

Chapitre 3

Eléments de Relativité restreinte

3.1 Bases de la Relativité restreinte

La théorie de la Relativité *restreinte* s'est développée à partir de l'observation expérimentale indéniable¹ que c , la célérité de la lumière dans le vide, est la même dans tous les référentiels dits galiléens. Cela a entraîné l'extension aux phénomènes électromagnétiques du concept d'*invariance galiléenne*, bien connu en Mécanique traditionnelle, et, partant, l'élargissement de la notion même de référentiel galiléen. Ceci a débouché sur l'émergence d'un principe général de relativité selon lequel *les lois de la Physique doivent s'exprimer de la même façon dans tous les référentiels galiléens*. Autrement dit, les formules exprimant les lois physiques doivent présenter une *covariance relativiste*. Ce principe a été initialement énoncé pour les référentiels galiléens seulement, conduisant à une théorie relativiste dite *restreinte*. Son extension à tous les référentiels fut l'objet de la Relativité *générale*.

Sous l'impulsion déterminée d'Einstein, concepteur de cette théorie, s'est parallèlement engagée une réflexion profonde sur notre perception spatio-temporelle des événements ayant lieu dans notre univers, provoquant une remise en cause de l'analyse traditionnelle de leur observation. Il en est résulté une présentation d'un protocole logique et irréfutable de la façon de procéder, d'une part, à la synchronisation des horloges et, d'autre part, à la comparaison des résultats d'une même observation obtenus dans des référentiels différents. Dans celle-ci, qui repose sur le principe que c est une constante universelle, chaque processus de comparaison fait intervenir un échange d'information par signaux électromagnétiques (lumineux en l'occurrence). A cet égard, on peut dire que la Relativité est une théorie élaborée de la mesure utilisant largement l'Electromagnétisme comme moyen d'information, précisément parce que la vitesse de propagation dans le vide des ondes électromagnétiques (dont fait partie la lumière) est la même pour tout observateur galiléen et que, de surcroît, la plupart des capteurs d'informations imaginés par la science, ou mis à notre disposition par la nature, ont pour base un phénomène électromagnétique (à commencer par la vision). De ce point de vue, Relativité et Electromagnétisme sont inséparables et ceci explique l'omniprésence de la constante c dans les équations de la Relativité²

De la théorie relativiste ont émergé deux autres principes fondamentaux. Le premier concerne la vitesse de propagation de nos moyens d'informations, ou celle de tout autre objet matériel : aucun signal, aucun objet matériel ne peut se propager avec une vitesse supérieure à c , vitesse de la lumière dans le vide³. A l'opposé de ce qu'il était admis en théorie classique, il n'existe pas de moyen d'information

1. Il s'agit des célèbres expériences de Michelson et Morley, menées de 1881 à 1887.

2. Il est difficile de faire un choix parmi la multitude d'excellents ouvrages traitant de la Relativité. Citons néanmoins les suivants : W. G. V. Rosser, "An Introduction to The Theory of Relativity", London Butterworths edts, third impression (revised), 1971 ; L. Landau et E. Lifchitz, "Théorie du Champ", Ed. Mir, Moscou, 1966 ; V. Ougarov, "Théorie de la Relativité Restreinte", Ed. Mir, Moscou, 1974 ; P. Lorrain et D.R. Corson, "Champs et Ondes Electromagnétiques", Librairie A. Colin, Paris, 1979, Chap. 5 et 6 ; et, sur l'engrenage des idées en Physique jusqu'au début du 20ème siècle, A. March, "La Physique Moderne et ses Théories", Ed. Gallimard, Paris, 1965.

3. Il est utile de préciser que la limite est bien la vitesse de la lumière *dans le vide*, car dans un milieu matériel

pouvant se propager à vitesse infinie. Le second est lié au premier et concerne la *causalité* entre deux événements. Si deux événements ne peuvent être joints par un signal quelconque de vitesse inférieure ou égale à c , ils sont sans lien causal direct, l'un ne peut être la cause de l'autre.

3.1.1 Synchronisation des horloges

Le problème est le suivant. Afin de repérer temporellement des événements, un observateur quelconque O doit se définir une *chronologie*. Celle-ci repose sur un mécanisme ayant une apparente périodicité dont l'observateur vérifie la régularité avec une précision acceptable. Le repérage temporel consiste alors à relever l'état dans lequel se trouve le mécanisme de référence dans son évolution lorsque se produit un événement, et à attribuer une date à cette coïncidence⁴. Les horloges sont construites sur ce principe de base. L'observateur dispose d'une batterie d'horloges toutes rigoureusement identiques qu'il va chercher à disséminer dans son propre repère d'espace pour dater les événements qui s'y produisent. Ce transport d'horloges sera confié à des collaborateurs qui emmèneront chacun une horloge depuis la position de référence de O jusqu'à un point donné du repère, après que toutes les horloges aient été réglées en O sur une même date origine. Deux questions importantes se posent alors. La première concerne l'immobilité des horloges par rapport à O lorsqu'elles auraient atteint ce qui est censé être leurs positions finales respectives : comment O peut-il être assuré de cette immobilité ? La seconde concerne le transport des horloges : comment O peut-il être assuré que ce transport n'a pas modifié la chronologie de chaque horloge ? Il est crucial de répondre à ces questions de façon irréprochable pour pouvoir faire état d'une chronologie commune à toutes les horloges d'un même repère. Les réponses seront bien sûr apportées par un transport d'information entre O et ses collaborateurs éloignés au moyen de signaux électromagnétiques.

En effet, l'invariance galiléenne de la vitesse de propagation des ondes électromagnétiques dans le vide ayant été dûment constatée, il devient évident que tous les observateurs galiléens devront utiliser ces ondes pour transmettre leurs informations. En effet, puisqu'il s'appuie sur une donnée physique universellement reconnue, ce procédé est finalement le seul auquel on puisse reconnaître une valeur universelle incontestable.

Décrivons maintenant ce processus de synchronisation. Initialement, l'observateur O et son collaborateur M préparent l'expérience au même point. Chacun est muni d'une horloge d'un type donné et d'un miroir. Les horloges sont réglées sur une même date conventionnellement prise comme date origine $t = 0$. A ce moment précis, M s'éloigne de O le long d'un axe $x'Ox$ à la vitesse constante $u > 0$ ⁵. Supposons que lorsque l'horloge de O indique la date t_0 , O envoie vers M un signal lumineux⁶. Comme la vitesse c de ce signal est toujours plus grande que u , pour O , il atteindra le miroir de M à la date t_1 telle que

$$c(t_1 - t_0) = ut_1, \quad \text{soit} \quad t_1 = \frac{t_0}{1 - \frac{u}{c}} \quad (3.1)$$

Nous supposons que M a orienté son miroir de telle sorte que, par réflexion, un signal de retour revienne vers O . Nous admettons que cette réflexion s'effectue instantanément. Lorsque le signal de retour parvient en O , ce dernier note la date t_2 qu'il évalue comme

la lumière se propage avec une vitesse inférieure. Dans un milieu matériel, une particule massive peut ainsi se propager plus vite que la lumière, ce qui donne lieu à un phénomène appelé *effet Čerenkov*.

4. Etablir une chronologie d'une évolution consiste donc à la comparer à une évolution de référence.

5. Nous écartons ici d'éventuelles critiques sur le fait que la vitesse de M ne peut passer instantanément de zéro à la valeur finie u . Pallier cette difficulté par un calcul plus complet n'apporterait rien de plus au raisonnement qui suit.

6. Ce pourrait être aussi bien un signal électromagnétique d'un autre domaine de fréquences que le domaine visible. Cependant, la physique des phénomènes d'émission, de réflexion ou de réception de signaux s'interprète plus aisément lorsqu'ils sont du domaine hertzien ou du domaine visible. C'est pourquoi la théorie fait volontiers usage de signaux lumineux dont la manipulation est plus familière.

$$t_2 = t_1 + \frac{ut_1}{c}, \quad \text{soit} \quad t_2 = t_0 \frac{1 + \frac{u}{c}}{1 - \frac{u}{c}} \quad (3.2)$$

Si O émet vers M des signaux se succédant avec une période T_E , il recevra les signaux de retour correspondants avec la période

$$T_R = T_E \frac{1 + \frac{u}{c}}{1 - \frac{u}{c}} \quad (3.3)$$

Tant que M est en mouvement par rapport à O , la période de réception observée par O est donc plus grande que la période d'émission. L'observateur M , qui voit O s'éloigner de lui, également à la vitesse u ⁷, peut d'ailleurs procéder lui-même à l'envoi de signaux vers le miroir porté par O et observer lui aussi une période de réception des signaux de retour plus grande que la période d'émission. Dans le référentiel de O comme dans celui de M , les deux périodes ne sont égales que si et seulement si O et M sont immobiles l'un par rapport à l'autre. C'est donc l'égalité des périodes qui permet à O et M d'affirmer qu'ils sont bien immobiles l'un par rapport à l'autre.

Ayant constaté leur immobilité relative, O et M peuvent savoir de quelle distance L ils sont séparés en s'envoyant l'un vers l'autre un signal. En effet, il leur suffira de noter le laps de temps Δt séparant l'émission du signal de la réception du signal de retour obtenu après réflexion sur le miroir du comparse : comme la propagation se fait à la vitesse c , la distance couverte par le signal pendant ce laps de temps est $c\Delta t = 2L$, d'où L .

Vient alors la synchronisation proprement dite. L'observateur O prévient M (par signaux) qu'il va envoyer un signal qui, pour lui, correspondra à la date origine $t = 0$. Lorsque M le reçoit, si son horloge indique la date t' , il devra régler celle-ci de telle sorte que la date $t' - L/c$ corresponde aussi à sa nouvelle date origine. Par la suite, en comparant leurs observations au moyen de signaux, O et M vérifieront que les temps donnés par leurs horloges respectives s'écoulent de la même manière.

Par ce procédé, on peut accorder toutes les horloges d'un repère donné. On constitue ainsi un référentiel qui apparaît alors comme un "tapis d'horloges", immobiles les unes par rapport et synchronisées comme ci-dessus.

3.1.2 Comparaison de mesures de longueurs

Envisageons deux observateurs galiléens O et O' se déplaçant l'un par rapport à l'autre le long d'un axe commun Δ . Nous supposons que leurs horloges respectives auront été réglées de telle sorte que leur rencontre corresponde, pour l'un comme pour l'autre, à la date origine. Chacun se sera muni de miroirs lui permettant de renvoyer vers l'autre les signaux lumineux que ce dernier aura émis vers lui. Une possibilité pour O d'analyser le mouvement de O' sans faire appel à des observateurs liés à son référentiel est la suivante.

L'axe Δ sera choisi comme axe $x'Ox$, orienté dans le sens du mouvement de O' par rapport à O . Vue depuis O , l'abscisse de O' est donc croissante, et nulle pour $t = 0$. A la date $t_1 > 0$, O envoie un signal lumineux vers O' qui se trouve alors à l'abscisse x_1 . Le signal parvient à O' à la date t_2 alors qu'il se trouve à l'abscisse x_2 et l'on a

$$c(t_2 - t_1) = x_2$$

Grâce à son miroir, O' renvoie instantanément un signal qui parvient à O à la date t_3 telle que

$$c(t_3 - t_2) = x_2$$

7. Selon le principe de relativité, si $u > 0$ est la vitesse relative de M par rapport à O , la vitesse de O par rapport à M ne peut qu'être égale, en valeur absolue, à u .

et l'on a donc

$$t_3 - t_1 = \frac{2x_2}{c} = \frac{2}{c} \int_0^{t_2} v(t) dt$$

où $v(t)$ est la vitesse instantanée de O' dans le référentiel de O . Si O émet un nouveau signal vers O' à la date $t_1 + dt_1$, celui-ci sera reçu par O' à la date $t_2 + dt_2$ et O recevra un signal de retour à la date $t_3 + dt_3$. On a

$$dt_3 = dt_2 + \frac{v(t_2)}{c} dt_2, \quad \text{et} \quad dt_2 = dt_1 + \frac{v(t_2)}{c} dt_1$$

d'où, avec $v_2 = v(t_2)$,

$$\frac{dt_3}{dt_1} = \frac{c + v_2}{c - v_2} \tag{3.4}$$

qui est une quantité constante si la vitesse de O' est constante. Ainsi, en traçant une courbe de variations de t_3 en fonction de t_1 , O peut vérifier que O' a bien un mouvement rectiligne uniforme par rapport à lui. Nous noterons encore u la vitesse de O' par rapport à O . Par ce même procédé, O' peut également s'assurer que O a bien une vitesse constante par rapport à lui. Les mouvements des deux observateurs O et O' l'un par rapport à l'autre étant *relatifs*, il est clair que O doit avoir la vitesse algébrique $-u$ par rapport à O' .

Supposons maintenant que l'observateur O' soit muni d'une règle qu'il dispose perpendiculairement à l'axe Δ , en se plaçant en son milieu. Cette règle, dont on note A' et B' les extrémités, est maintenue au repos dans le référentiel de O' . Par raison de symétrie, les positions respectives occupées par A' et B' telles qu'elles sont vues au cours du temps par O restent symétriques par rapport à Δ ⁸.

La mesure de la longueur d'une règle en mouvement par rapport à un référentiel donné se doit d'être faite *instantanément* dans ce référentiel. Pour la règle ici en question, cette opération peut être menée de la façon suivante. A une date donnée, O' envoie simultanément deux signaux lumineux, l'un vers A' et l'autre vers B' . Lorsqu'ils atteignent A' et B' , ces signaux déclenchent instantanément l'émission de signaux lumineux dans toutes les directions, depuis A' et B' , respectivement. De par la symétrie évoquée plus haut, et compte tenu de l'invariance de la vitesse de la lumière, il est évident que les nouveaux signaux atteindront en même temps O' et O , mais à des dates différentes. Cependant, l'important ici est que O comme O' seront d'accord pour dire que l'émission de signaux par A' et B' s'est faite *simultanément*. Pour l'un comme pour l'autre l'occurrence simultanée de ces deux événements revient à une mesure de la distance $A'B'$. Supposons que les résultats des mesures faites par O et O' conduisent à une distance L_O mesurée par O plus grande que celle, $L_{O'}$, mesurée par O' . Mais, pour O' , cette expérience peut aussi être envisagée comme la mesure de la distance séparant deux points donnés A et B du référentiel de O et coïncidant avec A' et B' au moment de l'émission des signaux par A' et B' . D'après le principe de relativité, O' devrait alors conclure que la distance $L_{O'}$ qu'il mesure entre ces deux points qui pour lui sont en mouvement, est certainement plus grande que la distance L_O mesurée dans le référentiel de O dans lequel ils sont au repos. La compatibilité de ces deux points de vue impose donc que l'on ait égalité des mesures :

$$L_O = L_{O'} \tag{3.5}$$

autrement dit, *une règle disposée perpendiculairement à la vitesse relative de deux référentiels galiléens est vue avec la même longueur dans ces deux référentiels, longueur qui est aussi celle de la règle dans le référentiel où elle est au repos.*

Considérons ensuite le dispositif suivant. L'observateur O' a placé sur son axe $O'y'$ et perpendiculairement à celui-ci un miroir M , à la distance h de O' . Lorsque O' passe en O à la date $t' = 0$, il émet un signal lumineux en direction du miroir (fig 3.1). Pour O' , le signal atteint le miroir à la date $t'_1 = h/c$ (fig 3.2). Une réflexion de ce signal s'opère instantanément et le signal de retour parvient en

8. Et, pour O , la règle reste bien sûr perpendiculaire en O' à cet axe.

O' à la date $t'_2 = t'_1 + h/c = 2h/c$ (fig 3.3). Toujours pour O' , la propagation du signal émis et celle du signal de retour se sont effectuées le long de l'axe $O'y'$. Il en va tout autrement pour l'observateur O qui voit le repère de O' filer parallèlement à l'axe Ox à la vitesse u .

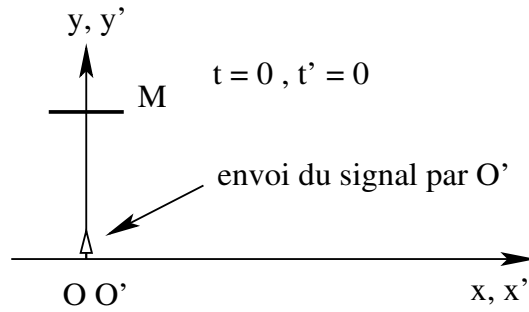


Figure 3.1

Pour O , le signal lumineux qui a été émis par O' lors de son passage en O à la date $t = 0$ se propage obliquement comme indiqué à la figure 3.2. Il atteint le miroir à la date $t_1 = d/c$ alors que O' est à la distance $x_1 = ut_1$ de O .

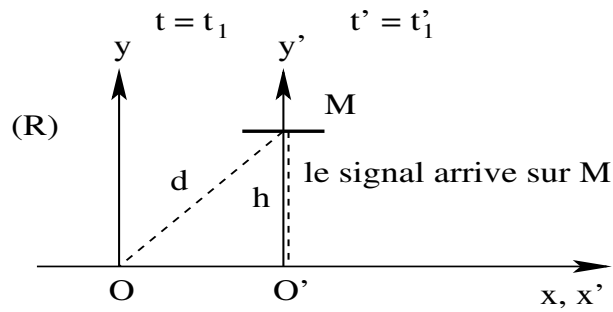


Figure 3.2

Toujours pour O , la réflexion sur le miroir se fait obliquement, et le signal de retour rejoint O' à la date $t_2 = t_1 + d/c = 2d/c$ (fig 3.3), alors que O' est à la distance $x_2 = x_1 + u(t_2 - t_1) = 2ut_1$.

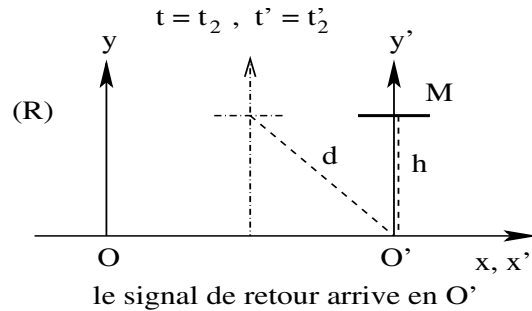


Figure 3.3

D'après ce qui précède, la distance entre O' et le miroir M est la même, soit h , pour O' et O . On peut donc écrire

$$d^2 = h^2 + (ut_1)^2 = c^2 t_1^2 = c^2 t_1'^2 + u^2 t_1'^2, \text{ soit } t_1' = t_1 \sqrt{1 - \frac{u^2}{c^2}}$$

d'où

$$t_2' = t_2 \sqrt{1 - \frac{u^2}{c^2}} \tag{3.6}$$

Or, t_2' est le laps de temps qui sépare deux événements ayant eu lieu *au même endroit* dans le référentiel de O' , à savoir, l'émission d'un signal par O' , et la réception d'un signal de retour au même endroit O' : il s'agit d'une *durée propre* à l'horloge attachée à O' . La relation ci-dessus montre que pour O , le laps de temps t_2 séparant ces deux événements est *plus grand* que t_2' . Il s'agit là d'un résultat général : la durée séparant deux événements qui se produisent au même endroit dans un référentiel galiléen donné est mesurée plus grande dans tout autre référentiel galiléen. On dit qu'il y a une *dilatation des durées* par rapport à une *durée propre*⁹.

