

DEMONSTRATION DE LA TRANSFORMATION DE LORENTZ EN RELATIVITE RESTREINTE

En mécanique classique newtonienne, une loi fondamentale est le *principe d'inertie*. Il stipule que : *tout point matériel suffisamment éloigné de tout autre pour n'en subir aucune influence persiste dans son état de repos ou de mouvement rectiligne uniforme* . Tout mouvement étant relatif à un système de référence, ce principe permet de sélectionner certains de ces systèmes, appelés *repères galiléens*.

Un repère galiléen est un système de référence par rapport auquel le principe d'inertie est vérifié.

Il existe une infinité possible de tels repères et ils ont, les uns par rapport aux autres, soit l'immobilité pour certains, soit des mouvements de translation rectiligne uniforme. La géométrie de l'espace où ils évoluent est admise comme **euclidienne**.

Considérons deux de ces repères que nous désignerons par K et K' et dont les axes homologues sont parallèles et de même sens. Par souci de simplicité, nous choisirons ces repères **orthonormés** et de telle façon que la translation uniforme de l'un par rapport à l'autre se fasse par glissement, l'un sur l'autre, des axes des abscisses. En outre, nous conviendrons qu'ils sont munis de la même norme métrique, c'est-à-dire qu'une unité sur l'axe des abscisses de K, mesurée selon K', devra être rigoureusement égale à une unité sur l'axe des abscisses de K', mesurée selon K, (ce qui n'implique pas nécessairement que la mesure d'une unité d'un repère faite selon l'autre donne encore une unité, comme nous le verrons).



Enfin, nous conviendrons que les temps t et t' , mesurés respectivement dans les repères K et K', sont donnés par des horloges de fabrications parfaitement identiques, lesquelles pourraient donc être parfaitement synchronisées si elles étaient immobiles les unes par rapport aux autres au même point, et nous prendrons pour valeur zéro des temps t et t' , aux origines O et O' des repères, l'instant où ces origines sont en coïncidence.

Pour de tels repères, la mécanique classique newtonienne a relevé

plusieurs principes dont certains n'ont, semble-t-il, jamais été pris en défaut par les innombrables observations et expériences faites jusqu'ici. Ce sont 4 de ces principes que nous prendrons comme **axiomes**, à savoir :

- 1) Si vecteur v est la vitesse de l'origine du repère K' par rapport à K :
 - a) Tous les points fixes de K' ont la vitesse v par rapport à K .
 - b) L'origine et tous les points fixes du repère K ont la vitesse $-v$ par rapport à K' .
- 2) Il y a totale symétrie (au signe près de leur vitesse relative) entre ces repères. (La même métrique que nous avons définie pour K et K' est en conformité avec cet axiome).
- 3) La coïncidence spatio-temporelle (qui consiste à ce que deux évènements aient lieu au même point au même instant) est vraie pour tous ces repères. (L'initialisation que nous avons faite des temps t et t' est en conformité avec cet axiome).
- 4) Des points fixes alignés dans l'un de ces repères sont aussi des points alignés pour l'autre repère et dans les mêmes rapports algébriques de leurs distances relatives, à tout instant pris dans cet autre repère.

En tenant compte de ces 4 axiomes, on peut passer des coordonnées spatio-temporelles (x, y, z, t) dans le repère K aux coordonnées (x', y', z', t') dans le repère K' par les relations suivantes :

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

Ces relations, qui ne sont pas les seules qu'on puisse tirer de ces 4 axiomes comme nous le verrons, sont désignées sous le nom de *transformation de Galilée*. Or, d'après cette transformation, couramment utilisée en mécanique classique, la loi de composition des vitesses est la suivante : Si, dans K' , un mobile se déplace sur $O'x'$ avec une vitesse w (mesurée dans K') cette vitesse est, par rapport à K , $v + w$. Ceci est si conforme au bon sens qu'il nous paraît évident. Les tapis roulants de certaines stations de métro à Paris, par exemple, utilisent cette loi d'addition des vitesses pour déplacer plus vite les piétons d'un quai à un autre.

