

Chapitre Trois

LA RELATIVITÉ RESTREINTE : CINÉMATIQUE

1. Le Principe de Relativité restreinte. - Puisque les équations de Maxwell supposées justes ne sont pas covariantes par la transformation de Galilée, celle-ci doit être fautive. Nous allons donc chercher à la modifier. Nous ne nous limitons pas à l'électromagnétisme et supposons que toutes les lois de la physique sont les mêmes dans tous les référentiel galiléens. Nous arrivons ainsi au *Principe de relativité restreinte* d'EINSTEIN, plus simplement appelé *principe de relativité*, qui généralise le principe de relativité de Galilée : “*Les lois de la physique sont identiques dans tous les référentiels galiléens*”. Aucune expérience de physique ne permet donc de mesurer d'une manière absolue le mouvement puisque tous les référentiels sont équivalents. La vitesse de la lumière qui se déduit des lois de l'électromagnétisme est donc la même dans tous les référentiels galiléens. Le mot “*restreint*” vient du fait que le Principe de relativité est limité aux mouvements de translations rectilignes uniformes et ne s'applique pas aux mouvements accélérés, en particulier aux mouvements de rotation. Ainsi, par exemple, un référentiel en rotation est non galiléen; il s'y développe des forces centrifuges : une particule libre ne s'y déplace pas en ligne droite.

Nous pouvons reprendre ici la remarque de la fin du §2 du chapitre 1 et celle de la fin du §2 du chapitre 2 en les appliquant à toutes les lois de la physique et en particulier aux lois de l'électromagnétisme : il nous paraît tout à fait naturel que tous nos appareils électriques et électroniques fonctionnent exactement de la même manière quelle que soit la période de l'année ou de la journée. Et pourtant, compte tenu des mouvements différents de ces appareils par rapport au référentiel de COPERNIC à ces différents moments, cela suppose la covariance des équations de MAXWELL.

Plus encore, la constitution des corps solides est d'origine électromagnétique. La modification des lois de l'électromagnétisme pourrait entraîner une déformation des solides qui serait a priori variable suivant la matière dont ils sont constitués. Le changement de vitesse de la Terre suivant les périodes de l'année

casserait, par déformations différentes des parties, les solides non homogènes!

2. Coordonnées d'un évènement. - Nous allons voir par la suite qu'il n'y a pas de temps absolu. Il faut donc définir avec précision comment le temps est mesuré. Dans chaque référentiel galiléen il y aura un réseau d'horloges immobiles synchronisées, aussi proches qu'on veut les unes des autres, appelées *horloges étalons*. Nous omettrons souvent, quand il n'y aura pas d'ambiguïté, et pour raison de simplicité, ce dernier adjectif.

Le mot étalon précise que le temps donné par l'horloge est le temps exact, compte tenu de l'unité choisie. L'horloge étalon ne présente ni avance ni retard par rapport à ce temps exact. Le fonctionnement d'une horloge étalon est basé sur un phénomène régulier permettant de mesurer le temps qui s'écoule. Pour plus de précision, voir § 10 du chapitre 7 .

Pour synchroniser ces horloges, nous disposons de deux méthodes équivalentes. Nous pouvons transporter une horloge étalon à une vitesse faible devant celle de la lumière. Nous supposons que la Mécanique newtonienne s'applique pour de tels objets se déplaçant lentement dans un référentiel galiléen. Dans ce cas, l'horloge indique constamment le temps absolu du référentiel. Elle permettra de mettre à l'heure toutes les horloges qu'elle rencontrera. Une autre méthode est d'utiliser la lumière. À l'instant $t_{(1)}$ nous envoyons, de l'horloge étalon (1) une impulsion lumineuse vers l'horloge (2). Elle arrive à l'instant $t_{(2)}$ et elle est renvoyée par un miroir vers l'horloge (1) où elle arrive à l'instant $t'_{(1)}$. Le principe de relativité restreinte implique que la lumière va à la même vitesse dans les deux sens. Nous avons donc :

$$t_{(2)} = \frac{t_{(1)} + t'_{(1)}}{2}$$

Il suffit alors de retarder ou d'avancer l'horloge (2) du décalage qu'elle avait avec le temps théorique $t_{(2)}$ au moment de l'arrivée du rayon lumineux sur le miroir, pour la mettre à l'heure.

Nous supposons alors que les horloges d'un référentiel galiléen restent synchronisées; elles mesurent le temps du référentiel.

