

Une dérivation de plus des transformations de Lorentz

Jean-Marc Lévy-Leblond

*Laboratoire de Physique Théorique et Hautes Energies, Université Paris VII, 75221
Paris Cedex 05, France*

(Reçu le 30 Janvier 1975 ; revu le 30 Avril 1975)

Traduit de l'anglais par Olivier Castéra

Après avoir critiquer l'accent mis sur l'invariance de la vitesse de la lumière dans la dérivation standard des transformations de Lorentz, une nouvelle approche de la relativité restreinte est proposée. Elle consiste en une version élémentaire d'arguments généraux sur la structure de l'espace-temps, issus de la théorie des groupes, et utilise seulement des techniques mathématiques simples. Le *principe* de relativité est d'abord énoncé en des termes généraux, amenant l'idée de référentiels équivalents, connectés par des transformations « inertielles » obéissant à une loi de groupe. La *théorie* de la relativité est alors construite en contraignant les transformations au travers de quatre hypothèses successives : homogénéité de l'espace-temps, isotropie de l'espace-temps, structure de groupe, condition de causalité. Seules les transformations de Lorentz et leur limite dégénérée Galiléenne obéissent à ces contraintes. Le rôle et la signification de chacune de ces hypothèses sont mis en évidence en montrant et en discutant de contre-exemples, c'est à dire, de transformations obéissant à toutes les hypothèses sauf une.

INTRODUCTION

Un grand nombre de dérivations de la transformation de Lorentz a déjà été donné, et le sujet, à cause de son importance pédagogique, continue de recevoir de l'attention, en particulier dans ce journal¹. La plupart de ces analyses, suivant celle originale d'Einstein², repose sur l'invariance de la vitesse de la lumière c comme hypothèse centrale. On ne peut nier que cette hypothèse, fermement établie sur des bases expérimentales, a eu un rôle historique crucial. Quoiqu'il en soit, la construction chronologique d'une théorie physique coïncide rarement avec sa structure logique. A partir de ce point de vue, je vais critiquer la trop grande importance donnée à la vitesse de la lumière dans les fondements de la relativité restreinte, et proposer une approche de ces fondements qui dispense de l'hypothèse de l'invariance de c . En établissant la relativité restreinte sur la propriété de la vitesse de la lumière, il semble que l'on relie cette théorie à une classe restreinte de phénomènes naturels que sont les radiations électromagnétiques. Quoiqu'il en soit, la leçon qu'il faut tirer depuis plus d'un demi siècle est que la relativité restreinte, jusqu'à présent, semble s'appliquer à *toutes* les classes de phénomènes naturels, qu'ils dépendent des interactions électromagnétiques, faibles, fortes, ou même gravitationnelles³. Cette théorie ne dérive pas de l'utilisation de signaux électromagnétiques pour synchroniser des horloges, par exemple, comme une lecture ultra-positiviste des écrits² d'Einstein pourrait nous faire croire ; au contraire, c'est la validité de la théorie qui contraint les signaux électromagnétiques à avoir leurs propriétés spécifiques de propagation. Nous pensons qu'aujourd'hui, la relativité restreinte se présente comme une théorie universelle décrivant la structure d'une arène commune de l'espace-temps, dans laquelle tous les processus fondamentaux prennent place. Toutes les lois de la physique sont

contraintes par la relativité restreinte, qui agit comme une sorte de « super loi⁴ », et les interactions électromagnétiques n'ont ici aucun privilège autre que celui historique et anthropocentrique. La théorie de la relativité, en fait, n'est autre que l'énoncé que toutes les lois de la physique sont invariantes sous le groupe de Poincaré (groupe de Lorentz inhomogène). La condition d'invariance, appliquée à la classification des particules fondamentales possibles⁵, permet mais ne nécessite pas l'existence d'objets de masse nulle⁶. La mise en évidence d'une masse non nulle pour le photon, ne ferait en aucun cas, en tant que telle, trembler la validité de la relativité restreinte. Cependant, cela annulerait les dérivations qui sont basées sur l'invariance de la vitesse du photon⁷. Par chance, on peut montrer qu'en partant de quelques hypothèses générales sur les propriétés de l'espace-temps, très peu de liberté est permise pour le choix d'un groupe de relativité, et que le groupe de Poincaré (ou le groupe de Galilée, sa limite singulière) est presque l'unique solution du problème⁸. Ces analyses utilisent des outils plus ou moins élaborés de la théorie des groupes, trop abstraits en général pour des objectifs didactiques⁹. C'est pourquoi je pense qu'il est utile d'offrir ici une version simple de ces arguments, reposant seulement sur des calculs assez élémentaires. La discussion sera réduite au cas uni-dimensionnel (dans l'espace).

