

# Principe d'équivalence et outils

## PLAN

- Principe d'équivalence
- Tenseurs contravariants et covariants
- Métrique
- Espace-temps
- Dérivée covariante



La relativité générale lie l'énergie (et donc la masse) à la géométrie de l'espace-temps.  
Tout part d'une simple hypothèse : l'équivalence entre accélération et gravitation...

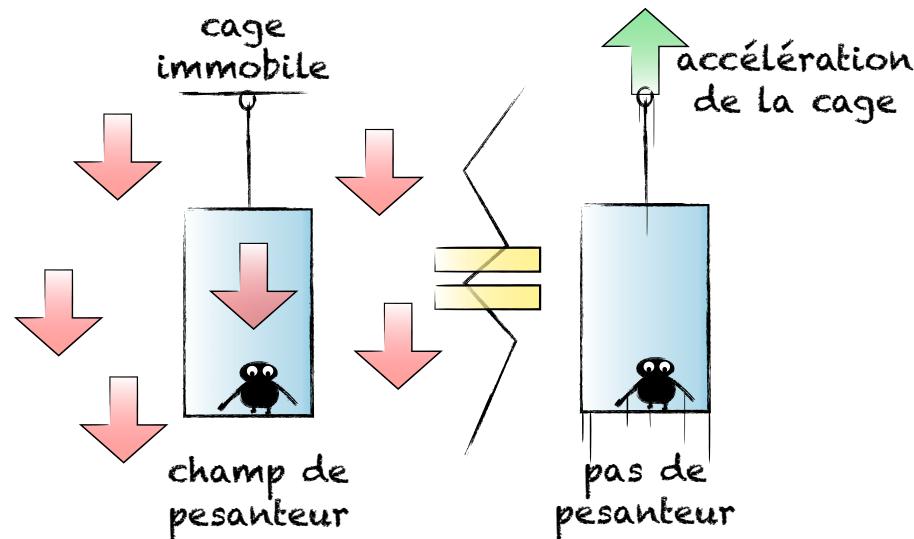
## SECTION 1

# Principe d'équivalence

Pour aboutir à la théorie de la relativité générale, Einstein est parti de constats étonnamment simples.

Le principal est sans doute **le principe d'équivalence** (dit faible pour ce qui nous intéresse ici) :

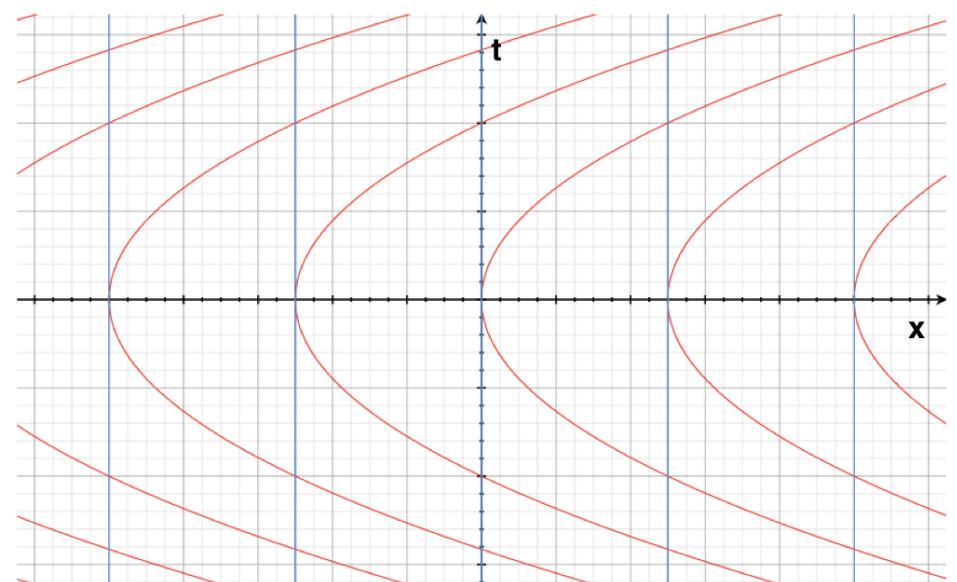
L'effet d'un champ de gravitation *uniforme* sur un observateur se confond pour lui avec l'effet d'un référentiel *uniformément accéléré*.



Supposons, dans un premier temps, qu'on nous mette dans une boîte, maintenue immobile, qui va tenir lieu de référentiel, et qu'on soumette la boîte (et son pauvre occupant) à un champ de gravitation uniforme. Bon ok, on se sent poussé sur la paroi vers laquelle le champ pointe.

Supposons maintenant qu'on enlève le champ et qu'à la place, on accélère uniformément la boîte, dans le sens contraire par rapport au champ précédent.

Pour une valeur bien choisie de l'accélération, nous serions dans l'incapacité complète de faire la différence entre les deux situations. C'est ça le principe d'équivalence.



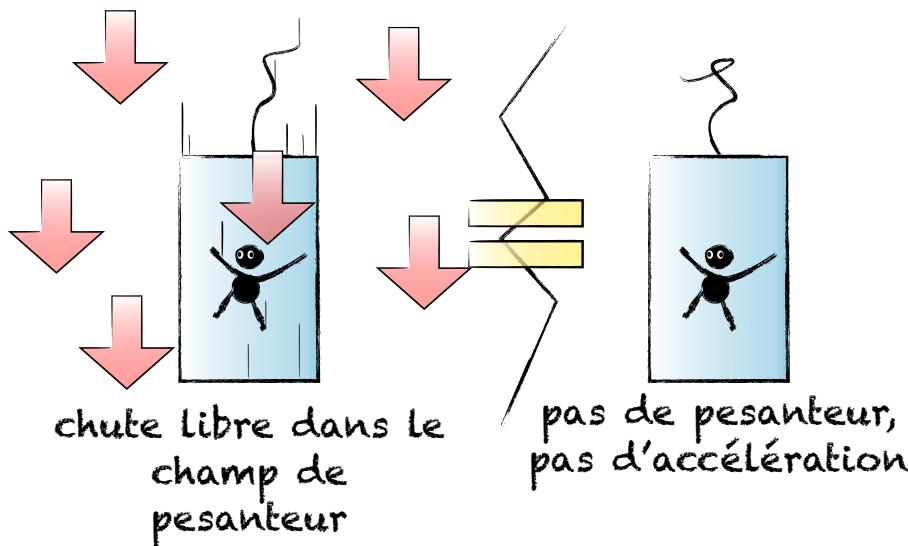
Guidé par cette équivalence, il devient tentant de décrire la gravitation **géométriquement** par un espace-temps uniformément accéléré. Or cet espace-temps fait de coordonnées courbes (des paraboles si on oublie un temps la relativité restreinte) peut se ramener à un espace cartésien

(droit) par simple changement de coordonnées global ( $x = \frac{1}{2}t^2$ ). Les deux systèmes de coordonnées décrivent donc le même espace, et cet espace est plat !

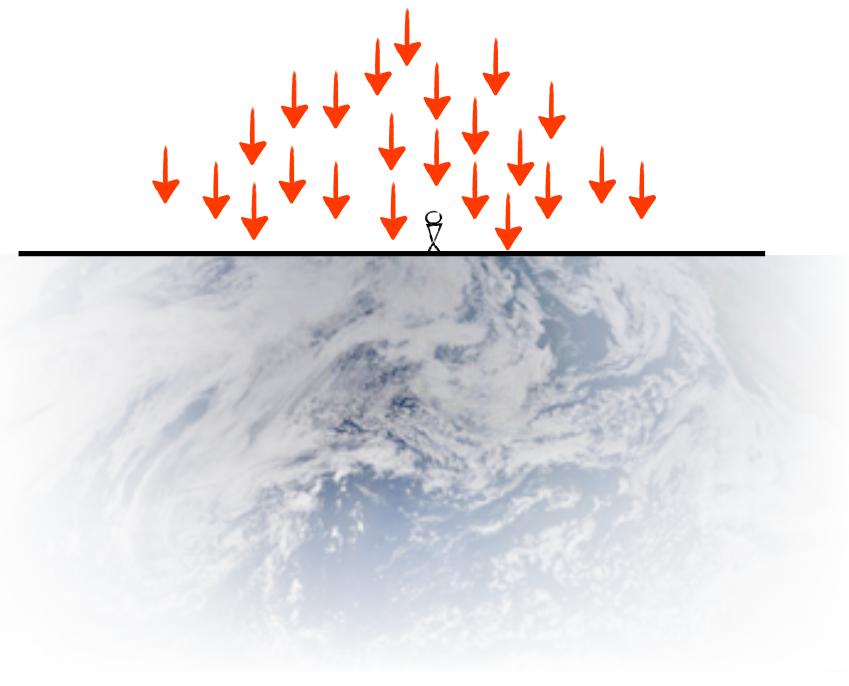
Retournons dans notre boîte immobile et soumise à un champ de gravitation uniforme pour s'en convaincre. Si on relâche la contrainte qui gardait la boîte immobile pour la laisser chuter librement dans le champ de gravitation, que ressentons-nous ?

On se retrouve en apesanteur ! En effet, rien ne distingue cette chute libre de l'absence totale de gravitation. Mes deux mains, mes pieds, ma tête, tout est accéléré pareil, tout comme la boîte. Et si tout bouge ensemble, comment se rendre compte que quelque chose bouge ?

C'est ça un espace plat, un espace identique pour mes deux mains, mes pieds et ma tête. Plat = partout identique.

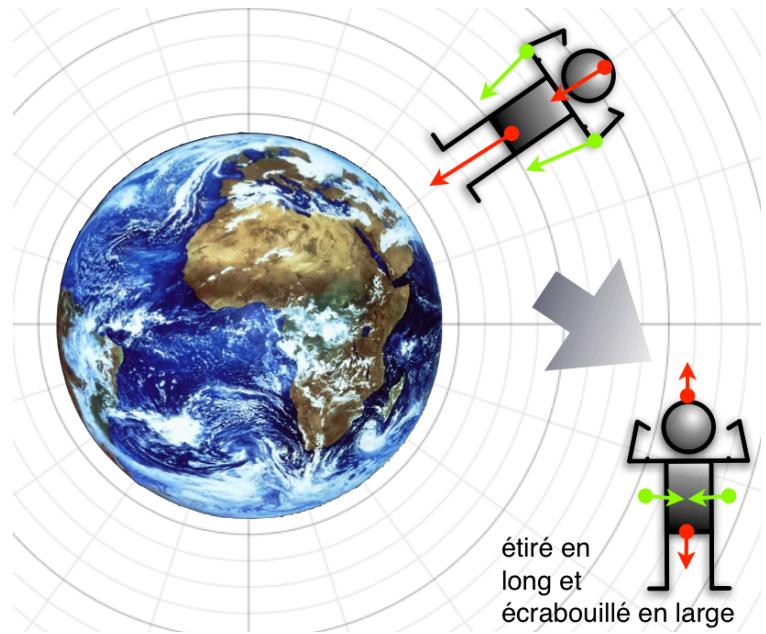


Jusque là, tout semble simple mais l'obstination à vouloir un champ de gravitation ou d'accélération *uniforme* laisse présager des complications. Qui a déjà vu un champ uniforme ? Ce n'est pas de notre monde. Un champ de gravitation uniforme devrait être créé par une terre plate et infinie, vision plutôt désuète...