Un *évènement* \mathcal{E} : désintégration d'une particule etc, définit un lieu dans l'espace à un moment donné. Un tel évènement est un point de l'*espace-temps*; Il sera repéré par les trois coordonnées x, y, z du lieu où il se produit et le temps t indiqué par l'horloge du référentiel située à cet endroit. Le lieu à un instant donné est un point de l'espace géométrique à trois dimensions appelé simplement *espace* lorsqu'il n'y a pas d'ambiguïté. Les quatre nombres

notés :

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = (x^\alpha)$$

sont les *coordonnées de l'évènement* \mathcal{E} dans le référentiel choisi. Un indice écrit en lettres grecques ira de 0 à 3 et correspondra aux *variables* d'espace-temps. un indice écrit en lettres latines ira de 1 à 3 et correspondra aux variables d'espace :

$$(x^\alpha) = \begin{pmatrix} x^0 \\ x^i \end{pmatrix}$$

3. La transformation spéciale de Lorentz. - Il nous faut maintenant voir le lien entre les coordonnées $(x^{\bar{\beta}})$ de l'évènement \mathcal{E} dans $\bar{\mathcal{R}}$ et (x^α) dans \mathcal{R} ; la barre au dessus d'un indice signifie que la coordonnée correspondante est dans $\bar{\mathcal{R}}$. Les référentiels \mathcal{R} et $\bar{\mathcal{R}}$ sont disposés comme sur la figure 1.1 . Nous réglons les horloges de \mathcal{R} et $\bar{\mathcal{R}}$ de façon à ce que lorsque 0 et $\bar{0}$ sont confondus, les deux horloges respectivement liées à 0 et $\bar{0}$ indiquent toutes les deux le temps 0 : $t = \bar{t} = 0$, comme cela a été fait implicitement au §2 du chapitre 1 .

Les variables t, x, y, z doivent être des fonctions linéaires de $\bar{t}, \bar{x}, \bar{y}, \bar{z}$, pour qu'une particule libre ayant un mouvement rectiligne uniforme dans $\bar{\mathcal{R}}$ ait également un mouvement rectiligne uniforme dans \mathcal{R} .

De plus $y = \bar{y}$ et $z = \bar{z}$ pour tout évènement, sinon on pourrait distinguer un sens absolu sur la droite des x : le sens de la vitesse du référentiel qui correspondrait à la coordonnée la plus petite par exemple (si on avait $y < \bar{y}$ cela impliquerait également $z < \bar{z}$ par symétrie de rotation autour de l'axe des x). L'espace étant supposé isotrope cela est impossible.

Considérons un rayon lumineux émis en O au moment de la rencontre avec \bar{O} . L'évènement \mathcal{E} est ici l'arrivée du rayon lumineux en un point. Le rayon se propage à la vitesse C dans \mathcal{R} et dans $\bar{\mathcal{R}}$. Sa direction est quelconque. Les repères d'espace étant supposés orthonormés, nous avons :

$$s^2 = C^2 t^2 - x^2 - y^2 - z^2 = 0 \quad \text{et} \quad \bar{s}^2 = C^2 \bar{t}^2 - \bar{x}^2 - \bar{y}^2 - \bar{z}^2 = 0$$

Les symboles 2 de s^2 et de \bar{s}^2 viennent du fait que ces nombres seront interprétés comme les carrés scalaires de vecteurs. Donc $s^2 = 0 \Leftrightarrow \bar{s}^2 = 0$; t, x, y, z étant des fonctions linéaires de $\bar{t}, \bar{x}, \bar{y}, \bar{z}$, cela implique $s^2 = k \bar{s}^2$ pour tout évènement qui n'annule pas ces nombres. Par raison de symétrie entre \mathcal{R} et $\bar{\mathcal{R}}$, nécessairement $k = 1$: $s^2 = \bar{s}^2$. Il vient : $C^2 t^2 - x^2 = C^2 \bar{t}^2 - \bar{x}^2$. Puisque cette relation doit être vraie quels que soient y et z, \bar{y} et \bar{z}, x et t

étant fixés ainsi que \bar{x} et \bar{t} , cela implique que x et t sont fonctions linéaires uniquement de \bar{x} et \bar{t} :

$$\begin{cases} x = a\bar{x} + b\bar{t} \\ t = c\bar{x} + d\bar{t} \end{cases}$$

Les coordonnées de \bar{O} vérifient $\bar{x} = 0$ soit :

$$x = b\bar{t} = \frac{bt}{d} = vt$$

La vitesse de $\bar{\mathcal{R}}$ étant ainsi maintenant notée par un petit v nous avons pour O :

$$\begin{aligned} x &= 0 \\ \frac{\bar{x}}{\bar{t}} &= -\frac{b}{a} = -v \end{aligned}$$

Il vient :

$$b = dv = av \quad \Rightarrow \quad a = d$$

$$\begin{cases} x = a\bar{x} + av\bar{t} \\ t = c\bar{x} + a\bar{t} \end{cases}$$

$$\begin{aligned} C^2t^2 - x^2 &= C^2c^2\bar{x}^2 + 2C^2ca\bar{x}\bar{t} + \\ C^2a^2\bar{t}^2 - a^2\bar{x}^2 - 2a^2v\bar{x}\bar{t} - a^2v^2\bar{t}^2 &= \\ C^2\bar{t}^2 - \bar{x}^2 & \end{aligned}$$

Il vient :

$$C^2ca - a^2v = 0$$

$$c = \frac{av}{C^2} ; C^2a^2 - a^2v^2 = C^2 \quad \Rightarrow \quad a^2 = \frac{C^2}{C^2 - v^2}$$