PRINCIPE DE RELATIVITÉ : RÉFÉRENTIELS D'INERTIE ET TRANSFORMATIONS

Je vais prendre comme point de départ l'énoncer du *principe de relativité* dans un cadre très général : il existe une infinie continue classe de référentiels dans l'espace-temps qui sont physiquement équivalents. En d'autres termes, les lois de la physique prennent la même forme quand on les réfère à n'importe lequel de ces repères, et aucun effet physique ne permet de les distinguer entre eux. Cela ne signifie pas, bien sûr, que les quantités physiques ont les mêmes valeurs dans chacun de ces référentiels : seules les relations entre eux sont invariantes. Le principe abstrait de relativité est donc a priori ouvert à de nombreuses réalisations concrètes de *théories de la relativité*. Une théorie de la relativité nous dit comment relier deux expressions de la même quantité physique référencée dans deux de ces repères équivalents ; une telle théorie devrait alors s'exprimer exactement par les « formules de transformation » qui connectent les différents repères équivalents. Une théorie de la relativité limite alors les formes possibles des lois physiques qui relient différentes quantités physiques d'un repère choisi ; cela nécessite que le même rapport existe dans différents repères par l'utilisation des formules de transformation. Pour des considérations physiques bien connues, j'ai trouvé commode d'appeler « repère d'inertie » et « transformations inertielles » les référentiels équivalents et les transformations les reliant. En effet, la véritable existence de ces référentiels équivalents correspond à la validité du principe d'inertie, à savoir, qu'un objet physique n'a pas d'état de mouvement absolu ou de repos ; par exemple, un corps libre (sur lequel aucune force n'agit) est caractérisé par un « mouvement inertiel » qui n'est pas entièrement déterminé, puisqu'il dépend de « conditions initiales », c'est à dire, du référentiel considéré.

Il est assez naturel, quand on essaye d'établir la nature d'une transformation inertielle, de considérer les formules de transformation pour des quantités physiques spécifiques, à savoir, les coordonnées spatio-temporelles (x, t) d'un événement dans un référentiel d'inertie. Donc, puisque nous avons supposé l'existence d'une classe infinie continue de référentiels d'inertie, le rapport entre deux d'entre eux dépend d'un certain

nombre de paramètres $\{a_1, a_2, \dots, a_N\}$, dont les valeurs caractérisent toute transformation inertielle particulière. En notant (x', t') les coordonnées du même évènement dans un autre référentiel d'inertie, on écrit les transformations inertielles reliant ces deux ensembles de coordonnées par la forme générale

$$\begin{aligned}x' &= f(x, t; a_1, \dots, a_N), \\t' &= g(x, t; a_1, \dots, a_N).\end{aligned}\tag{1}$$

Nous faisons maintenant appel à l'existence de deux classes particulières de transformations d'inertie, à savoir les translations d'espace et de temps, pour distinguer deux des paramètres $\{a\}$ et les transformations inertielles associées. En effet, les transformations

$$\begin{aligned}x' &= x + \xi, \\t' &= t + \tau,\end{aligned}\tag{2}$$

se résumant à un simple déplacement dans l'espace et le temps, sont supposées laisser invariantes les lois de la physique; les origines de l'espace et du temps pouvant être choisies arbitrairement. Utilisant cette liberté, à partir de maintenant nous limitons notre attention sur les classes de référentiels inertiels ayant des origines d'espace-temps communes. Se faisant, nous disposons de deux de nos paramètres $\{a_1, a_2, \dots, a_N\}$ —précisément ceux que l'on appelle ξ et τ dans (2). Il reste $n = N - 2$ paramètres et formules de transformation :

$$\begin{aligned}x' &= F(x, t; a_1, \dots, a_n), \\t' &= G(x, t; a_1, \dots, a_n),\end{aligned}\tag{3}$$

telles que $x' = 0, t' = 0$ si $x = 0, t = 0$; c'est à dire,

$$\begin{aligned}0 &= F(0, 0; a_1, \dots, a_n), \\0 &= G(0, 0; a_1, \dots, a_n).\end{aligned}\tag{4}$$

