratique sans signification substantielle. Un exemple est l'identité bien connue exprimant la valeur de la somme de tous les carrés des entiers de 1 à m

$$\frac{m(m+1)(2m+1)}{6} = \sum_{n=1}^{n=m} n^2$$

Ici la variable m a un sens substantiel : c'est le nombre jusqu'auquel on additionne. Mais la variable n du côté droit n'a pas de signification substantielle. On pourrait aussi bien écrire

$$\frac{m(m+1)(2m+1)}{6} = \sum_{k=1}^{k=m} k^2$$

Ce serait exactement la même équation.

---

21. Nous la donnons ici sous sa forme moderne. Galilée (1564–1642) a seulement écrit que la distance de chute était proportionnelle au carré de sa durée – ce qui, si l'on y songe, est une découverte extraordinaire. Cela veut dire que dans la nature les équations du second degré jouent un rôle aussi naturel et fondamental que les lignes droites comme la corde d'un arc tendu.

La variable n – ou la variable k – est appelée une *variable muette*. Elle n'est utilisée que pour noter commodément la somme.

On rencontrera de nombreuses formules contenant une ou plusieurs variables muettes, exprimant le plus souvent des sommes, en relativité générale. Elles sont si fréquentes qu'Einstein a proposé une règle pour simplifier les notations de sommes, se débarrassant en particulier du signe Σ. Sa règle ou convention s'est avérée non seulement être une grande simplification, mais aussi être une technique très utile pour écrire des équations de relativité générale. Elle fournit en effet un guide, ou garde-fou pourrait-on dire, pour ne pas faire d'erreur, tout en ayant une signification en soi. Cette convention est le sujet de l'interlude mathématique suivant. Plus loin dans cette leçon ainsi que dans le reste du livre, nous découvrirons à quel point elle est utile.

Interlude mathématique : Convention de sommation d'Einstein

À mesure que nous progressons, nous découvrirons que certaines expressions ayant une forme typique apparaissent fréquemment dans les équations. Ce sera une expression avec un index μ pour écrire une somme de termes de la même forme. Voici un exemple. Pour le moment, peu importe ce que l'expression signifie ; c'est juste une de ces formes que nous rencontrerons encore et encore.

$$\sum_{\mu} V^{\mu} U_{\mu}$$

Il y a plusieurs points à noter. Tout d'abord, il y a une sommation sur μ, donc μ est un index muet. Le terme *index muet* est juste un autre nom, dans le contexte spécifique des vecteurs et des tenseurs, pour *variable muette*. Par conséquent, la lettre que nous utilisons n'a pas d'importance. Les expressions avec μ, comme ci-dessus, ou avec ν, comme ci-dessous, représentent exactement la même chose, d'où, comme nous l'avons vu, le terme *muet*.

$$\sum_{\nu} V^{\nu} U_{\nu}$$

Deuxièmement, l'index muet apparaît deux fois dans la même expression – pas une fois, pas trois fois, deux fois.

Enfin, l'index répété deux fois apparaît une fois en *exposant* et une fois en *indice*. On dit familièrement qu'il apparaît en haut et en bas. C'est la forme typique : une somme sur un index qui apparaît une fois en haut et une fois en bas.

La célèbre astuce d'Einstein pour alléger les notations – ce qu'on appelle la *convention de sommation d'Einstein* – consiste simplement à omettre le signe Σ. Et la règle est que chaque fois que nous rencontrons un terme comme $V^{\mu}U_{\mu}$, *avec un index en haut et le même index en bas*, nous devons automatiquement effectuer une somme sur l'index μ.

Nous pouvons dès maintenant appliquer la convention à la formule (15) que nous avons vue plus haut exprimant la forme générale de la métrique dans un espace riemannien (ou d'ailleurs aussi dans un espace de Minkowski, comme nous le verrons). C'était

$$dS^2 = \sum_{m,\ n} g_{mn}(X)\ dX^m dX^n$$

Avec la convention d'Einstein elle devient

$$dS^2 = g_{mn}(X)\ dX^m dX^n$$

Plus simple ! N'est-ce pas ?

Habituellement, si l'on n'oublie pas que la collection de composantes g_{mn} dépend de X, c'est-à-dire en se rappelant que le tenseur métrique est en réalité un champ tensoriel, on simplifie encore davantage l'expression de la métrique en l'écrivant

$$dS^2 = g_{mn}\ dX^m dX^n$$

Andy : *Il a fallu vraiment attendre Einstein pour inventer la convention de sommation ?*

Lenny : *Eh bien oui. Quand j'étais étudiant, j'ai lu le célèbre article d'Einstein de 1916 intitulé « Les fondations de la théorie de la relativité générale ». J'avais pour habitude quand j'apprenais de la physique de pousser moi-même les équations à mesure que je lisais, et de vérifier que j'arrivais au même résultat que l'auteur. Au début* *de son article, Einstein écrivait les équations comme tout le monde avec le signe* Σ. *Voici par exemple son équation (2) :*

$$dX_\nu = \sum_\sigma a_{\nu\sigma} dx_\sigma$$

Puis soudainement, après son équation (7), Einstein fait remarquer qu'il y a toujours une somme lorsque le même index apparaît deux fois[22]*. Donc à partir de maintenant, dit-il, nous allons le garder à l'esprit et dans ce cas arrêter d'écrire le signe somme. Il est clair qu'il en avait juste assez de les écrire chaque fois. J'étais aussi fatigué de les écrire. Quel soulagement ce fut !*

Fin de l'interlude sur la convention de sommation d'Einstein.

Revenons à la métrique et ses différentes formes dans des systèmes de coordonnées différents. Trouver un ensemble de coordonnées qui transforme l'équation (15) en l'équation (14) est une procédure plus compliquée que simplement diagonaliser la matrice g_{mn}. La raison en est qu'il n'y a pas une seule matrice. Comme nous l'avons rappelé à plusieurs reprises, les composantes g_{mn} dépendent de X.

C'est un champ de tenseurs avec une matrice différente en chaque point[23]. Vous ne pouvez en général pas toutes les diagonaliser en même temps. En *un point X donné*, on peut effectivement diagonaliser $g_{mn}(X)$ même si la surface n'est pas plane. Cela revient à travailler localement dans le plan tangent à la surface en X, et à y orthogonaliser les axes de coordonnées. Mais vous ne pouvez pas conclure qu'une surface est intrinsèquement plate parce qu'en chaque point, localement, elle ressemble à un plan euclidien. De même une courbe lisse a une tangente en chaque point, mais ça ne veut pas dire que la courbe est une droite !

22. Plus tard, Einstein a conçu les notations en exposant ou en indice pour les composantes des tenseurs. Sa règle s'est appliquée désormais uniquement à des paires d'index muets, avec la même lettre, en haut et en bas.

23. Pour un ensemble de coordonnées donné, nous avons une collection de matrices – une en chaque point. Pour un autre ensemble de coordonnées, nous aurons une autre collection de matrices. En chaque point, les *composantes* du tenseur dépendent des coordonnées, mais le tenseur lui-même est un objet abstrait qui ne dépend pas du système de coordonnées. Nous avons déjà rencontré la même distinction avec un vecteur 3D – même si on ne considère pas un champ de vecteurs, mais un seul vecteur attaché à un point.

Regardons de plus près l'équation (14) page 28. On peut l'écrire à l'aide d'une matrice particulière dont les composantes sont le *symbole de Kronecker* δ_{mn} défini de la manière suivante [24].