Remarquez, à notre pauvre échelle humaine, ça marche relativement bien. Mais que se passerait-il si on était plus gros ? Aïe, les forces de marée ! Dès qu'un objet est plongé dans un champ gravitationnel divergent (ils le sont tous puisque les masses qui créent le champ sont nécessairement localisées), et si l'objet n'est pas ponctuel, il va subir des forces de marée... Ma main gauche n'est plus accélérée comme ma main droite,

mes pieds et ma tête veulent se détacher ! J'en déduis que le monde n'est pas partout le même ; il n'est plus plat, il est courbe.



Les forces de marée empêchent de transformer la gravitation en une simple accélération uniforme. Et de fait, la chute libre n'est plus équivalente à l'absence de gravitation. Sauf bien sûr si on est assez petit et sur des temps très courts, car à cette échelle, les différences deviendront suffisamment faibles pour être négligées.

Cela revient à dire que **le principe d'équivalence n'est plus valable que localement**.

Cette validité locale permet de continuer à associer une accélération au champ gravitationnel, mais l'espace-temps que ces coordonnées accélérées vont décrire ne sera plus plat !

Aucun changement de coordonnées global ne pourra plus ramener en tout point l'espace-temps à un espace plat et reposant (pour nos cerveaux s'entend). La courbure d'un espace est une caractéristique fondamentale qu'il est impossible de gommer (on ne peut pas, par exemple, projeter entièrement la surface terrestre sur un planisphère, il manque au moins un point).

Les forces de marées empêchent de se débarrasser d'un champ gravitationnel en se plaçant dans un référentiel uniformément accéléré de la même manière que la courbure empêche de retourner à un espace plat par changement de coordonnées. Ça ne devient plus possible que localement. Il y a une obstruction de même nature dans les deux cas ; courbure de l'espace-temps et champ de gravitation réel (= dû à une masse localisée) semblent donc intimement liés. La masse courbe l'espace-temps !

Au moment de se pencher sur ces histoires de courbure, Einstein avait déjà fait savoir au monde que masse et énergie étaient équivalentes. Plutôt que la masse, c'est donc plus généralement **l'énergie**, quelle que soit sa forme, qui **courbe l'espace-temps**.

D'autre part, l'existence de référentiels spéciaux, dits galiléens, chiffonnait grandement le moustachu épris d'universel. Pourquoi devrait-on changer les lois de la physique suivant que le référentiel soit galiléen ou non ? Ça manque d'élégance...

Concomitamment à l'achèvement d'une théorie «géométrique» de la gravitation, Einstein a su libérer les équations physiques du carcan référentiel et les rendre ainsi «relativité générale - compatible» (synonyme d'élégance). On a appelé ça **principe de covariance générale**. Et cela passe par l'usage intensif du raffiné, quoiqu'un peu effrayant, calcul tensoriel, un tenseur ayant le bon goût d'être indépendant du système de coordonnées.

Notre route est tracée : réussir à relier proprement la courbure de l'espace à l'énergie, tout en restant indépendant du référentiel.

Mais cela ne va pas sans quelques complications mathématiques... Et dans un premier temps, il va falloir apprivoiser certains outils comme les tenseurs, la métrique, les dérivées covariantes, etc.

## SECTION 2

# Tenseurs contravariants et covariants

Un tenseur est l'objet physique rêvé d'Einstein. Il reste invariant lors des changements de système de coordonnées. Par exemple, si un tenseur est nul dans un système de coordonnées, alors il est nul dans tous. Et de fait, **deux tenseurs égaux dans un système restent toujours égaux** (puisque leur différence doit rester nulle). Quelles que soient les coordonnées dans lesquelles il est décrit, le tenseur traduit toujours la même réalité physique.

Ça ne veut pas dire que le tenseur n'est pas transformé, au contraire. Comment faire pour ne pas être modifié par un changement de coordonnées ? Il faut changer avec lui, lui emboiter le pas comme son ombre !

Prenons un stylo (exemple de tenseur contravariant comme on le verra plus tard). Si on regarde son image dans un miroir déformant, on s'attend évidemment à ce que le stylo soit déformé (le contraire serait bizarre). Le miroir modifie le système de coordonnées, le stylo est lui aussi modifié.

Pourtant, d'un côté ou de l'autre du miroir, la réalité physique du stylo n'a pas changé. Seule sa description est transformée. Et pas n'importe comment : la prise en compte du détail de la

transformation doit nous permettre, dans l'image déformée, de retrouver les dimensions réelles du stylo.

Les tenseurs vont donc être définis par la façon dont ils se transforment lors d'un changement de système de coordonnées, cette transformation devant les laisser invariants.

## Le déplacement

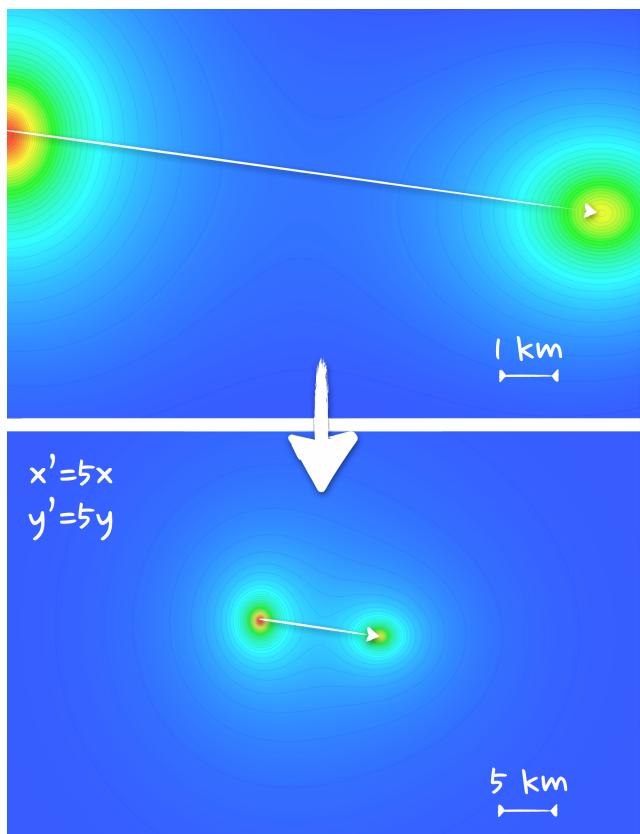
On va partir d'un exemple assez familier : le changement d'échelle sur une carte.

Prenons une carte plane représentant deux sommets montagneux. Le premier type de grandeur qu'on s'imagine être invariant lors d'un changement de coordonnées est le **déplacement «réel» entre deux points** de la carte. S'il y a 10 km entre le plus haut sommet et le plus bas, quel que soit le changement de coordonnées, on s'attend à ce qu'il y ait toujours 10 km entre les deux sommets.

On représente ce déplacement par un vecteur allant d'un point à un autre. Pourquoi un vecteur et pas juste un segment ? Si on inverse l'est et l'ouest sur la carte, le déplacement doit lui aussi être inversé. Ce n'est pas la même chose d'aller du plus haut au plus bas que du plus bas au plus haut. Et nous, on veut la même chose.

Place à la transformation ! On modifie les coordonnées sur la carte comme suit :  $x' = 5x$  et  $y' = 5y$ .

Cela revient à s'éloigner ou encore à diviser l'unité de longueur (par 5). Résultat : la distance correspondant à 1 km sur l'ancienne carte en vaut maintenant 5 et notre vecteur déplacement se voit du coup tout ratatiné puisqu'il doit continuer à représenter ses 10 km. Si on le mesure sans prendre garde à la transformation (au changement d'échelle), on a un vecteur 5 fois trop petit.



Pour qu'il retrouve de sa superbe, pas d'autres choix que de s'occuper de l'échelle et donc de la transformation !

Mathématiquement, en oubliant un temps les  $y$  pour simplifier, le nouveau vecteur  $\vec{V}'(x')$  est égal au vecteur mesuré sur la nouvelle carte avec les anciennes coordonnées  $\vec{V}(x)$  (pas de prime quand on utilise la vieille échelle, *ie* la même règle qu'au départ) corrigé de la variation des nouvelles coordonnées par rapport aux anciennes  $\partial x'/\partial x$  (c'est la prise en compte du changement d'échelle).

On ne fait que retrouver la formule liant les petits déplacements :  $dx' = \partial x'/\partial x \cdot dx$ .

En rétablissant les deux dimensions  $x$  et  $y$ , on calcule séparément chacune des composantes de  $\vec{V}'$  ( $V'_x$  et  $V'_y$ ) comme ci-dessus excepté que pour chacune, il faut également prendre en compte la possibilité que les  $x'$  (respectivement les  $y'$ ) varie dans la direction  $y$  (respectivement  $x$ ), même si ce n'est pas le cas dans l'exemple (cela correspondrait à un changement de coordonnées «tordu»). La formule générale serait alors :

$$V'_x(x', y') = \frac{\partial x'}{\partial x} V_x(x, y) + \frac{\partial x'}{\partial y} V_y(x, y)$$

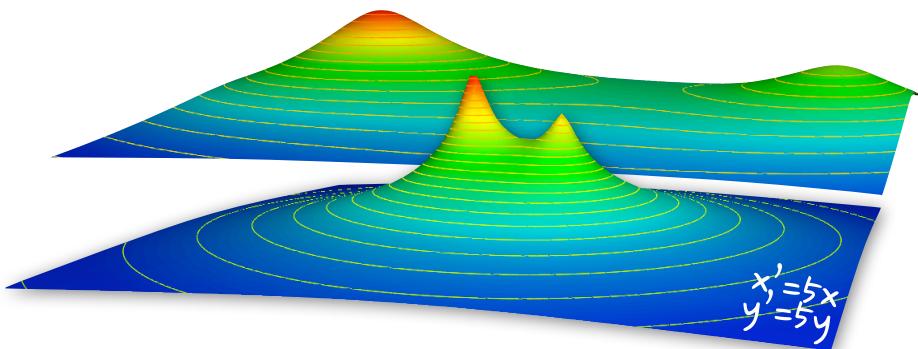
$$V'_y(x', y') = \frac{\partial y'}{\partial x} V_x(x, y) + \frac{\partial y'}{\partial y} V_y(x, y)$$

## Le gradient

Les déplacements ne sont pas les seules grandeurs intéressantes. **Le gradient** d'altitude entre les deux

sommets, mesure de la densité de lignes de niveau traversées entre les pics, a lui aussi droit à son invariance.