On sait que $a > 0$ de façon à retrouver la transformation de Galilée pour $v \ll C$; donc :

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

La transformation des coordonnées correspondant à la situation particulière envisagée s'appelle la *transformation spéciale de LORENTZ*. Elle s'écrit donc :

$$\begin{cases} x = \frac{1}{\sqrt{1 - \frac{v^2}{C^2}}} \bar{x} + \frac{v}{\sqrt{1 - \frac{v^2}{C^2}}} \bar{t} \\ t = \frac{\frac{v}{C^2}}{\sqrt{1 - \frac{v^2}{C^2}}} \bar{x} + \frac{1}{\sqrt{1 - \frac{v^2}{C^2}}} \bar{t} \end{cases} \quad (3, 1)$$

Soit $\tanh \varphi$ la tangente hyperbolique de φ ; posons $\tanh \varphi = \frac{v}{C}$; on sait que $\cosh^2 \varphi - \sinh^2 \varphi = 1$

$$1 - \tanh^2 \varphi = \frac{1}{\cosh^2 \varphi} ; \quad \cosh \varphi = \frac{1}{\sqrt{1 - \tanh^2 \varphi}}$$

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

$$\sinh \varphi = \tanh \varphi \cosh \varphi = \frac{\frac{v}{C}}{\sqrt{1 - \frac{v^2}{C^2}}}$$

On arrive à

$$\begin{cases} x = \cosh \varphi \bar{x} + \sinh \varphi C \bar{t} \\ C t = \sinh \varphi \bar{x} + \cosh \varphi C \bar{t} \end{cases} \quad (3, 2)$$

Ces équations sont très symétriques; l'utilisation de $C t$ et $C \bar{t}$ permet cette symétrie en ayant la même unité pour toutes les coordonnées.

Nous noterons désormais les coordonnées d'un évènement :

$$\begin{pmatrix} C t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = (x^\alpha) \quad (3, 3)$$

Remarquons la grande analogie de la transformation spéciale de LORENTZ avec les équations d'un changement d'axes par une rotation d'angle θ en géométrie euclidienne :

$$\begin{cases} x = \cos \theta \bar{x} - \sin \theta \bar{y} \\ y = \sin \theta \bar{x} + \cos \theta \bar{y} \end{cases}$$

Le passage des fonctions trigonométriques aux fonctions hyperboliques vient du fait que l'invariant n'est plus la distance usuelle : $x^2 + y^2$; mais l'intervalle : $C^2 t^2 - x^2$ avec l'apparition d'un signe moins.

4. Relativité de la simultanéité. - Nous voyons tout de suite que deux évènements simultanés dans $\bar{\mathcal{R}}$ (même \bar{t}) ne le sont pas dans \mathcal{R} dès qu'ils se produisent en deux endroits différents (\bar{x} différent). La Relativité restreinte établit ainsi une symétrie entre l'espace et le temps : la phrase "*Ces deux évènements se sont produits en même temps en deux endroits différents*" n'a

pas de sens si on ne précise pas le référentiel; de même que la phrase “*Ces deux évènements se sont produits au même endroit à deux instants différents*” (voir § 2 du chapitre 1).

5. Dilatation des temps. - Considérons une horloge étalon fixée en \bar{O} . Ses coordonnées sont : $\bar{x} = \bar{y} = \bar{z} = 0$; la transformation de LORENTZ donne $Ct = \cosh \varphi C\bar{t}$; soit :

$$t = \frac{1}{\sqrt{1 - \frac{v^2}{C^2}}} \bar{t} \quad (3, 4)$$

Le temps t croît donc plus vite que le temps \bar{t} . Il semble donc, vu du référentiel \mathcal{R} , que l’horloge fixée en \bar{O} retarde, indiquant ainsi un temps plus petit que celui de \mathcal{R} . Autrement dit, le temps de $\bar{\mathcal{R}}$, vu de \mathcal{R} , semble dilaté; c’est à dire qu’il s’écoule plus lentement.

Il faut insister sur le fait que ce phénomène correspond à une propriété physique de l’espace-temps et affecte de la même manière tous les processus physiques réguliers pouvant servir d’horloges (voir § 10 du chapitre 7).

Le fait que les horloges de $\bar{\mathcal{R}}$, vues de \mathcal{R} , retardent semble contredire la symétrie parfaite qui doit exister entre tous les référentiels galiléens, compte tenu du principe de relativité. Il semble impossible que les horloges de \mathcal{R} puissent également retarder, vues de $\bar{\mathcal{R}}$. Le paradoxe est résolu lorsqu’on se rend compte que la dissymétrie est introduite par le processus de mesure. En effet, une seule horloge de $\bar{\mathcal{R}}$ mesure le temps \bar{t} , tandis qu’une multitude d’horloges de \mathcal{R} donnent le temps t , les horloges coïncidant avec les différentes positions dans \mathcal{R} de l’horloge de $\bar{\mathcal{R}}$ étudiée.