Tout d'abord, δ_{mn} vaut zéro quand $m \neq n$. Par exemple, en trois dimensions, δ_{12}, δ_{13}, et δ_{23} sont tous nuls, et les symétriques aussi, mais δ_{11}, δ_{22}, et δ_{33} sont différents de zéro. Autrement dit, en chaque point X la matrice construite avec le symbole Kronecker est diagonale.

Deuxièmement, les éléments diagonaux sont tous égaux à 1 :

$$\delta_{11} = \delta_{22} = \delta_{33} = 1$$

Avec le symbole de Kronecker et la convention de sommation d'Einstein, nous pouvons réécrire l'équation (14) sous la forme compacte

$$dS^2 = \delta_{mn}\, dX^m dX^n \tag{17}$$

Pour déterminer si un espace est plat, nous recherchons un changement de coordonnées, $X \to Y$, qui transforme g_{mn} en δ_{mn} *partout*. Rappelez-vous que X et Y représentent le *même point* P. Ce point P est simplement localisé avec deux référentiels différents, qui, comme nous l'avons souligné, ne sont rien de plus qu'une procédure d'étiquetage des points de l'espace géométrique.

Plus loin, les points P seront des événements dans l'espace-temps, et la matrice unitaire construite avec le symbole de Kronecker sera remplacée par une matrice diagonale légèrement plus compliquée dans la géométrie de Minkowski (également appelée géométrie minkowskienne ou einsteinienne), mais de nombreuses idées resteront les mêmes. Toutefois n'allons pas trop vite, et restons pour l'instant en géométrie riemannienne. La géométrie riemannienne est partout localement euclidienne. On dit parfois que c'est une « géométrie euclidienne sur un morceau de caoutchouc ».

Pour la plupart des métriques, il n'est pas possible de trouver un changement de coordonnées qui transforme partout en même temps g_{mn} en δ_{mn}. Ce n'est que lorsque l'espace est intrinsèquement plat que c'est possible.

24. Nommé en l'honneur du mathématicien allemand Leopold Kronecker (1823–1891).

En résumé, je vous donne le tenseur métrique de ma surface, c'est-à-dire les g_{mn} de l'équation (15) page 30. Nous écrivons la métrique avec la formule fondamentale

$$dS^2 = g_{mn}(X)\ dX^m dX^n$$

La question que je vous pose est : pouvez-vous, par un changement de coordonnées $X \to Y$, simplifier cette formule en l'équation (17) ? Autrement dit, dans le système Y, la formule de la métrique doit prendre la forme beaucoup plus simple

$$dS^2 = \delta_{mn}\, dY^m dY^n$$

Noter qu'il n'est pas nécessaire d'écrire $\delta_{mn}(Y)$, puisque le symbole de Kronecker a par définition une forme unique. Cependant, par souci de clarté, nous écrirons parfois encore $\delta_{mn}(Y)$ car cela nous rappellera quel système de coordonnées nous utilisons.

Si la réponse à ma question est oui, on dit que l'espace est *plat*. Sinon on dit qu'il est *courbe*. Bien sûr, l'espace peut avoir des parties plates. Il peut exister un système de coordonnées tel que dans toute une région le tenseur métrique soit le symbole de Kronecker. Mais l'espace n'est dit plat que s'il est plat partout.

Cela devient un problème purement mathématique : étant donné un champ tensoriel $g_{mn}(X)$ sur un espace multidimensionnel (que les mathématiciens appellent une *variété*), comment savoir s'il existe un changement de coordonnées qui transforme les composantes du champ tensoriel en le symbole de Kronecker partout ?

Pour répondre à cette question, nous devons mieux comprendre comment les choses se transforment lorsque nous effectuons des changements de coordonnées. C'est le sujet de l'*analyse tensorielle*. Nous commençons à présenter le sujet dans la suite de cette leçon, puis le traiterons de manière plus approfondie dans la leçon 2.

L'analogie entre les forces de marée et la courbure n'est en fait pas une analogie, c'est une équivalence précise. Dans la théorie de la relativité générale, la façon dont vous diagnostiquez les forces de marée (ou, plus exactement, leur généralisation) consiste à calculer le tenseur de courbure. Un espace plat est défini comme un espace où le tenseur de courbure est nul partout.

Dit simplement :

La gravité est la courbure.

Nous parviendrons à cette conclusion après avoir étudié l'analyse tensorielle. Évidemment, en essayant de déterminer si nous pouvons transformer $g_{mn}(X)$ en la forme triviale $\delta_{mn}(Y)$, la première question qui se pose est : comment $g_{mn}(X)$ se transforme quand on change de coordonnées ? Il faut introduire des notions d'analyse tensorielle qui sont mathématiquement fort simples.

Dans ce qui suit, nous exprimerons la première règle tensorielle, puis présenterons un interlude mathématique expliquant quelques faits généraux sur les vecteurs et les tenseurs. Ensuite nous présenterons la deuxième règle tensorielle.

Enfin nous terminerons cette copieuse première leçon par quelques considérations générales sur les composantes covariantes et contravariantes des vecteurs et des tenseurs.

Première règle : Composantes contravariantes des vecteurs

Il faut admettre que les notations tensorielles, avec leurs nuées d'indices, peuvent parfois paraître compliquées. La première fois qu'on les rencontre, elles peuvent effrayer. Mais bientôt nous découvrons que la manipulation des tenseurs obéit à des règles strictes et que celles-ci sont en définitive plutôt simples.

Nous allons commencer par quelque chose de plus simple que $g_{mn}(X)$. Supposons qu'on ait deux systèmes de coordonnées sur notre surface : un système X^m, et un deuxième système que nous pourrions dénoter avec le signe prime comme nous l'avons fait précédemment. Mais nous aurions alors des expressions épouvantables comme X'^m. Aussi nous allons plutôt dénoter le second système avec Y^m. Pour être très explicite, si on est dans un espace de dimension N, le même point P a les coordonnées

$$[\ X^1(P),\ X^2(P),\ \ldots\ , X^N(P)\ ]$$

et aussi les coordonnées

$$[\ Y^1(P),\ Y^2(P),\ \ldots\ , Y^N(P)\ ]$$

Les X et Y sont liés entre eux car si vous connaissez les coordonnées d'un point P dans un système de coordonnées, alors en principe vous savez où il se trouve. Par conséquent, vous connaissez aussi ses coordonnées dans l'autre système. Ainsi chaque coordonnée X^m est une fonction de toutes les coordonnées Y^n. Nous nous efforçons d'utiliser les index muets qui créent le moins de confusion possible. Nous écrirons simplement

$$X^m(Y)$$

De même, chaque Y^m est supposé être une fonction connue de tous les X^n :

$$Y^m(X)$$

En bref, nous avons deux systèmes de coordonnées, chacun fonction de l'autre. La correspondance est biunivoque puisqu'il s'agit de systèmes de coordonnées. Et nous supposons, selon la formule consacrée, que les fonctions « se comportent bien », c'est-à-dire sont continues et différentiables.

Maintenant, nous nous demandons : comment les éléments différentiels dX^m se transforment-ils ? La collection d'éléments différentiels dX^m est un vecteur, figure 16. Rappelez-vous qu'un vecteur est une paire de points dans l'espace (une origine et une extrémité). *Il est indépendant du système de coordonnées.* Mais pour travailler avec des vecteurs, il est commode de les exprimer en utilisant leurs composantes dans un système de coordonnées.