À l'inverse du déplacement, le gradient est visiblement augmenté sur la deuxième carte quand on ne prend pas l'échelle en compte. Les sommets sont plus pentus et on s'attend donc à une correction allant cette fois-ci dans l'autre sens : il faut prévenir le gradient que les vieux  $x$  qu'on continue à utiliser pour le calculer sont moins espacés qu'ils le devraient.



La règle de dérivation des fonctions composées nous souffle encore la réponse :

en appelant  $\Phi$  la hauteur, le nouveau gradient vaut  $\partial\Phi'(x')/\partial x' = \partial\Phi'(x')/\partial x \cdot \partial x/\partial x' = \partial\Phi(x)/\partial x \cdot \partial x/\partial x'$  (où on utilise  $\Phi'(x')=\Phi(x)$  puisque la hauteur est invariante lors de ce changement de coordonnées, la hauteur étant ici un tenseur de rang 0 évoqué plus loin).

Et à deux dimensions, en appelant à nouveau  $V$  le gradient (il s'agit comme précédemment d'un vecteur bien qu'il se transforme autrement et n'ayons crainte des confusions, on va

apprendre à différencier ces deux sortes de vecteurs bien plus fondamentalement que par une lettre), on a :

$$V_{x'}(x',y') = \partial x/\partial x' \cdot V_x(x,y) + \partial y/\partial x' \cdot V_y(x,y)$$

$$V_{y'}(x',y') = \partial x/\partial y' \cdot V_x(x,y) + \partial y/\partial y' \cdot V_y(x,y)$$

Finalement, on a dégagé **deux types de grandeurs physiques se transformant à l'inverse l'une de l'autre** pour rester invariante lors d'un changement de coordonnées. Il y a donc deux types de tenseurs : les **contravariants** (qui varient comme le **vecteur déplacement**) et les **covariants** (qui varient comme le **gradient**).

## La position des indices

Afin de différencier efficacement chaque type de tenseur, on va jouer sur la position des indices (désignant la direction ou plutôt la dimension) : indices haut pour les contravariants et bas pour les covariants.

$(V^x, V^y, V^z)$  est ainsi un vecteur contravariant, alors que  $(V_x, V_y, V_z)$  est covariant.

La **convention de sommation implicite d'Einstein** va permettre d'alléger grandement les notations tout en respectant automatiquement la nature des tenseurs sans avoir à y réfléchir :

$$a_i b^i = \sum_{i=1}^3 a_i b^i = a_1 b^1 + a_2 b^2 + a_3 b^3$$

On somme sur chaque dimension quand des indices haut et bas sont répétés. L'indice est alors dit **muet**. On a pris ici un espace à 3 dimensions qui pourraient être  $x, y$  et  $z$  en cartésien.

### Lois de transformation des tenseurs de rang 1

Appliquons ces nouvelles règles sur les relations trouvées plus haut en généralisant à un nombre  $d$  quelconque de dimensions.

En appelant maintenant  $y$  les nouvelles coordonnées (à la place des primes) et  $x$  les anciennes (ça clarifie les formules, même si, dans le même temps, ça brouille le lien avec l'exemple), les lois de transformation par changement de coordonnées des **tenseurs de rang 1** (avec un seul indice) contra et covariants sont alors :

$$\begin{cases} V'^n(y) = \frac{\partial y^n}{\partial x^m} V^m(x) & \text{tenseur contravariant type petit déplacement} \\ V'_n(y) = \frac{\partial x^m}{\partial y^n} V_m(x) & \text{tenseur covariant type gradient} \end{cases}$$

$n$  (et  $m$ ) allant de 0 à  $d-1$ , la loi pour chaque type de tenseur est en réalité constituée de  $d$  lignes, une par dimension.

Et fort heureusement, avec  $d=2$ , on retrouve les 2 paires de transformations de l'exemple.

Résumons ce que l'on sait sur les deux familles de tenseurs de rang 1 :

- **les tenseurs contravariants sont de type longueur (ou longueur  $\times$  quelque chose) et varient comme les coordonnées. Il représente un objet physique «ordinaire» comme un stylo. L'archétype de tout vecteur est le petit déplacement  $dx^m$ .**
- **les tenseurs covariant sont de type densité et varient à l'inverse des coordonnées (il varie en fait comme les vecteurs de base de l'espace, comme l'échelle dans notre exemple).**

*Remarque importante :*

un tenseur covariant et un tenseur contravariant ne peuvent être égaux d'un point de vue physique. Sinon cette égalité elle-même ne serait pas tensorielle. En effet, lors d'un changement de coordonnées, la différence de transformation entre les deux grandeurs rendrait impossible le maintien de l'égalité ! Par conséquent, une égalité entre deux grandeurs physiques se doit d'être «**tensoriellement homogène**». Et comme on le verra plus en détail par la suite, la convention sur la position des indices et sur la sommation rend très simple et puissante cette vérification d'une importance capitale.

## Aparté mathématique expliquant les noms «co» et «contra» :

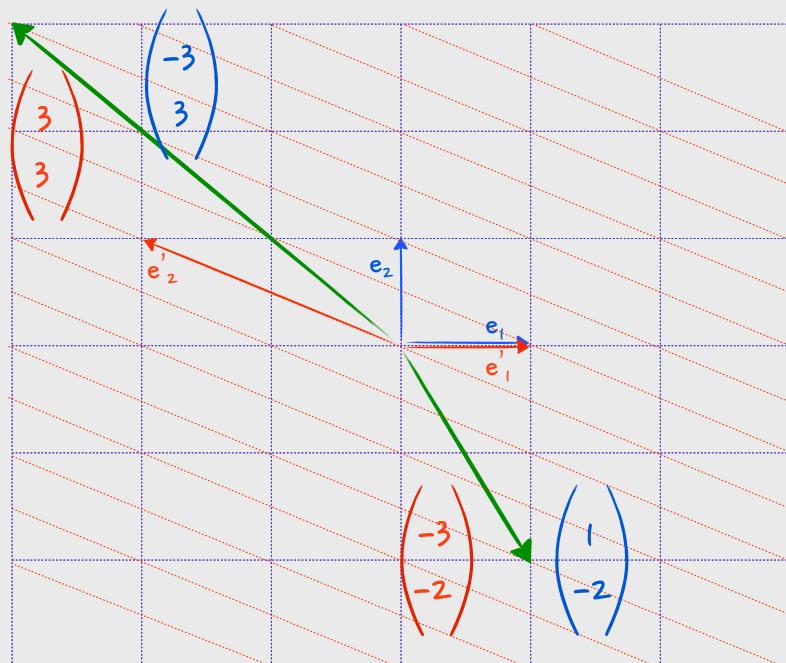
les vecteurs de base agissent comme l'échelle d'une carte, ils nous disent qu'une unité dans cette base fera leur longueur. En ça, ils ont la dimension d'une densité : ils valent **1 par longueur** du vecteur de base. Ils sont covariants !

La matrice de passage d'une base à une autre porte sur chaque colonne l'expression du nouveau vecteur de base en fonction des anciens.

Prenons l'exemple du changement de base suivant :

$$e'_1 = e_1$$

$$e'_2 = -2e_1 + e_2$$



On a représenté deux vecteurs (verts) et leurs coordonnées dans l'ancienne base (bleue) et dans la nouvelle (rouge).

La matrice de passage s'écrit :

$$P = \begin{pmatrix} e'_1 & e'_2 \\ 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

Décomposons  $y^i$ , vecteur de la base  $e'_i$ , sur la base  $e_i$  :  
 $y^1 e'_1 + y^2 e'_2 = y^1 e_1 + y^2 (-2e_1 + e_2) = x^1 e_1 + x^2 e_2$

D'où

$$\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y^1 \\ y^2 \end{pmatrix} = P \begin{pmatrix} y^1 \\ y^2 \end{pmatrix}$$

et donc

$$\begin{pmatrix} y^1 \\ y^2 \end{pmatrix} = P^{-1} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}$$

Conclusion : les vecteurs de base sont transformés suivant  $P$  et les coordonnées suivant  $P^{-1}$ . Par suite, tout tenseur se transformant selon  $P$  est dit covariant car il varie *avec* la base (comme elle). À l'inverse, les vecteurs *contravariants* varient *contre* la base (selon la matrice inverse).

De même, dans l'exemple introductif, les vecteurs de base de la carte initiale (ou, ce qui revient au même, l'échelle) se trouvent réduits 5 fois dans la transformation vers la nouvelle carte tout comme doit l'être le gradient pour rester invariant, alors qu'on doit multiplier par 5 les coordonnées (les distances).

Un **tenseur de rang 0** est beaucoup plus simple, c'est un **scalaire** ; pas de transformation quand on change de coordonnées :  $\Phi'(\mathbf{x}') = \Phi(\mathbf{x})$ .

Le produit d'un vecteur covariant  $A$  avec un vecteur contravariant  $B$  donne un tenseur de rang 0. Étant donné que  $A$  et  $B$  varient à l'inverse l'un de l'autre lors d'un changement de coordonnées, leur produit, lui, ne varie pas.

On le vérifie bien dans notre exemple : la distance (contravariante) multipliée par le gradient de hauteur (covariante) donne la variation de hauteur, complètement indépendante des changements de coordonnées dans le plan  $(x, y)$  considéré.

$A_i B^i \equiv \sum_i A_i B_i$  n'est rien d'autre que **le produit scalaire** entre  $A$  et  $B$  et la convention de sommation d'Einstein oblige ce produit scalaire à impliquer deux vecteurs de nature opposée, un covariant et un contravariant, pour qu'il puisse être considéré d'un point de vue tensoriel comme un «vrai» produit scalaire.