Or, la lumière déroge constamment à cette loi d'addition des vitesses. De nombreuses expériences de physique, méticuleuses et précises, ont montré que si un repère se précipite au devant d'un rayon lumineux (dans le vide), celui-ci a toujours, par rapport à ce repère, la même vitesse c . Même chose si le repère fuit ce rayon. Cette surprenante constatation heurte évidemment notre bon sens et a créé un grand embarras chez les physiciens du XIX^{ème} siècle.

Dès lors, on peut se demander s'il ne faut pas remettre en cause un ou plusieurs des 4 axiomes précédemment cités. Mais ceux-ci, comme nous l'avons déjà dit, n'ont jamais été pris en défaut jusqu'ici.

Une autre manière de procéder est de se demander si ces 4 axiomes doivent nécessairement engendrer la transformation de Galilée ou, sinon, si l'on peut trouver une autre transformation (ensemble de 4 relations) qui, non seulement satisfait aux 4 axiomes précédents, mais aussi à ce cinquième axiome :

- 5) La vitesse de la lumière dans le vide est invariante quel que soit le repère galiléen où on la mesure.

Ce cinquième axiome n'est pas incompatible avec les 4 autres comme nous allons le montrer et conduit à d'autres relations de transformation que celles de Galilée.

- 6) Pour démontrer cela, nous admettrons que la propriété qu'ont les 4 relations de Galilée – et d'innombrables fonctions quantifiant des phénomènes physiques – d'être différentiables s'étend aussi aux nouvelles relations que nous cherchons, ce qui constitue notre sixième et dernier axiome.

*

Préambule :

L'axiome 4 a pour première conséquence qu'une droite fixe dans un repère doit être également perçue comme droite dans l'autre repère à tout instant et quelle que soit la vitesse v de l'un par rapport à l'autre. Par conséquent, les axes de l'un quelconque de ces repères n'apparaissent pas déformés aux observateurs en repos dans l'autre repère. D'autre part, l'orthogonalité des axes d'un repère doit être également perçue comme telle dans l'autre repère, pour la raison qu'une inclinaison quelconque correspondrait – eu égard à l'isotropie de l'espace euclidien – à une direction aléatoire non acceptable dans le cadre du déterminisme admis et observé en cinématique. Il s'ensuit que deux plans respectivement perpendiculaires à l'axe des abscisses de chaque repère doivent se retrouver en totale coïncidence lorsque leurs points d'intersection avec les axes des abscisses le sont eux-mêmes, ce qui signifie que, pour tous les observateurs d'un même repère situés dans l'un de ces plans, la coïncidence sera simultanée, au sens absolu que l'on peut donner à ce terme. Par conséquent, leurs horloges devront indiquer le même instant lors de cette coïncidence et, en particulier, le même instant que celui repéré au point d'intersection du plan considéré avec l'axe des abscisses.

Première étape :

Considérons un évènement ayant, par rapport au repère K , les coordonnées spatio-temporelles (x, y, z, t) . Occupons-nous d'abord de trouver, par rapport à K' , les coordonnées x' et t' de cet évènement. Pour embrasser toutes les possibilités, nous considérerons que x' est une fonction des 4 coordonnées spatio-temporelles de cet évènement par rapport à K . Même chose pour t' . Nous posons donc : $x' = f(x, y, z, t)$ et $t' = g(x, y, z, t)$. La question est alors de déterminer ces fonctions f et g qui, nécessairement, doivent être partout définies sur l'ensemble des nombres réels.

Soit alors un point fixe de K ayant pour coordonnées spatiales dans K : (x, y, z) . D'après l'axiome 1 il est animé de la vitesse $-v$ par rapport à K' . Le vecteur $-v$ étant colinéaire à v , lui-même colinéaire à l'axe des abscisses, la vitesse de ce point par rapport à K' a donc pour seule composante scalaire non nulle $\Delta x'/\Delta t'$. Les fonctions f et g étant supposées différentiables en vertu du 6^{ème} axiome, on peut écrire :