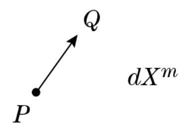


Figure 16 : Petit vecteur $\overrightarrow{PQ}$ exprimé dans le premier système de coordonnées.

Les quantités dX^m sont utilisées pour exprimer le petit vecteur

$$dX^m = [\ dX^1,\ dX^2,\ \ldots\ , dX^N\ ]$$

Dit autrement, lorsque nous modifions un peu chaque coordonnée X^m, le point P se déplace vers un point voisin Q, et le déplacement est dénoté avec la collection de composantes dX^m.

Regardons le même déplacement, exprimé maintenant dans le système de coordonnées Y. Nous voulons savoir comment les dY^m peuvent être exprimés en fonction des dX^p. C'est un résultat élémentaire de calcul différentiel que

$$dY^m = \sum_p \frac{\partial Y^m}{\partial X^p}\, dX^p$$

ou, en utilisant la convention de sommation,

$$dY^m = \frac{\partial Y^m}{\partial X^p}\, dX^p \tag{18}$$

Détaillons ce que veut dire l'équation (18) : la variation totale de la composante particulière Y^m (pensez à un index quelconque m fixe) est la somme du taux de variation de Y^m lorsque vous ne faites bouger que X^1, multiplié par le petit changement de X^1, à savoir dX^1, plus le taux de variation de Y^m lorsque vous ne faites bouger que X^2, multiplié par le petit changement de X^2, à savoir dX^2, et ainsi de suite jusqu'à X^N et dX^N car le membre de droite de l'équation (18) signifie une somme sur l'index muet p allant de 1 à N.

Quand on a bien compris l'équation (18), on a déjà compris une bonne partie de l'analyse tensorielle !

Passons maintenant à quelques considérations générales sur les vecteurs et les tenseurs.

Jusqu'à présent, nous avons utilisé à plusieurs reprise le terme *tenseur* (calcul tensoriel, tenseur métrique, tenseur de courbure, première règle sur les tenseurs, etc.), sans expliquer ce qu'est un tenseur ! Comme le lecteur et la lectrice l'ont compris, c'est un outil mathématique fondamental en relativité générale. Vous vous souvenez peut-être même que « cela étend le concept de vecteur ». Mais ce n'est certainement pas une explication suffisante pour saisir de quoi il s'agit :-)

Nous n'allons pas faire un exposé complet de l'algèbre linéaire et des tenseurs – que la lectrice et le lecteur peuvent trouver dans n'importe quel bon manuel sur le sujet. Cependant, comme je l'ai déjà fait plusieurs fois dans la série Le Minimum Théorique, par exemple, lorsque j'ai osé expliquer dans le volume 1 le calcul intégral ou la différentiation partielle dans de brefs interludes de quelques pages – car nous avions besoin de ces outils pour la mécanique classique –, il est temps d'introduire un troisième interlude mathématique présentant en détail les vecteurs et les tenseurs.

Interlude mathématique : Vecteurs et tenseurs

Commençons par le type de tenseur le plus simple, à savoir un scalaire. Un scalaire $S(X)$ est une fonction de la position, repérée ici par le vecteur X dans le premier système de coordonnées, avec la propriété qu'elle a la même valeur dans chaque système de coordonnées. Pour cette raison, nous pourrions aussi noter $S(P)$. Mais nous voulons insister sur le système de coordonnées que nous avons choisi d'utiliser, aussi écrivons-nous $S(X)$. Pour le même scalaire dans le système de coordonnées Y, nous utiliserons temporairement la notation $S'(Y)$. Plus tard, nous utiliserons même $S(Y)$ et $S(X)$ pour les deux, car ce sera plus clair lorsque nous parlerons de la règle de différentiation d'une fonction composée.

Ses propriétés de transformation sont triviales : il ne se transforme pas ! Un exemple tiré de la météorologie est la température en un point de l'espace. La formule de transformation d'un scalaire reflète cette règle particulièrement simple,

$$S'(Y) = S(X)$$

Dans le cas de la température, cela signifie que la température en un point est simplement un nombre [25]. Elle ne dépend pas du système de coordonnées en ce point. Notez également que les scalaires n'ont pas de composantes, ou peut-être plus précisément, ils n'ont qu'une seule composante : la valeur du scalaire elle-même.

Passons au type le plus simple suivant de tenseur, à savoir les vecteurs. Nous allons voir qu'il en existe deux sortes ou variantes.

25. *Nombre* et *scalaire* sont deux termes équivalents pour la même chose. Pourquoi parle-t-on de « scalaires » ? Les nombres sont souvent appelés scalaires car un nombre peut toujours être obtenu à partir d'un autre nombre avec un changement d'échelle : par exemple, vous pouvez obtenir 7 à partir de 2, simplement en multipliant 2 par 3,5. Vous ne pouvez pas faire cela avec n'importe quelle paire de vecteurs. Ce n'est possible qu'avec des vecteurs colinéaires. Strictement parlant, le terme *scalaire* est réservé aux nombres réels. Mais il arrive que nous appelions aussi les nombres complexes des scalaires.

Nous avons tous une idée intuitive de ce qu'est un vecteur en géométrie. C'est une petite flèche, généralement attachée à un point de l'espace. Elle pointe dans une direction et a une magnitude. En météorologie, l'exemple classique est la vitesse du vent.

En géométrie riemannienne, un vecteur *est une chose un peu plus abstraite qu'un vecteur ordinaire comme celui dont nous venons de parler.*

De même qu'un vecteur ordinaire, c'est une « chose » qui a des représentations avec des composantes. Mais, étant donné *un système de coordonnées et une métrique*, un vecteur, dans le nouveau sens que nous donnons à ce terme[26], peut être décrit par *deux jeux* de composantes : ses composantes contravariantes ou ses covariantes.

Précisons tout de suite que ce qu'on appelle les *composantes contravariantes* d'un vecteur abstrait sont les bonnes vieilles composantes avec lesquelles nous construisons un vecteur ordinaire à l'aide d'une combinaison linéaire des vecteurs de base.

Nous allons voir que nous pouvons aussi attacher à un vecteur un autre jeu de nombres, appelé ses *composantes covariantes.* Ce ne sont pas ses composantes contravariantes ordinaires, mais quelque chose d'autre, dont la signification géométrique sera expliquée dans la leçon 2. Les composantes contravariantes et les composantes covariantes d'un *même* vecteur abstrait seront liées entre elles très simplement à l'aide de la métrique.

Ces deux jeux de composantes, de même que les composantes de la métrique elle-même, changent quand on effectue un changement de coordonnées. Pour le moment, cependant, ne pensons pas à une métrique, mais seulement à un même vecteur abstrait dans un système de coordonnées X et dans un système Y. On se positionne en un point P de l'espace. Nous considérons un vecteur abstrait attaché au point P. Le vecteur a deux jeux de composantes qui dépendent chacun du système de coordonnées. Le vecteur a des composantes contravariantes et des composantes covariantes dans le système X. Et il a d'autres composantes contravariantes et d'autres composantes covariantes dans le système Y.

26. Pendant quelques paragraphes, nous allons l'appeler un *vecteur abstrait.* Puis nous abandonnerons le qualificatif abstrait. Un vecteur sera pour nous cette chose qui a deux jeux de composantes, chacun unidimensionnel : un jeu de composantes contravariantes, et un jeu de composantes covariantes.