Un produit scalaire entre deux vecteurs de même nature tensorielle,  $A^i B^i$  par exemple, n'aurait pas de sens puisqu'il donnerait une grandeur sensée varier lors d'un changement de coordonnées ( $A^i$  et  $B^i$  variant dans le même sens), donc pas un tenseur de rang 0, ou scalaire. Or par définition, un produit scalaire doit donner... un scalaire.

Et c'est là le génie de la convention d'Einstein, elle rend ces raffinements implicites : une grandeur sans indice (avec une seule composante), tenseur de rang 0, ne peut être obtenue

qu'avec une association co et contra, haut et bas. Notons bien en effet que  $A^i$  est une grandeur à  $d$  composantes (collection ordonnée de  $d$  scalaires) alors que  $A^i A_i$  n'en a plus qu'une seule.

Par conséquent, la convention sur la position des indices alliée à celle sur la sommation d'Einstein permet de tester très facilement l'homogénéité d'une équation du point de vue tensoriel.

Exemples : grâce à la convention d'Einstein, on sait automatiquement que l'opération  $A_m B^m C^n$  doit accoucher d'un tenseur de rang 1 contravariant  $D^n$ .

Et on pourra fanfaronner devant l'ineptie de l'équation  $\partial_\nu \delta_\mu^\sigma \Delta_\sigma^{\mu\nu} + \nabla_\tau \chi^\lambda \chi^\tau = 2$  alors qu'on n'aurait rien trouvé de choquant à  $\partial_\nu \delta_\mu^\sigma \Delta_\sigma^{\mu\nu} + \nabla_\tau \chi^\lambda \partial_\lambda \chi^\tau = 2$  (l'ordre n'importe pas ; si un indice haut et un indice bas sont répétés n'importe où dans un produit, alors ils sont contractés en une somme).  
Bref, cet enfantin jeu d'indices peut se révéler très fertile.

*Remarque :*

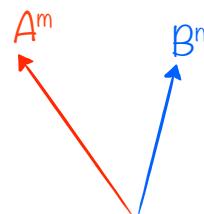
une collection de plusieurs scalaires ne fait pas automatiquement un tenseur de rang 1 (ou supérieur) ! Ce sont les opérations de symétrie de l'espace qui le définissent (elles disent comment il doit tourner, être translaté, etc.). Les composantes du tenseur doivent varier de façon coordonnée et ce sont les lois de transformation d'un tenseur lors d'un

changement de coordonnées qui se chargent de nous assurer de la nature de ce qui est transformé : si notre collection de scalaires se transforme comme la loi le dicte, alors c'est bien un tenseur ! Il fera du même coup tout ce qu'on attend d'un tenseur (ou d'un vecteur courant pour un tenseur de rang 1) comme rester invariant lors d'une translation dans un espace plat par exemple.

Prenons un morceau d'atmosphère considéré comme un espace euclidien à 3 dimensions, et formons en tout point  $x$  un objet 3D avec la température en ce point  $T(x)$ , la pression  $P(x)$ , et la masse volumique  $\rho(x)$ . Il y a peu de chance que cet objet se transforme comme il faut lors d'un changement de coordonnées ou qu'il reste identique lors d'une translation. Ce n'est pas un tenseur de rang 1...

Et les **tenseurs de rang 2** ?

Ce sont des objets ayant 2 directions différentes. La manière la plus simple d'en fabriquer un est de souder ensemble 2 vecteurs  $A^m$  et  $B^n$  (ce qu'on note  $A^m \otimes B^n$ , où  $\otimes$  désigne un produit tensoriel).



Pour fixer les choses, les  $A^m$  sont les coordonnées d'une particule A et les  $B^n$  celles d'une particule B. Le système des deux particules est donc repéré par l'ensemble des  $A^m B^n$  que l'on peut ranger dans un seul objet : le tenseur de rang 2. Il est alors facile de voir comment un tel tenseur se transforme :

$$\begin{aligned} T'^{mn}(y) &= A'^m(y)B'^n(y) \\ &= \frac{\partial y^m}{\partial x^r} A^r(x) \frac{\partial y^n}{\partial x^s} B^s(x) \\ &= \frac{\partial y^m}{\partial x^r} \frac{\partial y^n}{\partial x^s} T^{rs}(x) \end{aligned}$$

On fait de même avec un tenseur doublement covariant. Et on obtient les règles de transformation pour les tenseurs de rang 2 :

$$\begin{cases} T'^{mn}(y) = \frac{\partial y^m}{\partial x^r} \frac{\partial y^n}{\partial x^s} T^{rs}(x) \\ T'_{mn}(y) = \frac{\partial x^r}{\partial y^m} \frac{\partial x^s}{\partial y^n} T_{rs}(x) \end{cases}$$

**Tout ce qui se transforme de la sorte est un tenseur puisque ces transformations sont construites pour laisser l'objet invariant.**

*Remarque :*

le produit tensoriel additionne les dimensions quand le produit scalaire les écrabouille et cela se lit directement sur les indices.

## SECTION 3

# Métrique

En permettant d'exprimer les distances dans un espace courbe, la métrique constitue une brique fondamentale de la relativité générale.

Que vaut la distance infinitésimale  $ds$  dans un espace plat euclidien à trois dimensions (notre bon vieil espace classique auquel on est si habitué) muni de coordonnées cartésiennes ? Pythagore a la réponse :

$$\begin{aligned} ds^2 &= (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \\ &= \sum_m dx^m dx^m \\ &= \delta_{mn} dx^m dx^n \end{aligned}$$

Le symbole de kronecker  $\delta$  permet de préserver la règle de sommation implicite ( $\delta_{ij} = 1$  si  $i=j$  et 0 sinon)

On ne fait là que redéfinir l'espace euclidien, espace de dimension finie muni d'un produit scalaire (forme bilinéaire définie positive).

Exprimons maintenant  $ds^2$  dans un nouveau système de coordonnées  $y$  :

$$\begin{aligned} ds^2 &= \delta_{mn} \frac{\partial x^m}{\partial y^r} \frac{\partial x^n}{\partial y^s} dy^r dy^s \\ &= g'_{rs}(y) dy^r dy^s \end{aligned}$$

$g$  est appelé le **tenseur métrique** ou plus succinctement, **la métrique**. C'est lui qui s'occupe d'obtenir une distance à partir des petits déplacements.

Preuve qu'il s'agit bien d'un tenseur ? Il se transforme comme un tenseur (un covariant) !

En effet,  $\delta_{mn}$  est alors le tenseur métrique d'un espace plat en coordonnées cartésiennes qu'on peut généraliser au tenseur métrique d'un espace quelconque  $\delta_{mn} \rightarrow g_{mn}(x)$ . Le passage des coordonnées  $x$  aux coordonnées  $y$  s'écrivant donc :

$$g'_{rs}(y) = g_{mn}(x) \frac{\partial x^m}{\partial y^r} \frac{\partial x^n}{\partial y^s} \text{ CQFD}$$

Le tenseur métrique regarde comment les directions d'un système de coordonnées sont liées les unes aux autres. Il permet d'opérer le produit scalaire puisque, par définition, il s'occupe de projeter orthogonalement les vecteurs les uns sur les autres.

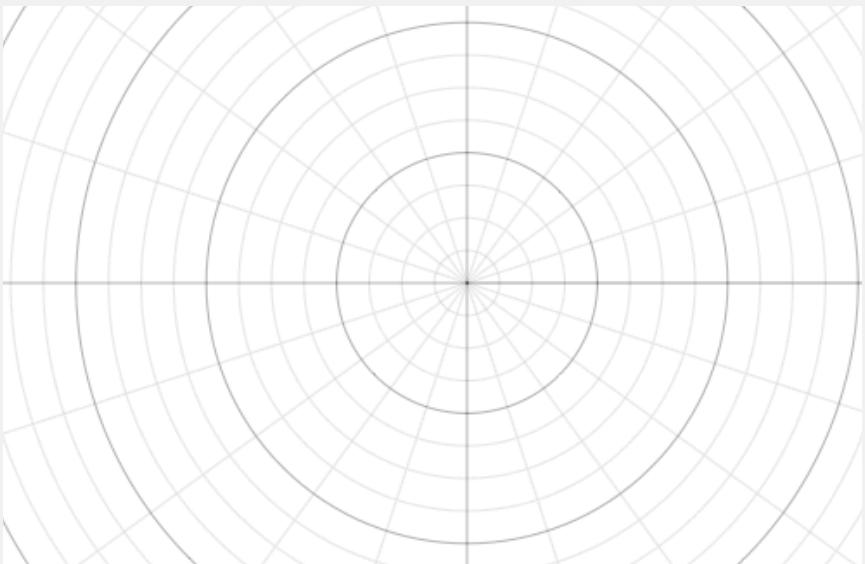
Le tenseur métrique nous offre aussi un «critère de platitude» : un **espace** est dit **plat** quand la métrique peut se ramener à un symbole de kronecker (= une matrice unité). Les vecteurs de bases sont alors tous unitaires et perpendiculaires les uns aux autres.

Petit tour chez les polaires :

En coordonnées polaires, la distance infinitésimale s'exprime :

$$ds^2 = dr^2 + r^2 d\theta^2$$

donc  $g_{11} = 1$ ,  $g_{12} = g_{21} = 0$  et  $g_{22} = r^2$  (avec l'indice 1 pour  $r$  et 2 pour  $\theta$ )



Remarque : les valeurs nulles des termes non diagonaux traduisent l'orthogonalité des directions  $r$  et  $\theta$ .