Les instant origines étant maintenant quelconques, le point \bar{O} ne coïncide pas avec O à $t = 0$, la relation (3,4) n’est plus valable. Cependant, nous pouvons toujours écrire, utilisant les différentielles :

$$d\bar{t} = \sqrt{1 - \frac{v^2}{C^2}} dt \quad (3, 5)$$

6. Temps propre. - Considérons maintenant une horloge étalon ayant un mouvement quelconque, c’est à dire ayant une accélération non nulle par rapport aux référentiels galiléens. Cette accélération peut être elle même variable.

Nous faisons ici l’hypothèse que la relation (3,5) est toujours valable. Ainsi, nous supposons que l’accélération d’une horloge étalon par rapport à un système inertiel n’a aucune influence sur le fonctionnement de cette horloge (voir § 10 du chapitre 7). L’accroissement du temps de cette horloge donné par l’équation (3,5) est ainsi égal à l’accroissement de temps dans le référentiel

inertiel \mathcal{R}^0 dans lequel l'horloge est immobile à l'instant considéré. Cet accroissement de temps sera appelé accroissement du *temps propre* de l'horloge et noté $d\tau$. Nous arrivons donc à :

$$d\tau = \sqrt{1 - \frac{v^2}{C^2}} dt \quad (3, 6)$$

Pour un mouvement quelconque de l'horloge entre les évènements (1) et (2), l'accroissement de temps propre de l'horloge sera alors :

$$\tau_2 - \tau_1 = \int_{t_1}^{t_2} \sqrt{1 - \frac{v^2}{C^2}} dt \quad (3, 7)$$

Bien sûr, le mouvement complet entre les évènements (1) et (2) doit être repéré toujours avec le même référentiel inertiel.

Nous devons insister sur le fait que l'hypothèse que nous venons de faire a une très grande importance et représente un saut dans l'inconnu. La Relativité restreinte traite des mouvements des référentiels galiléens les uns par rapport aux autres, donc elle ne traite que les mouvements de translation rectilignes uniformes. Nous venons ici de passer aux référentiels accélérés, donc de quitter la Relativité restreinte proprement dite. Nous verrons dans l'étude de la Relativité générale, que l'action d'une accélération est la même qu'une action gravitationnelle. Donc, avec cette hypothèse, nous venons en fait de faire un premier pas dans la Relativité générale.

7. Contraction des longueurs. - Cherchons quelle est la longueur l dans \mathcal{R} d'une barre rigide de longueur \bar{l} dans $\bar{\mathcal{R}}$. Pour mesurer sa longueur dans un référentiel, il faut considérer ses deux extrémités au même instant suivant le temps du référentiel. Les extrémités sont par exemple $\bar{x} = 0$ et $\bar{x} = \bar{l}$ dans $\bar{\mathcal{R}}$; supposons que à $t = \bar{t} = 0$, l'extrémité gauche soit en $x = 0$; l'extrémité droite est en $x = \cosh \varphi \bar{x} + \sinh \varphi C\bar{t}$ avec :

$$\begin{aligned} 0 &= Ct = \sinh \varphi \bar{x} + \cosh \varphi C\bar{t} \\ x &= \cosh \varphi \bar{l} + \sinh \varphi \left(-\frac{\sinh \varphi}{\cosh \varphi} \right) \bar{l} \\ &= (\cosh^2 \varphi - \sinh^2 \varphi) \frac{\bar{l}}{\cosh \varphi} = \frac{\bar{l}}{\cosh \varphi} < \bar{l} \\ l &= \frac{\bar{l}}{\cosh \varphi} \end{aligned} \quad (3, 8)$$

Ainsi, $l < \bar{l}$. Dans ce phénomène de contraction des longueurs, il s'agit encore d'une propriété physique, géométrique, de l'espace-temps et non pas d'une compression de la barre qui est supposée parfaitement rigide (voir §10 du

chapitre 7) Ainsi un solide animé d'une grande vitesse tient moins de place qu'immobile.

Nous faisons maintenant ici une hypothèse analogue à celle du paragraphe précédent. Nous supposons que l'accélération d'un corps n'a pas d'effet sur les longueurs. Ainsi, pour connaître la longueur d'un petit objet en accélération, il suffit de connaître la longueur qu'aurait le même objet fixé dans son système d'inertie \mathcal{R}^0 (donc avec la même vitesse, mais sans accélération) au moment considéré. Pour un objet étendu, la longueur sera mesurée par intégration des longueurs élémentaires obtenues grâce à l'hypothèse ci-dessus.