Nous devons maintenant regarder (3) comme donnant les formules de transformation de *l'intervalle* spatio-temporel entre les évènements de coordonnées (x, t) et l'évènement localisé à l'origine (commune) de l'espace et du temps du référentiel d'inertie. Demandons-nous si cet intervalle, de coordonnées (x, t) , peut avoir des coordonnées (x', t') dans un autre référentiel d'inertie; ce qui équivaut à se demander s'il existe un ensemble de paramètres $\{a_1, a_2, \dots, a_N\}$ tels que (3) existe bien. Autrement dit, on considère (3) comme donnant (x, t) et (x', t') en tant qu'ensemble de deux équations de n inconnues $\{a_1, a_2, \dots, a_n\}$. Il est clair que si $n \geq 2$, ces équations auront en général des solutions; un intervalle entre deux évènements physiques aurait alors des coordonnées arbitraires dans un référentiel d'inertie approprié, ce qui va à l'encontre de notre connaissance de la physique. En particulier, l'arbitraire sur la coordonnée de temps semble écarter toute notion raisonnable de causalité. D'autre part, si $n = 0$, il n'y aurait de transformations inertielles autre que les translations d'espace et de temps, et par conséquent aucune théorie de la relativité. De cet argument on conclue que $n = 1$; c'est à dire, la transformation inertielle entre référentiel ayant une origine commune ne dépend que d'un seul paramètre, et doit s'écrire

$$\begin{aligned}x' &= F(x, t; a), \\t' &= G(x, t; a),\end{aligned}\tag{5}$$

avec la condition correspondant à (4) :

$$\begin{aligned} 0 &= F(0, 0; a), \\ 0 &= G(0, 0; a). \end{aligned} \tag{6}$$

Un autre argument menant au même résultat serait le suivant. Considérons un objet se déplaçant avec pour équation du mouvement $x = \varphi(t)$ dans le premier référentiel d'inertie, et passant par l'origine de l'espace-temps, choisie de sorte que $\varphi(0) = 0$. Son mouvement, considéré dans le second référentiel, obtenu par les transformations d'inertie (3), sera de la forme $x' = \varphi'(t'; a_1, \dots, a_n)$, avec $0 = \varphi'(0; a_1, \dots, a_n)$ d'après (4). Sa vitesse, accélération, et dérivées supérieures par rapport au temps de sa position sont obtenues en différenciant φ' . Elles dépendront de la vitesse, de l'accélération, etc., de l'objet à l'origine du premier référentiel d'inertie et des n paramètres $\{a_1, \dots, a_n\}$. Au contraire, pour un mouvement donné, dans un référentiel donné, les paramètres $\{a_1, \dots, a_n\}$ pourraient être choisis de telle sorte que l'on puisse trouver un autre référentiel dans lequel l'objet aurait des valeurs arbitrairement assignées de sa vitesse, de son accélération, et des dérivées supérieures par rapport au temps jusqu'à l'ordre n de sa position. Nous savons en effet, par une simple expérience de physique, que la vitesse est seulement relative et peut varier d'un référentiel d'inertie à un autre; ceci étant la base empirique du principe de relativité. Nous savons toutefois, que ceci n'est pas vrai pour l'accélération, qui est associée aux effets physiques différenciant les nombreux référentiels. Il s'ensuit que $n = 1$.

Je vais maintenant dériver une forme fonctionnelle précise des transformations inertielles (5) en me servant d'une succession d'hypothèses simples et générales de la physique. J'annonce dès à présent que la clef de la question sera l'exigence d'une structure de groupe pour l'ensemble de toutes les transformations inertielles.