Certains auteurs parlent de deux types de vecteurs : des vecteurs covariants et des vecteurs contravariants. Mais il nous semble plus clair – et correct – de parler d'un seul type de vecteur un peu plus abstrait qu'un vecteur ordinaire, et qui a deux jeux de composantes : ses composantes contravariantes et ses composantes covariantes [27].

Quand nous aurons expliqué leur lien par la métrique, il deviendra plus clair pourquoi un vecteur (la flèche intuitive) peut être exprimé de deux façons.

Qu'est-ce qui fait qu'une collection de nombres comme les déplacements infinitésimaux dX^m sont les composantes *contravariantes* d'un vecteur, plutôt qu'une simple collection de nombres ? Réponse : c'est la façon dont elles se transforment lors d'un changement de système de coordonnées. L'équation (18), page 45, est le prototype de la transformation des composantes contravariantes d'un vecteur.

Plus généralement, les composantes contravariantes d'un vecteur sont une collection à un seul index de nombres, V^m, qui se transforme de la manière suivante quand on passe du système X au système Y :

$$(V')^m = \frac{\partial Y^m}{\partial X^p} V^p \tag{19}$$

Dans cette équation, les variables V sont les composantes du vecteur dans le système de coordonnées X et les variables (V') sont les composantes du même vecteur dans le système Y. Si nous retournons à l'équation (18), nous voyons que les composantes dX^m du déplacement différentiel de P vers Q, figure 16, sont les composantes contravariantes d'un vecteur infinitésimal.

Il y a plusieurs points à noter. Tout d'abord, nous avons utilisé la convention de sommation pour signifier qu'on doit faire une somme sur l'index p. Nous avons mentionné dans l'interlude sur la convention de sommation d'Einstein qu'on doit faire une somme sur l'index lorsqu'il apparaît en haut et en bas. Dans l'équation (19), nous

27. Nous pourrons nous-mêmes nous laisser parfois aller à parler de vecteur contravariant. Il faudra alors comprendre que nous voulons parler des composantes contravariantes d'un vecteur.

sommes dans un cas un peu particulier : la première apparition de p doit être vue comme étant en bas. C'est un aspect de la convention de sommation que le lecteur et la lectrice doivent retenir : lorsqu'un index supérieur apparaît dans un terme au dénominateur d'une expression, il compte comme un index inférieur.

De manière générale, dans une expression où tout est au même niveau (nous voulons dire par là, sans dénominateur) ou bien si on regarde le numérateur d'une fraction, un index en haut est appelé un *index contravariant*. Et un index en bas est appelé un *index covariant*. Mais, comme on vient de le voir, toujours selon la convention de sommation, un index en haut sur un terme au dénominateur d'une fraction compte comme un index en bas, ou covariant.

Passons au deuxième type de collection de composantes – les composantes covariantes. Si le vecteur iconique dont les composantes ordinaires sont contravariantes est le déplacement dX^m, le vecteur iconique dont les composantes ordinaires sont covariantes est le gradient d'un scalaire $S(X)$.

Ses composantes sont données par les dérivées du scalaire dans des petits déplacements le long des axes de coordonnées :

$$\frac{\partial S(X)}{\partial X^p} \tag{20}$$

Il est clair que ces composantes dépendent du choix du système de coordonnées et se transforment lorsque l'on change de système. Par exemple, supposons que nous passions du système X au système Y. Pour calculer les composantes du gradient dans le système Y, nous utilisons une version de la règle de différentiation des fonctions composées (voir la leçon 2 du volume 1 de LMT où elle est expliquée) :

$$\frac{\partial S}{\partial Y^m} = \frac{\partial S}{\partial X^p}\ \frac{\partial X^p}{\partial Y^m} \tag{21}$$

Nous pouvons en déduire la règle générale pour la transformation des composantes covariantes d'un vecteur :

$$(W')_m = W_p\ \frac{\partial X^p}{\partial Y_m} \tag{22}$$

Prenons un peu de recul : avec l'équation (18), on a rencontré le *premier exemple de transformation d'un tenseur*, car un vecteur ordinaire, correspondant par exemple à la position d'un point, ou un déplacement (autrement dit, une translation), ou une vitesse, ou une accélération, ou une force, etc., est un vecteur contravariant, qui est un type simple de tenseur.

En effet, nous avons maintenant les expressions dans deux repères différents du petit déplacement d'un point dans un espace, figure 16 et équation (18). Ce sont dX^m et dY^m. Répétons que les dX^m et dY^m sont deux jeux de composantes *contravariantes* pour le *même* déplacement. Et on sait passer d'un jeu à l'autre.

La figure 17, qui complète la figure 16, montre le petit déplacement, et montre aussi localement les deux systèmes de coordonnés.

La lectrice et le lecteur ont maintenant compris que l'équation (18) est simplement la formule de transformation des composantes d'un vecteur ordinaire – qui est lui-même un objet indépendant du système de coordonnées [28] – lorsqu'il est exprimé dans le système X et dans le système Y.

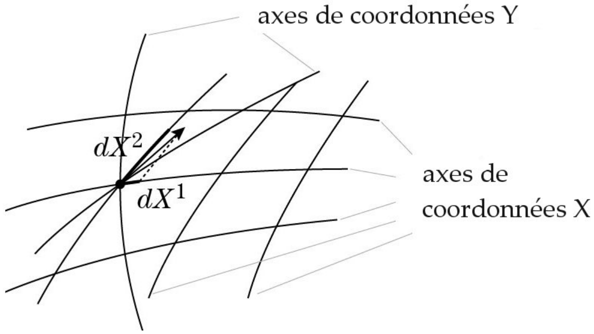


Figure 17 : Petit déplacement vu dans deux systèmes de coordonnées. Le petit vecteur a des composantes $(dX^1,\ dX^2)$ montrées, mais aussi $(dY^1,\ dY^2)$ pas montrées.

28. Remarquez, cependant, qu'il est difficile de parler de concepts géométriques sans au moins une sorte de système de repérage. Les deux célèbres géomètres américains Oswald Veblen (1880–1960) et John Henry Whitehead (1904 –1960), conscients de la difficulté de définir ce qu'est la géométrie, ont écrit dans leur livre *The Foundations of Differential Geometry* que la géométrie est ce que les experts appellent la géométrie :-)

Les mathématiciens russes Andreï Kolmogorov (1903–1987), Alexandre Alexandrov (1912–1999) et Mikhaïl Lavrentiev (1900–1980) ont exprimé dans leur célèbre ouvrage en trois volumes, *Mathématiques*, Les Éditions du Bec de l'Aigle, 2020-2021, leur indignation devant cette affirmation.

Remarque sur la terminologie : nous nous sommes brièvement autorisés page précédente à parler de « vecteur contravariant » pour les vecteurs ordinaires, car c'est un usage répandu. Cependant comme nous traiterons dans cet ouvrage de vecteurs abstraits pouvant avoir des expressions contravariantes mais aussi des expressions covariantes, nous préférerons généralement parler des composantes contravariantes ou des composantes covariantes d'un même vecteur abstrait.

En bref, le terme *contravariant* vient du fait que si nous changeons les vecteurs unitaires dans le système de coordonnées, par exemple si nous *divisons* simplement la longueur de chacun d'eux par dix, les composantes d'un vecteur ordinaire représentant une translation seront *multipliées* par dix.