Une autre conséquence est la suivante. Entre les deux événements précités, l'observateur O' aura parcouru dans le référentiel de O la distance $L = ut_2$, tandis que dans le référentiel de O' on aura vu l'observateur O parcourir la distance $L' = ut_2'$ ¹⁰. D'après le résultat précédent, la relation entre ces deux distances est donc

$$L' = L \sqrt{1 - \frac{u^2}{c^2}} \tag{3.7}$$

Or, pour O , la distance L peut être considérée comme étant la longueur d'une règle au repos dans son référentiel. Cette longueur est donc une "longueur propre" pour O . La relation ci-dessus montre que pour O' la mesure L' de cette règle est plus petite que la longueur propre L : il y a "contraction des longueurs". C'est aussi un résultat général : une règle au repos dans un référentiel galiléen est mesurée plus courte par tout autre observateur galiléen qui se meut parallèlement à la règle par rapport à ce référentiel.

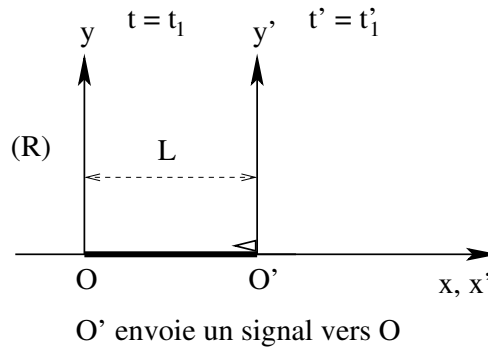


Figure 3.4

Un procédé plus immédiat de mesure de longueur à l'aide de signaux lumineux est le suivant. L'observateur O a disposé le long de son axe Ox une règle de longueur L comme indiqué à la figure

9. On retiendra bien que ce résultat ne concerne que des événements ayant lieu au même endroit dans un référentiel donné. Il explique notamment le fait que la durée de vie de particules instables en mouvement dans un référentiel est plus grande que celle mesurée dans un référentiel où elles sont au repos. Ce phénomène a été confirmé dans les expériences de détection de muons de haute énergie (voir les articles suivants : B. Rossi, D. B. Hall, Physical Review, 59, 223 (1941) ; D.H. Frisch, J.H. Smith, Am. J. Phys., 31, 342-355 (1963)). Pour que l'effet puisse être mis en évidence, la vitesse relative doit être très élevée : même pour des engins spatiaux dont la vitesse nous semble déjà vertigineuse (elle peut atteindre $10 \text{ km/s} = 36000 \text{ km/h} \approx 3 \cdot 10^{-5} c$), l'effet reste indécélable. Ceci permet de comprendre pourquoi des vérifications claires de la théorie sont couramment apportées par la physique des particules où les objets étudiés ont des vitesses voisines de c .

10. Rappelons que O et O' ont la même vitesse relative (en valeur absolue) l'un par rapport à l'autre.

3.4. La règle reste fixe dans ce référentiel et O en est l'une des extrémités. Lorsque O' passe en O , les observateurs conviennent que cet événement correspond à $t = 0$ et $t' = 0$. Lorsque O' passe à l'autre extrémité de la règle, il envoie un signal lumineux vers O . Cet événement correspond à la date $t_1 = L/u$ pour O et à la date t'_1 pour O' . Pour O' , la longueur de la règle sera $L' = ut'_1$. Pour O , le signal, pour lui parvenir, doit parcourir la distance L (fig. 3.5). Pour O' , il devra parcourir la distance L' . Le signal parvient en O à la date $t_2 = t_1 + L/c = \frac{L}{c} (1 + \frac{c}{u})$ pour O et à la date $t'_2 = t'_1 + L'/c = \frac{L'}{c} (1 + \frac{c}{u})$ pour O' . Mais pour O , l'événement "passage de O' en O " à la date $t = 0$ et l'événement "arrivée du signal en O " ont lieu au même endroit et le laps de temps t_2 qui les sépare est un temps propre pour O . On doit donc avoir

$$t'_2 = t_2 \sqrt{1 - \frac{u^2}{c^2}}$$

et par suite,

$$L' = L \sqrt{1 - \frac{u^2}{c^2}}$$

Il y a bien contraction des longueurs selon la direction du mouvement relatif des deux observateurs.

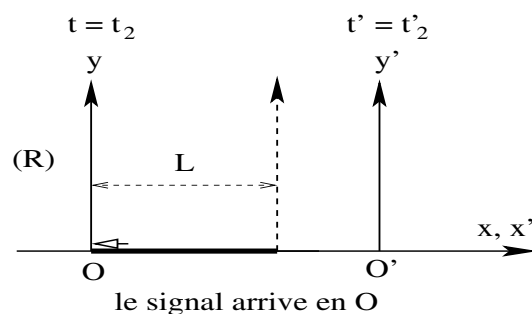


Figure 3.5

3.2 Changement de référentiels en Relativité Restreinte

3.2.1 Les transformations de Lorentz

Etablissons maintenant les formules qui permettent aux deux observateurs précédents O et O' de comparer leurs données concernant un événement donné (E), lequel sera visualisé, pour simplifier l'exposé, comme le passage d'un objet ponctuel en un point donné M et à une date donnée. Ces données concernent au premier chef les coordonnées spatio-temporelles, telles qu'elles sont mesurées pour ce même événement par l'un et l'autre observateur. Pour commencer, nous supposons que l'événement a lieu sur l'axe joignant O à O' . Pour l'observateur O , à la date t , l'objet est au point M d'abscisse x , alors que O' est à l'abscisse $x_{O'} = ut$. On peut écrire

$$x = ut + (\overline{O'M})_R \quad (3.8)$$

où l'indice "R" signifie que la valeur attribuée à la grandeur concernée est celle mesurée dans le référentiel R attaché à O , différente de celle mesurée par O' . En effet, O' voit l'événement se produire en une position, fixe par rapport à lui, définie par l'abscisse x' . Pour O' , cette abscisse x' est une longueur

propre. Pour O , en mouvement par rapport à O' , la distance $(\overline{O'M})_R$ n'est pas une longueur propre et doit prendre en compte la contraction des longueurs dont il a été question précédemment :

$$(\overline{O'M})_R = x' \sqrt{1 - \frac{u^2}{c^2}} \quad (3.9)$$

On en déduit ainsi la relation

$$x' = \frac{x - ut}{\sqrt{1 - \frac{u^2}{c^2}}} \quad (3.10)$$

Symétriquement, on aura¹¹

$$x = \frac{x' + ut'}{\sqrt{1 - \frac{u^2}{c^2}}} \quad (3.11)$$

A partir des deux dernières formules, il est facile d'obtenir t' en fonction de x et t :

$$t' = \frac{t - \frac{ux}{c^2}}{\sqrt{1 - \frac{u^2}{c^2}}} \quad (3.12)$$

Considérons ensuite le cas plus général d'un événement ayant lieu en un point quelconque de l'espace, c'est-à-dire pour lequel $y \neq 0$ et $z \neq 0$.

Nous avons vu que dans le passage du référentiel de O à celui de O' , les distances mesurées dans des directions perpendiculaires à l'axe suivant lequel s'effectue le mouvement relatif ne sont pas changées. On doit donc avoir $y' = y$ et $z' = z$.

En conclusion, les formules permettant de passer des données spatio-temporelles enregistrées dans le référentiel de O à celles enregistrées dans le référentiel de O' sont les suivantes

$$x' = \frac{x - ut}{\sqrt{1 - \frac{u^2}{c^2}}} \quad , \quad t' = \frac{t - \frac{ux}{c^2}}{\sqrt{1 - \frac{u^2}{c^2}}} \quad , \quad y' = y \quad , \quad z' = z \quad (3.13)$$

Il s'agit d'un cas particulier d'une *transformation de Lorentz*. Les transformations de Lorentz¹² définissent en Relativité Restreinte les changements de référentiels galiléens. Elles remplacent les transformations dites *de Galilée* utilisées en Mécanique non relativiste pour effectuer ces mêmes changements de référentiels. On retrouve d'ailleurs les transformations de Galilée comme cas limite des transformations de Lorentz lorsque $|u| \ll c$. En effet, si l'on fait tendre c vers l'infini dans les formules (3.13), on obtient

$$x' = x - ut \quad , \quad t' = t \quad , \quad y' = y \quad , \quad z' = z \quad (3.14)$$

où l'égalité $t' = t$ traduit l'hypothèse d'un temps "absolu" en Mécanique galiléenne, hypothèse remise en cause en Relativité par le principe qu'il n'existe aucun moyen physique de transmettre des informations à vitesse infinie.

Dans l'étude qui précède, nous avons supposé une certaine disposition des repères d'espace attachés aux deux observateurs galiléens O et O' . Notamment, les "lignes de vol" de O et O' sont identiques. Considérons alors un nouveau repère galiléen R'' ayant une origine O'' au repos dans le repère R' de O' et dont les axes sont encore parallèles à ceux de R ou R' . Notons X' , Y' et Z' les coordonnées

11. En notant que la vitesse de O par rapport à O' est $-u$.

12. Qui portent le nom de leur découvreur, H. A. Lorentz.

de O'' dans le référentiel R' . Dans R , les coordonnées X , Y et Z de O' dans R' se verront attribuer les valeurs

$$X = X' \sqrt{1 - \frac{u^2}{c^2}} \quad , \quad Y = Y' \quad , \quad Z = Z'$$

Un événement vu dans R à la date t au point de coordonnées x , y et z sera vu dans R'' à la date t'' au point de coordonnées x'' , y'' et z'' avec

$$x'' = \frac{x - X - ut}{\sqrt{1 - \frac{u^2}{c^2}}} \quad , \quad t'' = t' = \frac{t - \frac{ux}{c^2}}{\sqrt{1 - \frac{u^2}{c^2}}} \quad , \quad y'' = y - Y \quad , \quad z'' = z - Z \quad (3.15)$$

en supposant ici encore que les dates origines correspondent à la coïncidence de O et O' .

En outre, nous avons considéré que les axes des référentiels ont des directions communes. Les transformations de Lorentz entre référentiels dont les axes sont parallèles sont qualifiées de transformations de Lorentz "pures". Lorsque les axes de deux référentiels ne sont pas parallèles, le passage de l'un à l'autre nécessite d'effectuer une rotation d'axes. Une transformation de Lorentz plus générale impliquera donc une rotation, suivie d'une transformation de Lorentz pure.

3.2.2 Deux conséquences remarquables

♣ Relativité de la simultanéité

Supposons que dans le référentiel R de O on perçoive au même moment deux événements E_1 et E_2 en des points M_1 et M_2 dont la séparation spatiale est caractérisée par $\Delta x > 0$, Δy et Δz . Dans R , les deux événements sont déclarés être *simultanés*. D'après (3.13), les deux événements correspondants E'_1 et E'_2 vus dans le référentiel R' de O' *ne peuvent être simultanés*. En effet, dans ce référentiel, le laps de temps les séparant est non nul :

$$\Delta t' = -\frac{u}{c^2} \frac{\Delta x}{\sqrt{1 - \frac{u^2}{c^2}}} < 0 \quad (3.16)$$

Dans R' , la séparation spatiale entre E'_1 et E'_2 est telle que

$$\Delta x' = \frac{\Delta x}{\sqrt{1 - \frac{u^2}{c^2}}} \quad , \quad \Delta y' = \Delta y \quad , \quad \Delta z' = \Delta z \quad (3.17)$$

Dans la forme, cet ensemble d'événements est à rapprocher d'un processus de mesure de distance par O . Supposons en effet que O ait l'intention de mesurer la longueur d'une règle au repos dans le référentiel de O' , disposée selon l'axe des x . Pour O , un procédé immédiat pour effectuer la mesure de cette règle en mouvement consiste à pointer simultanément les deux extrémités de celle-ci. Dans R , il s'agit de deux événements simultanés, distants de Δx . Pour O' , la longueur $\Delta x'$ de la règle est une longueur propre. Entre ces deux mesures, on doit avoir un facteur de contraction, et l'on obtient bien

$$\Delta x = \Delta x' \sqrt{1 - \frac{u^2}{c^2}} \quad (3.18)$$

Pour O' , ces deux pointages n'ont pas lieu au même moment, mais sont séparés dans le temps d'une durée égale à

$$\Delta t' = -\frac{u}{c^2} \Delta x' \quad (3.19)$$

Le rapport

$$\left| \frac{\Delta x'}{\Delta t'} \right| = \frac{c^2}{u} > c \quad (3.20)$$

peut d'ailleurs être interprété par O' comme la vitesse de propagation d'une éventuelle action à distance qui aurait provoqué les deux événements, le premier événement perçu pouvant alors être considéré comme la cause du second. Ce rapport apparaît être plus grand que c . Comme aucun signal ne peut se propager plus vite que c , l'observateur O' en conclura qu'il n'y a aucune relation de cause à effet entre ces deux événements.

♣ Causalité

Si, dans un référentiel galiléen, un événement est perçu comme étant la cause d'un autre, cette relation de cause à effet reste vraie dans tout autre référentiel galiléen.

Supposons que dans le référentiel R un événement E_1 se produisant à la date t_1 soit reconnu comme étant la cause d'un événement E_2 apparu à la date $t_2 > t_1$. Dans R' , les événements E'_1 et E'_2 correspondant respectivement à E_1 et E_2 se produisent aux dates respectives

$$t'_1 = \frac{t_1 - \frac{ux_1}{c^2}}{\sqrt{1 - \frac{u^2}{c^2}}}, \quad t'_2 = \frac{t_2 - \frac{ux_2}{c^2}}{\sqrt{1 - \frac{u^2}{c^2}}} \quad (3.21)$$

et l'on a

$$t'_2 - t'_1 = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} \left[t_2 - t_1 - \frac{u}{c^2} (x_2 - x_1) \right] \quad (3.22)$$

Dans R' , E'_1 ne pourra être considéré comme la cause de E'_2 que si $t'_2 > t'_1$. Ceci est vérifié si $x_2 < x_1$. Si $x_2 > x_1$, on obtient la condition

$$w = \frac{x_2 - x_1}{t_2 - t_1} < \frac{c^2}{u} \quad (3.23)$$

Ici encore, ce rapport peut être interprété comme étant, dans R , la vitesse de propagation depuis le point où a lieu E_1 à celui où a lieu E_2 (en supposant $\Delta y = 0$, $\Delta z = 0$), de l'action causale provoquant E_2 à la suite de E_1 . Cette vitesse est certainement inférieure à c et l'on a bien $c^2 > wu$, soit $t'_2 > t'_1$. La relation de causalité est donc bien préservée dans le changement de référentiel galiléen.

3.2.3 Structure de groupe des transformations de Lorentz

La structure en groupe des changements de référentiels est intuitivement évidente. Nous allons préciser cette structure dans le cas des transformations de Lorentz pures introduites précédemment.

Posons

$$\cosh \chi = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}, \quad \sinh \chi = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} \frac{u}{c}, \quad x_0 = ct \quad (3.24)$$

A l'aide du paramètre χ , appelé *rapidité*, les relations (3.13) peuvent être réexprimées sous la forme

$$x'_0 = x_0 \cosh \chi - x \sinh \chi, \quad x' = x \cosh \chi - x_0 \sinh \chi, \quad y' = y, \quad z' = z \quad (3.25)$$

La transformation inverse correspond au paramètre $\chi' = -\chi$:

$$x_0 = x'_0 \cosh \chi + x' \sinh \chi \quad , \quad x = x' \cosh \chi + x'_0 \sinh \chi \quad , \quad y = y' \quad , \quad z = z' \quad (3.26)$$

Envisageons ensuite un troisième référentiel R_1 ayant des axes parallèles à ceux de R et de R' , et dont l'origine O_1 se déplace le long de l'axe commun Ox avec la vitesse u'_1 relativement à R' . Le changement de référentiel $R' \rightarrow R_1$ sera caractérisé par le paramètre χ'_1 tel que

$$x'_{10} = x'_0 \cosh \chi'_1 - x' \sinh \chi'_1 \quad , \quad x'_1 = x' \cosh \chi'_1 - x'_0 \sinh \chi'_1 \quad , \quad y_1 = y' \quad , \quad z_1 = z' \quad (3.27)$$

On obtient

$$\begin{aligned} x'_{10} &= x_0 \cosh(\chi'_1 + \chi) - x \sinh(\chi'_1 + \chi) \\ x'_1 &= x \cosh(\chi'_1 + \chi) - x_0 \sinh(\chi'_1 + \chi) \\ y_1 &= y \quad , \quad z_1 = z \end{aligned} \quad (3.28)$$

Ainsi, la composition des deux transformations de Lorentz est une transformation de Lorentz dont la rapidité est la somme algébrique de leurs rapidités. La structure de groupe est ici évidente : pour ces transformations de Lorentz particulières, on a affaire à un groupe abélien (commutatif). Ce fait peut être exprimé de façon plus évidente encore en utilisant une notation matricielle. Posons

$$X = \begin{pmatrix} x_0 \\ x \\ y \\ z \end{pmatrix} \quad , \quad \text{et} \quad X' = \begin{pmatrix} x'_0 \\ x' \\ y' \\ z' \end{pmatrix} \quad (3.29)$$

On a alors

$$X' = \Lambda_x(\chi)X \quad , \quad \text{avec} \quad \Lambda_x(\chi) = \begin{pmatrix} \cosh \chi & -\sinh \chi & 0 & 0 \\ -\sinh \chi & \cosh \chi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.30)$$

La matrice $\Lambda_x(\chi)$ décrivant la transformation de Lorentz peut être écrite sous la forme

$$\Lambda_x(\chi) = \cosh \chi - \sinh \chi T_x \quad , \quad \text{avec} \quad T_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (3.31)$$

et l'on a bien

$$\Lambda_x(\chi)\Lambda_x(\chi') = \Lambda_x(\chi')\Lambda_x(\chi) = \Lambda_x(\chi + \chi') \quad (3.32)$$

Il est d'ailleurs facile d'établir la relation exponentielle

$$\Lambda_x(\chi) = e^{-\chi T_x} \quad (3.33)$$

On notera que T_x joue le rôle de générateur pour les transformations de Lorentz s'effectuant le long de l'axe des x .