HYPOTHÈSE 1 : ESPACE-TEMPS HOMOGENÈNE ET LINEARITÉ DES TRANSFORMATIONS INERTIELLES

Nous supposons d'abord que l'espace-temps est *homogène*, en ce sens qu'il a « partout et toujours » les mêmes propriétés. Plus précisément, les propriétés de transformation de l'intervalle spatio-temporel $(\Delta x, \Delta t)$ dépendent seulement de cet intervalle et pas de l'endroit où sont localisées ses extrémités (dans le référentiel considéré). En d'autres termes, l'intervalle transformé $(\Delta x', \Delta t')$ obtenu par une transformation inertielle (5) est indépendante de ces extrémités. En regardant un intervalle infinitésimal (dx, dt) , pour lequel

$$\begin{aligned} dx' &= \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial t} dt, \\ dt' &= \frac{\partial G}{\partial x} dx + \frac{\partial G}{\partial t} dt, \end{aligned} \tag{7}$$

cette exigence signifie que les coefficients de dx et dt dans (7) doivent être indépendants de x et t , donc que F et G sont des fonctions linéaires de x et t . Nous écrivons donc

$$x' = H(a)x - K(a)t, \tag{8a}$$

$$t' = L(a)t - M(a)x, \tag{8b}$$

où le signe moins a été introduit pour de futures commodités. La condition de coïncidence des origines a été utilisé pour écrire une transformation *homogène*, suivant (6).

La linéarité des transformations inertielles a une conséquence physique importante. Appelons *mouvements inertiels* les mouvements qui sont obtenus à partir du repos par une transformation inertielle ; un objet a un mouvement inertielle dans un référentiel s'il est au repos dans un référentiel équivalent. Les mouvements inertiels sont alors caractérisés par le même paramètre que les transformations inertielles. Leur équation générale du mouvement, d'après (8a), est

$$H(a)x - K(a)t = C^{ste}. \quad (9)$$

Les mouvements inertiels sont donc des mouvements uniformes avec une vitesse¹⁰ $v = K(a)/H(a)$, excepté pour la situation pathologique pour laquelle K est identiquement nulle¹¹. Cela suggère d'utiliser le paramètre v , comme étant une vitesse, au lieu du précédent paramètre indéfini¹² a . Avec un changement de notation adéquat¹³, nos formules de transformation générales peuvent se réécrire¹⁴

$$\begin{aligned} x' &= \gamma(v)(x - vt), \\ t' &= \gamma(v)[\lambda(v)t - \mu(v)x], \end{aligned} \quad (10)$$

dépendant de trois fonctions inconnues γ , λ , μ .

Contre-exemple 1

Il est intéressant d'insister sur la rigueur de l'Hypothèse 1. L'homogénéité de l'espace-temps, en particulier en ce qui concerne le temps, n'a *pas* lieu dans toute théorie physique concevable. Les modèles d'univers évolutionnaires n'ont pas un temps homogène, leurs transformations inertielles, si elles existent, ne sont pas linéaires, et leurs mouvements inertiels ne sont pas uniformes. Tel est le cas de l'espace-temps de De Sitter ou de son approximation « non relativiste⁸ » ; les formules de transformations de ce dernier s'écrivant

$$\begin{aligned} x' &= x - vT \sinh(t/T), \\ t' &= t, \end{aligned} \quad (11)$$

Où T est une échelle de temps cosmologique (dans un modèle oscillant, un « sin » remplacerait le « sinh » de cet univers en perpétuelle expansion). Ces transformations, il faut le noter, satisfont toutes les hypothèses à venir, comme on peut le vérifier aisément.

En termes abstraits de la théorie des groupes, les remarques précédentes sont en rapport avec la connexion qui existe entre le groupe des transformations inertielles et le groupe des translations de l'espace-temps à l'intérieur du groupe relativiste complet de la théorie. C'est uniquement lorsque le groupe des translations est un sous-groupe invariant du groupe relativiste complet (ce qui est le cas pour les groupes de Galilée et de Poincaré, mais *pas* pour celui de De Sitter) que la linéarité des transformations inertielles devient vraie.