8. Les quadrivecteurs. - La transformation spéciale de LORENTZ peut s'écrire, avec un produit de matrices, de la manière suivante :

$$\begin{pmatrix} Ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cosh \varphi & \sinh \varphi & 0 & 0 \\ \sinh \varphi & \cosh \varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} C\bar{t} \\ \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix} \quad (3,9)$$

Soit : $(x^\alpha) = \Lambda(x^{\bar{\beta}})$

ou

$$\begin{pmatrix} | \\ x^\alpha \\ | \end{pmatrix} = \Lambda \begin{pmatrix} | \\ x^{\bar{\beta}} \\ | \end{pmatrix}$$

$\Lambda = (\Lambda^\alpha_{\bar{\beta}})$ que nous appellerons *matrice de LORENTZ* est la matrice d'élément général $\Lambda^\alpha_{\bar{\beta}}$ écrite en (3,9). On écrit :

$$x^\alpha = \sum_{\bar{\beta}} \Lambda^\alpha_{\bar{\beta}} x^{\bar{\beta}} = \Lambda^\alpha_{\bar{\beta}} x^{\bar{\beta}} \quad (3,10)$$

avec la convention d'EINSTEIN : la sommation est sous entendue quand le même indice apparaît une fois en position haute, une fois en position basse. On appelle un tel indice : indice *muet*, car on peut changer sa notation sans changer le sens de l'expression mathématique écrite. Pour la matrice $(\Lambda^\alpha_{\bar{\beta}})$, l'indice en position haute est un indice de ligne, l'indice en position basse est un indice de colonne. C'est également le cas pour les matrices colonnes (x^α) et $(x^{\bar{\beta}})$; elles ont bien des indices en position haute (indices de lignes). Nous verrons que, une fois la position haute choisie pour l'indice des composantes d'un vecteur, la convention d'Einstein impose la position de tous les indices qui interviendront. Ce qui est remarquable, c'est qu'il y aura toujours une solution et une seule pour cette position. Ainsi, nous devons mettre les indices des vecteurs de base de l'espace en position basse de façon à pouvoir écrire :

$$\mathbf{V} = V^i \mathbf{e}_i \quad (3,11)$$

pour tout vecteur \mathbf{V} . Les nombres V^i sont les composantes de \mathbf{V} , la base étant notée $\{\mathbf{e}_i\}$.

L'équation (3,11) peut s'écrire matriciellement comme produit d'une matrice ligne par une matrice colonne placée à sa droite :

$$\mathbf{V} = (\dots \mathbf{e}_i \dots) \begin{pmatrix} | \\ x^j \\ | \end{pmatrix} = (e_i) (x^j)$$

Nous avons posé : $e_i x^i = x^i e_i$ (ssi). (ssi) veut dire : sans sommation sur l'indice i .

Dans ce dernier cas, la matrice ligne a bien encore des indices en position basse (indices de colonnes). Nous verrons cependant au chapitre 5 (calcul tensoriel), que cette convention de la position des indices des lignes et des colonnes pour une matrice n'est pas générale.

Si les axes de \mathcal{R} ne sont plus parallèles aux axes de $\bar{\mathcal{R}}$, les origines des temps et des coordonnées étant également quelconques, la transformation des coordonnées s'appelle simplement la *transformation de LORENTZ*. Pour une telle transformation, nous avons :

$$x^\alpha = \Lambda^\alpha_{\bar{\beta}} x^{\bar{\beta}} + c^\alpha \quad (3,12)$$

Les c^α sont des constantes quelconques. Λ est le produit de la matrice de la transformation spéciale de LORENTZ (3,9) par une matrice de la forme (3,13) :

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \boxed{\mathbf{U}} \\ 0 & & & \\ 0 & & & \end{pmatrix} \quad (3,13)$$

U est une matrice inversible quelconque. U est une matrice orthogonale dans le cas où les axes de \mathcal{R} sont orthonormés.

Supposons dans ce qui suit que les axes de \mathcal{R} soient quelconques :

Soient deux évènements \mathcal{A} et \mathcal{B} de coordonnées a^α et b^α dans \mathcal{R} et $a^{\bar{\alpha}}$ et $b^{\bar{\beta}}$ dans $\bar{\mathcal{R}}$:

$$\begin{aligned} a^\alpha &= \Lambda^\alpha_{\bar{\beta}} a^{\bar{\beta}} + c^\alpha \quad ; \quad b^\alpha = \Lambda^\alpha_{\bar{\beta}} b^{\bar{\beta}} + c^\alpha \\ \Rightarrow \quad \Delta x^\alpha &= b^\alpha - a^\alpha \end{aligned} \quad (3,14)$$

$$\begin{aligned} &= \Lambda^\alpha_{\bar{\beta}} b^{\bar{\beta}} - \Lambda^\alpha_{\bar{\beta}} a^{\bar{\beta}} = \Lambda^\alpha_{\bar{\beta}} (b^{\bar{\beta}} - a^{\bar{\beta}}) = \Lambda^\alpha_{\bar{\beta}} \Delta x^{\bar{\beta}} \\ \Delta x^\alpha &= \Lambda^\alpha_{\bar{\beta}} \Delta x^{\bar{\beta}} \end{aligned} \quad (3,15)$$

Revenons aux vecteurs du chapitre 1, le vecteur force par exemple. Ce vecteur était visualisé géométriquement par une flèche dans l'espace à trois dimensions, mais pour des calculs numériques on utilise les composantes qui sont un ensemble de trois nombres se transformant suivant : (3,16) dans le changement de repère défini par (3,17) :

$$F^i = U^i_{\bar{j}} F^{\bar{j}} \quad (3,16)$$

$$e_{\bar{j}} = U^i_{\bar{j}} e_i \quad (3,17)$$

On dit qu'un vecteur est *contravariant* car la loi de transformation (3,16) est opposée d'une certaine manière à la loi définissant le changement de base (3,17). Cette dernière loi est dite *covariante*. La matrice U donne la nouvelle base $(e_{\bar{j}})$ en fonction de l'ancienne (e_i) , mais donne les anciennes composantes en fonction des nouvelles; de plus dans (3,16) il y a sommation sur les colonnes et dans (3,17) sur les lignes. Remarquons que la convention d'Einstein permet d'écrire automatiquement les formules (3,16) et (3,17) sans aucune ambiguïté.