La forme des relations (3.25) permet également d'établir de façon simple une propriété fondamentale des transformations de Lorentz. On a en effet

$$x'_0 + x' = e^{-\chi} (x_0 + x) \quad , \quad x'_0 - x' = e^{\chi} (x_0 - x) \quad \text{donc} \quad (x'_0)^2 - (x')^2 = x_0^2 - x^2 \quad (3.34)$$

Introduisons alors la forme quadratique

$$X^2 = x_0^2 - x^2 - y^2 - z^2 \quad (3.35)$$

Les relations précédentes montrent que cette forme est *invariante* sous les transformations de Lorentz. Autrement dit, il s'agit d'un *invariant relativiste*. Notons tout de suite que la notation X^2 ne signifie nullement que cette forme soit toujours positive : elle peut être positive, négative, ou même nulle. Rappelons que les grandeurs x, y, z et x_0 représentent les coordonnées spatio-temporelles d'un événement, telles qu'elles sont enregistrées dans le référentiel d'un observateur O . Elles sont définies par rapport à celles, prises égales à zéro, d'un événement "origine" (il s'agissait alors de la coïncidence de O et O'). Plus généralement, l'écart spatio-temporel ΔX entre deux événements, mesuré dans un référentiel donné, sera défini par l'ensemble des écarts $\Delta x_0, \Delta x, \Delta y$ et Δz , et on appelle *intervalle* entre les deux événements la quantité

$$(\Delta X)^2 = (\Delta x_0)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 \quad (3.36)$$

Cet intervalle a la même valeur pour tous les observateurs galiléens. L'ensemble des transformations de Lorentz peut alors se définir comme un groupe continu de transformations qui laisse l'intervalle invariant. Ce groupe est appelé *groupe de Lorentz*. Ce groupe contient plus de transformations qu'il n'a été considéré jusqu'à présent. Il contient notamment comme sous-groupe le groupe des rotations spatiales, qui décrit l'éventuelle nécessité d'effectuer une rotation d'axes pour comparer les observations de deux observateurs dont les repères auraient des axes de différentes orientations (tout en gardant leur orientation relative fixée). Mais il contient aussi :

♣ les transformations de Lorentz $\Lambda_y(\chi)$ le long de l'axe des y , générées par

$$T_y = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (3.37)$$

♣ les transformations de Lorentz $\Lambda_z(\chi)$ le long de l'axe des z , générées par

$$T_z = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad (3.38)$$

Plus généralement, une transformation de Lorentz de rapidité χ effectuée le long d'un axe d'orientation définie par un vecteur unitaire \vec{n} prend la forme¹³

$$\Lambda_{\vec{n}}(\chi) = e^{-\chi \vec{n} \cdot \vec{T}} = \begin{pmatrix} \cosh \chi & -n_x \sinh \chi & -n_y \sinh \chi & -n_z \sinh \chi \\ -n_x \sinh \chi & 1 + n_x^2 (\cosh \chi - 1) & n_x n_y (\cosh \chi - 1) & n_x n_z (\cosh \chi - 1) \\ -n_y \sinh \chi & n_x n_y (\cosh \chi - 1) & 1 + n_y^2 (\cosh \chi - 1) & n_y n_z (\cosh \chi - 1) \\ -n_z \sinh \chi & n_x n_z (\cosh \chi - 1) & n_y n_z (\cosh \chi - 1) & 1 + n_z^2 (\cosh \chi - 1) \end{pmatrix} \quad (3.39)$$

où $\vec{T} = T_x \vec{e}_x + T_y \vec{e}_y + T_z \vec{e}_z$, $\vec{n} = n_x \vec{e}_x + n_y \vec{e}_y + n_z \vec{e}_z$, \vec{e}_x, \vec{e}_y et \vec{e}_z étant les vecteurs unitaires d'un repère galiléen de référence. Cette transformation est couramment appelée *boost*. Elle dépend de trois paramètres indépendants : χ et des deux angles définissant l'orientation du vecteur unitaire \vec{n} . Comme

13. A démontrer.

$$\left[\vec{n} \cdot \vec{T}, \vec{n}' \cdot \vec{T} \right] = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & n_x n'_y - n_y n'_x & n_x n'_z - n_z n'_x \\ 0 & n_y n'_x - n_x n'_y & 0 & n_y n'_z - n_z n'_y \\ 0 & n_z n'_x - n_x n'_z & n_z n'_y - n_y n'_z & 0 \end{pmatrix} \quad (3.40)$$

les transformations de Lorentz ne commutent généralement pas entre elles : le groupe de Lorentz n'est pas abélien¹⁴.

3.3 Les quadri-vecteurs et quadri-tenseurs

3.3.1 Bases d'espace-temps, tenseur métrique

En Relativité, le temps n'est pas une donnée absolue. Comme nous l'avons vu, données temporelles et données spatiales se mélangent dans les changements de référentiels. Ceci incite à envisager ceux-ci comme des transformations dans un espace à quatre dimensions, dont l'une est temporelle ($x_0 = ct$), les trois autres étant les trois dimensions spatiales ordinaires (x , y et z). Dans cet espace, un référentiel galiléen correspondra à un choix particulier d'une base de vecteurs. Une base de référence sera constituée par les vecteurs

$$e_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad e_x = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_y = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_z = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (3.41)$$

le premier vecteur e_0 correspondant à la base "temporelle" (chronologie). A un observateur P correspondra une base différente pouvant être notée $e_0(P)$, $e_x(P)$, $e_y(P)$, $e_z(P)$. A un événement donné \mathcal{E} , l'observateur galiléen P attribuera le vecteur

$$E = x_0 e_0(P) + x e_x(P) + y e_y(P) + z e_z(P) \quad (3.42)$$

A ce même événement un autre observateur galiléen P' attribuera le vecteur

$$E' = x'_0 e_0(P') + x' e_x(P') + y' e_y(P') + z' e_z(P') \quad (3.43)$$

Comme il s'agit du même événement, on doit avoir $E = E'$. Les coordonnées attribuées par P' sont liées à celles attribuées par P par une transformation de Lorentz Λ , ce que nous écrivons sous la forme condensée

$$x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu} \quad (3.44)$$

en adoptant la convention de sommation d'Einstein. Ici, les indices μ et ν prennent les valeurs 0 (pour la base temps), 1 pour x , 2 pour y et 3 pour z . La base de P' s'obtient donc à partir de celle de P au moyen de la transformation inverse :

$$e_{\mu}(P') = (\Lambda^{-1})^{\nu}_{\mu} e_{\nu}(P) \quad (3.45)$$

L'objet mathématique (3.42), qui représente un phénomène physique, est l'exemple type de ce qu'on appelle un quadri-vecteur (souvent écrit 4-vecteur). On peut définir un 4-vecteur comme un ensemble de quatre quantités physiques dont les mesures dans différents référentiels galiléens sont liées par les transformations de Lorentz permettant de passer d'un référentiel galiléen à un autre. Cette notion est donc bien liée au groupe de Lorentz, tout comme dans l'espace à trois dimensions, la notion de vecteur est liée au groupe des rotations dans cet espace. Il serait donc erroné de croire que tout vecteur mathématique de l'espace à quatre dimensions peut être considéré comme un 4-vecteur. Il ne mérite le

14. Une étude plus complète du groupe de Lorentz fait l'objet d'un chapitre ultérieur.

label “4-vecteur” que si dans un changement de base, ses composantes attribuées dans les deux bases sont effectivement liées par une transformation de Lorentz¹⁵. Par la suite, nous utiliserons la notation spéciale \underline{x} , déjà utilisée au chapitre 2, pour désigner, dans la mesure du possible, un 4-vecteur. La distinction est encore plus nette lorsque, s'appuyant sur la propriété des transformations de Lorentz de conserver l'intervalle (3.36), on munit l'espace quadri-dimensionnel de la *pseudo-norme*

$$\underline{x}^2 = x_0^2 - x^2 - y^2 - z^2 \quad (3.46)$$

qui est un invariant relativiste pour tout 4-vecteur \underline{x} de composantes (x_0, x, y, z) dans un référentiel donné. Le *pseudo-produit scalaire* de deux 4-vecteurs \underline{x} et \underline{x}' dont les composantes dans un référentiel donné sont (x_0, x, y, z) et (x'_0, x', y', z') respectivement, sera défini comme

$$\underline{x} \cdot \underline{x}' = x_0 x'_0 - x x' - y y' - z z' \quad (3.47)$$

Il s'agit encore d'un invariant relativiste si \underline{x} et \underline{x}' sont bien des 4-vecteurs. Malgré la notation “carré” dans (3.46), la pseudo-norme peut être positive, négative, voire nulle, selon le 4-vecteur considéré. On classe habituellement les 4-vecteurs en trois genres. Un 4-vecteur \underline{x} est dit :

- ♣ du genre “temps” si sa pseudo-norme est positive ;
- ♣ du genre “espace” si sa pseudo-norme est négative ;
- ♣ du genre “lumière”, ou isotrope, si sa pseudo-norme est nulle (cette circonstance peut exister pour des 4-vecteurs non nuls).

Cette classification est invariante relativiste : un 4-vecteur catalogué du genre espace dans un référentiel galiléen donné le sera dans tout autre référentiel galiléen. On voit ainsi que dans toute base de 4-vecteurs $e_0(P)$, $e_x(P)$, $e_y(P)$, $e_z(P)$ attachée à un observateur P , le 4-vecteur $e_0(P)$ est toujours du genre temps, alors que les 4-vecteurs $e_x(P)$, $e_y(P)$, $e_z(P)$ de la triade spatiale sont toujours du genre espace :

$$e_0(P)^2 = 1, \quad e_x(P)^2 = e_y(P)^2 = e_z(P)^2 = -1 \quad (3.48)$$

Le pseudo-produit scalaire et la pseudo-norme peuvent s'exprimer au moyen d'un *tenseur métrique* $g_{\mu\nu}$, possédant a priori 16 composantes mais dont finalement 4 seulement sont non nulles et qui sont

$$g_{00} = 1, \quad g_{xx} = g_{yy} = g_{zz} = -1 \quad (3.49)$$

Nous utiliserons le plus souvent les indices 1, 2, et 3 à la place des indices x , y et z respectivement. Ainsi, le tenseur métrique sera tel que

$$g_{00} = 1, \quad g_{0i} = g_{i0} = 0, \quad g_{ij} = -\delta_{ij} \quad \text{avec } i, j = 1, 2, 3 \quad (3.50)$$

Les composantes *contravariantes* x^0 , x^1 , x^2 et x^3 du 4-vecteur \underline{x} (attention, ici 1, 2 et 3 sont des indices “en haut”) seront les composantes usuelles : $x^0 = x_0$, $x^1 = x$, $x^2 = y$ et $x^3 = z$. Les composantes *covariantes* de ce 4-vecteur seront définies à l'aide du tenseur métrique par (avec la convention de sommation d'Einstein)

$$x_\mu = g_{\mu\nu} x^\nu \quad (3.51)$$

On a donc

$$x_0 = x^0, \quad x_i = -x^i, \quad i = 1, 2, 3 \quad (3.52)$$

15. Voir à ce sujet la discussion de la section 7 du chapitre 2.

Avec ces notations, le pseudo-produit scalaire et la pseudo-norme s'écrivent

$$\underline{x} \cdot \underline{x}' = g_{\mu\nu} x^\mu x'^\nu = x_\mu x'^\mu = x^\mu x'_\mu, \quad \underline{x}^2 = g_{\mu\nu} x^\mu x^\nu = x_\mu x^\mu \quad (3.53)$$

On passe des composantes covariantes aux composantes contravariantes au moyen du tenseur métrique inverse $g^{\mu\nu}$:

$$x^\mu = g^{\mu\nu} x_\nu \quad (3.54)$$

Comme il a été fait au chapitre 2, on montre que sous l'effet d'une transformation de Lorentz Λ , les composantes covariantes d'un 4-vecteur \underline{v} se transforment selon la matrice ${}^t\Lambda^{-1}$

$$v'_\mu = (\Lambda^{-1})^\nu_\mu v_\nu \quad (3.55)$$

et l'invariance du tenseur métrique sous les transformations de Lorentz s'exprime sous forme matricielle comme¹⁶

$$g = {}^t\Lambda g \Lambda \quad (3.56)$$

De cette relation on déduit notamment que

$$\det\Lambda = \pm 1 \quad (3.57)$$

Une base e_0, e_1, e_2 et e_3 de 4-vecteurs sera dite *orthonormée* au sens du pseudo-produit scalaire si

$$e_\mu \cdot e_\nu = g_{\mu\nu}, \quad \mu, \nu = 0, 1, 2, 3 \quad (3.58)$$

Une telle base est donc définie par 10 relations indépendantes. Comme l'ensemble des quatre 4-vecteurs comportait a priori 16 paramètres, on voit qu'une base dépend de 6 paramètres indépendants. C'est exactement la dimension de l'algèbre de Lie du groupe de Lorentz. La notion de 4-vecteur et du pseudo-produit scalaire étant bien comprises, nous omettrons désormais le terme "pseudo", et parlerons simplement de "produit scalaire" à la place de "pseudo-produit scalaire" et de "carré scalaire" à la place de "pseudo-norme".

Nous laissons au lecteur le soin de vérifier la relation de fermeture

$$(e_0)_\mu (e_0)_\nu - (e_1)_\mu (e_1)_\nu - (e_2)_\mu (e_2)_\nu - (e_3)_\mu (e_3)_\nu = g_{\mu\nu} \quad (3.59)$$

valable pour toute base "orthonormée", et de démontrer les théorèmes suivants :

- ♠ *Un 4-vecteur orthogonal à un 4-vecteur du genre temps est nécessairement du genre espace.*
- ♠ *Un 4-vecteur orthogonal à un vecteur du genre lumière est soit du genre espace, soit colinéaire à ce 4-vecteur.*

L'ensemble orthogonal à un 4-vecteur constitue ce qu'on appelle un *hyperplan*. Ainsi, l'hyperplan orthogonal à un 4-vecteur du genre temps est du genre espace, isomorphe à l'espace ordinaire à trois dimensions.

Dans l'espace des 4-vecteurs, les 4-vecteurs du genre lumière forment un cône, appelé *cône de lumière*. Il sépare cet espace en trois régions distinctes, conduisant ainsi à la classification suivante :

- 4-vecteur nul ;
- 4-vecteurs du genre espace ;

16. A démontrer.

- 4-vecteurs du genre temps futur pour lesquels $x_0 > 0$;
- 4-vecteurs du genre temps passé ($x_0 < 0$) ;
- 4-vecteurs du genre lumière futur ($x_0 > 0$) ;
- 4-vecteurs du genre lumière passé ($x_0 < 0$).

Un repère attaché à un observateur (repère de Lorentz) est dit *propre* si son 4-vecteur du genre temps pointe vers le demi-cône futur. Nous admettrons qu'il en sera ainsi pour tout repère de Lorentz.

Dans l'espace *ponctuel* à quatre dimensions, un observateur galiléen parcourt une *droite d'univers* Δ parallèle à son 4-vecteur du genre temps $e_0(P)$. Tout objet qui lui paraît au repos aura pour *courbe d'univers* une droite parallèle à Δ , indépendamment du choix de la triade de 4-vecteurs du genre espace associée à $e_0(P)$. Il en résulte que deux repères de Lorentz propres dont les 4-vecteurs du genre temps associés sont parallèles, sont *équivalents*¹⁷. Un observateur galiléen sera donc plutôt caractérisé par la donnée d'un 4-vecteur du genre temps futur unitaire e_0 (4-vecteur "libre"). Cette équivalence rend compte du fait qu'un observateur quelconque a toujours le choix de l'origine des dates, de l'origine de l'espace, et de la triade d'espace e_1, e_2 et e_3 .

Un observateur de Lorentz analysera un 4-vecteur quelconque \underline{v} en le projetant d'une part sur son 4-vecteur de base du genre temps e_0 et d'autre part dans l'hyperplan orthogonal à e_0 pour obtenir la décomposition

$$\underline{v} = (v_0, \vec{v})_{e_0}, \quad \text{avec } v_0 = \underline{v} \cdot e_0, \quad \vec{v} = \underline{v} - v_0 e_0 \quad (3.60)$$

qui est en fait une réduction du 4-vecteur \underline{v} par rapport au sous-groupe des rotations dans l'hyperplan orthogonal à e_0 , en une partie scalaire v_0 et un 3-vecteur \vec{v} . Il opérera des réductions analogues sur des tenseurs d'ordre quelconque.

3.3.2 Tenseurs d'espace-temps

Un tenseur d'espace-temps T d'ordre n est une entité mathématique que l'on peut caractériser par ses composantes contravariantes. Ces grandeurs, notées $T^{\mu_1 \mu_2 \dots \mu_n}$, comportent n indices $\mu_1, \mu_2, \dots, \mu_n$ prenant chacun les valeurs 0, 1, 2, 3, et, sous l'effet d'une transformation de Lorentz Λ , se transforment selon la loi

$$T'^{\nu_1 \nu_2 \dots \nu_n} = \Lambda^{\nu_1}_{\mu_1} \Lambda^{\nu_2}_{\mu_2} \dots \Lambda^{\nu_n}_{\mu_n} T^{\mu_1 \mu_2 \dots \mu_n} \quad (3.61)$$

Ce tenseur possède des composantes complètement *covariantes*, obtenues par *contraction* des composantes complètement contravariantes avec le tenseur métrique :

$$T_{\mu_1 \mu_2 \dots \mu_n} = g_{\mu_1 \nu_1} g_{\mu_2 \nu_2} \dots g_{\mu_n \nu_n} T^{\nu_1 \nu_2 \dots \nu_n} \quad (3.62)$$

et qui se transforment ainsi :

$$T'_{\mu_1 \mu_2 \dots \mu_n} = (\Lambda^{-1})^{\nu_1}_{\mu_1} (\Lambda^{-1})^{\nu_2}_{\mu_2} \dots (\Lambda^{-1})^{\nu_n}_{\mu_n} T_{\nu_1 \nu_2 \dots \nu_n} \quad (3.63)$$

Il possède aussi des composantes *mixtes* ayant des indices contravariants et des indices covariants, qui ont des lois de transformation mixtes (voir chapitre 2). Deux tenseurs particuliers ont la propriété exceptionnelle d'être invariants par transformation de Lorentz. Il s'agit tout d'abord du tenseur métrique. Puisque

$$g_{\mu\nu} = \Lambda^{\rho}_{\mu} \Lambda^{\sigma}_{\nu} g_{\rho\sigma} \quad (3.64)$$

17. Les classes d'équivalence de cette relation sont précisément les "observateurs galiléens".

les composantes de ce tenseur (ainsi que celles de son inverse) sont les mêmes dans tout référentiel galiléen. Le tenseur de *Levi-Civita*, noté $\epsilon_{\mu\nu\rho\sigma}$, possède aussi cette propriété. Il s'agit du tenseur à quatre indices, complètement antisymétrique par transposition de deux indices quelconques, et dont les composantes non nulles sont égales à ± 1 . On se rend compte facilement que seules les composantes avec 4 indices différents sont non nulles. On pose généralement

$$\epsilon_{0123} = -\epsilon^{0123} = +1 \quad (3.65)$$

On a ainsi

$$\epsilon_{0123} = \epsilon_{1230} = \epsilon_{2301} = \epsilon_{3012} = -\epsilon_{1023} = -\epsilon_{0321} , \text{ etc} \quad (3.66)$$

Comme

$$\epsilon_{\alpha\beta\gamma\delta} A_{\mu}^{\alpha} A_{\nu}^{\beta} A_{\rho}^{\gamma} A_{\sigma}^{\delta} = \det A \epsilon_{\mu\nu\rho\sigma} \quad (3.67)$$

pour toute matrice 4×4 A , et que $\det A = \pm 1$ si A est une transformation de Lorentz, ce tenseur est inchangé si la transformation de Lorentz laisse aussi invariante l'orientation de l'espace-temps, auquel cas son déterminant vaut $+1$ (ce que nous supposerons, au moins pour le moment)¹⁸.