HYPOTHÈSE 2 : ISOTROPIE DE L'ESPACE

Revenons à la forme générale (10) de nos transformations inertielles. On suppose que l'espace est non-directionnel, si bien que toutes les orientations des axes de l'espace sont physiquement équivalentes. Supposons que (x, t) et (x', t') soient deux

ensembles de coordonnées d'un évènement donné, reliés par une transformation inertielle (10) de paramètre v . Si la direction de l'axe spatial est arbitraire, $(-x, t)$ et $(-x', t')$ sont autant qualifiés pour des coordonnées équivalentes du même évènement, et doivent aussi être reliés par une transformation inertielle de forme générale (10) mais dépendant d'un paramètre u , inconnu pour le moment. En d'autres termes,

$$\begin{aligned} -x' &= \gamma(u)(-x - ut), \\ t' &= \gamma(u)[\lambda(u)t + \mu(u)x]. \end{aligned} \quad (12)$$

En comparant (12) avec (10), on obtient

$$\gamma(u) = \gamma(v), \quad (13a)$$

$$u\gamma(u) = -v\gamma(v), \quad (13b)$$

$$\lambda(u)\gamma(u) = \lambda(v)\gamma(v), \quad (13c)$$

$$-\mu(u)\gamma(u) = \mu(v)\gamma(v). \quad (13d)$$

A partir de (13a) et (13b), nous voyons d'abord que

$$u = -v. \quad (14)$$

Un résultat si naturel, exprimant la vitesse relative des référentiels « inversés » comme l'opposée de la vitesse relative des référentiels initiaux, aurait pu être pris pour acquis. Néanmoins, il est satisfaisant de le dériver à partir de principes premiers. Les relations restantes dans (13) expriment maintenant les propriétés de parité de nos fonctions inconnues γ , λ , μ :

$$\gamma(-v) = \gamma(v), \quad (15a)$$

$$\lambda(-v) = \lambda(v), \quad (15b)$$

$$\mu(-v) = -\mu(v). \quad (15c)$$

Il est facile de vérifier qu'exactement le même résultat aurait pu être obtenu en imposant à nos formules de transformation une condition similaire de symétrie par renversement de l'axe du temps. Alternativement, cette symétrie peut maintenant être prise comme une conséquence de celle spatiale¹⁵.

Contre-exemples 2

Le rôle de l'Hypothèse 2 est mieux illustré lorsqu'on l'abandonne. Un exemple simple de transformation inertielle respectant l'homogénéité de l'espace-temps (de telle sorte qu'elles soient linéaires) et formant un groupe de transformation (ce qui sera notre prochaine condition) est la suivante :

$$\begin{aligned} x' &= \exp(\sigma v)(x - vt), \\ t' &= \exp(\sigma v)t, \end{aligned} \quad (16)$$

pour une constante σ . Ces transformations approximent les transformations usuelles de Galilée pour $v \ll \sigma^{-1}$. Un autre exemple de groupe de transformations linéaires ne satisfaisant pas l'Hypothèse 2 est donné par

$$\begin{aligned}x' &= \frac{x - vt}{1 + \rho v}, \\t' &= \frac{t - \rho^2 vx}{1 + \rho v},\end{aligned}\tag{17}$$

où ρ est une constante caractéristique. C'est un exercice amusant que de vérifier les propriétés de groupe qui sont nécessaires par la suite et de trouver la loi de composition des vitesses¹⁶.

L'importance des propriétés d'inversion de l'espace et/ou du temps ne devrait pas être surprenante. En effet, l'inversion d'espace dans un espace à une dimension joue le rôle important des rotations en deux dimensions ou plus. Cela exprime l'*isotropie* de l'espace, c'est à dire, l'équivalence physique de toutes les orientations possibles, réduites ici au nombre de 2. Il est peut-être opportun de rappeler que cette hypothèse avait déjà été faite dans la première dérivation des formules de transformations de Lorentz par Einstein².

HYPOTHÈSE 3 : LA LOI DE GROUPE

L'équivalence physique des référentiels inertiels implique une structure de groupe pour l'ensemble de toutes les transformations inertielles (10). Cela exige de remplir les conditions suivantes :

(a) *Transformation identique*. Il doit exister un paramètre v tel que $x' = x$ et $t' = t$. Clairement il s'agit de $v = 0$, et nécessite

$$\gamma(0) = 1, \tag{18a}$$

$$\lambda(0) = 1, \tag{18b}$$

$$\mu(0) = 0. \tag{18c}$$

Remarquez que (18c) est déjà obtenu par (15c).