Nous noterons également :

$$F^{\bar{j}} = U^{\bar{j}}_i F^i \quad \text{et} \quad e_i = U^{\bar{j}}_i e_{\bar{j}}$$

Ainsi :

$$(U^{\bar{j}}_i) = (U^i_{\bar{j}})^{-1} \quad (3,18)$$

Les deux matrices $(U^{\bar{j}}_i)$ et $(U^i_{\bar{j}})$ sont inverses l'une de l'autre. Aucune ambiguïté n'existe avec cette notation entre U et U^{-1} . Nous pouvons écrire :

$$U^i_{\bar{k}} U^{\bar{k}}_j = \delta^i_j$$

δ^i_j est le *symbole de KRONECKER*. Il vaut 1 lorsque $i = j$ et 0 lorsque $i \neq j$. Suivant les cas, et pour avoir la même position des indices non muets dans les deux membres d'une équation, nous noterons également ce symbole :

$$\delta^{ij} \quad \text{ou} \quad \delta_{ij}$$

Remarquons que dans toutes les formules ci-dessus, nous pouvons remplacer sans ambiguïté \bar{j} par \bar{i} , en ayant ainsi les indices j et \bar{i} (idem $\bar{\beta}$ par $\bar{\alpha}$). Nous utiliserons parfois cette notation par la suite. La structure mathématique représentée par la loi de transformation (3,15) est la même que celle de la loi (3,16) sauf qu'il y a un nombre en plus. On peut donc considérer l'ensemble des quatre nombres Δx^α dans un référentiel muni de la loi de transformation (3,15) comme un vecteur d'un espace à quatre dimensions que nous appellerons l'*espace de MINKOWSKI*. On dit qu'on a un *quadrivecteur*; c'est ici le *quadrivecteur déplacement*. On a bien une structure d'espace vectoriel. La somme consiste à ajouter terme à terme les composantes, le produit externe par un nombre consiste à multiplier toutes les composantes par ce nombre.

Ces opérations peuvent être effectuées dans n'importe quel référentiel. On peut dire qu'un vecteur est la classe d'équivalence des couples (Référentiels, Composantes) munis de la relation d'équivalence définie par (3,15). Cette relation d'équivalence est bien stable pour l'addition interne, et la multiplication externe par un scalaire. La stabilité pour l'addition est détaillée dans les équations (4,1). Dorénavant tout ensemble de quatre nombres se transformant suivant (3,15) dans un changement de référentiel représentera un quadrivecteur. Les quadrivecteurs seront symbolisés par une lettre surmontée d'une flèche. Le référentiel auquel correspondent les composantes sera mis en indice lorsque nous écrirons le quadrivecteur comme vecteur colonne de ses composantes; nous omettrons cet indice lorsqu'il n'y aura pas d'ambiguïté. On écrit :

$$\Delta\vec{x} = (\Delta x^\alpha)_{\mathcal{R}} = (\Delta x^{\bar{\beta}})_{\bar{\mathcal{R}}}$$

Posons :

$$\vec{e}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}_{\mathcal{R}} ; \vec{e}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}_{\mathcal{R}} ; \vec{e}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}_{\mathcal{R}} ; \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}_{\mathcal{R}} ; \quad (3,19)$$

Ces quatre vecteurs forment une base de l'espace de MINKOWSKI. On a :

$$\Delta\vec{x} = \Delta x^\alpha \vec{e}_\alpha \quad (3,20)$$

Avec la même convention pour les vecteurs de base obtenus au moyen de $\bar{\mathcal{R}}$, nous avons :

$$\begin{aligned} \Delta\vec{x} &= \Delta x^\alpha \vec{e}_\alpha = \Lambda^\alpha_{\bar{\beta}} \Delta x^{\bar{\beta}} \vec{e}_\alpha = \Delta x^{\bar{\beta}} \vec{e}_{\bar{\beta}} \\ &\Rightarrow \vec{e}_{\bar{\beta}} = \Lambda^\alpha_{\bar{\beta}} \vec{e}_\alpha \end{aligned} \quad (3,21)$$

Cette dernière équation est l'analogue dans l'espace de MINKOWSKI de (3,17). Résumons nous : un référentiel galiléen est muni d'un repère orthonormé de l'espace à trois dimensions et d'un réseau d'horloges donnant le temps du référentiel. Il lui correspond un repère de *l'espace-temps* constitué de quatre vecteurs de base de l'espace de MINKOWSKI et d'une origine, l'évènement qui consiste en l'existence du point 0 au temps $t = 0$. Cet évènement est noté \mathcal{O}