On peut se ramener à un kronecker avec la transformation inverse de celle des cartésiens en polaires :

$$dx = \cos(\theta) dr - r \sin(\theta) d\theta$$

$$dy = \sin(\theta) dr + r \cos(\theta) d\theta$$

Donc

$$dr = \cos(\theta) dx + \sin(\theta) dy$$

$$d\theta = -\frac{\sin(\theta)}{r} dx + \frac{\cos(\theta)}{r} dy$$

d'où

$$dr^2 + r^2 d\theta^2 = dx^2 + dy^2$$

Tout va bien... Et heureusement puisque les coordonnées polaires décrivent aussi le plan (donc un espace plat à 2D) ! On peut réécrire tout ça avec des matrices (en appelant maintenant  $x$  les coordonnées cartésiennes et  $y$  les coordonnées polaires) :

$$\frac{\partial x^m}{\partial y^n} = A_n^m = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$$

donc

$$\frac{\partial y^m}{\partial x^n} = B_n^m = (A_n^m)^{-1} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta/r & \cos \theta/r \end{pmatrix}$$

et

$$\begin{aligned} ds^2 &= g'_{mn}(y) dy^m dy^n \\ &= g'_{mn}(y) \frac{\partial y^m}{\partial x^r} \frac{\partial y^n}{\partial x^s} dx^r dx^s \\ &= g'_{mn}(y) B_r^m B_s^n dx^r dx^s \\ &= dx^r B_r^m g'_{mn}(y) B_s^n dx^s \end{aligned}$$

Pour faire apparaître des produits matriciels (entre lignes et colonnes), il faut transposer  $B_r^m$  et  $dx^r$  (mais ça nous fait perdre la règle sur les indices haut et bas) :

$$ds^2 = dx_r {}^t B_m^r g'_{mn}(y) B_s^n dx^s$$

On retrouve sans surprise la formule de changement de base (pour des matrices orthogonales)  ${}^t\mathbf{B} \times \mathbf{g}' \times \mathbf{B}$  au milieu. Et on obtient :

$$\begin{aligned} ds^2 &= g'_{mn}(y) dy^m dy^n \\ &= \delta_{rs}(x) dx^r dx^s \end{aligned}$$

Explicitons les matrices pour s'en convaincre :

$$\begin{aligned} (x_1 &\quad x_2) \begin{pmatrix} \cos \theta & -\sin \theta/r \\ \sin \theta & \cos \theta/r \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta/r & \cos \theta/r \end{pmatrix} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} \\ &= (x_1 \quad x_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} \end{aligned}$$

Par contre, aucune transformation linéaire, aucun changement de base ne permet de passer des coordonnées cartésiennes aux coordonnées d'une sphère de rayon unité. Prenons par exemple comme système de coordonnées la longitude  $\phi$  et la colatitude  $\theta$  sur la sphère. On a alors :

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

$d\phi$  n'a pas de variation linéaire en fonction de  $\theta$  contrairement à  $d\theta$  en fonction de  $r$  dans le cas des polaires. Mais si on se contente de petits déplacements, le  $\sin(\theta)$  varie presque comme  $\theta$  et on retrouve la formule des polaires (et la linéarité), avec  $\theta$  jouant le rôle de  $r$  et  $\phi$  le rôle de  $\theta$ .

Cela prouve que l'on peut passer, dans cette approximation, de  $g$  au  $\delta$ . Et c'est cohérent avec l'aspect de la sphère sur une petite partie de sa surface : elle paraît plate.

Ce résultat est général : **quelle que soit la métrique, si on ne regarde que des petits déplacements, l'espace peut être approximé localement comme plat.**

Le tenseur métrique  $g_{rn}$  est inversible et son inverse  $g^{mr}$  est aussi un tenseur (pas besoin de  $(\cdot)^{-1}$ , la position des indices suffit à savoir si on est en présence de la métrique ou de la métrique inverse).

$$g^{mr} g_{rn} = \delta_n^m$$

*preuve (dans le cas particulier où  $g$  est obtenu par changement de coordonnées depuis un espace plat):*

$$\begin{aligned} g_{rn} &= \delta_{st} \frac{\partial x^s}{\partial y^r} \frac{\partial x^t}{\partial y^n} \\ g^{mr} &= \delta^{uv} \frac{\partial y^m}{\partial x^u} \frac{\partial y^r}{\partial x^v} \\ g^{mr} g_{rn} &= \delta_{st} \delta^{uv} \frac{\partial y^r}{\partial x^v} \frac{\partial x^s}{\partial y^r} \frac{\partial y^m}{\partial x^u} \frac{\partial x^t}{\partial y^n} \\ &= \delta_{st} \delta^{uv} \delta_v^s \frac{\partial y^m}{\partial x^u} \frac{\partial x^t}{\partial y^n} \\ &= \delta_t^u \frac{\partial y^m}{\partial x^u} \frac{\partial x^t}{\partial y^n} \\ &= \frac{\partial y^m}{\partial x^t} \frac{\partial x^t}{\partial y^n} \\ &= \delta_n^m \end{aligned}$$

**La métrique permet d'abaisser l'indice d'un vecteur contravariant** (le transformer en covariant) :

$$V^m g_{mn} = V_n$$

$$(puisque ds^2 = dy_s dy^s = g_{rs} dy^r dy^s)$$

C'est d'ailleurs la **vraie définition d'un vecteur covariant**, et cela met en lumière un point important : la covariance est liée aux projections orthogonales de vecteurs.

La métrique inverse permet, elle, de monter les indices :

$$V_n g^{nm} = V^m$$

(s'obtient en multipliant par  $g^{nm}$  l'équation précédente...).

Dans les calculs tensoriels, on peut se contenter de voir les métriques comme des ascenseurs qui montent et descendent les indices...

Utiliser la métrique permet de ne manipuler que des vecteurs petits déplacements  $dx$  qui le sont réellement (tensoriellement parlant). C'est la cuisine interne de la métrique que de transformer ces tenseurs contravariants en une distance scalaire :

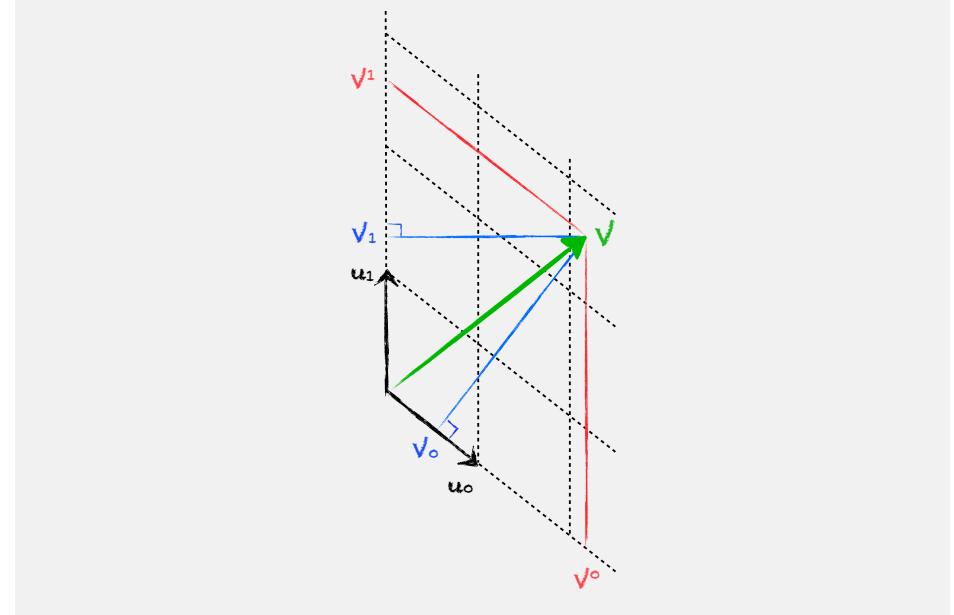
$$ds^2 = dx^m dx_m = dx^m dx^n g_{nm}$$

La métrique permet aussi de **contracter un tenseur** de rang  $\geq 2$ . La contraction généralise la trace d'une matrice à des rangs supérieurs à 2 en faisant passer le tenseur d'un rang  $p$  à un rang  $p-2$ .

Par exemple :  $A^{mnp} g_{pn} = A^{mn}_n = B^m$  où  $B$  n'est plus qu'un tenseur de rang 1 (si on est dans un espace de dimension 3,  $A^{mn}_n$  est une somme de 3 tenseurs de rang 1 :  $(A^m)_1^1 + (A^m)_2^2 + (A^m)_3^3$  ).

*Retour sur la distinction entre vecteur contravariant et vecteur covariant :*

les deux portent la même information ou plus techniquement, l'espace vectoriel et son dual (l'espace des formes linéaires que sont réellement les vecteurs covariants) ont même dimension. Mais alors que le vecteur contravariant est une somme des vecteurs de base (un développement sur la base :  $\vec{V} = V^i \vec{u}_i$ ), le covariant est une **projection** sur ces mêmes vecteurs ( $V_i = V^*(\vec{u}_i) = \vec{V} \cdot \vec{u}_i$ ).



Les coordonnées covariantes correspondent en effet à des projections orthogonales sur les vecteurs de base (et il faut la métrique pour préciser l'orthogonalité), alors que les coordonnées contravariantes correspondent à des projections sur l'axe d'un vecteur parallèlement à tous les autres (la métrique ne sert pas).

Mais une coordonnée contravariante dépend du coup de tous les vecteurs de base tandis qu'une coordonnée covariante ne dépend que du vecteur sur lequel on projette !

# Espace-temps

Depuis 1905, l'espace du physicien est promu espace-temps. Le point devient alors évènement puisque l'instant  $t$  vient compléter les trois coordonnées spatiales. La métrique en est toute chamboulée mais commençons par une frivolité : les lettres grecques remplacent les romaines pour les indices (indice en lettres romaines : varie de 1 à 3 ; lettres grecques : 0 à 3 ; 0 = le temps), car maintenant on joue sur 4 dimensions.

Un vecteur petit déplacement s'écrit donc maintenant :

$$dx^\mu = \begin{pmatrix} dx^0 \\ dx^1 \\ dx^2 \\ dx^3 \end{pmatrix}$$

Et on parle alors de **quadrivecteur position**.

Beaucoup plus profond : la petite distance  $ds$  est remplacée par le petit temps propre  $d\tau$ .

Et la métrique d'espace plat  $\delta_{mn}$  devient  $\eta_{\mu\nu}$ , métrique

d'**espace-temps plat**, ou de **Minkowski**, s'écrivant :

$$\eta_{\mu\nu} = g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Cette métrique n'est pas unique, car on peut la translater et la faire tourner (et comme en espace plat, ça ne change rien). On peut aussi lui appliquer une transformation de Lorentz (ça, c'est nouveau).

L'espace-temps peut bien évidemment aussi devenir courbe. Il suffit par exemple d'avoir un jeu de coordonnées accélérées comme on l'a vu précédemment (courbure et accélération sont très liées) !

Lors d'un changement de système de coordonnées, la métrique minkowskienne se transforme comme les métriques euclidiennes (c'est toujours un tenseur).

À noter que la **signature** (signe des valeurs propres du tenseur) reste la même lors de ces changements de coordonnées, gardant trace de la nature spatiotemporelle de la métrique : quelque soit la transformation, la signature sera constituée d'1 signe plus et 3 signes moins (+---) (dans cette convention tout au moins sachant que la convention inverse (-+++), avec un signe moins pour le temps, était tout autant envisageable). C'est là la grande différence avec l'espace ordinaire, ou euclidien, dont les signatures des métriques se bornent aux signes plus (+++).