Quant au terme *covariant*, il vient du fait que, dans le même genre de changement de coordonnées, les composantes d'un gradient seront en revanche aussi *divisées* par dix.

L'interlude a présenté les types de tenseurs les plus simples : les tenseurs de rang 0, qui sont simplement des scalaires ; et les tenseurs de rang 1, qui sont des vecteurs abstraits avec des composantes contravariantes et aussi des composantes covariantes. Les types suivants de tenseurs, de rang 2 ou plus, seront présentés dans la dernière section de la leçon.

Deuxième règle :
Composantes covariantes des vecteurs

Bien que nous l'ayons déjà mentionnée brièvement dans l'interlude mathématique précédent, par souci de symétrie, précisons la deuxième règle tensorielle concernant les composantes covariantes des vecteurs. Ces vecteurs sont utilisés pour représenter autre chose qu'une position, ou une translation, ou une vitesse, ou une accélération, etc. Pour soutenir l'intuition, le lecteur et la lectrice peuvent penser au gradient d'un champ scalaire.

Des exemples de champs scalaires sont la température, la pression atmosphérique, le champ de Higgs, tout ce qui a, en tout point de l'espace, une valeur qui n'est pas multidimensionnelle mais simplement un nombre, et qui ne change pas si nous changeons de coordonnées.

La vitesse du vent n'est pas un champ scalaire car en tout point c'est une donnée vectorielle. C'est donc un *champ vectoriel.* Il est important de noter le point suivant, qui devrait clarifier les choses :

Si on ne considérait que la première composante du vecteur représentant le vent, on n'obtiendrait pas un champ scalaire, car il ne serait pas invariant par changement de coordonnées.

Ainsi le gradient d'une fonction scalaire est un vecteur (au sens d'une collection unidimensionnelle de composantes). Mais ce n'est pas un vecteur ordinaire. En effet, ses composantes ne se transforment pas de la même manière que les composantes contravariantes des vecteurs ordinaires.

Nous avons vu précédemment qu'une application de la règle de différentiation des fonctions composées donnait l'équation (21), reproduite ci-dessous :

$$\frac{\partial S}{\partial Y^m} = \frac{\partial S}{\partial X^p} \frac{\partial X^p}{\partial Y^m}$$

En dénotant par (W') le gradient de S par rapport aux Y, et par W son gradient par rapport aux X, il peut être réécrit comme l'équation (22) , que nous reproduisons également ci-après en lui attribuant un nouveau numéro :

$$(W')_m = \frac{\partial X^p}{\partial Y^m} W_p \tag{23}$$

L'équation (23) ne s'applique pas seulement aux gradients ; c'est l'équation fondamentale reliant les versions primes et non primes des composantes covariantes d'un vecteur, c'est-à-dire ses composantes covariantes dans le système Y et dans le système X.

Observez que les index m de W' et p de W sont en bas. L'index p est un index muet indiquant qu'il faut faire une somme dessus, car il apparaît également en haut dans ∂X^p. C'est un bel exemple du fonctionnement harmonieux de la convention de sommation d'Einstein et de sa grande utilité.

Réécrivons les équations (19) et (23) l'une à côté de l'autre avec de nouveaux numéros :

Composantes contravariantes

$$(V')^m = \frac{\partial Y^m}{\partial X^p}\; V^p \tag{24a}$$

Composantes covariantes

$$(W')_m = \frac{\partial X^p}{\partial Y^m}\; W_p \tag{24b}$$

Elles se ressemblent beaucoup, sauf que $\partial Y^m / \partial X^p$ apparaît dans l'une, et son inverse, $\partial X^p / \partial Y^m$, dans l'autre.

Rappelons une dernière fois que des déplacements, ou des positions, ou des vitesses, etc., sont décrits ordinairement par des vecteurs ayant des composantes contravariantes. Nous avons vu que celles-ci changent *contrairement* au changement de base.

Les gradients, par contre, sont décrits ordinairement par des vecteurs dont les composantes changent *comme* la base change. C'est pourquoi leurs composantes sont dites covariantes.

En mathématiques, les vecteurs à composantes covariantes sont parfois vus comme des vecteurs dans l'espace dual de l'espace vectoriel initial considéré. Ce sont alors des *vecteurs duaux* comme le sont les formes linéaires. Mais nous n'adopterons pas cette approche. Pour nous, les vecteurs seront des objets qui ont une collection à un index de composantes contravariantes et aussi une collection à un index de composantes covariantes.

Les équations (24a) et (24b) sont des équations fondamentales de ce cours. Le lecteur et la lectrice doivent bien les comprendre et être à l'aise avec elles, car elles sont absolument centrales dans toute la relativité générale. Vous devez savoir où vont les indices pour différents types d'objets et comment ces objets se transforment. Dans une vision superficielle des choses, on peut dire que la théorie de la relativité générale se ramène à une théorie des propriétés de transformation de divers types d'objets.

Composantes covariantes et contravariantes des vecteurs et des tenseurs

Nous avons vu deux façons de penser à un vecteur *ordinaire*. Tout d'abord, on peut y penser comme on l'a appris au lycée : c'est un déplacement avec une magnitude et une direction, c'est-à-dire *une flèche* dans un espace. C'est un objet géométrique bien défini avant même que nous considérions une base quelconque.

Nous pouvons aussi, de manière plus abstraite, le considérer comme un objet abstrait qui a des composantes. Ces composantes dépendent de la base. Si les composantes se transforment d'une certaine manière quand on change de base, à savoir selon l'équation (24a), alors l'objet se comporte exactement comme nos bons vieux vecteurs. Par conséquent, nous pouvons également assimiler l'objet à un vecteur ordinaire. En analyse tensorielle, nous les appelons des *vecteurs dont les composantes sont contravariantes.*

De même, certains autres objets ont des composantes qui se transforment selon l'équation (24b). Ils ne peuvent pas être assimilés à nos bons vieux vecteurs ordinaires. Mais on peut encore leur donner un sens géométrique. Nous avons mentionné que les mathématiciens les considèrent comme des vecteurs duaux. En analyse tensorielle, nous appelons ce second type d'objet des *vecteurs dont les composantes sont covariantes.*

En fait, nous avons dit plus simplement que ce sont les *mêmes* objets abstraits – que nous avons appelés des vecteurs abstraits – qui ont une version contravariante et une version covariante. Nous approfondirons ce point de vue dans la leçon suivante.

En calcul tensoriel, dont la relativité générale fait grand usage[29],

---

29. Lors de ses travaux en relativité restreinte, Einstein a développé ses idées sans utiliser le calcul tensoriel ni même la géométrie de Minkowski, que Minkowski, qui avait été le professeur d'Einstein à Zurich, n'a du reste introduite qu'en 1908. Poincaré a également fait quelques travaux préliminaires du même genre que la géométrie de Minkowski.