$dx' = f'_x dx + f'_y dy + f'_z dz + f'_t dt$ (f'_i étant la dérivée partielle de f par rapport à i). Comme le point considéré est fixe par rapport à K , $dx=dy=dz=0$. Donc : $dx' = f'_t dt$. D'autre part : $dt' = g'_x dx + g'_y dy + g'_z dz + g'_t dt$. Pour la même raison que précédemment, nous obtenons $dt' = g'_t dt$. Par conséquent, la vitesse de ce point par rapport à K' est :

$$\frac{dx'}{dt'} = \frac{f'_t dt}{g'_t dt} = \frac{f'_t}{g'_t} = -v$$

quels que soient x, y, z, t

Donc : $f'_t = -vg'_t \Leftrightarrow f = -vg + C_t$. Mais C_t ne peut être affirmée constante que par rapport au temps t et dépend sans doute de x, y et z . Désignons alors C_t par la fonction $F(x, y, z)$, d'où :

$$f(x, y, z, t) = F(x, y, z) - vt'$$

Ce qui peut aussi s'écrire :

$x' = f(x, y, z, t) = F(x, y, z) - vt'$	relation (A)
---	--------------

La fonction F n'a pour variables que les trois coordonnées spatiales. Comme f et g sont nécessairement définies sur tout \mathbb{R} , F l'est donc aussi. Cherchons alors à la déterminer :

Tout d'abord, lorsque $O = O'$ et que les plans (Oy, Oz) et $(O'y', O'z')$ sont confondus, tout point M de K , d'abscisse $x = 0$, donc dans les plans précédemment cités, a aussi pour abscisse dans K' : $x' = 0$. D'autre part, lorsque $O' = O$, $t'(O') = 0$ et, selon la simultanéité évoquée en fin du préambule,

l'instant $t'(M)$, repéré en M , est aussi égal à zéro. Nous pouvons donc écrire, d'après la relation (A) : $0 = F(0, y, z) - v.0$, ce qui entraîne $F(0, y, z) = 0$, quels que soient y et z . Ceci est une première propriété intrinsèque de la fonction F .

Considérons maintenant deux points fixes M_1 ($M_1 \neq O$) et M_2 dans K , alignés avec O . Il existe un nombre réel λ tel que : mesure algébrique de $OM_2 = \lambda$ fois mesure algébrique de OM_1 , ce qui implique : $x_2 = \lambda x_1$; $y_2 = \lambda y_1$ et $z_2 = \lambda z_1$, et ces trois dernières égalités sont vraies quelle que soit la direction dans K de la droite sur laquelle sont alignés O, M_1, M_2 (même si celle-ci est confondue avec l'un des axes du repère K). D'après l'axiome 4, les points M_1 et M_2 sont aussi alignés avec O , repérés dans K' à tout instant t' , et, à l'instant $t' = 0$ (O et O' confondus), nous avons : mesure algébrique de $O'M_2 = \lambda$ fois mesure algébrique de $O'M_1$, ce qui implique : $x'_2 = \lambda x'_1$; $y'_2 = \lambda y'_1$ et $z'_2 = \lambda z'_1$, et là aussi quelle que soit la direction par rapport à K' de la droite où sont alignés ces trois points O', M_1 et M_2 .

Ne nous occupons pour l'instant que de x'_1 et x'_2 . En vertu de la relation (A), nous pouvons écrire :

$$\begin{aligned} x'_1 &= F(x_1, y_1, z_1) - v.0 = F(x_1, y_1, z_1) \\ x'_2 &= F(x_2, y_2, z_2) - v.0 = F(x_2, y_2, z_2) \end{aligned}$$

D'où : $\lambda x'_1 = F(\lambda x_1, \lambda y_1, \lambda z_1)$, quels que soient les réels x_1, y_1, z_1 et λ . On en déduit que : $\lambda F(x, y, z) = F(\lambda x, \lambda y, \lambda z)$, quels que soient x, y, z et λ , ce qui est donc une deuxième propriété intrinsèque de la fonction F (homogénéité de degré 1).