9. Le produit scalaire. - Dans ce paragraphe, nous supposons les axes d'espace de \mathcal{R} orthonormés. Posons :

$$\vec{a}^2 = (a^0)^2 - (a^1)^2 - (a^2)^2 - (a^3)^2 \quad (3,22)$$

Cette quantité est invariante par changement de référentiel. En effet (3,15) appliquée à \vec{a} donne :

$$a^\alpha = \Lambda^\alpha_{\bar{\beta}} a^{\bar{\beta}} \quad (3,23)$$

ce qui implique que nous avons également :

$$\vec{a}^2 = (a^0)^2 - (a^1)^2 - (a^2)^2 - (a^3)^2$$

En effet, la matrice de la transformation (3,23) est le produit de la matrice (3,13) écrite avec U orthogonale, par la matrice écrite en (3,9). Les deux transformations correspondantes laissent d'une manière évidente la quantité (3,22) invariante.

On a donc bien un nombre associé d'une manière intrinsèque au quadrivecteur \vec{a} . Posons maintenant :

$$\vec{a} \bullet \vec{b} = \frac{1}{2} [(\vec{a} + \vec{b})^2 - \vec{a}^2 - \vec{b}^2]$$

On a bien une quantité invariante, tous les termes du membre de droite étant invariants. Un calcul simple montre que :

$$\vec{a} \bullet \vec{b} = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3 \quad (3, 24)$$

$$\vec{a} \bullet \vec{b} = a^0 b^0 - \mathbf{a} \mathbf{b} \quad (3, 25)$$

avec :

$$\mathbf{a} = \begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix} = (a^i)$$

La relation (3,25) permet de définir le produit scalaire lorsque le repère d'espace n'est pas orthonormé.

Il est facile de vérifier que $\vec{a} \bullet \vec{b}$ vérifie les axiomes d'un produit scalaire :

$$\vec{a} \bullet (\lambda \vec{b}) = (\lambda \vec{a}) \bullet \vec{b} = \lambda (\vec{a} \bullet \vec{b}) \quad (3, 26)$$

$$\vec{a} \bullet \vec{b} = \vec{b} \bullet \vec{a} \quad (3, 27)$$

$$\vec{a} \bullet (\vec{c} + \vec{d}) = \vec{a} \bullet \vec{c} + \vec{a} \bullet \vec{d} \quad (3, 28)$$

La seule différence avec un espace euclidien est que $\vec{a}^2 = \vec{a} \bullet \vec{a}$ peut être négatif ou nul avec $\vec{a} \neq \vec{0}$, en particulier $\vec{e}_i^2 = -1$ (dans un espace euclidien, un tel nombre est strictement positif pour un vecteur non nul); on dit qu'on a un espace *pseudo-euclidien*.

\vec{a}^2 est le *carré scalaire* de \vec{a} ; d'où la justification du symbole 2 dans s^2 car : $s^2 = (\Delta \vec{x})^2$; avec $\mathcal{B} = \mathcal{E}$ et $\mathcal{A} = \mathcal{O}$. Nous appellerons, comme dans le cas des espaces euclidiens *produit scalaire* des deux vecteurs \vec{a} et \vec{b} le nombre $\vec{a} \bullet \vec{b}$.

La base $\{\vec{e}_\alpha\}$ est telle que :

$$\vec{e}_\alpha \bullet \vec{e}_\beta = \pm \delta_{\alpha\beta} \quad (3, 29)$$

C'est ce que les mathématiciens appellent une *base type* dans un espace pseudo-euclidien. Cette notion généralise celle de base orthonormée en espace euclidien.

La mesure du temps dans un référentiel galiléen donne automatiquement un axe des temps dans l'espace de MINKOWSKI de vecteur de base normé ($\vec{e}_0 \bullet \vec{e}_0 = 1$) et perpendiculaire aux axes d'espace.

10. Conclusion. - Nous avons donc construit un espace vectoriel à quatre dimensions, l'espace de MINKOWSKI muni d'un produit scalaire, et ayant ainsi la structure d'espace pseudo-euclidien. À deux évènements quelconques \mathcal{A} et \mathcal{B} de l'espace-temps de coordonnées a^α et b^β on associe un vecteur unique de l'espace de MINKOWSKI, le vecteur déplacement :

$$\overrightarrow{\mathcal{AB}} = (b^\alpha - a^\alpha)\vec{e}_\alpha \quad (3, 30)$$

$$\overrightarrow{\mathcal{AB}} = \Delta x^\alpha \vec{e}_\alpha = \Delta \vec{x} \quad (3, 31)$$

D'autre part, il est facile de vérifier que $\overrightarrow{\mathcal{AB}} = -\overrightarrow{\mathcal{BA}}$ et que $\overrightarrow{\mathcal{AB}} + \overrightarrow{\mathcal{BC}} = \overrightarrow{\mathcal{AC}}$; enfin, quel que soit le quadrivecteur \vec{V} , il existe un évènement unique \mathcal{M} tel que $\overrightarrow{\mathcal{OM}} = \vec{V}$. Ces relations correspondent aux axiomes d'un espace affine sur un espace vectoriel.