Quand il prit connaissance des travaux de son ancien professeur, Einstein jugea d'abord que cette refonte mathématique de la théorie de la relativité était inutilement compliquée sans apporter grand-chose. Mais il a vite changé d'avis. Lorsqu'il a achevé la relativité générale, en 1915, Einstein a déclaré que cela n'aurait pas été possible sans la géométrie abstraite non euclidienne et le calcul tensoriel. Hermann Minkowski (1864–1909) n'a pas participé au développement de la relativité générale car il est mort en 1909. Son ami David Hilbert (1862–1943) a cependant joué un rôle en 1915 ; voir leçon 9.

paradoxalement pour les personnes qui ont un esprit ou une intuition géométrique, il est souvent utile, du moins au début, d'oublier l'interprétation géométrique des objets que nous manipulons, et de se concentrer sur la façon dont les collections de nombres attachées aux points de notre espace se comportent lorsque nous changeons de système de coordonnées. De même qu'on n'apprend pas à faire du vélo en lisant une description des propriétés d'un vélo, mais en montant dessus, d'abord avec une draisienne ou des petites roues, nous allons apprendre les tenseurs *en nous en servant*.

Un vecteur – que ses composantes soient contravariantes ou covariantes – est un cas particulier de tenseur. Nous n'allons pas donner une description géométrique, indépendante de la base, des tenseurs. Pour nous, au début, un tenseur sera une collection multidimensionnelle de composantes définie par la façon dont ces composantes se transforment lorsqu'on passe d'un système de coordonnées à un autre. Plus tard nous donnerons une interprétation géométrique de certains tenseurs. Nous approfondirons également les composantes contravariantes et covariantes. Nous verrons à nouveau comment une collection à un index de composantes peut avoir une version contravariante et une version covariante. Tout cela sera développé dans la leçon suivante. Pour l'instant, continuons à avancer pas à pas dans notre construction des outils mathématiques nécessaires à la relativité générale.

La prochaine étape, pour nous maintenant, est de parler des tenseurs avec plus qu'un seul indice.

La meilleure façon d'aborder les tenseurs à plusieurs indices est de commencer par considérer un cas particulier très simple. Imaginons le « produit » de deux vecteurs ayant chacun des composantes contravariantes[30]. On considère les deux vecteurs à composantes contravariantes, V et U, et on considère le produit suivant :

$$V^m U^n$$

Sans plus nous étendre chaque fois en commentaires, nous utiliserons désormais systématiquement la convention que les indices de composantes contravariantes sont notés dans la position d'exposant, c'est-à-dire en haut.

30. Ce n'est pas le produit scalaire ni le produit vectoriel. Il s'appelle le *produit extérieur* ou *produit tensoriel*. Comme dans les autres cas, c'est une *opération* qui à deux choses en associe une troisième.

Il n'est pas nécessaire que les vecteurs V et U proviennent du même espace vectoriel. Si la dimension de l'espace de V est M, et celle de l'espace de U est N, il y a $M \times N$ produits $V^m U^n$.

Comme d'habitude, nous utilisons la même notation $V^m U^n$ pour désigner à la fois un produit particulier, mais aussi la collection de tous les produits. Le contexte permet de savoir de quoi on parle. On a rencontré la même notation habituelle avec V^m qui désigne aussi bien une composante particulière du vecteur V que le vecteur V dans son ensemble.

Définissons T^{mn} par

$$T^{mn} = V^m U^n \tag{25}$$

Prêtez attention au fait que la position et l'ordre dans lesquels nous écrivons les indices de T^{mn} sont importants. Par exemple, T^{mn} n'est pas la même chose que T^{nm}. Il est laissé en exercice d'expliquer pourquoi. Bientôt, nous verrons également des combinaisons d'indices positionnés en haut et en bas.

Le produit T^{mn} est un cas particulier de tenseur de rang 2. Le terme « rang 2 » signifie que c'est une collection de produits de composantes, avec en tout deux indices. Les indices parcourent chacun une plage d'indice : m va de 1 à M et n va de 1 à N. Par exemple, si V et U proviennent d'un espace à quatre dimensions, il y aura 16 composantes $V^m U^n$. Insistons sur le fait que le terme T^{mn} représente à la fois, comme on vient de le dire, une composante particulière du produit tensoriel mais aussi l'ensemble des 16 composantes.

Comment T^{mn} se transforme-t-il ?

Par exemple V^m et U^n pourraient être les composantes des vecteurs V et U dans le référentiel non prime, c'est-à-dire le référentiel utilisant le système de coordonnées X. Puisque nous savons comment les composantes individuelles se transforment, lorsque nous passons au système de coordonnées Y, nous pouvons nous figurer immédiatement comment T se transforme. Appelons $(T')^{mn}$ la composante mn-ième du tenseur dans le référentiel prime :

$$(T')^{mn} = (V')^m (U')^n$$

En utilisant l'équation (24a) deux fois, on peut réécrire cette expression comme ceci

$$(T')^{mn} = \frac{\partial Y^m}{\partial X^p}\, V^p\, \frac{\partial Y^n}{\partial X^q}\, U^q$$

Les quatre termes du côté droit sont juste quatre nombres dont on prend le produit, nous pouvons donc les réécrire dans un autre ordre :

$$(T')^{mn} = \frac{\partial Y^m}{\partial X^p}\, \frac{\partial Y^n}{\partial X^q}\, V^p\, U^q$$

Enfin, $V^p\, U^q$ est simplement T^{pq}. Par conséquent la façon dont T se transforme est

$$(T')^{mn} = \frac{\partial Y^m}{\partial X^p}\, \frac{\partial Y^n}{\partial X^q}\, T^{pq} \tag{26}$$

Nous avons établit dans le cas particulier du produit tensoriel de vecteurs ordinaires comment le tenseur produit T se transforme. Cela nous conduit à la définition suivante :

N'importe quelle collection de composantes, indexée avec deux indices, qui se transforme selon la formule (26), est appelée un tenseur de rang 2 à deux indices contravariants.

Quand il y a plus d'indices en haut, la règle est adaptée de la manière évidente. Un tenseur de rang 3, dont tous les indices sont contravariants, se transforme comme ceci :

$$(T')^{lmn} = \frac{\partial Y^l}{\partial X^p}\, \frac{\partial Y^m}{\partial X^q}\, \frac{\partial Y^n}{\partial X^r}\, T^{pqr}$$

Quelles choses sont des tenseurs de ce genre ? Réponse : beaucoup de choses. Les produits tensoriels de vecteurs sont des exemples particuliers, mais il y a d'autres objets qui ne sont pas des produits tensoriels et qui sont cependant des tenseurs selon notre définition.

Pour commencer, nous allons voir que la collection des g_{mn} qui sert à définir la métrique de l'espace est un tenseur de rang 2. Mais c'est un tenseur avec des indices covariants. Donc, pour terminer cette leçon, voyons comment les tenseurs avec des indices covariants se transforment. L'équation (24b) montre comment un vecteur abstrait avec un seul indice covariant se transforme. C'est un tenseur de rang 1 de type covariant.

Reprenons le cas particulier du produit tensoriel de deux vecteurs, cette fois-ci covariants, W et Z. Pour parler de manière moins cavalière, on va regarder le produit tensoriel de deux vecteurs abstraits lorsque l'on travaille avec leurs composantes covariantes. Leur produit se transforme comme ceci :

$$(W')_m (Z')_n = \frac{\partial X^p}{\partial Y^m} \frac{\partial X^q}{\partial Y^n} W_p Z_q$$

Ici, nous avons découvert la formule de transformation d'une chose avec deux indices covariants, c'est-à-dire deux indices en bas.