Considérons enfin trois points fixes de K : M, M_1 ($M_1 \neq M$) et M_2 , alignés sur une droite parallèle à l'axe Ox et tels que M soit dans le plan (Oy, Oz) , donc d'abscisse nulle. Là encore il existe λ tel que : mesure algébrique de $MM_2 = \lambda$ fois mesure algébrique de MM_1 , ce qui implique : $x_2 = \lambda x_1$, mais, dans ce cas particulier, $y_2 = y_1$ et $z_2 = z_1$. A l'instant $t' = 0$, O et O' sont confondus ainsi que les plans (Oy, Oz) et $(O'y', O'z')$. D'après l'axiome 4, nous pouvons encore écrire : $x'_2 = \lambda x'_1$, ce qui entraîne dans ce cas : $\lambda x'_1 = F(\lambda x_1, y_1, z_1)$, quels que soient les réels $x_1 \neq 0, y_1, z_1$ et λ . Dans le cas particulier où $x_1 = 0$, $\lambda F(0, y_1, z_1) = \lambda.0 = 0 = F(\lambda.0, y_1, z_1)$, ceci d'après la première propriété intrinsèque de F . On déduit de ces deux derniers cas que : $\lambda F(x, y, z) = F(\lambda x, y, z)$, quels que soient x, y, z et λ , ce qui est une troisième propriété intrinsèque de la fonction F .

De ces propriétés nous tirons : $F(x, y, z) = xF(1, y, z)$. La fonction $F(1, y, z)$ ne dépend plus que des variables y et z (puisque x est remplacé par une valeur constante). On peut désigner cette nouvelle fonction par $G(y, z)$ et, par conséquent :

$$F(x, y, z) = xG(y, z).$$

F étant définie sur tout l'ensemble des réels, $G(y, z)$ l'est aussi.

Etant donné que, selon l'axiome 6, $f(x, y, z, t)$ et $g(x, y, z, t)$ sont différentiables, elles sont donc continues et, d'après la relation (A), $F(x, y, z)$ est elle aussi continue et, par conséquent, est aussi continue $G(y, z) = F(1, y, z)$.

D'autre part : $F(\lambda x, \lambda y, \lambda z) = \lambda x G(\lambda y, \lambda z) = F(\lambda x, y, z) = \lambda x G(y, z)$, ce qui entraîne, pour λ et x non nuls :

$$G(\lambda y, \lambda z) = G(y, z), \text{ quels que soient } y, z \text{ (et } \lambda \neq 0).$$

Puisque G est continue, si, pour des valeurs données de y et z , nous faisons tendre λ vers 0, $G(\lambda y, \lambda z)$ tend vers $G(0, 0)$ et, par conséquent, la valeur ici constante de $G(y, z)$ tend aussi vers $G(0, 0)$, c'est-à-dire que : $G(y, z) = G(0, 0)$, quels que soient y et z . Donc $G(y, z)$ est une constante. Désignons-la par « a » et ainsi :

$$F(x, y, z) = ax \quad (\text{fonction linéaire}).$$

Dans le cas particulier où $x = 0$ cette fonction linéaire $F = a \cdot 0 = 0$. Or nous avons déjà démontré que $F(0, y, z) = 0$ (première propriété intrinsèque de F). Il n'y a donc pas contradiction ni, par conséquent, de valeur de x à exclure.

Donc la relation (A) peut s'écrire maintenant :

$x' = ax - vt'$ relation (B)

« a » est une constante qui dépend certainement de v et ne peut être nulle pour la raison suivante : Si a était nul, la relation (B) deviendrait $x' = -vt'$, ce qui signifierait que deux événements qui auraient lieu sur l'axe des abscisses en des points x_1 et x_2 distincts dans K et qui seraient simultanés pour K' auraient la même abscisse ($-vt'$) dans K' , c'est-à-dire qu'il s'agirait dans K' d'une coïncidence spatio-temporelle. Or, en vertu de l'axiome 3, cette coïncidence serait aussi vraie dans K , contredisant ainsi l'hypothèse $x_1 \neq x_2$.