On notera les formules suivantes

$$\begin{aligned} \epsilon_{\mu\nu\rho\sigma} \epsilon^{\alpha\beta\gamma\sigma} &= - \left[\delta_{\mu}^{\alpha} \delta_{\nu}^{\beta} \delta_{\rho}^{\gamma} + \delta_{\nu}^{\alpha} \delta_{\rho}^{\beta} \delta_{\mu}^{\gamma} + \delta_{\rho}^{\alpha} \delta_{\mu}^{\beta} \delta_{\nu}^{\gamma} - \delta_{\mu}^{\alpha} \delta_{\rho}^{\beta} \delta_{\nu}^{\gamma} - \delta_{\nu}^{\alpha} \delta_{\mu}^{\beta} \delta_{\rho}^{\gamma} - \delta_{\rho}^{\alpha} \delta_{\nu}^{\beta} \delta_{\mu}^{\gamma} \right] \\ \epsilon_{\mu\nu\rho\sigma} \epsilon^{\alpha\beta\rho\sigma} &= -2 \left[\delta_{\mu}^{\alpha} \delta_{\nu}^{\beta} - \delta_{\nu}^{\alpha} \delta_{\mu}^{\beta} \right] \\ \epsilon_{\mu\nu\rho\sigma} \epsilon^{\alpha\nu\rho\sigma} &= -6\delta_{\mu}^{\alpha} , \quad \epsilon_{\mu\nu\rho\sigma} \epsilon^{\mu\nu\rho\sigma} = -24 \end{aligned} \quad (3.68)$$

On sait que par des multiplications entre eux de tenseurs ou par le jeu des contractions d'indices, on peut construire de nouveaux tenseurs (voir chapitre 2). Considérons alors un tenseur d'espace-temps F antisymétrique, de composantes contravariantes $F^{\mu\nu}$. Par définition, son *tenseur dual* $*F$ a pour composantes

$$*F_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} F^{\rho\sigma} \quad (3.69)$$

En utilisant la seconde équation de (3.68), on trouve que le dual de ce tenseur dual est

$$**F = -F \quad (3.70)$$

Le *produit extérieur* $w = \underline{u} \wedge \underline{v}$ de deux 4-vecteurs distincts \underline{u} et \underline{v} est un tenseur antisymétrique d'ordre 2 de composantes covariantes $u_{\mu} v_{\nu} - u_{\nu} v_{\mu}$. Un tel tenseur est aussi qualifié de *tenseur plan*¹⁹. Tout vecteur c orthogonal à la fois à \underline{u} et \underline{v} est dit orthogonal à w car on a $c^{\mu} w_{\mu\nu} = 0$ et $w_{\mu\nu} c^{\nu} = 0$. On rencontre notamment des produits extérieurs et leurs tenseurs duaux dans les formules suivantes, qui généralisent des formules bien connues relatives aux bases de l'espace à trois dimensions.

Pour toute base orthonormée d'orientation directe e_0, e_1, e_2, e_3 , on a les relations²⁰

18. Le fait que le tenseur de Levi-Civita change de signe lorsque l'orientation de l'espace est inversée lui confère le qualificatif de *pseudo-tenseur*.

19. Un tenseur plan peut être représenté géométriquement par une surface plane orientée.

20. A vérifier!

$$(e_0)_\mu = \epsilon_{\mu\nu\rho\sigma} (e_1)^\nu (e_2)^\rho (e_3)^\sigma, \quad (e_1)_\mu = \epsilon_{\mu\nu\rho\sigma} (e_0)^\nu (e_2)^\rho (e_3)^\sigma \quad (3.71)$$

$$(e_2)_\mu = -\epsilon_{\mu\nu\rho\sigma} (e_0)^\nu (e_1)^\rho (e_3)^\sigma, \quad (e_3)_\mu = \epsilon_{\mu\nu\rho\sigma} (e_0)^\nu (e_1)^\rho (e_2)^\sigma$$

$$\epsilon_{\mu\nu\rho\sigma} (e_0)^\mu (e_1)^\nu (e_2)^\rho (e_3)^\sigma = 1 \quad (3.72)$$

et

$$\begin{aligned} (e_0)_\mu (e_1)_\nu - (e_0)_\nu (e_1)_\mu &= -\epsilon_{\mu\nu\rho\sigma} (e_2)^\rho (e_3)^\sigma \\ (e_0)_\mu (e_2)_\nu - (e_0)_\nu (e_2)_\mu &= \epsilon_{\mu\nu\rho\sigma} (e_1)^\rho (e_3)^\sigma \\ (e_0)_\mu (e_3)_\nu - (e_0)_\nu (e_3)_\mu &= -\epsilon_{\mu\nu\rho\sigma} (e_1)^\rho (e_2)^\sigma \\ (e_1)_\mu (e_2)_\nu - (e_1)_\nu (e_2)_\mu &= \epsilon_{\mu\nu\rho\sigma} (e_0)^\rho (e_3)^\sigma \\ (e_2)_\mu (e_3)_\nu - (e_2)_\nu (e_3)_\mu &= \epsilon_{\mu\nu\rho\sigma} (e_0)^\rho (e_1)^\sigma \\ (e_3)_\mu (e_1)_\nu - (e_3)_\nu (e_1)_\mu &= \epsilon_{\mu\nu\rho\sigma} (e_0)^\rho (e_2)^\sigma \end{aligned} \quad (3.73)$$

Ces derniers tenseurs constituent des *projecteurs* dans des 2-plans.

♣ Il est facile de montrer que le tenseur dual d'un tenseur plan lui est orthogonal. En effet

$$*(\underline{u} \wedge \underline{v}) \cdot (\underline{u} \wedge \underline{v}) = 2\epsilon_{\mu\nu\rho\sigma} u^\rho v^\sigma u^\mu v^\nu = 0 \quad (3.74)$$

Avant de démontrer une réciproque, envisageons l'observation par l'observateur P d'un tenseur antisymétrique F d'ordre 2. Pour simplifier l'écriture, soit n, e_x, e_y, e_z la tétrade de 4-vecteurs de base attachée à P , n étant le 4-vecteur du genre temps futur (par la suite nous appellerons tout simplement "tétrade" une telle base). Posons donc

$$\begin{aligned} \underline{E} &= -F \cdot n, \quad \text{soit} \quad E_\mu = -F_{\mu\nu} n^\nu \\ \underline{B} &= *F \cdot n, \quad \text{soit} \quad B_\mu = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} n^\nu F^{\rho\sigma} \end{aligned} \quad (3.75)$$

Les deux 4-vecteurs \underline{E} et \underline{B} sont orthogonaux à n et sont donc du genre espace. Calculons $*(n \wedge \underline{B})$. Il vient

$$*(n \wedge \underline{B})_{\mu\nu} = \epsilon_{\mu\nu\rho\sigma} n^\rho B^\sigma = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} n^\rho \epsilon^{\sigma\alpha\beta\gamma} n_\alpha F_{\beta\gamma}$$

soit, en utilisant la première équation en (3.68),

$$*(n \wedge \underline{B})_{\mu\nu} = F_{\mu\nu} - n_\mu F_{\rho\nu} n^\rho + n_\nu F_{\rho\mu} n^\rho$$

On en déduit

$$F = n \wedge \underline{E} + *(n \wedge \underline{B}) \quad (3.76)$$

et le résultat que *tout tenseur d'ordre 2 antisymétrique est décomposable par l'observateur P en deux tenseurs plans*.

En notant que si \underline{B} n'est pas nul, étant orthogonal à n , on peut l'utiliser pour définir un nouveau vecteur de base dans l'espace à trois dimensions orthogonal à n ; on voit alors, d'après (3.73) que $*(n \wedge \underline{B})$ s'exprime bien comme un tenseur plan.

Dans la base associée à P , \underline{E} et \underline{B} n'ont que des composantes spatiales et l'on a

$$E^i = F^{0i}, \quad B^i = F^{jk} \quad (3.77)$$

i, j, k prenant les valeurs 1, 2 et 3, et (i, j, k) étant une permutation circulaire de 1, 2, 3. Lorsqu'on étudie le champ électromagnétique, on est amené à considérer le tenseur antisymétrique

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (3.78)$$

où A_μ sont les composantes covariantes du champ de 4-vecteurs "quadri-potentiel", incluant le potentiel électrique et le potentiel-vecteur au sens usuel. Il est facile de se rendre compte que les parties F^{0i} et F^{ij} de ce tenseur concernent le champ électrique et le champ magnétique, respectivement²¹. D'où la notation utilisée plus haut, et les appellations courantes de *partie électrique* et de *partie magnétique* d'un tenseur antisymétrique d'ordre 2 pour désigner \underline{E} et \underline{B} , respectivement. On notera que ces deux parties ont des parités différentes vis-à-vis des symétries opérant dans l'espace à trois dimensions orthogonal à n , la partie électrique se comportant comme un "vrai" 3-vecteur, alors que la partie magnétique se comporte comme un "pseudo - 3-vecteur". Il est cependant important de préciser que l'ensemble de ces deux parties représente bien, d'un point de vue global dans l'espace à 4 dimensions, un vrai quadri-tenseur antisymétrique d'ordre 2.

En utilisant les formules (3.68) et compte-tenu de $n \cdot \underline{E} = n \cdot \underline{B} = 0$, on démontre aisément que

$$\begin{aligned} \Phi(F) &= \frac{1}{2} *F \cdot F = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma} = -2 \underline{E} \cdot \underline{B} = 2 \vec{E} \cdot \vec{B} \\ \Delta(F) &= -\frac{1}{2} F \cdot F = \underline{B}^2 - \underline{E}^2 = \vec{E}^2 - \vec{B}^2 \end{aligned} \quad (3.79)$$

Ces grandeurs sont invariantes relativistes et sont les *seuls* invariants indépendants que l'on peut construire à partir de F ²²; $\Delta(F)$ est appelé *discriminant* de F ²³. Suivant le signe de Δ , on classe de la façon suivante les tenseurs antisymétriques d'ordre 2.

- ♣ Si $\Delta(F) > 0$, le tenseur est du genre *électrique*;
- ♣ Si $\Delta(F) < 0$, il est du genre *magnétique*;
- ♣ Si $\Delta(F) = 0$, il est du genre *lumière*.

Considérons le cas où $\Phi(F) = 0$. Si $\underline{E} = 0$, ou bien si $\underline{B} = 0$, il est évident, d'après (3.76), que F ainsi que son dual prennent la forme de tenseurs plans. Supposons alors les deux parties non nulles. Comme ce sont deux 4-vecteurs orthogonaux à n et orthogonaux entre eux, on peut les utiliser pour définir une nouvelle base :

21. En fait, pour le champ électromagnétique, on a $E^i = -cF^{0i}$, $B^i = -F^{jk}$, voir plus loin.

22. Voir plus loin.

23. Lorsque F est un tenseur plan $\underline{u} \wedge \underline{v}$, $\Delta(F) = (\underline{u} \cdot \underline{v})^2 - \underline{u}^2 \underline{v}^2$ est le discriminant des deux 4-vecteurs \underline{u} et \underline{v} . Voir H. Bacry, "Leçons sur la Théorie des Groupes et les Symétries des Particules Élémentaires", Gordon and Breach, 1967 (distribué par Dunod ed.), p. 250.

$$\underline{\mathbf{E}} = \lambda_1 e_1, \quad \underline{\mathbf{B}} = \lambda_2 e_2$$

λ_1 et λ_2 étant deux réels, le dernier vecteur e_3 étant défini comme en (3.71). On a alors

$$\epsilon_{\mu\nu\rho\sigma} n^\rho B^\sigma = \lambda_2 \epsilon_{\mu\nu\rho\sigma} n^\rho (e_2)^\sigma = -\lambda_2 [(e_1)_\mu (e_3)_\nu - (e_1)_\nu (e_3)_\mu]$$

d'où, d'après (3.76)

$$F_{\mu\nu} = [\lambda_1 n_\nu + \lambda_2 (e_3)_\mu] (e_1)_\nu - [\lambda_1 n_\nu + \lambda_2 (e_3)_\nu] (e_1)_\mu \quad (3.80)$$

et F s'exprime alors comme un tenseur plan et il en est de même pour son dual. D'où le théorème réciproque

♣ Si un tenseur antisymétrique d'ordre deux et son dual sont orthogonaux entre eux, alors ce sont des tenseurs plans.

Dans ce cas,

♣ Si $\Delta(F) > 0$, il existe des observateurs pour lesquels la partie magnétique du tenseur est nulle ;

♣ Si $\Delta(F) < 0$, il existe des observateurs pour lesquels sa partie électrique est nulle ;

♣ Si $\Delta(F) = 0$, il n'existe aucun observateur pour lequel l'une ou l'autre partie est nulle car on a alors $|\vec{E}| = |\vec{B}|$.

Notons encore le résultat suivant

♠ un tenseur antisymétrique d'ordre 2 orthogonal à un 4-vecteur $\underline{\mathbf{a}}$ est plan.

Posons

$$\underline{\mathbf{u}} = \frac{\underline{\mathbf{a}} - (\underline{\mathbf{a}} \cdot \underline{\mathbf{n}}) \underline{\mathbf{n}}}{\sqrt{(\underline{\mathbf{a}} \cdot \underline{\mathbf{n}})^2 - \underline{\mathbf{a}}^2}}$$

où $\underline{\mathbf{n}}$ est un 4-vecteur du genre temps futur, et unitaire. Ce 4-vecteur $\underline{\mathbf{u}}$, orthogonal à $\underline{\mathbf{n}}$, est du genre espace et unitaire ($\underline{\mathbf{u}}^2 = -1$). A ces deux 4-vecteurs $\underline{\mathbf{n}}$ et $\underline{\mathbf{u}}$ on peut adjoindre deux 4-vecteurs orthogonaux et unitaires e_1 et e_2 , appartenant au sous-espace de dimension 2, du genre espace, orthogonal à la fois à $\underline{\mathbf{n}}$ et $\underline{\mathbf{u}}$, pour constituer une base du 4-espace. Relativement à cette base, on trouve que la seule composante non nulle dudit tenseur F est F_{12} . Par conséquent

$$F_{\mu\nu} = F_{12} [(e_1)_\mu (e_2)_\nu - (e_1)_\nu (e_2)_\mu]$$

et le tenseur est bien plan. De façon équivalente, il prend aussi les formes

$$F_{\mu\nu} = \frac{F_{12}}{\sqrt{(\underline{\mathbf{a}} \cdot \underline{\mathbf{n}})^2 - \underline{\mathbf{a}}^2}} \epsilon_{\mu\nu\rho\sigma} \underline{\mathbf{n}}^\rho \underline{\mathbf{a}}^\sigma = \frac{1}{2 \underline{\mathbf{a}}^2} \epsilon_{\mu\nu\alpha\beta} \underline{\mathbf{a}}^\alpha \epsilon^{\beta\lambda\rho\sigma} \underline{\mathbf{a}}_\lambda F_{\rho\sigma}$$

la dernière n'étant valable que si $\underline{\mathbf{a}}$ n'est pas du genre lumière.

Montrons maintenant comment les parties électrique et magnétique d'un tenseur antisymétrique d'ordre 2 se transforment sous une transformation de Lorentz telle que (3.13), faisant passer d'une base n, e_x, e_y, e_z à la base n', e_x, e_y, e_z' . On a bien sûr

$$E'_\mu = -F_{\mu\nu} n'^\nu, \quad B'_\mu = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} n'^\nu F^{\rho\sigma}$$

avec

$$n' = n \cosh \chi + e_z \sinh \chi, \quad e'_z = e_z \cosh \chi + n \sinh \chi, \quad e'_x = e_x, \quad e'_y = e_y$$

Posant $E'^x = +e'_x{}^\mu F_{\mu\nu} n'^\nu$, $E'^y = +e'_y{}^\mu F_{\mu\nu} n'^\nu$, $E'^z = +e'_z{}^\mu F_{\mu\nu} n'^\nu$, etc, et tenant compte des relations

$$\begin{aligned} n'^\rho e'_z{}^\sigma - n'^\sigma e'_z{}^\rho &= n^\rho e_z{}^\sigma - n^\sigma e_z{}^\rho, & \epsilon_{\mu\nu\rho\sigma} n'^\rho e'_z{}^\sigma &= \epsilon_{\mu\nu\rho\sigma} n^\rho e_z{}^\sigma \\ \epsilon_{\mu\nu\rho\sigma} e'_{(x,y)}{}^\mu n'^\nu &= \cosh \chi (\epsilon_{\mu\nu\rho\sigma} e_{(x,y)}{}^\mu n^\nu) \pm \sinh \chi \left[n_\rho e_{(y,x)\sigma} - n_\sigma e_{(y,x)\rho} \right] \\ n'^\rho e'_{(x,y)}{}^\sigma - n'^\sigma e'_{(x,y)}{}^\rho &= \cosh \chi \left[n^\rho e_{(x,y)}{}^\sigma - n^\sigma e_{(x,y)}{}^\rho \right] \pm \sinh \chi \epsilon^{\rho\sigma\alpha\beta} n_\alpha e_{(y,x)\beta} \end{aligned} \quad (3.81)$$

on obtient

$$\begin{aligned} E'^z &= E^z, & B'^z &= B^z \\ E'^x &= E^x \cosh \chi + B^y \sinh \chi, & E'^y &= E^y \cosh \chi - B^x \sinh \chi \\ B'^x &= B^x \cosh \chi - E^y \sinh \chi, & B'^y &= B^y \cosh \chi + E^x \sinh \chi \end{aligned} \quad (3.82)$$

Ces dernières formules permettent de retrouver simplement quels sont les invariants relativistes pouvant être construit à partir des composantes d'un tenseur antisymétrique d'ordre 2. Envisageons le 3-vecteur *complexe*

$$\vec{F} = \vec{E} + i \vec{B} \quad (3.83)$$

La transformation (3.82) opère de la façon suivante sur ce vecteur

$$F'^x = F^x \cos i\chi + F^y \sin i\chi, \quad F'^y = F^y \cos i\chi - F^x \sin i\chi, \quad F'^z = F^z \quad (3.84)$$

Autrement dit, pour ce vecteur, la transformation est équivalente à une "rotation" d'un angle imaginaire pur dans le plan x, y . L'ensemble de toutes les transformations de Lorentz dans l'espace à 4 dimensions, comprenant les rotations usuelles, est en fait équivalent à l'ensemble de toutes les rotations possibles d'angles complexes dans un espace complexe à trois dimensions (aux six paramètres des transformations dans le 4-espace correspondent trois angles *complexes* des rotations de cet espace à trois dimensions). Le *seul* invariant du vecteur \vec{F} vis-à-vis de ces rotations est son "carré"

$$\vec{F}^2 = \vec{E}^2 - \vec{B}^2 + 2i \vec{E} \cdot \vec{B} \quad (3.85)$$

On en déduit que les grandeurs $\Delta(F) = \vec{E}^2 - \vec{B}^2$ et $\Phi(F) = 2 \vec{E} \cdot \vec{B}$ sont bien les seuls invariants réels indépendants que l'on peut construire à partir du tenseur $F_{\mu\nu}$.