Plus généralement considérons un objet que nous noterons T_{mn}. Ce n'est plus simplement un produit de vecteurs. Ça peut être quelque chose d'autre. Cependant, la lettre T signale que cette chose est un tenseur. Quand il a deux indices en bas, il se transforme par la formule suivante :

$$T'_{mn} = \frac{\partial X^p}{\partial Y^m} \frac{\partial X^q}{\partial Y^n} T_{pq} \tag{27}$$

Encore une fois, tout ce qui se transforme selon cette équation (27) est appelé un tenseur de rang 2 avec deux indices covariants.

Nous laissons au lecteur et à la lectrice le soin de déterminer comment un tenseur avec un indice supérieur et un indice inférieur doit se transformer.

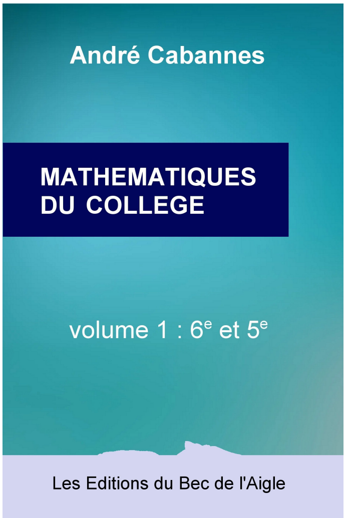
Dans la prochaine leçon, nous verrons aussi comment la collection des composantes g_{mn} servant à définir la métrique à l'aide de l'équation (15), page 30, se transforme. Comme nous l'avons annoncé, nous verrons qu'il s'agit d'un tenseur à deux indices covariants.

Alors la question que nous nous poserons est la suivante : étant donné que l'équation (27) est la formule de transformation du tenseur métrique g_{mn} quand on passe d'un référentiel X à un référentiel Y, pouvons-nous trouver un changement de référentiel qui transforme partout (c'est-à-dire en tous les points de la variété) le tenseur métrique g_{mn} en δ_{mn} ?

C'est une question mathématique. Elle est difficile en général. Mais nous trouverons la condition.

Catalogue des

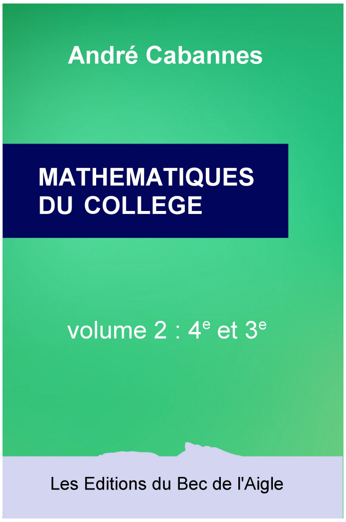
ÉDITIONS DU BEC DE L'AIGLE



www.amazon.fr/dp/2957239159
Cours de mathématiques du collège.

Volume 1 : 6e et 5e.

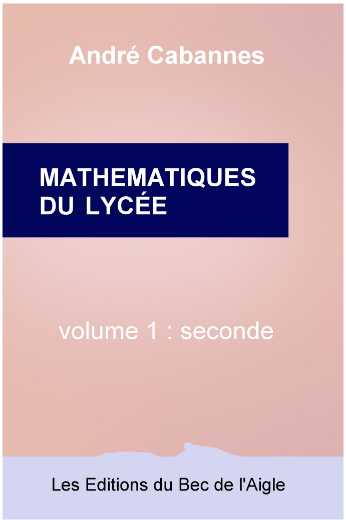
à l'intention des collégiens et de leurs parents



www.amazon.fr/dp/2957239167
Cours de mathématiques du collège.

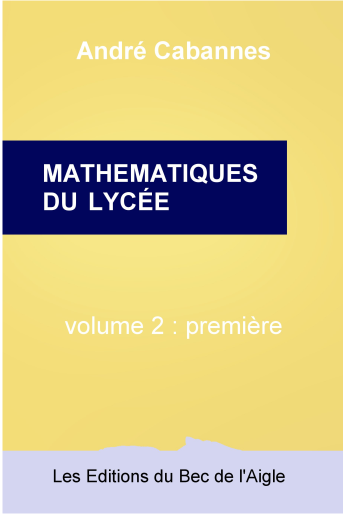
Volume 2 : 4e et 3e.

à l'intention des collégiens et de leurs parents



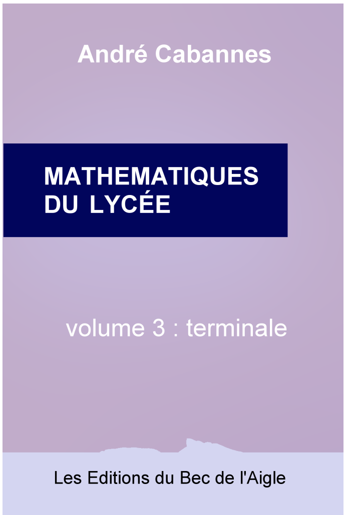
www.amazon.fr/dp/2957239183
Cours de mathématiques de seconde

à l'intention des lycéens et de leurs parents



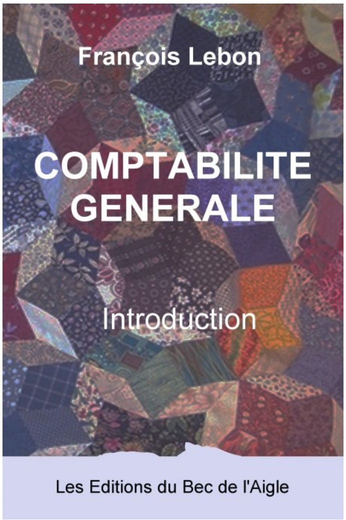
www.amazon.fr/dp/2957239191
Cours de mathématiques de première

à l'intention des lycéens et de leurs parents

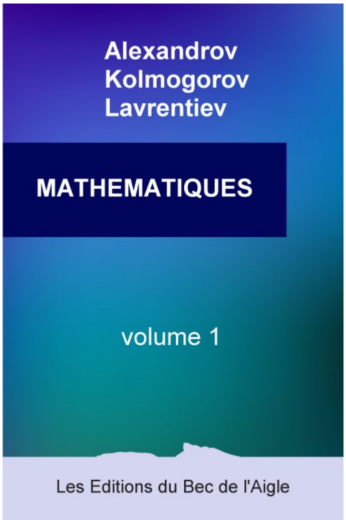


www.amazon.fr/dp/2958738507
Cours de mathématiques de terminale

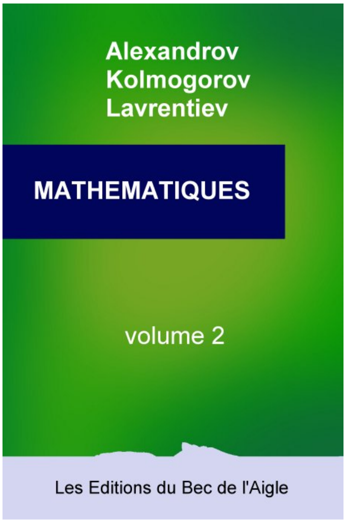
à l'intention des lycéens et de leurs parents



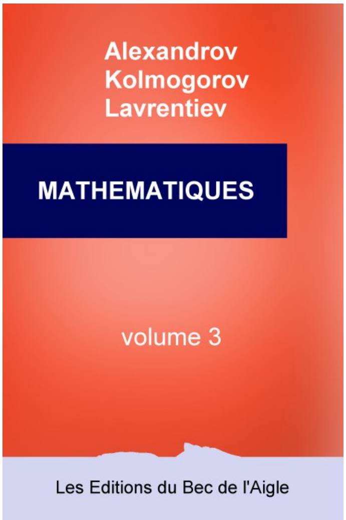
www.amazon.fr/dp/2957239140
Cours de comptabilité (niveau baccalauréat)