(b) *Transformation inverse*. Si (x', t') provient de (x, t) par la transformation (10) paramétrisée par v , la transformation inverse, de (x', t') à (x, t) , doit être de même forme fonctionnelle que (10), avec un paramètre w différent ; c'est à dire,

$$\begin{aligned}x &= \gamma(w)(x' - wt'), \\t &= \gamma(w)[\lambda(w)t' - \mu(w)x'].\end{aligned}\tag{19}$$

En inversant (10), on obtient aussi¹⁷

$$\begin{aligned}x &= \frac{1}{\gamma(v)} \left(1 - \frac{v\mu(v)}{\lambda(v)}\right)^{-1} \left(x' + \frac{v}{\lambda(v)}t'\right), \\t &= \frac{1}{\gamma(v)} \left(1 - \frac{v\mu(v)}{\lambda(v)}\right)^{-1} \left(\frac{1}{\lambda(v)}t' + \frac{\mu(v)}{\lambda(v)}x'\right).\end{aligned}\tag{20}$$

En identifiant (20) et (19) nous dérivons les équations suivantes qui lient l'inconnue w

à v et en même temps nous contraignons les fonctions γ , λ , μ :

$$w = -v/\lambda(v), \quad (21a)$$

$$\lambda(w) = 1/\lambda(v), \quad (21b)$$

$$\mu(w) = -\mu(v)/\lambda(v) \quad (21c)$$

$$\gamma(w) = [1/\gamma(v)][1 - v\mu(v)/\lambda(v)]^{-1}. \quad (21d)$$

Portons notre attention sur les deux premières équations (21a) et (21b). Elles donnent une équation fonctionnelle pour λ ,

$$\lambda[-v/\lambda(v)] = 1/\lambda(v). \quad (22)$$

Pour rendre plus acceptable l'apparente bizarrerie de l'équation fonctionnelle (22), définissons la fonction associée :

$$\zeta(v) \stackrel{df}{=} v/\lambda(v). \quad (23)$$

La condition (22) s'écrit maintenant¹⁸

$$\zeta[-\zeta(v)] = -v \quad (24)$$

ou

$$\zeta^{-1}(-v) = -\zeta(v). \quad (25)$$

Mais les courbes de $\zeta(v)$ et $\zeta^{-1}(v)$, quand elles sont dessinées sur le même repère cartésien, sont symétriques par rapport à la droite $\zeta = v$. La condition (25) requiert alors que le graphe de $\zeta(v)$ soit symétrique¹⁹ par rapport à la droite $\zeta = -v$.

D'autre part, puisque $\lambda(0) = 1$ [voir (18b)], d'après (23), nous avons²⁰

$$\left. \frac{d\zeta}{dv} \right|_{v=0} = 1. \quad (26)$$

N'importe quel ζ dont le graphe est symétrique par rapport à la droite $\zeta = -v$ et est tangente à l'origine à la droite $\zeta = v$ donnera $\lambda(v) = v/\zeta(v)$, qui obéit à (22). Comme exemple le plus simple, le lecteur pourra vérifier que $\lambda(v) = 1 - kv$ est une solution quel que soit le nombre k .

Enfin, la conséquence (15b) de l'Hypothèse 2, qui requiert que λ soit une fonction paire, réduit radicalement cet arbitraire. En prenant cela en compte, l'équation fonctionnelle en λ s'écrit

$$\lambda[v/\lambda(v)] = 1/\lambda(v). \quad (27)$$

Où, en utilisant la fonction associée ζ (23)²¹,

$$\zeta[\zeta(v)] = v; \quad (28)$$

c'est à dire,

$$\zeta^{-1}(v) = \zeta(v). \quad (29)$$

Le graphe de $\zeta(v)$ doit donc être symétrique par rapport à la droite $\zeta = v$. De plus, de part (18b), il doit aussi respecter (26); cela signifie qu'il doit être tangent à l'origine à cette même droite. Ainsi, en vue de la continuité de λ et ζ^{22} , le graphe de $\zeta(v)$ doit en fait être identique à la droite $\zeta = v$, soit $\zeta(v) = v$. Nous pouvons conclure que

$$\lambda(v) = 1. \quad (30)$$

Revenant à (21a), il résulte que le paramètre de la transformation (19) inverse de (10) est, assez naturellement, donnée par