Signalons enfin que les opérations de dérivation par rapport aux coordonnées doivent être considérées comme les composantes d'un 4-vecteur. En effet, si $\Psi(x_0, x, y, z)$ est un invariant relativiste (un *scalaire* du point de vue des transformations de Lorentz), sa différentielle, pour des variations infinitésimales dx^μ des coordonnées, l'est aussi. Ecrivant

$$d\Psi = \sum_{\mu} \frac{\partial \Psi}{\partial x^\mu} dx^\mu$$

et, reconnaissant que les dx^μ sont les composantes contravariantes d'un 4-vecteur, force est d'admettre que les dérivées partielles $\frac{\partial \Psi}{\partial x^\mu}$ sont les composantes *covariantes* d'un 4-vecteur, le 4-gradient de Ψ ,

dont les composantes *contravariantes* sont les dérivées partielles de Ψ par rapport aux coordonnées *covariantes* :

$$\frac{\partial \Psi}{\partial x_0} = \frac{\partial \Psi}{\partial x^0}, \quad \frac{\partial \Psi}{\partial x_i} = -\frac{\partial \Psi}{\partial x^i}, \quad i = 1, 2, 3$$

La *quadri-divergence* d'un champ de 4-vecteurs de composantes covariantes A_μ est l'invariant relativiste

$$\partial^\mu A_\mu = \partial_\mu A^\mu \equiv \sum_\mu \frac{\partial A_\mu}{\partial x_\mu} = \sum_\mu \frac{\partial A^\mu}{\partial x^\mu}$$

3.3.3 Intégration dans l'espace à quatre dimensions

Dans l'espace à quatre dimensions, on peut envisager quatre types d'intégration.

♣ Intégration le long d'une courbe d'espace-temps.

Elle fait intervenir les formes différentielles de degré 1 que sont les différentielles dx^μ . Une intégration invariante impliquera la forme invariante $A_\mu dx^\mu$ où les grandeurs A_μ sont les composantes covariantes d'un champ de 4-vecteurs.

♣ Intégration sur une surface à deux dimensions d'espace-temps.

On se rappelle que dans l'espace ordinaire à trois dimensions, un élément de surface construit à l'aide de deux vecteurs infinitésimaux \vec{du} et \vec{dv} peut être caractérisé par le produit vectoriel $\vec{du} \wedge \vec{dv}$ qui donne à la fois, par son module, l'aire de l'élément de surface ainsi construit, et par son orientation, la normale orientée de cette surface. Ses projections sur les axes de coordonnées sont les quantités $du^i dv^j - du^j dv^i$. De façon similaire, pour le 4-espace, un élément de surface sera caractérisé par un tenseur antisymétrique d'ordre 2 : $dS^{\mu\nu} = du^\mu dv^\nu - du^\nu dv^\mu$. Cependant, alors que dans l'espace à trois dimensions l'élément de surface peut être défini au moyen du produit vectoriel qui est un vecteur, ce qui est une particularité de la dimension 3, dans le 4-espace on ne peut construire un tel vecteur. L'équivalent est le tenseur dual d'ordre deux

$$d^*S^{\mu\nu} = -\frac{1}{2} \epsilon^{\mu\nu\alpha\beta} dS_{\alpha\beta} \quad (3.86)$$

Comme $dS_{\alpha\beta}$ est plan, son dual $d^*S^{\mu\nu}$ lui est orthogonal. Ce dernier est également orthogonal à tout segment construit dans l'élément de surface, et l'on a

$$d^*S^{\mu\nu} d^*S_{\mu\nu} = -dS_{\alpha\beta} dS^{\alpha\beta} \quad (3.87)$$

♣ Intégration sur une hyper-surface à trois dimensions d'espace-temps.

Dans l'espace à trois dimensions, le volume d'un parallélépipède construit à l'aide de trois vecteurs est donné par le produit mixte de ces trois vecteurs, qui est aussi leur déterminant. L'analogue pour l'espace à quatre dimensions est le tenseur d'ordre trois antisymétrique, dont les composantes sont données par les déterminants

$$dS^{\mu\nu\rho} = \begin{vmatrix} du^\mu & dv^\mu & dw^\mu \\ du^\nu & dv^\nu & dw^\nu \\ du^\rho & dv^\rho & dw^\rho \end{vmatrix} \quad (3.88)$$

des composantes de trois 4-vecteurs infinitésimaux $d\underline{u}$, $d\underline{v}$ et $d\underline{w}$ constituant l'élément d'hypersurface. L'intégration correspondante fait plutôt intervenir le dual de ce tenseur, qui est un 4-vecteur :

$$d^*S^\mu = \frac{1}{6} \epsilon^{\mu\nu\alpha\beta} dS_{\nu\alpha\beta} \quad (3.89)$$

qui est orthogonal à tout segment dans l'élément d'hyper-surface, et l'on a, dans un repère quelconque,

$$d^*S^0 = dx dy dz, \quad d^*S^1 = -dx^0 dy dz, \quad d^*S^2 = -dx^0 dx dz, \quad d^*S^3 = -dx^0 dx dy \quad (3.90)$$

où l'on voit que la composante temporelle de ce 4-vecteur est l'élément de volume tri-dimensionnel habituel.

♣ Intégration dans un volume d'espace-temps.

L'élément de volume à quatre dimensions est bien sûr invariant et a pour expression

$$d^4x \equiv dx^0 dx^1 dx^2 dx^3 \equiv c dt dx dy dz \quad (3.91)$$

Pour le 4-espace, il existe des théorèmes similaires aux théorème de Stokes ou de Green-Ostrogradski. Rappelons tout d'abord le théorème de Stokes *généralisé*²⁴

♠ Soit ω une forme différentielle extérieure de degré $r - 1$, différentiable dans une variété différentielle orientable \mathcal{D} , de dimension r et de bord $\partial\mathcal{D}$. On a

$$\int_{\mathcal{D}} d\omega = \int_{\partial\mathcal{D}} \omega$$

L'application de ce théorème à l'espace de la Relativité s'effectue selon les règles suivantes.

♣ L'intégration le long d'une courbe *fermée* d'espace-temps Γ peut être transformée en une intégration sur la surface Σ qu'elle délimite en faisant la substitution

$$\oint_{\Gamma} dx^\mu \rightarrow \int_{\Sigma} dS^{\nu\mu} \partial_\nu \quad (3.92)$$

La circulation d'un champ de 4-vecteurs le long d'une courbe fermée devient ainsi

$$\oint_{\Gamma} dx^\mu A_\mu = \frac{1}{2} \int_{\Sigma} dS^{\alpha\beta} \left[\frac{\partial A_\beta}{\partial x^\alpha} - \frac{\partial A_\alpha}{\partial x^\beta} \right] \equiv \frac{1}{4} \int_{\Sigma} d^*S_{\mu\nu} \epsilon^{\mu\nu\rho\sigma} [\partial_\rho A_\sigma - \partial_\sigma A_\rho] \quad (3.93)$$

Exprimé en termes de différentiation extérieure, cela donne

$$\oint_{\Gamma} dx^\mu A_\mu = \int_{\Sigma} dA_\mu \wedge dx^\mu = \frac{1}{2} \int_{\Sigma} [\partial_\rho A_\mu - \partial_\mu A_\rho] dx^\rho \wedge dx^\mu \quad (3.94)$$

Par exemple, si la surface se trouve dans le plan x, y , seule la composante dS^{12} est non nulle, et l'on a

$$\oint_{\Gamma} dx^\mu A_\mu = - \oint_{\Gamma} \overrightarrow{dr} \cdot \overrightarrow{A} = \frac{1}{2} \int_{\Sigma} dS^{12} \left[\frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2} \right] = - \int_{\Sigma} d\Sigma \left(\overrightarrow{\text{rot}} \overrightarrow{A} \right)_z$$

résultat bien conforme au théorème de Stokes habituel.

24. Voir par exemple, J. Bass, Cours de Mathématiques, Tome I, Masson et Cie Ed., Paris, 1968, § 23.8 et 27.16 ; V. Arnold, "Méthodes Mathématiques de la Mécanique Classique", Ed. Mir, Moscou, 1976, § 36 ; et, pour une présentation plus mathématique : C. Godbillon "Géométrie Différentielle et Mécanique Analytique", Hermann Ed. Paris, 1969.

♣ L'intégration sur une surface fermée à deux dimensions Σ_2 est transformée en intégrale sur l'hyper-surface (à trois dimensions) Σ_3 qu'elle délimite, en faisant la substitution

$$\oint_{\Sigma_2} d^*S_{\mu\nu} \rightarrow - \int_{\Sigma_3} [d^*S_\mu \partial_\nu - d^*S_\nu \partial_\mu] \quad (3.95)$$

Du point de vue de la différentiation extérieure, on a en effet

$$\oint_{\Sigma_2} \Psi dx^\rho \wedge dx^\sigma = \int_{\Sigma_3} d\Psi \wedge dx^\rho \wedge dx^\sigma = \int_{\Sigma_3} \partial_\alpha \Psi dx^\alpha \wedge dx^\rho \wedge dx^\sigma$$

Or

$$d^*S_\mu = \frac{1}{6} \epsilon_{\mu\nu\alpha\beta} dx^\nu \wedge dx^\alpha \wedge dx^\beta, \quad \text{et} \quad dx^\alpha \wedge dx^\beta \wedge dx^\gamma = -\epsilon^{\mu\alpha\beta\gamma} d^*S_\mu$$

d'où

$$\epsilon_{\theta\lambda\beta\gamma} \epsilon^{\mu\alpha\beta\gamma} d^*S_\mu = -2 [d^*S_\theta \delta_\lambda^\alpha - d^*S_\lambda \delta_\theta^\alpha]$$

et

$$\frac{1}{2} \epsilon_{\theta\lambda\beta\gamma} dx^\alpha \wedge dx^\beta \wedge dx^\gamma = d^*S_\theta \delta_\lambda^\alpha - d^*S_\lambda \delta_\theta^\alpha$$

Par suite,

$$\frac{1}{2} \epsilon_{\theta\lambda\rho\sigma} \oint_{\Sigma_2} \Psi dx^\rho \wedge dx^\sigma = \int_{\Sigma_3} \partial_\alpha \Psi [d^*S_\theta \delta_\lambda^\alpha - d^*S_\lambda \delta_\theta^\alpha] = \int_{\Sigma_3} [d^*S_\theta \partial_\lambda - d^*S_\lambda \partial_\theta] \Psi$$

d'où la règle annoncée. Par exemple, l'intégrale d'un tenseur antisymétrique prend les deux formes

$$\frac{1}{2} \oint_{\Sigma_2} d^*S_{\mu\nu} A^{\mu\nu} = - \int_{\Sigma_3} d^*S_\mu \partial_\nu A^{\mu\nu} \quad (3.96)$$

A titre de vérification, prenons $A^{\mu\nu} = n_\mu A_\nu - n_\nu A_\mu$ où n_μ sont les composantes covariantes d'un vecteur du genre temps futur \underline{n} , et A_μ celles d'un 4-vecteur orthogonal à \underline{n} . Alors, dans un repère lié à \underline{n} , on a $A_0 = 0$, et

$$-\frac{1}{2} \oint_{\Sigma_2} d^*S_{\mu\nu} A^{\mu\nu} = - \oint_{\Sigma_2} d^*S_{0i} A^i = \oint_{\Sigma_2} \sum_{i=1}^3 A^i dx^j \wedge dx^k$$

(i, j, k) étant une permutation circulaire de 1, 2, 3. D'un autre côté,

$$d^*S_\mu \partial_\nu A^{\mu\nu} = d^*S_0 \partial_\nu A^\nu - d^*S_\mu \partial_0 A^\mu$$

Supposons de plus que, dans ce repère, les composantes A^μ ne dépendent pas de $x_0 = n^\mu x_\mu$. Alors $\partial_0 A^\mu = 0$ et il vient

$$d^*S_\mu \partial_\nu A^{\mu\nu} = d^*S_0 \partial_\nu A^\nu = \operatorname{div} \vec{A} \, dx \wedge dy \wedge dz$$

En identifiant, on trouve

$$\oint_{\Sigma_2} \sum_{i=1}^3 A^i dx^j \wedge dx^k = \int_{\Sigma_3} \operatorname{div} \vec{A} dx \wedge dy \wedge dz$$

c'est-à-dire, le théorème de Green-Ostrogradski pour trois dimensions.

♣ Enfin, l'intégration sur une hyper-surface fermée Σ se transforme comme une intégrale dans le 4-volume Ω qu'elle enferme, en faisant la substitution

$$\oint_{\Sigma} d^*S_{\mu} \rightarrow \oint_{\Omega} d^4x \partial_{\mu} \quad (3.97)$$

Cette règle repose sur l'identité

$$\frac{1}{6} \epsilon_{\lambda\nu\rho\sigma} dx^{\alpha} \wedge dx^{\nu} \wedge dx^{\rho} \wedge dx^{\sigma} = \delta_{\lambda}^{\alpha} dx^0 \wedge dx \wedge dy \wedge dz$$

L'intégrale d'un champ de 4-vecteurs devient ainsi

$$\oint_{\Sigma} d^*S_{\mu} A^{\mu} = \oint_{\Omega} d^4x \partial_{\mu} A^{\mu} \quad (3.98)$$

C'est l'équivalent, pour quatre dimensions, du théorème de Green-Ostrogradski pour trois dimensions, rappelé plus haut.

3.4 Le quadri-vecteur Energie-Impulsion

3.4.1 Transformation des vitesses

Supposons que les coordonnées apparaissant dans (3.25) soient celles d'un objet en mouvement, et voyons comment se comparent les vitesses de cet objet, telles qu'elles sont mesurées dans deux référentiels galiléens en mouvement relatif le long d'un axe commun $x'x$. Dans le premier référentiel, la vitesse de l'objet a pour composantes (en notations usuelles)

$$v_x = \frac{dx}{dt}, \quad v_y = \frac{dy}{dt}, \quad v_z = \frac{dz}{dt}$$

alors que dans le second référentiel, ce sont

$$v'_x = \frac{dx'}{dt'}, \quad v'_y = \frac{dy'}{dt'}, \quad v'_z = \frac{dz'}{dt'}$$

Utilisant (3.25), on trouve

$$\begin{aligned} v'_x &= \frac{dx'}{dt'} = c \frac{dx \cosh \chi - c dt \sinh \chi}{c dt \cosh \chi - dx \sinh \chi} = c \frac{v_x \cosh \chi - c \sinh \chi}{c \cosh \chi - v_x \sinh \chi} \\ v'_y &= \frac{dy'}{dt'} = c \frac{dy}{c dt \cosh \chi - dx \sinh \chi} = c \frac{v_y}{c \cosh \chi - v_x \sinh \chi} \\ v'_z &= c \frac{v_z}{c \cosh \chi - v_x \sinh \chi} \end{aligned} \quad (3.99)$$

Manifestement, ces lois de transformation ne correspondent en rien à celles des composantes d'un 4-vecteur : telle que, la vitesse ne peut donc constituer un 4-vecteur. Recherchons alors quelles combinaisons des composantes de la vitesse pourraient conduire à un 4-vecteur pouvant la généraliser.

Dans le schéma de la Mécanique non relativiste, on relie habituellement le vecteur vitesse aux paramètres intrinsèques de la trajectoire de l'objet que sont l'abscisse curviligne et le vecteur tangent. Ce dernier est défini comme

$$\vec{T} = \frac{\vec{dM}}{ds}$$

s étant l'abscisse curviligne. On sait que cet objet mathématique est un 3-vecteur parce que \vec{dM} , déplacement élémentaire, se comporte comme un 3-vecteur sous les rotations spatiales, et que ds , élément de longueur le long de la trajectoire, est un invariant par rotation. Le rapport de ces deux grandeurs ne peut donc être autre chose qu'un 3-vecteur.