Le nombre de signes + et - de la signature donne la dimension du plus grand sous-espace défini respectivement positif ou négatif (donc munis d'une norme positive ou négative).

Mais d'où vient ce temps propre et que change l'apparition du signe moins dans la métrique ?

Comme pour l'espace euclidien, on souhaite trouver une norme à la fois quadratique et positive. Mais le temps ne s'ajoute pas comme une simple nouvelle dimension indépendante. Il existe en effet un lien entre temps et espace : l'existence d'une vitesse indépassable que l'on va appeler vitesse de la lumière  $c$ .

La constante  $c$  va déjà nous permettre de donner la même dimension à toutes les composantes du vecteur petit déplacement (on va choisir le temps, ici, comme dimension) :

$$dx^\mu = \begin{pmatrix} dt \\ dx/c \\ dy/c \\ dz/c \end{pmatrix}$$

Bien que compliquant la notion de distance, cet invariant indépassable n'en fournit pas moins la clé pour définir une norme positive sur l'espace-temps :

$$\begin{aligned} v \leq c &\Leftrightarrow \frac{dr}{dt} \leq c \Leftrightarrow \frac{dr}{c} \leq dt \\ &\Rightarrow \frac{dr^2}{c^2} \leq dt^2 \Leftrightarrow dt^2 - \frac{dr^2}{c^2} \geq 0 \end{aligned}$$

Donc en appelant le petit intervalle entre deux événements,

**temps propre**  $d\tau$ , défini comme :

$$d\tau = \sqrt{dt^2 - \frac{dr^2}{c^2}}$$

on se retrouve bien avec l'équivalent d'une «distance» (ayant ici la dimension d'un temps mais c'est arbitraire ; on pourrait tout aussi bien multiplier  $dx^\mu$  par  $c$  et se retrouver avec des distances) toujours positive (du moins pour des événements reliés physiquement comme on va le voir tout de suite).

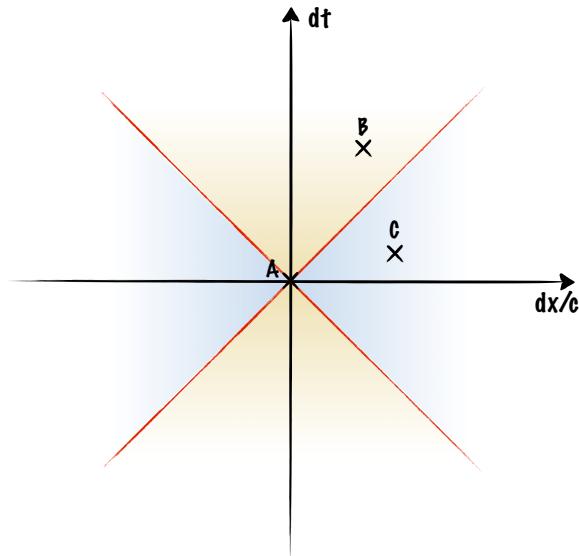
Et de la définition de  $d\tau$ , on déduit la métrique de Minkowski avec ses curieux signes moins.

Mais attention, la condition  $dt^2 - \frac{dr^2}{c^2} \geq 0$  ne couvre pas tout l'espace, mais seulement un double cône limité par les génératrices  $dt = \pm \frac{dr}{c}$ .

Dans ce cône se trouvent les événements pouvant être liés causalement puisque le temps propre qui les sépare respecte la condition d'indépassabilité de la vitesse de la lumière. En dehors du cône, une relation causale n'est plus possible.

L'espace de Minkowski contient donc trois types de vecteurs suivant le signe du carré de leur norme :

- les vecteurs de type temps, de norme positive ;
- les vecteurs de type espace, de norme négative ;
- et les vecteurs de type lumière, de norme nulle, ces derniers correspondants à des événements reliés à la vitesse de la lumière.



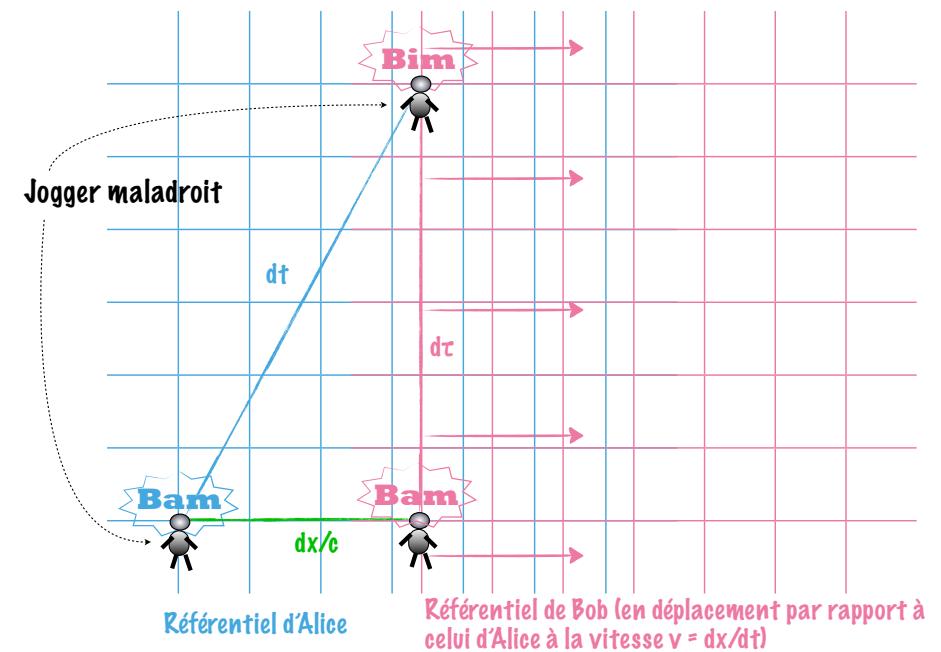
Sur le graphique précédent, le vecteur  $\vec{AB}$  est de type temps, sa norme est positive, tandis que le vecteur  $\vec{AC}$  est de type espace. Les événements A et B peuvent être reliés par un signal se déplaçant moins vite que la lumière tandis que c'est impossible pour les événements A et C.

Pythagore peut nous aider à mieux comprendre le temps propre. Écrire  $dt^2 = d\tau^2 + dx^2/c^2$  met en évidence les deux composantes orthogonales du laps de temps  $dt$  entre deux événements mesuré dans un référentiel donné :

- un temps lié au décalage spatial entre les événements et correspondant au temps mis par la lumière pour parcourir cet espace,
- et un temps lié au décalage temporel entre les événements pour un décalage spatial nul.

Le temps propre peut ainsi être défini comme le temps s'écoulant dans le référentiel «propre» des événements, le référentiel dans lequel l'évènement A et l'évènement B se déroulent aux mêmes coordonnées spatiales.

Alice voit un jogger cogner une pierre devant elle, «bam», puis, plus loin, se prendre une branche «bim». Pour Alice, il s'écoule un laps de temps  $dt$  entre ces deux évènements. Mais pour Bob, la fourmi coincée dans la capuche du jogger, la mesure donne  $d\tau$  entre «bam» et «bim» car il n'y a alors, pour Bob, aucun décalage spatial entre les évènements.



On en conclut plusieurs choses :

- L'espace, par l'intermédiaire du déplacement, crée du temps ! L'inverse nous est plus familier : sauf incident, on sera en des lieux différents entre notre montée et notre descente du train, mais cela aura pris du temps. On ne fait donc que symétriser les rôles du temps et de l'espace. D'où l'«espace-temps».



- Le temps propre, lié à aucun décalage spatial, et ne bénéficiant donc d'aucune «fabrication» de temps par le déplacement, correspond au temps minimum liant deux événements.
- Enfin, de l'orthogonalité entre  $d\tau$  et  $dx/c$ , on déduit qu'un objet se déplaçant à la vitesse de la lumière est vu par un spectateur comme ayant un temps propre nul, une vie éternelle d'un instant étiré à l'infini. Mais pour l'objet, principe de relativité oblige, le battement du temps persiste

identiquement, c'est le reste de l'univers qui est figé devant ses yeux.

#### Aparté :

Dans un espace-temps en expansion, le temps propre s'écrit :  $d\tau^2 = dt^2 - a(t)ds^2$

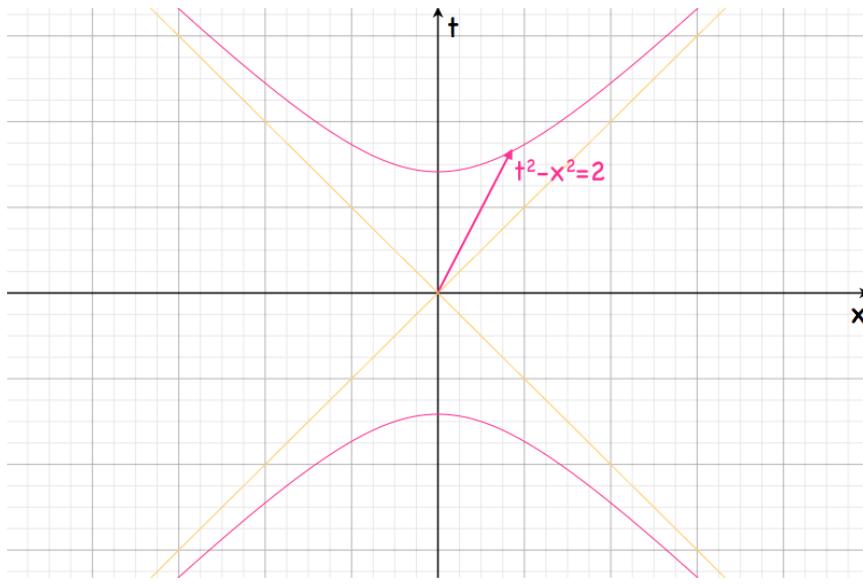
La ressemblance avec la métrique d'un espace 2D permet d'en tirer une analogie.

$ds^2 = dr^2 + r^2d\theta^2$  décrit un espace plat avec un  $d\theta$  qui ne cesse d'augmenter quand  $r$  augmente.

$ds^2 = d\theta^2 + (\sin \theta)^2 d\phi^2$  décrit une sphère avec un  $d\phi$  qui augmente pour les petits  $\theta$  puis diminue. Et si à la place de  $\sin \theta$ , on a un terme  $e^\theta$ , on se retrouve avec un encornet.