$$w = -v, \quad (31)$$

mais ceci a été prouvé et non supposé. Finalement, (21d) et (15a) donnent une relation entre les deux fonctions restantes γ et μ :

$$[\gamma(v)]^2[1 - v\mu(v)] = 1. \quad (32)$$

(c) *Loi de composition.* Effectuons maintenant successivement deux transformations de la forme (10), en prenant en compte notre résultat précédent (30) :

$$\begin{aligned} x_1 &= \gamma(v_1)(x - v_1 t), \\ t_1 &= \gamma(v_1)[t - \mu(v_1)x], \end{aligned} \quad (33)$$

$$\begin{aligned} x_2 &= \gamma(v_2)(x_1 - v_2 t_1), \\ t_2 &= \gamma(v_2)[t_1 - \mu(v_1)x_1]. \end{aligned} \quad (34)$$

La transformation qui en résulte,

$$x_2 = \gamma(v_1)\gamma(v_2)[1 + \mu(v_1)v_2] \left(x - \frac{v_1 + v_2}{1 + \mu(v_1)v_2} t \right), \quad (35a)$$

$$t_2 = \gamma(v_1)\gamma(v_2)[1 + v_1\mu(v_2)] \left(t - \frac{v_1 + v_2}{1 + v_1\mu(v_2)} x \right), \quad (35b)$$

doit être identifiée à une transformation générale de la forme (10), dépendante d'un nouveau paramètre V (une fonction de v_1 et v_2) :

$$\begin{aligned} x_2 &= \gamma(V)[x - Vt] \\ t_2 &= \gamma(V)[t - \mu(V)x]. \end{aligned} \quad (36)$$

En identifiant le facteur $\gamma(V)$ dans (35a) et (35b), on obtient la condition

$$\mu(v_1)v_2 = v_1\mu(v_2). \quad (37)$$

D'où,

$$\mu(v) = \alpha v, \quad (38)$$

où α est une constante²³.

D'après (32), la dernière fonction inconnue est finalement

$$\gamma(v) = (1 - \alpha v^2)^{-1/2}, \quad (39)$$

où le signe de la racine carrée a été choisi en accord avec (18a). De plus, « la loi de composition des vitesses » dérive directement de (35) et (36) :

$$V = (v_1 + v_2)/(1 + \alpha v_1 v_2). \quad (40)$$

Il est clair que les trois différents cas suivants se présentent maintenant, dépendant du signe de la constante α ou de sa nullité.

(i) $\alpha < 0$. Nous pouvons écrire $\alpha = -\kappa^{-2}$, où κ a les dimensions d'une vitesse. La loi de transformation s'écrit

$$\begin{aligned} x' &= \frac{x - vt}{(1 + v^2/\kappa^2)^{-1/2}}, \\ t' &= \frac{t + vx/\kappa^2}{(1 + v^2/\kappa^2)^{-1/2}}. \end{aligned} \quad (41)$$

Notez que les valeurs de la vitesse v sont permises sur l'ensemble des réels. La loi de composition des vitesses est

$$V = \frac{v_1 + v_2}{1 - v_1 v_2 / \kappa^2}. \quad (42)$$

(ii) $\alpha = 0$. Les formules correspondantes sont

$$\begin{aligned} x' &= x - vt, \\ t' &= t, \end{aligned} \quad (43)$$

(Galiléen)

et

$$V = v_1 + v_2. \quad (44)$$

(iii) $\alpha > 0$. Nous écrivons donc $\alpha = c^{-2}$, où c est une constante ayant les dimensions d'une vitesse, et nous avons

$$x' = \frac{x - vt}{(1 - v^2/c^2)^{1/2}}, \quad (45)$$

$$t' = \frac{t - vx/c^2}{(1 - v^2/c^2)^{1/2}}, \quad (\text{Lorentz})$$

et

$$V = \frac{v_1 + v_2}{1 + v_1 v_2 / c^2}. \quad (46)$$

Ici les vitesses (en tant que paramètres de la transformation de Lorentz²⁴) sont limitées à l'intervalle $-c \leq v \leq c$. Clairement, comme dans le cas (i), la valeur numérique de α dépend du choix initial des unités des coordonnées d'espace et de temps, si bien que physiquement, il n'y a ici qu'une seule situation. Avant d'examiner ces trois cas à la lumière de notre dernière hypothèse physique, il peut être utile de mentionner ce qui suit.