En relativité, l'objet suit une "ligne d'univers", courbe dans l'espace à quatre dimensions. Chaque "position" de l'objet sur cette courbe correspond à un événement. Deux positions infiniment voisines sont séparées d'un intervalle infinitésimal

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = c^2 dt^2 \left(1 - \frac{v^2}{c^2}\right) > 0 \quad (3.100)$$

Positif, cet intervalle peut servir à définir l'équivalent d'une abscisse curviligne

$$ds = \sqrt{ds^2} = c dt \sqrt{1 - \frac{v^2}{c^2}} \quad (3.101)$$

sur la ligne d'univers de l'objet. Comme il s'agit d'une grandeur invariante relativiste, les quantités dx^μ/ds représenteront manifestement les composantes d'un 4-vecteur : celles du 4-vecteur *tangent* à la ligne d'univers

$$T^\mu = \frac{dx^\mu}{ds} \quad (3.102)$$

Dans le repère où la vitesse de l'objet est \vec{v} de module v , les composantes du 4-vecteur tangent sont

$$T^0 = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad \vec{T} = \frac{\vec{v}}{c \sqrt{1 - \frac{v^2}{c^2}}} \quad (3.103)$$

On vérifie que le 4-vecteur tangent est unitaire, du genre temps futur :

$$\underline{T}^2 = 1, \quad T^0 > 0 \quad (3.104)$$

La *4-vitesse* de l'objet, qui généralise la vitesse usuelle, est alors définie comme étant le 4-vecteur

$$\underline{V} = c \underline{T} \quad (3.105)$$

3.4.2 Principe de moindre Action et 4-vecteur énergie-impulsion

Même si globalement un objet n'est pas en mouvement rectiligne uniforme dans un référentiel galiléen, on peut néanmoins considérer qu'entre deux positions infiniment voisines, alors que sa vitesse a peu évolué, son référentiel propre est assimilable à un référentiel galiléen, en supposant toutefois que son repère d'espace garde une orientation fixe par rapport aux axes du référentiel galiléen. Il est alors possible d'exprimer l'intervalle séparant les deux "événements positions" infiniment voisins au moyen des données propres à l'objet. Celui-ci, au repos dans son propre référentiel, aura constaté un intervalle

de temps $d\tau$ entre les deux événements qu'il aura bien sûr situés au même endroit $dX = dY = dZ = 0$, tandis que l'observateur galiléen P aura mesuré la durée dt , et les écarts de coordonnées $dx = v_x dt$, $dy = v_y dt$, $dz = v_z dt$, v_x , v_y et v_z étant les composantes du vecteur vitesse *instantanée* qu'il attribue au mouvement de l'objet entre les deux événements. Appliquant une invariance galiléenne "locale", écrivons alors que l'intervalle séparant les deux événements doit avoir la même valeur pour l'objet et l'observateur galiléen P :

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = c^2 dt^2 \left(1 - \frac{v^2}{c^2}\right) = c^2 d\tau^2 \quad (3.106)$$

où $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$. La durée $d\tau$ est une durée propre pour l'objet. On en déduit la durée propre séparant deux positions quelconques de l'objet sur sa trajectoire d'univers, en fonction des mesures effectuées par l'observateur galiléen, comme l'intégrale

$$\tau_2 - \tau_1 = \Delta\tau = c \int_{t_1}^{t_2} dt \sqrt{1 - \frac{v^2(t)}{c^2}} \quad (3.107)$$

Nous exprimerons aussi la durée propre séparant deux positions d'espace-temps A et B de l'objet et de son horloge comme l'intégrale

$$\Delta\tau = \frac{1}{c} \int_A^B ds \quad (3.108)$$

calculée le long de la ligne d'univers de l'objet, avec $ds = \sqrt{ds^2}$. Il est clair, d'une part, que la durée propre a une valeur inhérente à l'objet, indépendante d'un observateur galiléen externe ; et que, d'autre part, comme cela est évident d'après (3.107), elle est plus courte que la durée $t_2 - t_1 = \Delta t$ mesurée par tout observateur galiléen P . Or, les deux "événements positions" A et B étant séparés par un intervalle du genre temps, il existera toujours un observateur galiléen P_0 pour lequel ils ont lieu au même endroit dans son repère d'espace, par exemple à l'endroit même où se situe P_0 . Pour cet observateur P_0 , la durée qui sépare les deux événements est aussi une durée propre. D'après ce qui précède, cette dernière sera plus grande que la durée propre de l'objet, quelle que soit la forme prise par l'évolution de ce dernier dans le repère de P_0 . Il est important de noter ici que la situation n'est pas symétrique vis-à-vis de l'objet et de P_0 . En effet, P_0 est galiléen et voit l'objet partir puis revenir vers lui. L'objet ne peut donc être en translation rectiligne uniforme par rapport à P_0 et ne peut de ce fait constituer lui-même un référentiel galiléen.

Pour tout autre observateur galiléen, la trajectoire d'univers de P_0 est la *droite* d'univers entre les deux points A et B . On conclut donc de cette discussion que, étant donnés deux points d'univers A et B , l'intégrale

$$\int_A^B ds \quad (3.109)$$

atteint sa valeur *maximum* lorsqu'elle est effectuée le long de la droite d'univers entre ces deux points.

On sait, de la Mécanique Classique, qu'un objet en évolution ne peut se prévaloir d'avoir un référentiel propre galiléen que s'il n'est soumis à aucune interaction avec d'autres systèmes. Ce qui précède incite alors à utiliser l'intégrale (3.109) pour définir l'Action d'un objet isolé, afin de déduire les équations d'évolution de cet objet, via un principe variationnel. L'Action de l'objet M sera définie comme

$$W = -\kappa \int_A^B ds \quad (3.110)$$

où κ est supposée être une constante *positive*, caractéristique de M . Il s'agit d'un invariant relativiste. La présence du signe "-" se justifie par le fait qu'avec ce signe, l'Action prend sa valeur *minimum*,

plutôt que *maximum*, lorsque la ligne d'univers est une droite d'univers. Dans un référentiel galiléen donné, l'Action peut être réécrite à l'aide d'une fonction de Lagrange

$$\mathcal{L} = -\kappa c \sqrt{1 - \frac{v^2}{c^2}} \quad (3.111)$$

de sorte que

$$W = \int_{t_1}^{t_2} \mathcal{L} dt \quad (3.112)$$

La quantité de mouvement, ou 3-impulsion, de l'objet M a pour composantes cartésiennes

$$p_i = \frac{\partial \mathcal{L}}{\partial v_i} = \frac{\kappa}{c} \frac{v_i}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad i = x, y, z \quad (3.113)$$

et les équations d'Euler-Lagrange, donnant les équations de l'évolution effectivement réalisée, conduisent à

$$\frac{dp_i}{dt} = 0, \quad \text{soit } v_i = \text{constante} \quad (3.114)$$

Ce principe variationnel permet bien de retrouver le fait qu'un objet matériel isolé dans un référentiel galiléen y est au repos ou en mouvement rectiligne uniforme, constituant lui-même un nouveau repère galiléen. Sa trajectoire d'univers est une droite d'univers. Il y a cependant une nouveauté de taille par rapport au résultat classique : l'expression relativiste (3.113) de la quantité de mouvement en fonction de la vitesse est différente. Cependant, pour $v \ll c$, on doit retrouver la formule classique $p_i = mv_i$, où m est la masse de l'objet. Comme dans ce cas (3.113) donne $p_i \approx \kappa v_i/c$, on est conduit à faire l'identification

$$\kappa = mc \quad (3.115)$$

attribuant ainsi au paramètre masse m l'invariance relativiste. La 3-impulsion relativiste a donc pour expression

$$\vec{p} = \frac{m \vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} = mc \vec{T} \quad (3.116)$$

où l'on a introduit la partie 3-vecteur du 4-vecteur tangent à la droite d'univers de l'objet M . Cette dernière formule suggère fortement d'élargir cette relation à la quatrième composante temporelle et de définir les composantes covariantes d'un "4-vecteur impulsion" par la relation

$$p_\mu = mc T_\mu \quad (3.117)$$

Pour interpréter physiquement la quatrième composante,

$$p_0 = \frac{mc}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (3.118)$$

retournons au formalisme Lagrangien. Le Hamiltonien de l'objet, qui en fait représentera son énergie E , est défini par

$$E = \vec{p} \cdot \vec{v} - \mathcal{L} = \frac{mv^2}{\sqrt{1 - \frac{v^2}{c^2}}} + mc^2 \sqrt{1 - \frac{v^2}{c^2}}, \text{ soit}$$

$$E = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} = cp_0 \quad (3.119)$$

La quatrième composante p_0 est donc associée à l'énergie de l'objet, d'où l'appellation de 4-vecteur *énergie-impulsion* donnée au 4-vecteur (3.117). Lorsque $v \ll c$, son expression devient

$$E \approx mc^2 + \frac{mv^2}{2} \quad (3.120)$$

A une constante près, on retrouve l'expression classique de l'énergie cinétique. Lorsque $v = 0$, il subsiste un terme d'énergie E_0 , souvent appelé "énergie au repos" de l'objet, directement lié à sa masse. Son expression est donnée par la célèbre formule d'Einstein

$$E_0 = mc^2 \quad (3.121)$$

Que l'énergie soit associée à la composante temporelle d'un 4-vecteur peut être retrouvé de la façon suivante. On sait²⁵ que l'Action, considérée comme fonction des paramètres définissant la limite supérieure A d'intégration, a pour différentielle

$$dW = \vec{p}_A \cdot d\vec{r}_A - E_A dt_A$$

Cette expression possède ici le caractère d'invariance relativiste. Comme $d\vec{r}_A$ et cdt_A sont des composantes d'un 4-vecteur, \vec{p}_A et E_A/c doivent obligatoirement représenter, respectivement, les composantes spatiales et la composante temporelle d'un 4-vecteur.

On notera la relation

$$\vec{p} = \frac{E}{c^2} \vec{v} \quad (3.122)$$

entre la quantité de mouvement et l'énergie d'un objet relativiste. Elle subsiste même lorsque $v = c$. Dans ce cas, les formules (3.116) et (3.119) semblent ne plus être applicables à cause du dénominateur infini. Ceci révèle à nouveau qu'un objet massif ne peut se déplacer à la vitesse c . Cette vitesse ne peut être réalisée que par des "objets" de masse nulle, qui ne sont donc jamais au repos dans aucun référentiel galiléen. C'est le cas du photon. Pour cette particule, la relation quantité de mouvement - énergie est

$$\vec{p} = \frac{E}{c} \vec{u} \quad (3.123)$$

où \vec{u} est le 3-vecteur unitaire définissant la direction de propagation du photon dans le référentiel galiléen considéré.

Exprimée en fonction de l'impulsion, l'énergie d'un objet massif est

$$E = c \sqrt{p^2 + m^2 c^2} \quad (3.124)$$

et prend la forme

25. Voir le chapitre I.

$$E \approx mc^2 + \frac{p^2}{2m} \quad (3.125)$$

lorsque $v \ll c$.

Lors d'une transformation de Lorentz telle que (3.25), les composantes du 4-vecteur énergie-impulsion deviennent :

$$E' = E \cosh \chi - c p_x \sinh \chi, \quad p'_x = p_x \cosh \chi - \frac{E}{c} \sinh \chi, \quad p'_y = p_y, \quad p'_z = p_z \quad (3.126)$$

Dans le cas d'un photon (masse nulle), ces formules conduisent à

$$p' = p \cosh \chi - p_x \sinh \chi, \quad p'_x = p_x \cosh \chi - p \sinh \chi, \quad p'_y = p_y, \quad p'_z = p_z \quad (3.127)$$

où $p = \sqrt{p_x^2 + p_y^2 + p_z^2}$. Si la direction de propagation du photon est parallèle au plan xy en faisant l'angle θ avec le plan xz , on a $p_z = 0$, $p_x = p \cos \theta$, $p_y = p \sin \theta$ dans le premier référentiel, $p'_z = 0$, $p'_x = p' \cos \theta'$, $p'_y = p' \sin \theta'$ dans le second référentiel, et l'on a

$$p' = p [\cosh \chi - \cos \theta \sinh \chi], \quad p' \cos \theta' = p [\cos \theta \cosh \chi - \sinh \chi], \quad p' \sin \theta' = p \sin \theta$$

d'où

$$\begin{aligned} \frac{p'}{p} &= \cosh \chi - \cos \theta \sinh \chi \\ \cos \theta' &= \frac{\cos \theta \cosh \chi - \sinh \chi}{\cosh \chi - \cos \theta \sinh \chi} \\ \sin \theta' &= \frac{\sin \theta}{\cosh \chi - \cos \theta \sinh \chi} \end{aligned} \quad (3.128)$$

et, inversement,

$$\begin{aligned} \frac{p}{p'} &= \cosh \chi + \cos \theta' \sinh \chi \\ \cos \theta &= \frac{\cos \theta' \cosh \chi + \sinh \chi}{\cosh \chi + \cos \theta' \sinh \chi} \\ \sin \theta &= \frac{\sin \theta'}{\cosh \chi + \cos \theta' \sinh \chi} \end{aligned} \quad (3.129)$$

Ces relations montrent que d'un référentiel à l'autre, la direction de propagation du photon est différente. Si par exemple on a $\theta' = \pi/2$, alors

$$\cos \theta = \tanh \chi = \frac{u}{c}, \quad \sin \theta = e^{-\chi} \quad (3.130)$$

résultat conforme au schéma du paragraphe 3.1.2. En outre, l'énergie du photon diffère aussi d'un référentiel à l'autre. La Mécanique quantique nous enseigne que l'énergie E d'un photon est liée à la fréquence ν de l'onde à laquelle il est associé par la relation $E = h\nu$, h étant la constante de Planck. Dans le changement de référentiel, la fréquence de l'onde varie donc dans le rapport

$$\frac{\nu'}{\nu} = \cosh \chi - \cos \theta \sinh \chi \quad (3.131)$$

Imaginons que le photon soit émis par une source de lumière en mouvement de translation rectiligne uniforme, de vitesse u , selon l'axe des x d'un référentiel galiléen. La fréquence $\nu' = \nu_0$ représente alors une *fréquence propre*. L'observateur galiléen enregistrera pour sa part une onde de fréquence

$$\nu = \nu_0 \frac{\sqrt{1 - \frac{u^2}{c^2}}}{1 - \frac{u}{c} \cos \theta} \quad (3.132)$$

Cette formule décrit *l'effet Doppler relativiste*.

Pour terminer ce paragraphe, notons aussi que dans le cas relativiste, l'équation d'évolution d'une particule isolée peut s'écrire comme

$$\frac{dT^\mu}{ds} = 0 \quad (3.133)$$

le long de sa ligne d'univers. Mais cette équation n'est rien d'autre que celle d'une géodésique dans l'espace à quatre dimensions, muni de la métrique $g_{\mu\nu} = \text{constante}$, avec la condition $ds^2 > 0$, comme on peut s'en rendre compte facilement en utilisant les formules du complément 2.7. Un objet matériel isolé suit donc une géodésique dans le 4-espace, qui est ici une droite d'univers.

3.5 Interaction d'une particule chargée avec un champ électromagnétique

3.5.1 Lagrangien non-relativiste

Montrons que l'on décrit bien classiquement l'interaction électromagnétique d'une particule (ponctuelle) de charge q et de masse m , avec un champ électromagnétique dont le potentiel électrique est $V(x, y, z, t)$ et le potentiel vecteur $\vec{A}(x, y, z, t)$, avec le Lagrangien

$$\mathcal{L} = \frac{m}{2} \vec{v}^2 - qV + q \vec{v} \cdot \vec{A} \quad (3.134)$$

En dérivant ce Lagrangien par rapport aux composantes de la vitesse, on trouve l'impulsion :

$$\vec{p} = m \vec{v} + q \vec{A} \quad (3.135)$$

La dérivation par rapport aux coordonnées d'espace donne

$$\frac{\partial \mathcal{L}}{\partial x_k} = -q \frac{\partial V}{\partial x_k} + q \vec{v} \cdot \frac{\partial}{\partial x_k} \vec{A} \quad (3.136)$$

Appliquons alors l'équation d'Euler-Lagrange :

$$\frac{\partial \mathcal{L}}{\partial x_k} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial v_k}$$

Posant $a_k = \frac{dv_k}{dt}$ (composante de l'accélération), il vient

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial v_k} = ma_k + q \frac{dA_k}{dt} = ma_k + q \left[\frac{\partial A_k}{\partial t} + v_x \frac{\partial A_k}{\partial x} + v_y \frac{\partial A_k}{\partial y} + v_z \frac{\partial A_k}{\partial z} \right]$$

En prenant pour exemple l'équation concernant la coordonnée x , on obtient ainsi :

$$ma_x + q \left[\frac{\partial A_x}{\partial t} + v_x \frac{\partial A_x}{\partial x} + v_y \frac{\partial A_x}{\partial y} + v_z \frac{\partial A_x}{\partial z} \right] = -q \frac{\partial V}{\partial x} + q \left[v_x \frac{\partial A_x}{\partial x} + v_y \frac{\partial A_y}{\partial x} + v_z \frac{\partial A_z}{\partial x} \right]$$

D'où, en réarrangeant les termes,

$$ma_x = -q \frac{\partial V}{\partial x} - q \frac{\partial A_x}{\partial t} + q v_y \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) - q v_z \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right)$$

Les grandeurs apparaissant au second membre de cette équation sont liés aux composantes du champ électrique

$$\vec{E} = - \text{grad } V - \frac{\partial \vec{A}}{\partial t}$$

et du champ magnétique

$$\vec{B} = \text{rot } \vec{A}$$

On retrouve ainsi l'une des projections de la relation fondamentale classique, bien connue,

$$ma_x = F_x = q \left[E_x + (\vec{v} \wedge \vec{B})_x \right] \quad (3.137)$$

où F_x est la composante suivant l'axe des x de la *force de Lorentz* agissant sur la charge. On trouve des équations similaires concernant les coordonnées y et z . Le Lagrangien (3.134) permet donc bien de retrouver les équations d'évolution classique d'une charge plongée dans un champ électromagnétique.

On notera que, d'après (3.135), l'impulsion de la particule en interaction avec le champ électromagnétique n'est plus seulement formée par la "partie cinétique" usuelle $m \vec{v}$, mais contient aussi le terme $q \vec{A}$. Le Hamiltonien de la particule

$$\mathcal{H} = \vec{p} \cdot \vec{v} - \mathcal{L} = \frac{m \vec{v}^2}{2} + qV$$

doit être exprimé en fonction de cette impulsion, ce qui conduit à la formule

$$\mathcal{H} = \frac{1}{2m} \left(\vec{p} - q \vec{A} \right)^2 + qV \quad (3.138)$$

3.5.2 Lagrangien relativiste ; 4-potentiel du champ électromagnétique

Le problème consiste maintenant à généraliser le Lagrangien (3.134) au cas relativiste, avec pour objectif d'obtenir une intégrale d'Action qui soit invariante relativiste. Il semble tout d'abord assez

naturel de remplacer la partie "énergie cinétique" classique $\frac{m \vec{v}^2}{2}$ par (3.111), de façon à retrouver le Lagrangien relativiste d'une particule isolée dans le cas où le champ électromagnétique est inexistant.

Pour ce qui concerne la partie décrivant l'interaction électromagnétique de la charge, on note qu'elle conduit à une Action s'écrivant

$$W_{\text{em}} = -q \int_{t_1}^{t_2} dt \left[V - \vec{v} \cdot \vec{A} \right] \equiv -q \int_A^B \left[dx^0 \frac{V}{c} - \vec{dr} \cdot \vec{A} \right] \quad (3.139)$$

La dernière expression est assez suggestive car sous son signe somme intervient une combinaison de termes ayant la forme d'un produit scalaire relativiste. Aussi, puisque \vec{dr} et dx^0 sont des composantes d'un 4-vecteur, l'expression ci-dessus sera invariante relativiste à la condition que le potentiel vecteur \vec{A} et le potentiel électrique V divisé par c soient également les composantes d'un 4-vecteur. L'expérience montre qu'il en est bien ainsi. On appelle *quadri-potentiel* du champ électromagnétique le 4-vecteur \underline{A} dont les composantes contravariantes A^μ sont

$$A^0 = \frac{V}{c}, \quad A^k = (\vec{A})^k \quad (3.140)$$

Nous admettons donc que l'expression relativiste de l'Action d'une charge dans un champ électromagnétique est

$$W = \int_A^B [-m c ds - q A_\mu dx^\mu] \quad (3.141)$$

l'intégrale étant effectuée le long de la courbe d'univers de la particule (sur laquelle, du reste, on a $ds^2 > 0$). Il est intéressant de déduire les équations d'évolution directement à partir de cette expression, via sa variation, sans avoir recours à un Lagrangien dont l'expression serait liée à un référentiel particulier.