Extrapolé à l'espace-temps,  $ds$  joue le rôle de  $d\phi$ , et suivant l'allure de  $a(t)$ , on a un univers qui s'étend à l'infini ou se recontracte de diverses manières.

L'existence d'une vitesse indépassable va de pair avec une symétrie de l'espace de type **hyperbolique**. En effet, ce type de géométrie s'accorde parfaitement de l'existence d'asymptotes. Et de fait, les transformations de Lorentz qui permettent de passer d'un référentiel à l'autre (de vitesses différentes) sont en fait des rotations hyperboliques et c'est pour cela que le temps propre  $t^2 - x^2$ , invariant, dessine une hyperbole.



On remarque d'ailleurs qu'une vitesse «newtonienne» définie comme  $v^\mu = c \frac{dx^\mu}{dt} = \begin{pmatrix} c \\ v^i \end{pmatrix}$  ne convient pas puisque sa norme  $v_\mu v^\mu = c^2 - v^2$  n'est pas invariante.

Par contre, la **quadrivitesse**  $u^\mu = c \frac{dx^\mu}{d\tau} = \gamma \begin{pmatrix} c \\ v^i \end{pmatrix}$  réussit le test :  $u_\mu u^\mu = c^2$ .

Transformations de Lorentz :

$$\begin{pmatrix} dt' \\ dx'/c \\ dy'/c \\ dz'/c \end{pmatrix} = \Lambda(v) \begin{pmatrix} dt \\ dx/c \\ dy/c \\ dz/c \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} dt \\ dx/c \\ dy/c \\ dz/c \end{pmatrix} = \begin{pmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} dt \\ dx/c \\ dy/c \\ dz/c \end{pmatrix}$$

avec  $\gamma = \frac{1}{\sqrt{1 - \beta^2}}$ ,  $\beta = v/c$  et  $\cosh\phi = \gamma$

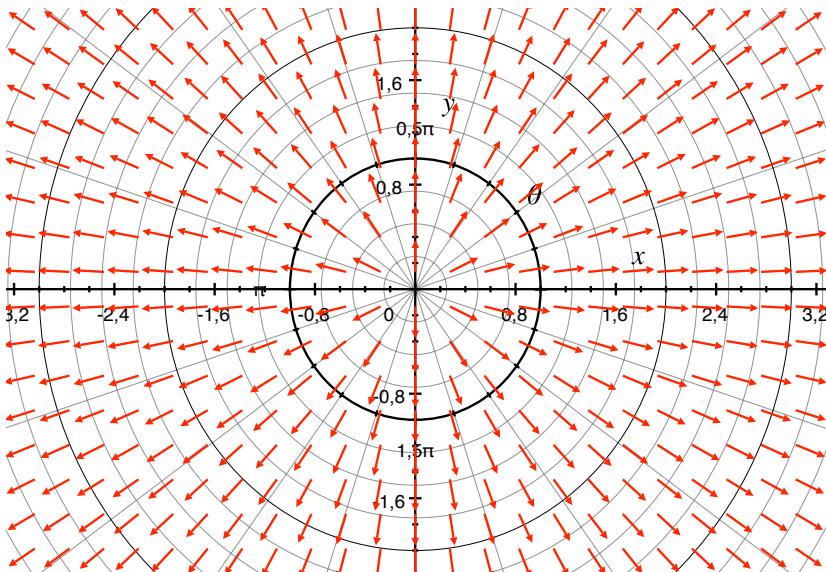
C'est donc le principe de relativité qui dicte aux quadri-vecteurs de conserver cette drôle de norme afin de paraître identiques aux yeux d'un observateur, quelle que soit sa vitesse (pour que le stylo dans la main de Bob n'apparaisse pas distordu si Bob se met à foncer à  $0,9c$ , il faut que le stylo subisse, comme Bob, une rotation hyperbolique).

## SECTION 5

# La dérivée covariante

Peut-on dériver des tenseurs ?

Savoir comment les composantes de ces objets physiques varient d'un point à l'autre de l'espace peut se révéler, on s'en doute, très intéressant.



Mais un détail nous chagrine vite : la dérivée d'un vecteur n'est pas un tenseur !

On peut facilement s'en rendre compte à l'aide d'un contre-exemple : en coordonnées cartésiennes, la divergence d'un champ de vecteur s'écrit  $\partial_i V^i$ , mais l'utilisation de la même formule en polaires amène à pas mal de désillusions.

Sur le champ de vecteurs représenté ci-contre, défini par  $V^r = 1$  et  $V^\theta = 0$ , on a bien  $\partial_i V^i \neq 0$  en cartésiennes mais

$$\partial_i V^i = \frac{\partial}{\partial r} V^r + \frac{1}{r} \frac{\partial}{\partial \theta} V^\theta = 0 \text{ en polaires.}$$

Donc clairement  $\frac{\partial V^m(x)}{\partial x^m} \neq \frac{\partial V^m(y)}{\partial y^m}$  !

Pourtant, on voit bien qu'il diverge (on peut toujours faire confiance aux cartésiennes). Source du problème : notre formule ne prend pas en compte la variation spatiale de  $\vec{u}_r$  ...

*Remarque :*

la formule de la divergence en polaire est effectivement plus compliquée que  $\partial_i V^i$  et s'écrit :

$$div(\vec{V}) = \frac{1}{r} \frac{\partial(rV^r)}{\partial r} + \frac{1}{r} \frac{\partial(V^\theta)}{\partial \theta}$$

Il faut trouver une autre manière de dériver (une manière qui ne donne pas une réponse nulle pour ce champ).

On va se mettre sur la voie en supposant que la dérivée est un tenseur et en regardant ce qui cloche. Prenons un vecteur covariant par commodité (afin d'avoir un tenseur doublement covariant) :

$$\frac{\partial V_m(x)}{\partial x^n} = T_{mn}(x)$$

Si  $T$  est bien un tenseur (notre naïve et erronée supposition), il doit se transformer comme un tenseur et donc :

$$\begin{aligned} T'_{mn}(y) &= \frac{\partial x^r}{\partial y^m} \frac{\partial x^s}{\partial y^n} T_{rs}(x) \\ &= \frac{\partial x^r}{\partial y^m} \frac{\partial x^s}{\partial y^n} \frac{\partial V_r(x)}{\partial x^s} \\ &= \frac{\partial x^r}{\partial y^m} \frac{\partial V_r(x)}{\partial y^n} \end{aligned}$$

Or par définition :

$$T'_{mn}(y) = \frac{\partial V'_m(y)}{\partial y^n}$$

La clé est dans ce qui différencie les deux expressions.

Passons d'abord de  $V_m(y)$  à  $V_r(x)$  dans la 2<sup>ème</sup> expression :

$$\begin{aligned} \frac{\partial V'_m(y)}{\partial y^n} &= \frac{\partial}{\partial y^n} \left( \frac{\partial x^r}{\partial y^m} V_r(x) \right) \\ &= \left( \frac{\partial}{\partial y^n} \frac{\partial x^r}{\partial y^m} \right) V_r(x) + \frac{\partial x^r}{\partial y^m} \frac{\partial V_r(x)}{\partial y^n} \end{aligned}$$

On retrouve l'expression espérée (en couleur) mais avec un terme supplémentaire ! Et ce terme est très intéressant, il est proportionnel au vecteur de départ (et non à sa dérivée) multiplié par une dérivée du lien entre systèmes  $x$  et  $y$ , c'est-à-dire un terme qui regarde comment le lien entre les

coordonnées  $x$  et  $y$  évolue dans l'espace. Si ce terme est nul, la dérivée «normale» se transforme bien comme on veut. Or si ce terme est nul, cela signifie que le rapport entre nos 2 systèmes de coordonnées est partout le même. Ouf ! On est content de retrouver la même dérivée en  $x$  et en  $y$  dans ce cas. Et on voit bien, à l'inverse, que c'est la possibilité de modification du rapport entre les 2 systèmes qui empêche la dérivée ordinaire d'être un tenseur.

Si on veut construire une dérivée qui soit bien un tenseur, il paraît donc naturel de lui ajouter un terme supplémentaire qui sera proportionnel au vecteur et à une combinaison linéaire de termes de type  $\frac{\partial}{\partial y^n} \frac{\partial x^r}{\partial y^m}$ , objet à 3 indices (2 co et un contra) qu'on va appeler  $\Gamma$ . Réinjecté dans le calcul de changement de coordonnées, on pourra ainsi rétablir l'égalité recherchée.

On définit alors une dérivée «vraie», dont le résultat est un tenseur. On l'appelle **dérivée covariante** et on la note  $\nabla_n$  :

$$\nabla_n V_m = \frac{\partial V_m}{\partial y^n} - \Gamma_{nm}^r V_r \quad (= T_{nm})$$

*Remarque :* le signe moins n'est qu'une convention (on réserve le plus à la dérivée covariante d'un tenseur contravariant)

Mathématiquement, la dérivée covariante est une **connexion affine** sur la variété pseudo-riemannienne représentant

l'espace-temps qui permet de transporter un vecteur parallèlement à lui-même le long d'une courbe de la variété. Rien que ça...

En langage moins barbare, la dérivée covariante est appelée connexion puisqu'elle connecte entre eux deux espaces tangents plats de la variété courbe (variété  $\approx$  espace).

Dans notre cas, il s'agit de la **connexion de Levi-Civita** (ou encore connexion métrique). Et les  $\Gamma$  sont les coefficients de la connexion qu'on appellera ici coefficient de Christoffel.

Reste à trouver  $\Gamma$ ... On sait déjà qu'il va comporter des termes de la forme «dérivée de lien entre coordonnées». Or si on remonte plus haut, on a une grandeur qui généralise le lien entre coordonnées : la métrique. Donc plutôt que de réinjecter la formule générale de dérivée covariante dans celle de changement de coordonnées, intéressons-nous à la dérivée covariante d'une métrique  $g$ .

Première chose :

la dérivée covariante d'un tenseur de rang 2 se calcule indice par indice = «vecteur par vecteur» (on applique la formule sur  $m$  puis sur  $n$ ), chaque rang étant un vecteur :

$$\nabla_p T_{mn} = \frac{\partial T_{mn}}{\partial y^p} - \Gamma_{pm}^r T_{rn} - \Gamma_{pn}^r T_{mr}$$

Deuxième chose :

la dérivée covariante de la métrique des coordonnées cartésiennes est nécessairement nulle puisque rien ne bouge.