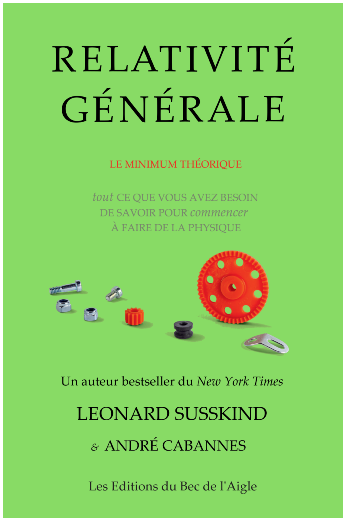
www.amazon.fr/dp/2957239124
Introduction aux mathématiques (niveau baccalauréat)



www.amazon.fr/dp/2957239116
Les mathématiques pour l'utilisateur (niveau première année d'université)

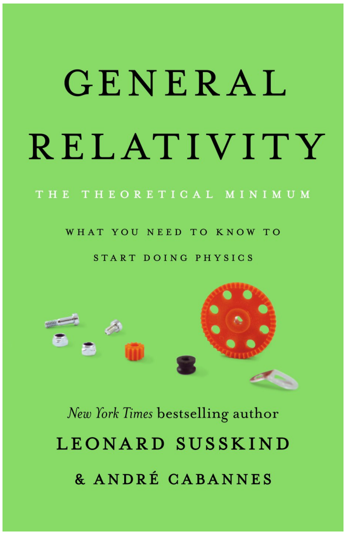


www.amazon.fr/dp/2957239132
Les mathématiques pour l'étudiant spécialisé et le chercheur (niveau licence)



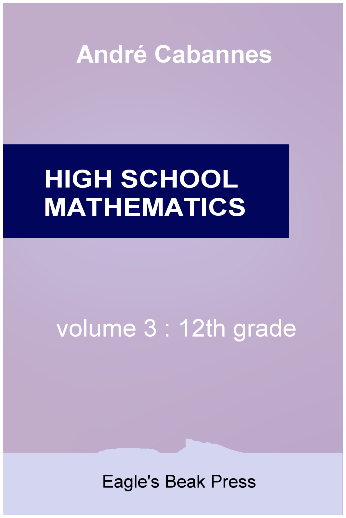
www.amazon.fr/dp/2957239175
Cours de physique
(niveau maîtrise)

English titles by André Cabannes



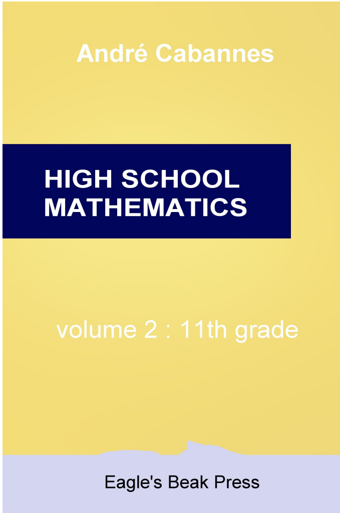
www.amazon.com/dp/B09ZB613QY
General Relativity

Graduate studies.



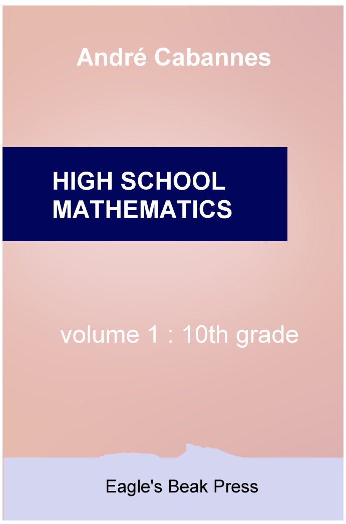
www.amazon.com/dp/2958738515
High school mathematics

Volume 3 : 12th grade



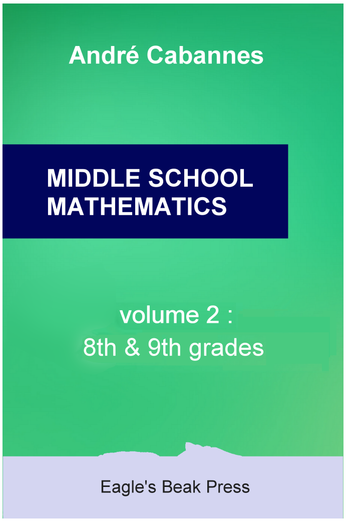
www.amazon.com/dp/2958738523
High school mathematics

Volume 2 : 11th grade



www.amazon.com/dp/2958738531
High school mathematics

Volume 1 : 10th grade

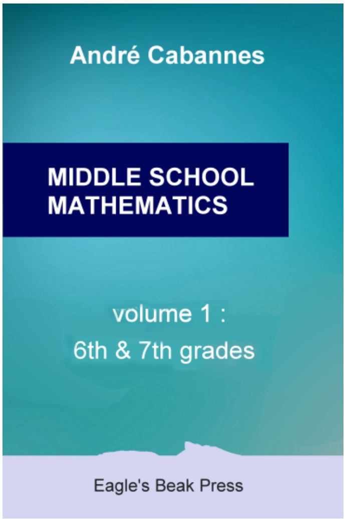


www.amazon.com/dp/295873854X

Middle school mathematics

Volume 2 : 8th & 9th grades

for middle school students and their parents

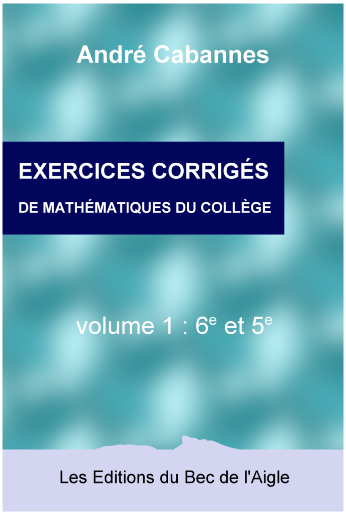


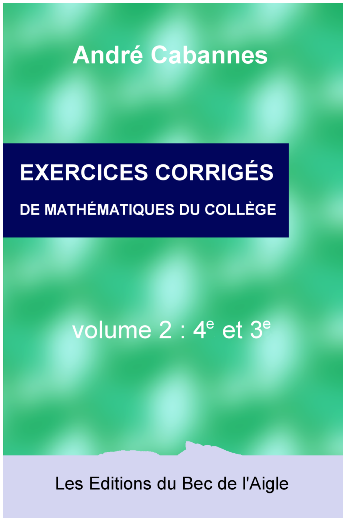
www.amazon.com/dp/2958738558

Middle school mathematics

Volume 1 : 6th & 7th grades

for middle school students and their parents


André Cabannes
EXERCICES CORRIGÉS
DE MATHÉMATIQUES DU COLLÈGE
volume 1 : 6e et 5e
Les Editions du Bec de l'Aigle


André Cabannes
EXERCICES CORRIGÉS
DE MATHÉMATIQUES DU COLLÈGE
volume 2 : 4e et 3e
Les Editions du Bec de l'Aigle