On a

$$\delta W = \int_A^B [-m c \delta(ds) - q(\delta A_\mu) dx^\mu - q A_\mu (\delta dx^\mu)]$$

Or, d'une part,

$$\delta(ds^2) = 2\delta(ds)ds = \delta(dx^\mu dx_\mu) = 2dx^\mu (\delta dx_\mu), \quad \text{donc} \quad \delta(ds) = T^\mu \delta(dx_\mu) \quad (T^\mu = \frac{dx^\mu}{ds})$$

et

$$T^\mu \delta(dx_\mu) = T^\mu d(\delta x_\mu) = d(T^\mu \delta x_\mu) - \delta x_\mu dT^\mu$$

D'autre part,

$$\delta A_\mu = \frac{\partial A_\mu}{\partial x^\nu} \delta x^\nu, \quad \text{et} \quad A_\mu (\delta dx^\mu) = d(A_\mu \delta x^\mu) - \delta x^\mu \frac{\partial A_\mu}{\partial x^\nu} dx^\nu$$

On en déduit la variation

$$\delta W = [-m c T^\mu - q A^\mu] \delta x_\mu \Big|_A^B + \int_A^B \delta x^\mu ds \left[m c \frac{dT^\mu}{ds} - q(\partial_\mu A_\nu - \partial_\nu A_\mu) T^\nu \right] \quad (3.142)$$

Le premier terme tout intégré est nul, puisque, dans l'esprit du principe variationnel, les paramètres définissant les limites d'intégration doivent être fixés, ce qui impose $\delta x_\mu = 0$. Les paramètres δx^μ étant arbitraires entre les limites d'intégration, le second terme est nul si et seulement si

$$mc \frac{dT_\mu}{ds} = \frac{dp_\mu}{ds} = q F_{\mu\nu} T^\nu \quad (3.143)$$

où

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (3.144)$$

est le *tenseur électromagnétique*, tenseur d'ordre deux, antisymétrique, et $p_\mu = mcT_\mu$ est le 4-vecteur énergie-impulsion de la particule. L'équation quadri-vectorielle (3.143) est l'écriture relativiste de la relation fondamentale de la Dynamique, dans le cas de l'interaction entre une particule chargée et le champ électromagnétique. Plaçons-nous dans un référentiel galiléen particulier. On a alors $ds = cdt\sqrt{1-v^2/c^2}$, v étant le module de la vitesse de la particule dans ce référentiel. Il vient

$$\frac{dp_\mu}{dt} = cq \sqrt{1 - \frac{v^2}{c^2}} F_{0k} T^k = cq F_{0k} v^k$$

v^k étant les composantes du 3-vecteur vitesse de la particule. Or,

$$c F_{0k} = \frac{\partial A_k}{\partial t} - \frac{\partial V}{\partial t} = -\frac{\partial A^k}{\partial t} - \frac{\partial V}{\partial t} = E^k$$

est la composante k du champ électrique. On trouve donc

$$\frac{dcp_0}{dt} = q \vec{v} \cdot \vec{E} \quad (3.145)$$

Cette équation ne fait que traduire la conservation de l'énergie, en interprétant à nouveau la quantité $E = cp_0$ comme l'énergie cinétique relativiste de la particule.

Les composantes spatiales de l'équation donnent

$$\frac{dp_i}{dt} = cq F_{i0} + q \sum_j F_{ij} v^j$$

Comme

$$F_{ij} = \frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j} = - \left[\frac{\partial A^j}{\partial x^i} - \frac{\partial A^i}{\partial x^j} \right] = -(\vec{\text{rot}} \vec{A})^k = -B^k$$

B^k étant l'une des composantes du champ magnétique, et (i, j, k) une permutation circulaire de 1, 2, 3, il vient

$$\sum_j F_{ij} v^j = -(\vec{v} \wedge \vec{B})^i$$

D'où l'équation 3-vectorielle

$$\frac{d\vec{p}}{dt} = q \vec{E} + q \vec{v} \wedge \vec{B} = \vec{F} \quad (3.146)$$

où \vec{F} est à nouveau la force de Lorentz.

Bien entendu, on retrouve tous ces résultats en utilisant le Lagrangien

$$\mathcal{L} = -mc^2 \sqrt{1 - \frac{v^2}{c^2}} - qV + q \vec{v} \cdot \vec{A} \quad (3.147)$$

Ce Lagrangien conduit d'ailleurs à définir l'impulsion par

$$\vec{P} = \vec{p} + q \vec{A}, \quad \text{avec} \quad \vec{p} = \frac{m \vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (3.148)$$

(\vec{p} est maintenant la *partie cinétique* de l'impulsion) et le Hamiltonien s'écrit

$$\mathcal{H} = c \sqrt{\left(\vec{P} - q \vec{A} \right)^2 + m^2 c^2} + qV \quad (3.149)$$

Pour terminer, faisons les remarques suivantes.

♣ La relation (3.146) se généralise à un objet quelconque soumis à l'action d'un système extérieur, caractérisée par une force \vec{F} . C'est l'équivalent en Relativité de la relation fondamentale de la Dynamique Classique.

♣ Les équations d'évolution étant satisfaites, la relation (3.142) permet d'obtenir la variation de l'Action consécutive à une modification infinitésimale des paramètres définissant les limites extrêmes. Fixant les paramètres de A , on voit que l'évolution de l'Action vis-à-vis des paramètres de B est gouvernée par le 4-vecteur impulsion généralisée $P_\mu = p_\mu + qA_\mu$:

$$\frac{\partial W}{\partial x^\mu} = -P_\mu \quad (3.150)$$

♣ On sait de la théorie électromagnétique que les potentiels ne sont pas définis de façon univoque. Les grandeurs physiques fondamentales du champ électromagnétique sont les champs électrique et magnétique, et ceux-ci restent inchangés si les potentiels subissent simultanément la transformation

$$\vec{A} \rightarrow \vec{A} + \text{grad } f, \quad V \rightarrow V - \frac{\partial f}{\partial t} \quad (3.151)$$

appelée *transformation de jauge*, $f(x, y, z, t)$ étant une fonction quelconque des coordonnées et du temps. En termes de composantes du 4-potentiel, cette transformation se traduit par

$$A_\mu \rightarrow A_\mu - \partial_\mu f \quad (3.152)$$

Autrement dit, la transformation de jauge consiste à rajouter au 4-potentiel le 4-gradient d'une fonction arbitraire.

Toute théorie électromagnétique se doit d'être insensible à un tel changement de jauge : on dit qu'elle doit posséder l'*invariance de jauge*²⁶. On vérifie que celle présentée plus haut a cette propriété. En effet, l'équation (3.143) fait intervenir le tenseur (3.144) qui est bien invariant de jauge puisque

$$F'_{\mu\nu} = \partial_\mu(A_\nu - \partial_\nu f) - \partial_\nu(A_\mu - \partial_\mu f) = F_{\mu\nu}$$

On le vérifie aussi en notant que le changement de jauge rajoute à l'intégrand dans (3.141) la différentielle totale $q dx^\mu \partial_\mu f = d(qf)$, ce qui n'a aucune incidence sur l'application du principe variationnel²⁷.

26. Ce point sera éclairci et généralisé dans un chapitre à venir.

27. Le changement de jauge se compare encore à une transformation canonique classique.

3.6 Compléments : les boosts, précession de Thomas

3.6.1 Les boosts

Lorsqu'on aborde l'étude des changements de référentiels galiléens en relativité, par souci de simplicité on considère généralement en tout premier lieu deux observateurs galiléens dont les repères d'espace respectifs ont un axe commun, pris selon la direction de leur ligne de vol supposée commune, les autres axes, tout en étant parallèles entre eux, sont perpendiculaires à cette direction. On obtient ainsi la transformation décrite par (3.25), que l'on appelle *transformation de Lorentz spéciale*. Que deviennent les formules de transformation si les axes des repères ont une orientation différente ?

Notons tout d'abord que si un observateur O voit, dans son repère d'espace R un observateur O' se déplacer avec la vitesse (constante) \vec{u} , selon le principe de relativité, l'observateur O' doit voir, dans son repère R' , O se déplacer avec la vitesse opposée. Les relations de transformation des vitesses (3.99) satisfont ce principe, puisque, prenant $v_x = v_y = v_z = 0$ (O au repos dans R), on trouve bien $v'_y = v'_z = 0$, $v'_x = -u$.

Forts de ce principe, nous pourrions affirmer que l'on passe d'un référentiel \mathcal{R} à un référentiel \mathcal{R}' sans qu'il y ait rotation d'axes si, O ayant mesuré les composantes de la vitesse de O' dans R comme étant u_x , u_y et u_z , O' trouve $-u_x$, $-u_y$ et $-u_z$ pour composantes de la vitesse de O dans R' . Posons $\vec{n} = \vec{u} / u$. Ce 3-vecteur unitaire aura donc les mêmes composantes dans R et dans R' .

Reconsidérant la transformation spéciale (3.25), on réalise que x' et x représentent en fait les projections sur \vec{n} des parties spatiales \vec{x}' et \vec{x} d'un même 4-vecteur \underline{x} , observées dans R et R' respectivement, tandis que y' et y ainsi que z' et z en sont des projections dans des directions perpendiculaires à \vec{n} . Ceci nous permet de récrire les relations (3.25) sous les formes suivantes :

$$x'_0 = x_0 \cosh \chi - (\vec{n} \cdot \vec{x}) \sinh \chi \quad (3.153)$$

pour la composante temporelle et

$$\vec{x}' \cdot \vec{n} = (\vec{n} \cdot \vec{x}) \cosh \chi - x_0 \sinh \chi \quad (3.154)$$

pour la composante "longitudinale" ; pour les composantes "transversales", nous utiliserons les composantes des produits vectoriels $\vec{x}' \wedge \vec{n}$ et $\vec{x} \wedge \vec{n}$

$$y'n_z - z'n_y = yn_z - zn_y, \quad z'n_x - x'n_z = zn_x - xn_z, \quad x'n_y - y'n_x = xn_y - yn_x \quad (3.155)$$

où x' , y' , z' d'une part et x , y , z sont les nouvelles composantes spatiales dans R' et R respectivement, ces repères étant maintenant munis de leurs nouveaux axes. Les relations

$$\begin{aligned} x'n_x + y'n_y + z'n_z &= -x_0 \sinh \chi + (xn_x + yn_y + zn_z) \cosh \chi \\ y'n_z - z'n_y &= yn_z - zn_y \\ z'n_x - x'n_z &= zn_x - xn_z \end{aligned} \quad (3.156)$$

constituent un système d'équations permettant d'obtenir x' , y' et z' en fonction de x_0 , x , y et z (à noter que seulement deux des trois relations (3.155) sont indépendantes, du fait que les produits vectoriels sont orthogonaux à \vec{n}). On trouve ainsi

$$\begin{aligned}
 x' &= x [1 + n_x^2 (\cosh \chi - 1)] + y n_x n_y (\cosh \chi - 1) + z n_x n_z (\cosh \chi - 1) - n_x x_0 \sinh \chi \\
 y' &= x n_y n_x (\cosh \chi - 1) + y [1 + n_y^2 (\cosh \chi - 1)] + z n_y n_z (\cosh \chi - 1) - n_y x_0 \sinh \chi \\
 z' &= x n_z n_x (\cosh \chi - 1) + y n_z n_y (\cosh \chi - 1) + z [1 + n_z^2 (\cosh \chi - 1)] - n_z x_0 \sinh \chi
 \end{aligned} \quad (3.157)$$

L'écriture matricielle de la transformation fait ainsi intervenir la matrice de boost (3.39). Un boost est donc une transformation de Lorentz ne faisant pas intervenir de matrice de rotation d'axes. On dit aussi que l'on a affaire à une transformation de Lorentz *pure*. On remarque à l'occasion qu'une matrice de boost est *symétrique*.

Considérons plus particulièrement le cas où $n_x = \sin \theta$, $n_y = 0$, $n_z = \cos \theta$. La matrice de boost correspondante s'écrit

$$B = \begin{pmatrix} \cosh \chi & -\sin \theta \sinh \chi & 0 & -\cos \theta \sinh \chi \\ -\sin \theta \sinh \chi & 1 + \sin^2 \theta (\cosh \chi - 1) & 0 & \sin \theta \cos \theta (\cosh \chi - 1) \\ 0 & 0 & 1 & 0 \\ -\cos \theta \sinh \chi & \sin \theta \cos \theta (\cosh \chi - 1) & 0 & 1 + \cos^2 \theta (\cosh \chi - 1) \end{pmatrix} \quad (3.158)$$

Cette matrice est en fait le produit de trois matrices :

$$B = R^{-1}(\theta)H(\chi)R(\theta) \quad (3.159)$$

où

$$H(\chi) = \begin{pmatrix} \cosh \chi & 0 & 0 & -\sinh \chi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh \chi & 0 & 0 & \cosh \chi \end{pmatrix} \quad (3.160)$$

est la matrice d'un boost de rapidité χ effectué le long de l'axe des z , et

$$R(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & 0 & -\sin \theta \\ 0 & 0 & 1 & 0 \\ 0 & \sin \theta & 0 & \cos \theta \end{pmatrix} \quad (3.161)$$

est la matrice d'une rotation spatiale d'angle θ effectuée autour de l'axe des y . On voit alors clairement comment agit un boost : la première rotation $R(\theta)$ transforme le vecteur unitaire $(\sin \theta, \cos \theta)$ de la vitesse relative en vecteur unitaire de l'axe des z . Le boost $H(\chi)$ effectué suivant l'axe des z assure le changement effectif de référentiel. La seconde rotation $R^{-1}(\theta)$ permet de remettre les axes spatiaux en position, de telle sorte que la vitesse relative ait la même orientation dans le nouveau repère d'espace. Au final, il n'y aura pas eu de rotation supplémentaire des axes²⁸.

Par ailleurs, la décomposition (3.159) met en évidence deux faits ; d'une part, celui, déjà connu, que les rotations doivent aussi faire partie du groupe de Lorentz ; d'autre part, puisque

$$R^{-1}(\theta)H(\chi)R(\theta) \neq H(\chi), \quad \text{soit} \quad R(\theta)H(\chi) \neq H(\chi)R(\theta) \quad (3.162)$$

on constate que les matrices $H(\chi)$ et $R(\theta)$ *ne commutent pas*, ce qui illustre une nouvelle fois le fait que le groupe de Lorentz n'est pas abélien.

²⁸. A noter que puisque l'orthogonalité spatiale n'est plus préservée dans les transformations de Lorentz, pour un observateur de R , les nouveaux axes d'espace de R' n'apparaissent généralement plus ni orthogonaux entre eux, ni parallèles aux axes d'espace de R . Pour ce qui concerne les directions spatiales, seules la direction de la vitesse relative et toute autre direction qui lui est perpendiculaire sont vues de la même façon par R et par R' .

Notons que les relations (3.157) peuvent être obtenues plus rapidement comme suit. L'ensemble des trois composantes x' , y' et z' peuvent servir à définir dans la base spatiale (X, Y, Z) de \mathcal{R} le 3-vecteur fictif

$$\vec{x}' = x'X + y'Y + z'Z \quad (3.163)$$

dont la partie \vec{x}'_{\perp} perpendiculaire à \vec{n} est égale à celle, \vec{x}_{\perp} , de \vec{x} . On a donc

$$\vec{x}'_{\perp} = \vec{x}' - (\vec{x}' \cdot \vec{n}) \vec{n} = \vec{x}_{\perp} = \vec{x} - (\vec{x} \cdot \vec{n}) \vec{n} \quad (3.164)$$

Utilisant (3.154), il vient

$$\vec{x}' = \vec{x} + (\cosh \chi - 1)(\vec{x} \cdot \vec{n}) \vec{n} - x_0 \sinh \chi \vec{n} \quad (3.165)$$

Une transformation de Lorentz correspond aussi à un changement de base d'espace-temps, la nouvelle base étant obtenue au moyen de la matrice inverse de la matrice de transformation des composantes. Notant E_0, X, Y et Z la base de 4-vecteurs associée à \mathcal{R} et n, e_x, e_y et e_z , celle associée à \mathcal{R}' et qui est déduite par le boost, on a²⁹

$$\begin{aligned} n &= \cosh \chi E_0 + \sinh \chi (n_x X + n_y Y + n_z Z) \\ e_x &= n_x \sinh \chi E_0 + X [1 + n_x^2 (\cosh \chi - 1)] + Y n_x n_y (\cosh \chi - 1) + Z n_x n_z (\cosh \chi - 1) \\ e_y &= n_y \sinh \chi E_0 + X n_y n_x (\cosh \chi - 1) + Y [1 + n_y^2 (\cosh \chi - 1)] + Z n_y n_z (\cosh \chi - 1) \\ e_z &= n_z \sinh \chi E_0 + X n_z n_x (\cosh \chi - 1) + Y n_z n_y (\cosh \chi - 1) + Z [1 + n_z^2 (\cosh \chi - 1)] \end{aligned} \quad (3.166)$$

On peut généraliser (3.165) aux quatre composantes x'_0, x', y' et z' en considérant le 4-vecteur fictif

$$\underline{x}' = x'_0 E_0 + x'X + y'Y + z'Z \quad (3.167)$$

Nous laissons au lecteur le soin de démontrer la relation

$$\underline{x}' = \underline{x} + E_0 \frac{2x'_0 \cosh \chi + x'_0 - x_0}{\cosh \chi + 1} - n \frac{x'_0 + x_0}{\cosh \chi + 1} \quad (3.168)$$

qui peut d'ailleurs se généraliser à tout boost transformant un 4-vecteur du genre temps unitaire p en un 4-vecteur du genre temps unitaire q par

$$x' = x + p \frac{2(x \cdot q)(p \cdot q) + x \cdot q - x \cdot p}{p \cdot q + 1} - q \frac{x \cdot q + x \cdot p}{p \cdot q + 1} \quad (3.169)$$

où, au second membre, x est le 4-vecteur dont on transforme les composantes. A l'aide de cette relation, le lecteur vérifiera aussi que dans une transformation infinitésimale $p \rightarrow p + d p$, on a

$$dx = x' - x = (p \wedge dp) \cdot x \quad (3.170)$$

29. Vérifier qu'il s'agit bien d'une base "orthonormée".

3.6.2 Produits de boosts

Considérons trois référentiels. Le premier, R_0 (observateur O), est un référentiel "de référence"³⁰. Ses axes ont pour 4-vecteurs unitaires respectifs $n_0 \equiv (1, 0, 0, 0)$, $X \equiv (0, 1, 0, 0)$, $Y \equiv (0, 0, 1, 0)$ et $Z \equiv (0, 0, 0, 1)$. Le second, R_1 (observateur O_1), est animé d'une vitesse de module v_1 et orientée suivant l'axe des z de R_0 . Le troisième référentiel R_2 (observateur O_2) se déplace dans le plan xz de R_0 avec une vitesse de module v_2 et faisant l'angle θ avec l'axe des z de R_0 . Exprimés dans la base de R_0 , les 4-vecteurs unitaires du genre temps associés respectivement à R_1 et à R_2 sont $n_1 = (\cosh \alpha_1, 0, 0, \sinh \alpha_1)$ et $n_2 = (\cosh \alpha_2, \sin \theta \sinh \alpha_2, 0, \cos \theta \sinh \alpha_2)$ avec $\cosh \alpha_1 = 1/\sqrt{1 - v_1^2/c^2}$ et $\cosh \alpha_2 = 1/\sqrt{1 - v_2^2/c^2}$.