$\overrightarrow{\mathcal{AB}}^2 = \Delta \vec{x}^2 = \Delta s^2$ est appelé *l'intervalle* entre les deux évènements \mathcal{A} et \mathcal{B} . L'intervalle généralise la notion de distance d entre deux points de l'espace (espace affine euclidien).

$$\Delta \vec{x} = (\Delta x^\alpha) = \begin{pmatrix} \Delta x^0 = C\Delta t \\ \Delta x^1 = \Delta x \\ \Delta x^2 = \Delta y \\ \Delta x^3 = \Delta z \end{pmatrix} \quad (3, 32)$$

Il vient donc, lorsque le repère d'espace est orthonormé :

$$\Delta \vec{x}^2 = C^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 \quad (3, 33)$$

Si il est quelconque, on a encore :

$$\Delta \vec{x}^2 = C^2t^2 - d^2 \quad (3, 34)$$

d est la distance spatiale des deux évènements.

Lorsque $\Delta s^2 > 0$ on parle d'intervalle du genre *temps*; lorsque $\Delta s^2 < 0$, on parle d'intervalle du genre *espace*; et lorsque $\Delta s^2 = 0$ on parle d'intervalle du genre *lumière*. Le même langage est utilisé pour le carré scalaire \vec{a}^2 : $\vec{a}^2 > 0 \Rightarrow \vec{a}$ quadrivecteur du genre temps. Cette notion est intrinsèque. Lorsque l'intervalle est du genre espace, il existe un référentiel galiléen dans lequel les deux évènements sont simultanés en deux endroits différents. Lorsque l'intervalle est du genre temps, il existe un référentiel galiléen dans lequel les deux évènements ont lieu au même endroit à des instants différents.

En conclusion, l'espace-temps a la structure d'un espace affine sur un espace vectoriel pseudo-euclidien à quatre dimensions, l'espace vectoriel des quadrivecteurs, ou espace de MINKOWSKI . La Relativité restreinte, qui correspond à cette structure, lie ainsi physiquement l'espace et le temps.

Précisons maintenant la terminologie que nous utiliserons par la suite : Un référentiel galiléen, donnée d'un repère d'espace (à priori quelconque, non nécessairement orthonormé) et d'un temps correspond à la donnée d'un *repère de l'espace-temps*. Les coordonnées correspondantes sont appelées *coordonnées galiléennes*. Lorsque le repère d'espace est orthonormé, et lorsqu'on mesure le temps avec la coordonnée : $x^0 = Ct$, Les équations (3,29) sont vérifiées; on a un *repère type de l'espace-temps*. Nous dirons alors qu'on a des *coordonnées galiléennes type*. Ceci généralise les repères orthonormés de l'espace euclidien et les coordonnées correspondantes au cas de l'espace-temps de la Relativité restreinte.

Les coordonnées galiléennes sont ce qu'on appelle d'une manière générale des *coordonnées rectilignes* ou *cartésiennes* (voir §3 chapitre 9). Les coordonnées galiléennes types sont ce qu'on appelle des *coordonnées rectilignes* ou *cartésiennes types*, ou *coordonnées types*. Elles ont pour équivalent en espace euclidien les *coordonnées orthonormées* (obtenues avec un repère orthonormé). Notons, d'après (3,34) .ue l'axe des temps est toujours orthogonal aux axes d'espace (voir également la dernière remarque du § 8).

L'espace-temps de la Relativité restreinte, espace affine sur l'espace de MINKOWSKI sera appelé *l'espace-temps plat*. ou *espace-temps plat pseudo-euclidien*. Nous verrons en effet au § 8 du chapitre 11 qu'un tel espace a une courbure nulle, donc est plat, comme le plan, en opposition à la surface d'une sphère par exemple.

Ceci sera également en opposition avec l'espace-temps de la Relativité générale dont nous verrons qu'il possède une courbure non nulle. Dans la Théorie de la relativité restreinte, nous avons un espace-temps plat global. En relativité générale, nous verrons que nous avons des espaces-temps plats locaux dans des régions suffisamment petites de l'espace-temps, dans lesquelles la Relativité restreinte s'applique (voir § 2 chapitre 12).

EXERCICE

3.1 $\bar{\mathcal{R}}$ est animé de la vitesse v_r par rapport à $\bar{\mathcal{R}}$, et $\bar{\mathcal{R}}$ est animé de la vitesse v_e par rapport à \mathcal{R} . v_e et v_r sont parallèles aux axes x , \bar{x} , et $\bar{\bar{x}}$.

1. Exprimez la formule de composition de ces vitesses colinéaires en utilisant l'équation (3,2) mise sous la forme matricielle correspondant à (3,9).
2. Que donne la formule pour $v_r \ll C$ et $v_e \ll C$; et pour $v_e \simeq C$?