Symétriquement : Supposons que, sans aucun autre changement, nous retournions le sens de l'axe $O'x'$ de K' . Pour tout événement, seule sa nouvelle abscisse X' sera modifiée et découlera de l'ancienne par la relation $X' = -x'$. Ainsi la relation (B) s'écrira : $X' = -ax + vt'$. Or, ce changement de sens de l'axe $O'x'$ aura aussi pour conséquence que la vitesse de K par rapport à K' aura les mêmes composantes scalaires que celles de K' par rapport à K , lesquelles n'ont pas été modifiées. Ainsi l'axiome 2 deviendra sans réserve du signe de la vitesse relative et nous pourrons donc écrire, par symétrie : $x = -aX' + vt$. En remplaçant X' par $-x'$ (c'est-à-dire en revenant au sens initial de $O'x'$), nous obtenons :

$x = ax' + vt$ relation (C)

A partir des deux relations (B) et (C), on peut exprimer x' et t' en fonction de x et t et, bien sûr, des constantes a et v . Pour cela il suffit de poser (B) et (C) sous forme d'un système de deux équations d'inconnues x' et t' et de résoudre ce système, par exemple par la méthode de Cramer :

$$\begin{cases} x' + vt' = ax \\ ax' + 0t' = x - vt \end{cases}$$

Le déterminant de ce système est :

$$D = \begin{vmatrix} 1 & v \\ a & 0 \end{vmatrix} = -av$$

v est évidemment supposée non nulle dans cette étude et a n'étant pas nul non plus, le déterminant du système précédent est différent de zéro, donc ce système est déterminé.

Solutions :

$$D_{x'} = \begin{vmatrix} ax & v \\ x-vt & 0 \end{vmatrix} = -vx + v^2t \quad D_{t'} = \begin{vmatrix} 1 & ax \\ a & x-vt \end{vmatrix} = x-vt-a^2x$$

$$x' = \frac{-vx + v^2t}{-av} \Leftrightarrow \boxed{x' = \frac{x-vt}{a} \text{ relation (D)}}$$

$$t' = \frac{x-vt-a^2x}{-av} \Leftrightarrow \boxed{t' = \frac{t - \frac{1-a^2}{v}x}{a} \text{ relation (E)}}$$

Supposons maintenant qu'à l'instant $t = t' = 0$ en O et O' confondus, un signal lumineux parte de $O = O'$ le long des axes des abscisses et dans le sens positif. Au bout d'un temps t , mesuré dans K , il atteint l'abscisse x telle que $x = ct$. D'après l'axiome 5, nous devons avoir $x' = ct'$ dans K' .

En remplaçant x' par ct' et x par ct dans les relations (D) et (E), nous obtenons :

* dans (D) : $ct' = (ct - vt)/a$ c'est-à-dire :

$$t' = \frac{(c - v)t}{ac} \quad \text{relation (F)}$$

* dans (E) :

$$t' = \frac{t - \frac{1 - a^2}{v} ct}{a} \quad \text{relation (G)}$$

Identifions les relations (F) et (G), ce qui est licite puisqu'elles se rapportent au même événement :

$$\frac{(c - v)t}{ac} = \frac{t - \frac{1 - a^2}{v} ct}{a} \Leftrightarrow \frac{c - v}{c} = \frac{1 - \frac{1 - a^2}{v} c^2}{a}$$

$$\Leftrightarrow v^2 = (1 - a^2)c^2 \Leftrightarrow 1 - a^2 = v^2/c^2 \Leftrightarrow a^2 = 1 - v^2/c^2 \quad (\text{ce qui impose } v < c)$$

Par conséquent :

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

Pour déterminer le signe de a reprenons l'exemple ci-dessus de notre rayon lumineux. Ce rayon se dirige dans le sens positif des axes Ox et $O'x'$. Donc, pour le point atteint, $x' > 0$ et $x > 0$. Puisque $x' = ct'$, en utilisant à nouveau la relation (B), nous obtenons : $x' = ax - v(x'/c) \Leftrightarrow x'(1 + v/c) = ax$. Comme x' , x et $1 + v/c$ sont positifs, a l'est aussi. Par conséquent :

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

(On obtient le même résultat si le rayon lumineux part sur la partie négative des axes des abscisses.)