Pour un référentiel donné dont l'axe des temps a pour 4-vecteur unitaire n , il existe une infinité de possibilités pour le choix de sa triade de base spatiale. Le passage d'une triade directe ou indirecte à une autre orientée de la même façon s'effectue au moyen d'une rotation qui laisse invariant le 4-vecteur n . L'ensemble de ces rotations qui laissent n invariant forment un groupe appelé le *petit groupe* de n . Pour les trois référentiels considérés, nous ferons en sorte que leurs axes des y soient les mêmes, ce qui nous autorisera à omettre dans les matrices la ligne et la colonne concernant cette coordonnée. Concernant R_1 , nous définirons sa base de telle sorte qu'elle soit reliée à celle de R_0 par le simple boost

$$B_{01} = H(\alpha_1) = \begin{pmatrix} \cosh \alpha_1 & 0 & -\sinh \alpha_1 \\ 0 & 1 & 0 \\ -\sinh \alpha_1 & 0 & \cosh \alpha_1 \end{pmatrix} \quad (3.171)$$

Exprimés dans la base de R_0 , on a ainsi $X_1 = X$, $Y_1 = Y$, $Z_1 = (\sinh \alpha_1, 0, 0, \cosh \alpha_1)$.

Concernant R_2 , un premier choix de base spatiale de R_2 consiste à effectuer un boost à partir de R_1 , selon la direction de la vitesse de O_2 dans R_1 . Ceci nécessite de connaître l'orientation de la partie spatiale de n_2 dans R_1 , donc, finalement, les composantes de ce 4-vecteur relativement à R_1 . Celles-ci sont obtenues à partir des composantes dans R_0 en appliquant le boost B_{01} . On obtient ainsi

$$n_2 = \cosh \chi n_1 + \sin \psi \sinh \chi X_1 + \cos \psi \sinh \chi Z_1 \quad (3.172)$$

avec³¹

$$\begin{aligned} \cosh \chi &= n_2 \cdot n_1 = \cosh \alpha_1 \cosh \alpha_2 - \cos \theta \sinh \alpha_1 \sinh \alpha_2 \\ \sin \psi \sinh \chi &= \sin \theta \sinh \alpha_2 \\ \cos \psi \sinh \chi &= \cosh \alpha_1 \sinh \alpha_2 \cos \theta - \sinh \alpha_1 \cosh \alpha_2 \end{aligned} \quad (3.173)$$

L'angle ψ définit l'orientation de la vitesse de O_2 dans le référentiel R_1 , tandis que le module V de cette vitesse est donné par $V = c \tanh \chi$. Appliquant alors à partir de R_1 le boost³²

$$\begin{aligned} B_{12} &= \begin{pmatrix} \cosh \chi & -\sin \psi \sinh \chi & -\cos \psi \sinh \chi \\ -\sin \psi \sinh \chi & 1 + \sin^2 \psi (\cosh \chi - 1) & \sin \psi \cos \psi (\cosh \chi - 1) \\ -\cos \psi \sinh \chi & \sin \psi \cos \psi (\cosh \chi - 1) & 1 + \cos^2 \psi (\cosh \chi - 1) \end{pmatrix} \\ &= R^{-1}(\psi)H(\chi)R(\psi) \end{aligned} \quad (3.174)$$

30. S'appelant souvent *référentiel du laboratoire*.

31. A vérifier.

32. Les matrices de rotations $R(\zeta)$ ont la même forme que celle donnée en (3.161) où la ligne et la colonne relatives à y ont été supprimées.

on trouve la base

$$\begin{aligned} X'_2 &= \sin \psi \sinh \chi n_1 + X_1 [1 + \sin^2 \psi (\cosh \chi - 1)] + Z_1 \sin \psi \cos \psi (\cosh \chi - 1) \\ Z'_2 &= \cos \psi \sinh \chi n_1 + X_1 \sin \psi \cos \psi (\cosh \chi - 1) + Z_1 [1 + \cos^2 \psi (\cosh \chi - 1)] \end{aligned} \quad (3.175)$$

Nous noterons R'_2 le troisième référentiel défini par la base (3.175). A partir de R'_2 , le retour vers R_0 doit se faire en trois étapes. La première consiste à tourner la base spatiale de R'_2 de telle sorte à faire coïncider le nouvel axe des z avec la direction de la vitesse de O , telle que celle-ci est vue dans R'_2 . Pour obtenir cette direction, il suffit d'appliquer le boost B_{12} aux composantes du 4-vecteur n_0 dans R_1 , c'est-à-dire, au 4-vecteur $n_0 = (\cosh \alpha_1, 0, 0, -\sinh \alpha_1)$. On trouve³³

$$\begin{aligned} (n_0)_{R'_2} &= (\cosh \alpha_2, -\sinh \alpha_2 \sin \theta' - \sinh \alpha_2 \cos \theta'), \text{ avec} \\ \cos \theta' &= \frac{1}{1 + \cosh \chi} ((1 + \cosh \alpha_1 \cosh \alpha_2) \cos \theta - \sinh \alpha_1 \sinh \alpha_2) \\ \sin \theta' &= \frac{\sin \theta}{1 + \cosh \chi} (\cosh \alpha_1 + \cosh \alpha_2) \end{aligned} \quad (3.176)$$

L'observateur effectuera donc tout d'abord une rotation d'angle θ' autour de Y . Pour retrouver un référentiel R'_0 lié à O , il appliquera dans une seconde étape le boost inverse du boost

$$H(\alpha_2) = \begin{pmatrix} \cosh \alpha_2 & 0 & -\sinh \alpha_2 \\ 0 & 1 & 0 \\ -\sinh \alpha_2 & 0 & \cosh \alpha_2 \end{pmatrix} \quad (3.177)$$

qui est effectué selon la direction de la vitesse relative de O et O_2 . De façon évidente, les axes X' et Z' de ce référentiel ont, pour O , tourné d'un angle θ par rapport à ceux de R_0 , si bien que pour retrouver complètement R_0 , une rotation d'angle $-\theta$ autour de Y s'avère nécessaire.

Pour passer de R'_2 à R_0 , il faut donc appliquer la transformation $R^{-1}(\theta)H^{-1}(\alpha_2)R(\theta')$. On en déduit la relation

$$R^{-1}(\theta)H^{-1}(\alpha_2)R(\theta')B_{12}B_{01} = 1 \quad (3.178)$$

Or, la transformation $R^{-1}(\theta')H^{-1}(\alpha_2)R(\theta') = [B'_{02}]^{-1}$ est un boost qui fait passer directement le référentiel R'_2 à un référentiel R'_0 lié à O , dont la base spatiale fait un angle $\eta = \theta - \theta'$ avec celle de R_0 . Il est facile d'obtenir les relations

$$\begin{aligned} \sin \eta &= \frac{\sin \theta}{1 + \cosh \chi} [\cos \theta (\cosh \alpha_1 - 1)(\cosh \alpha_2 - 1) - \sinh \alpha_1 \sinh \alpha_2] \\ \cos \eta &= 1 - \frac{\sin^2 \theta}{1 + \cosh \chi} (\cosh \alpha_1 - 1)(\cosh \alpha_2 - 1) \end{aligned} \quad (3.179)$$

et (3.178) conduit à

$$B_{12}B_{01} = B'_{02} R(\eta) \quad (3.180)$$

Cette dernière relation montre clairement que le produit de deux boosts de *directions différentes* n'est pas un boost, mais une transformation de Lorentz plus générale, produit d'un boost et d'une rotation. Il en résulte que la succession des boosts B_{01} , B_{12} puis $[B'_{02}]^{-1}$ ne redonne pas le référentiel de départ

33. Vérifier que la relation $\cos^2 \theta' + \sin^2 \theta' = 1$ est bien satisfaite.

R_0 , mais un référentiel R'_0 qui en diffère par une rotation. Ce résultat s'exprime aussi de la façon suivante.

A partir de R_0 , effectuons un boost selon la direction de la vitesse de O_2 . La matrice correspondante s'écrit

$$B_{02} = \begin{pmatrix} \cosh \alpha_2 & -\sin \theta \sinh \alpha_2 & -\cos \theta \sinh \alpha_2 \\ -\sin \theta \sinh \alpha_2 & 1 + \sin^2 \theta (\cosh \alpha_2 - 1) & \sin \theta \cos \theta (\cosh \alpha_2 - 1) \\ -\cos \theta \sinh \alpha_2 & \sin \theta \cos \theta (\cosh \alpha_2 - 1) & 1 + \cos^2 \theta (\cosh \alpha_2 - 1) \end{pmatrix} \quad (3.181)$$

$$= R^{-1}(\theta)H(\alpha_2)R(\theta)$$

Les 4-vecteurs des axes x et z du nouveau référentiel R_2 attaché à O_2 et obtenu de cette manière directe sont

$$X_2 = \sin \theta \sinh \alpha_2 E_0 + X [1 + \sin^2 \theta (\cosh \alpha_2 - 1)] + Z \sin \theta \cos \theta (\cosh \alpha_2 - 1)$$

$$Z_2 = \cos \theta \sinh \alpha_2 E_0 + X \sin \theta \cos \theta (\cosh \alpha_2 - 1) + Z [1 + \cos^2 \theta (\cosh \alpha_2 - 1)] \quad (3.182)$$

Relativement à cette base, la direction de la vitesse de O fait l'angle θ avec l'axe des z , alors qu'elle fait l'angle θ' avec l'axe des z de (3.175). Pour passer de R_2 à R'_2 il faut donc effectuer une rotation d'angle $-(\theta' - \theta) = \eta$. On en déduit

$$B_{12} B_{01} = R(\eta) B_{02} \quad (3.183)$$

relation que l'on pouvait déduire directement de (3.180) puisque

$$R(\eta) B_{02} = R(\theta - \theta')R(-\theta)H(\alpha_2)R(\theta) = B'_{02}R(\eta) \quad (3.184)$$

Pour un 4-vecteur A quelconque, on obtient

$$(A)_{R'_2} = B_{12} (A)_{R_1} = B_{12} B_{01} (A)_{R_0} = B_{12} B_{01} [B_{02}]^{-1} (A)_{R_2} = R(\eta) (A)_{R_2} \quad (3.185)$$

La différence des chemins de comparaison $R_0 \rightarrow R_1 \rightarrow R'_2$ et $R_0 \rightarrow R_2$ induit donc la rotation supplémentaire $R(\eta)$.

3.6.3 Précession de Thomas

Envisageons le mouvement varié d'une particule P dans le référentiel du laboratoire R_0 (observateur O), supposé galiléen. Cette particule possède certaines propriétés définies de façon univoque dans son référentiel propre. Pour l'observateur O , la façon la plus directe d'établir un pont entre ces propriétés et leur observation dans R_0 serait de faire appel à des boosts successifs reliant R_0 à un référentiel où la particule est au repos, en suivant le mouvement de celle-ci. Cependant, la particule ayant un mouvement a priori non uniforme, ce "référentiel propre" n'est pas galiléen. Pour pouvoir néanmoins utiliser les transformations de Lorentz, on imaginera une infinité de référentiels galiléens "tangents" dont les vitesses dans R_0 sont précisément les valeurs prises par la vitesse de la particule aux différentes dates. Ainsi, si $\vec{v}(t_1) = \vec{v}_1$ est la vitesse de la particule à la date t_1 , on imaginera un référentiel galiléen R_1 (observateur O_1) se mouvant à cette vitesse dans R_0 et tel que, à la date t_1 , la particule y sera vue au repos. A une date ultérieure t_2 , la particule sera vue au repos dans un autre référentiel galiléen R_2 (observateur O_2), dont la vitesse dans R_0 est $\vec{v}_2 = \vec{v}(t_2)$.

Les deux vecteurs \vec{v}_1 et \vec{v}_2 formant un plan, nous choisirons pour les trois référentiels R_0 , R_1 et R_2 un axe des y commun, perpendiculaire à ce plan. Comme au paragraphe précédent, les référentiels R_0 et R_1 auront un axe des z commun, orienté selon la vitesse \vec{v}_1 , et des axes des x perpendiculaires à cette vitesse, dans le plan formé par \vec{v}_1 et \vec{v}_2 .

Pour l'observateur O , la comparaison, au cours du temps, des données prises dans son référentiel R_0 et celles provenant des référentiels successifs R_i repose sur la relation

$$(A)_{R_i}(\tau_i) = B_{0i}(A)_{R_0}(t_i) \quad (3.186)$$

où τ_i est un temps propre, et B_{0i} le boost reliant R_0 au référentiel R_i de vitesse $\vec{v}_i = \vec{v}(t_i)$ dans R_0 , dont la matrice est de la forme (3.181), et qui n'induit pas de rotation supplémentaire des axes.

Toutefois, les orientations des référentiels R_i obtenus par ces boosts à partir de R_0 ne correspondent pas au protocole naturel d'échange d'information entre les divers observateurs O_i . Ayant fait un premier choix de référentiel propre de la particule pour un temps propre donné, ce sera, en l'occurrence, le référentiel R_1 correspondant à la date t_1 de R_0 , l'échange procèdera plutôt par boosts successifs B_{ij} d'un référentiel R'_i à un référentiel infiniment voisin R'_j , le long de leur vitesse relative, telle qu'ils la perçoivent, et ce, en prenant R_1 pour référence de départ du processus. Le référentiel propre de la particule sera défini par cette succession de référentiels R'_i . Il est important de noter ici encore que l'utilisation de boosts exclut des rotations supplémentaires des axes.

Le transfert d'information vers R_0 s'effectuera donc plutôt depuis ces référentiels R'_i . D'après la relation (3.185), ceci induit une rotation supplémentaire d'angle η_i , et la relation (3.186) devra être changée en

$$(A)_{R_0}(t_i) = [B_{0i}]^{-1} R^{-1}(\eta_i) (A)_{R'_i}(\tau_i) = R^{-1}(\eta_i) [B'_{0i}]^{-1} (A)_{R'_i}(\tau_i) \quad (3.187)$$

Dans cette relation, le boost $[B'_{0i}]^{-1}$ qui transfère des informations de R'_i vers l'observateur O n'induit pas de nouvelle rotation d'axe. Mais, du fait de la présence de la rotation $R^{-1}(\eta_i)$ l'observateur O conclura néanmoins que les axes du référentiel propre ont tourné d'un angle η_i . Pour un mouvement varié de la particule, cette rotation évolue au cours du temps³⁴. Ce mouvement de rotation est appelé *précession de Thomas*. Il est observé dans le référentiel du laboratoire dès lors que l'objet étudié (particule) y est en mouvement accéléré. C'est un effet de pure cinématique relativiste que l'on peut imputer au fait que, d'une façon générale, les boosts et les rotations ne commutent pas.

Imaginons que dans R_0 la particule effectue un mouvement plan, perpendiculairement à un axe fixe de vecteur unitaire \vec{j} et que la particule possède une propriété vectorielle, telle un éventuel vecteur de spin \vec{s} . Alors, indépendamment de tout autre type de variation, un observateur de R_0 verra ce vecteur tourner autour de l'axe \vec{j} . La rotation est associée à un vecteur rotation instantanée tel que $\vec{\Omega}_T \delta t = \Omega_T \vec{j} \delta t = \delta\eta \vec{j}$ que l'on peut déduire comme suit. D'une part, la variation d'un vecteur, due à la rotation infinitésimale $R(\delta\eta)$, s'écrit

$$\delta \vec{A} = \delta t \vec{\Omega}_T \wedge \vec{A} = \Omega_T \delta t (A_z \vec{i} - A_x \vec{k}) \quad (3.188)$$

D'autre part, comme $\delta A_x = -\delta\eta A_z$, $\delta A_z = \delta\eta A_x$, on obtient $\Omega_T = \dot{\eta}$, soit, en utilisant (3.179) avec $\alpha_2 \approx \alpha_1$, $\cosh \chi \approx 1$,

$$\Omega_T = -\dot{\theta}(\cosh \alpha_1 - 1) \quad (3.189)$$

Enfin, un calcul simple de cinématique permet d'écrire

$$\dot{\theta} \vec{j} = \frac{\vec{v} \wedge \vec{a}}{v^2} \quad (3.190)$$

34. Dans les cas les plus compliqués, l'axe de rotation peut d'ailleurs lui aussi évoluer au cours du temps.

où \vec{a} est l'accélération de la particule dans R_0 . Comme $v^2 = c^2 \tanh^2 \alpha_1$, on obtient ainsi

$$\vec{\Omega}_T = -\frac{\vec{v} \wedge \vec{a}}{c^2} \frac{\cosh^2 \alpha_1}{1 + \cosh \alpha_1} \quad (3.191)$$

où l'on voit que ladite précession apparaît dès que l'accélération de la particule a une composante perpendiculaire à la vitesse.

Cet effet de précession a été pour la première fois suggéré sur le plan théorique par L. Thomas en 1926³⁵. Avant même l'avènement de la théorie relativiste de l'électron de Dirac, Thomas a montré que l'inclusion de cet effet dans la théorie du spin de l'électron élaborée par Uhlenbeck et Goudsmit³⁶ permettait de rendre compte de façon correcte et simultanément des résultats expérimentaux concernant l'effet Zeeman anormal, la structure fine des raies d'émission des atomes, et le couplage "spin-orbite" des électrons atomiques.

35. L. H. Thomas, "Motion of the spinning electron", Nature 117, 514, 1926; Phil. Mag. 3, 1 (1927).

36. G.E. Uhlenbeck, S. Goudsmit, Naturwissenschaften 47 (1925) 953; Nature 117 (1926) 264.