Or il s'agit d'un tenseur, donc ça doit être vrai dans tout jeu de coordonnées !

On passe des  $x$  cartésiens aux  $y$  et  $\delta$  devient  $g(y)$  :

$$\nabla_r \delta_{mn} = 0 \Rightarrow \nabla_r g_{mn}(y) = 0$$

Ce résultat porte même un nom : **théorème de Ricci** (on l'a obtenu ici d'une façon un peu cavalière, mais bon...). On le voit parfois exprimé d'une autre façon : «**la dérivée covariante tue la métrique**».

Mathématiquement : on dit que la dérivée covariante est **compatible avec la métrique** quand  $\nabla_r g_{mn}(y) = 0$  .

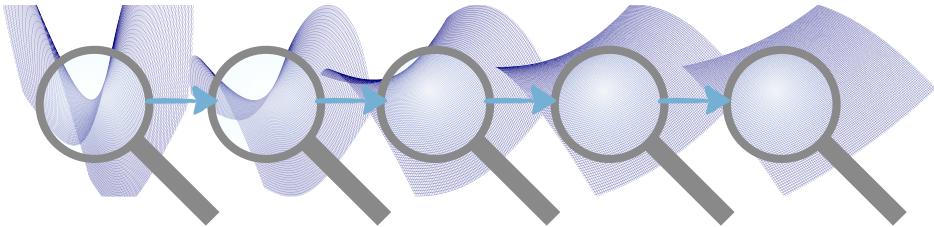
Ça nous dit donc que la dérivée covariante d'une métrique est nulle. Mais attention, la relation n'implique pas que la métrique soit partout constante ! Et heureusement puisqu'**une métrique constante signifierait un espace plat**.

En effet, si le rapport entre les directions de l'espace ne varie pas quand on se déplace, on obtient un quadrillage régulier. Et il suffit alors de quelques rotations et/ou dilatations sur les différentes directions pour vite se ramener à un petit  $\delta_n^m$  .

Donc, reprenons : on n'est pas en train d'affirmer que l'espace est forcément plat (ça gâcherait le plaisir), on dit juste qu'un petit décalage spatial conserve la métrique. Bref, **l'espace est**

**localemement plat** = on a **localemement**, en chaque point de l'espace, un petit repère cartésien.

C'est pour cette raison qu'on a l'impression de vivre sur une Terre plate, l'espace accessible au regard humain étant suffisamment restreint comparé au rayon de courbure de la sphère terrestre.



La métrique (qui lie les directions de l'espace entre elles) ne varie donc pas entre deux points infiniment proches à une correction près qui est au minimum du second ordre. Plus géométrique : tout endroit de l'espace épouse localement la base d'un paraboloïde.

Cette «platitude» locale semble logique dans un espace physique sans singularités (pas d'arrêtes, tout varie gentiment) et ne fait que généraliser la tangente à 1D. Une tangente est bien un rapport constant entre les directions  $x$  et  $y$ . Et partout, si on zoomé assez, une courbe (physique, c.-à-d. sans bizarrie) épouse sa tangente : une courbe curviligne est une suite de segments droits infiniment petits... Ce n'est pas vraiment une découverte ! En 2D, on le vit tous les jours (la Terre est plate)... Et à 3D ou plus, c'est pareil, un espace, aussi tordu qu'il puisse être, est irrémédiablement plat vu de près ;

son quadrillage (son pavage) peut être fait de petits cubes (ou hypercubes) réguliers.

*Aparté : autre façon de montrer que l'espace est localement plat.*

Soit une métrique  $g_{mn}(x)$  telle que sa dérivée par rapport à  $x$  ne soit pas nulle. Quelque soit l'espace, on peut toujours passer au système de coordonnées  $y$ , de même origine, et tel que la déviation entre les  $x$  et les  $y$  soit quadratique (du moment que les déplacements sont assez petits) :

$$y^m = x^m + C_{rs}^m x^r x^s$$

C comporte  $D(D + 1)/2$  termes indépendants puisque  $r$  et  $s$  sont symétriques ( $(D + 1)/2$  pour chaque tenseur symétrique de rang 2, et là, on en a  $D$ ).

Or, si on veut que l'espace soit localement plat, il faut que :

$$\frac{\partial g_{mr}}{\partial y^s} = 0$$

ce qui fait là encore  $D(D + 1)/2$  équations linéaires.

Autant de paramètres que d'équations, donc on peut toujours construire  $y$  afin d'être localement plat !

En développant le calcul de  $\nabla_r g_{mn} = 0$  et en jouant sur les indices, on peut isoler  $\Gamma$  et obtenir une formule le liant aux dérivées de la métrique :

$$\Gamma_{mn}^r(y) = \frac{1}{2} g^{rp}(y) \left( \frac{\partial g_{pm}}{\partial y^n} + \frac{\partial g_{pn}}{\partial y^m} - \frac{\partial g_{mn}}{\partial y^p} \right)$$

démonstration :

$$\nabla_p g_{mn} = \nabla_n g_{pm} = \nabla_m g_{pn} = 0$$

$$\Rightarrow \nabla_n g_{pm} + \nabla_m g_{pn} - \nabla_p g_{mn} = 0$$

$$\Leftrightarrow [\partial_n g_{pm} - \Gamma_{np}^r g_{rm} - \Gamma_{nm}^r g_{pr}] + [\partial_m g_{pn} - \Gamma_{mp}^r g_{rn} - \Gamma_{mn}^r g_{pr}] \\ - [\partial_p g_{mn} - \Gamma_{pm}^r g_{rn} - \Gamma_{pn}^r g_{mr}] = 0$$

$$2\Gamma_{mn}^r g_{pr} = \partial_n g_{pm} + \partial_m g_{pn} - \partial_p g_{mn}$$

Car les indices covariants des Gammas sont symétriques par construction (tout comme ceux de la métrique)

$$\Rightarrow \Gamma_{mn}^r = \frac{1}{2} g^{rp} (\partial_n g_{pm} + \partial_m g_{pn} - \partial_p g_{mn})$$

$\Gamma$  s'appelle **symbole de Christoffel**. C'est lui qui arrange la dérivée ordinaire pour prendre en compte le fait que le petit déplacement  $dx$  qu'opère la dérivée ne s'inscrit pas nécessairement dans un espace avec une métrique constante. Ce n'est par conséquent pas une surprise que cette correction dépende de variations de la métrique et donc de variations à la planéité.

On remarque que les  $\Gamma$  sont évidemment nuls en coordonnées cartésiennes.

Et la **dérivée covariante d'un tenseur contravariant** ?

On s'attend à quelque chose de ressemblant et effectivement :

$$\nabla_n V^m = \partial_n V^m + \Gamma_{nr}^m V^r$$

On a le droit au signe plus maintenant...

preuve :

Écrivons seulement que la dérivée covariante est donnée grâce à un terme à 3 indices indéterminé  $\tilde{\Gamma}_{nr}^m$  :

$$\nabla_n V^m = \partial_n V^m + \tilde{\Gamma}_{nr}^m V^r$$

Et calculons la dérivée covariante du tenseur  $v_m V^m$  :

$$\begin{aligned} \nabla_n (v_m V^m) &= (\nabla_n v_m) V^m + v_m (\nabla_n V^m) \\ &= \partial_n v_m V^m - \Gamma_{nm}^r v_r V^m + v_m \partial_n V^m + v_m \tilde{\Gamma}_{nr}^m V^r \end{aligned}$$

Or la dérivée covariante d'un scalaire doit se confondre avec sa dérivée partielle :

$$\nabla_n (v_m V^m) = \partial_n v_m V^m + v_m \partial_n V^m$$

$$\text{Donc } -\Gamma_{nm}^r v_r V^m + v_m \tilde{\Gamma}_{nr}^m V^r = 0$$

Et en jouant sur les indices muets ( $r$  et  $m$ ) :

$$\tilde{\Gamma}_{nr}^m v_m V^r = \Gamma_{nr}^m v_m V^r$$

$$\tilde{\Gamma}_{nr}^m = \Gamma_{nr}^m$$

On peut finalement vérifier que notre champ  $V^r = 1$  et  $V^\theta = 0$  admet maintenant bien une divergence non nulle en polaires (et la même qu'en cartésiennes qui plus est). Il diverge, on est sauvé !

preuve :

En polaire :

$$V^r = 1 \text{ et } V^\theta = 0$$

De plus :

$$g_{rr} = 1, g_{r\theta} = g_{\theta r} = 0, g_{\theta\theta} = r^2$$

$$\nabla_n V^m = \partial_n V^m + \Gamma_{np}^m V^p$$

$\partial_n V^m = 0$  (c'était notre souci au départ)

et  $\Gamma_{np}^m V^p = 0$  pour  $p \neq r$  puisque  $V^\theta = 0$ .

$$\text{Or } \Gamma_{nr}^m = \frac{1}{2} g^{ms} (\partial_r g_{sn} + \partial_n g_{sr} - \partial_s g_{nr})$$

On a alors :

$$\Gamma_{rr}^r = 0$$

$$\Gamma_{\theta r}^r = \Gamma_{r\theta}^r = 0$$

$$\Gamma_{\theta r}^\theta = \Gamma_{r\theta}^\theta = \frac{1}{2} g^{\theta\theta} \partial_r g_{\theta\theta} = \frac{1}{r}$$

$$\Gamma_{\theta\theta}^\theta = 0$$

$$\text{Où on a utilisé } g^{\theta\theta} = \frac{1}{r^2}$$

Par conséquent, la divergence du champ vaut :

$$\nabla_r V^r = \Gamma_{\theta r}^\theta V^r = \frac{1}{r}$$

Et en cartésien :

$$V^x = \frac{x}{\sqrt{x^2 + y^2}}, V^y = \frac{y}{\sqrt{x^2 + y^2}}$$

Le divergent vaut :

$$\frac{y^2}{(x^2 + y^2)^{3/2}} + \frac{x^2}{(x^2 + y^2)^{3/2}} = \frac{1}{\sqrt{x^2 + y^2}} = \frac{1}{r}$$

On a bien accord !

On voit ici clairement le rôle des symboles des Christoffel : ils prennent le relai des dérivées partielles des coordonnées cartésiennes pour tenir compte des variations du nouveau jeu de coordonnées par rapport aux cartésiennes.

Remarque sur les notations :

les économiseurs d'espace notent les dérivées partielles par des virgules et les dérivées covariantes par des points-virgules :

$$\partial_y V^m = V^m,_y$$

$$\nabla_y V^m = V^m_{;y}$$