Enfin, en remplaçant a par cette valeur dans les relations (D) et (E), nous obtenons :

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{relation (H)}$$

et :

$$t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{relation (I)}$$

*

Deuxième étape :

Maintenant, il nous reste, pour avoir la transformation complète, à trouver les deux relations qui lient y' et z' à x, y, z, t .

Nous allons d'abord exprimer y' et z' de la façon la plus générale comme des fonctions de x, y, z, t . Posons : $y' = l(x, y, z, t)$ et $z' = m(x, y, z, t)$, fonctions qui, elles aussi, doivent être définies sur tout \mathbb{R} , et différentiables en vertu du sixième axiome.

Intéressons-nous uniquement pour l'instant au cas de y' :

Soit un point fixe M dans K . D'après l'axiome 1, il a donc la vitesse $(-v)$ par rapport à K' , donc les composantes dy'/dt' et dz'/dt' sont nulles, ce qui signifie que $dy' = dz' = 0$. Or, $dy' = l'_x dx + l'_y dy + l'_z dz + l'_t dt$. M étant fixe dans K , $dx = dy = dz = 0$, quels que soient x, y, z, t , ce qui entraîne $dy' = l'_t dt = 0$. Comme dt n'est pas nul, on en déduit que $l'_t = 0$. Par conséquent, la fonction l est indépendante du temps t . On peut donc écrire : $y' = l(x, y, z)$.

En réutilisant l'axiome 4 et en faisant le même type de raisonnement que celui qui nous a permis de déterminer la fonction F dans la première étape, nous obtenons : $y' = \beta y$, (β étant une constante).

Mais β est une constante qui dépend peut-être de v ?

En réutilisant l'axiome 2, dans le même type de raisonnement que celui utilisé dans la première étape, nous obtenons $y = \beta y'$. Par conséquent : $y' = \beta \beta y'$, ce

qui entraîne $\beta^2 = 1$, d'où $\beta = \pm 1$.

Soit alors un point situé sur le demi-axe positif Oy . Lorsque O et O' sont confondus, le demi-axe positif $O'y'$ est confondu avec le demi-axe positif Oy et, par conséquent, le point situé sur ce dernier doit aussi apparaître sur le premier. Dans ce cas particulier, à $y > 0$ correspond donc $y' > 0$, ce qui entraîne $\beta = 1$. Comme β est une constante indépendante des coordonnées x, y, z, t , comme nous l'avons montré pour a , ce cas particulier est aussi le cas général pour β et nous aurons toujours $y' = y$.

Il est évident que la même démonstration peut être faite en considérant z' et aboutira à l'égalité : $z' = z$.

Ainsi sommes-nous parvenus aux 4 relations formant la nouvelle transformation qui tient compte, cette fois-ci, des six axiomes que nous avons pris pour point de départ en pages 2 et 3.

Cette transformation est la suivante :

$$\boxed{x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \quad y' = y \quad z' = z \quad t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Elle est connue sous le nom de *Transformation de Lorentz*, du nom du grand physicien néerlandais Hendrik Antoon LORENTZ (1853 – 1928, prix Nobel 1902) qui l'avait établie, non de cette façon, mais par des considérations d'électrodynamique et d'optique.

Toutefois, il ne faudrait pas conclure un peu trop vite que notre démonstration est terminée. En effet, nous n'avons utilisé l'axiome 5 que dans le cas particulier où un rayon lumineux se propageait le long des axes des abscisses. Nous devons donc vérifier qu'un rayon se propageant dans toute autre direction satisfait aussi à cet axiome 5. D'autre part, nous devons vérifier aussi que nos quatre relations respectent le principe d'inertie. Ceci ne pose pas de difficultés mathématiques et nous ne le ferons pas ici pour ne pas trop allonger ce texte.

Je remercie M. Jean Bouzon, professeur de mathématiques émérite des universités, pour l'aide amicale qu'il m'a apportée dans l'analyse de la fonction $F(x, y, z, t)$, rencontrée dans la première étape de cette démonstration, et pour ses remarques en général.