Contre-exemples 3

Il est aisé de présenter des formules de transformation avec un look « raisonnable » satisfaisant, par exemple, toutes les hypothèses précédentes mais n'ayant pas la propriété de groupe; deux d'entre elles, en les composant, ne donne pas une transformation qui appartient à la famille. Voici ces transformations

$$\begin{aligned} x' &= x - vt, \\ t' &= t - vx/c^2. \end{aligned} \quad (47)$$

Egalement, l'ensemble des transformations

$$\begin{aligned} x' &= (1 + \frac{1}{2}v^2/c^2)x - vt, \\ t' &= (1 + \frac{1}{2}v^2/c^2)t - vx/c^2, \end{aligned} \quad (48)$$

est un développement mathématique consistant de la formule de Lorentz (45) au premier ordre en c^2 et a été proposé comme « nouveau » groupe relativiste²⁵; malheureusement, ce n'est *pas* un groupe.

HYPOTHÈSE 4 : CAUSALITÉ

Exepté pour le cas particulier (ii) ci-dessus, l'intervalle de temps entre deux événements dépend du référentiel, puisqu'il varie avec les transformations (41) ou (45). A présent, dans le but de maintenir un *certain* ordre dans l'univers, nous aimerions pouvoir exiger l'existence d'au moins une classe d'évènements spatio-temporels tels que le *signe* de l'intervalle de temps, c'est à dire, la nature d'une possible relation causale, ne soit pas changé par les transformations inertielles.

Cette condition est évidemment présente dans les transformations de Galilée (43) pour *tous* les intervalles de temps. Elle est aussi présente dans le cas Lorentzien (iii) pour les intervalles tels que $|\Delta x / \Delta t| \leq c$ (« de genre temps ») à cause de l'éventail limité des vitesses possibles v . Toutefois, le cas (i) n'est pas en accord avec nos hypothèses, d'où ce qui suit.

Contre-exemple 4²⁶

En effet, puisque

$$\Delta t' = \frac{\Delta t + v\Delta x/\kappa^2}{(1 + v^2/\kappa^2)^{1/2}} \quad (49)$$

pour tout $(\Delta t, \Delta x)$, nous devrions toujours trouver une vitesse v , dans l'éventail illimité, tel que $\Delta t'$ ait un signe opposé à celui de Δt , ce qui interdit l'existence même d'une relation de causalité. Une autre caractéristique paradoxale du cas (i) apparaît lorsque l'on compose deux vitesses positives, par exemple $v_1 = 2\kappa$ et $v_2 = \kappa$, ce qui donne une vitesse $V = -3\kappa, \dots$ dans la direction opposée²⁷ !

CONCLUSION

Nos quatre hypothèses générales suffisent donc à mettre en avant les transformations de Lorentz et leur limite dégénérée galiléenne comme les seules transformations inertielles possibles. Le cas Lorentzien est caractérisé par un paramètre ayant les dimensions d'une vitesse qui est une constante universelle associée à la structure même de l'espace-temps. Une analyse plus poussée des objets pouvant se déplacer dans un tel espace-temps, montre que cette constante *se trouve* être la vitesse (invariante) des objets de masse nulle.

Note ajoutée : Ce travail avait déjà été soumis pour publication quand une recherche quasi similaire de A. R. Lee et T. M. Kalotas est apparue dans ce journal [Am. J. Phys. **43**, 434 (1975)]. Ces auteurs ont aussi mis en évidence les racines très anciennes de telles considérations, en revenant à de nombreux écrits ayant plus de soixante ans, et depuis longtemps oubliés ou négligés. Parmi les références historiques qu'ils citent, il faudrait ajouter celle-ci : L. A. Pars, Philos. Mag. **43**, 249 (1921). Malgré la similarité de ce présent écrit avec l'article de Lee et Kalotas, la généralité quelque peu supérieure de ces hypothèses, ainsi que les contre-exemples offerts, continuent de justifier sa publication.

REMERCIEMENTS

C'est un plaisir de remercier F. Balibar, C. Godrèche, B. Jamin, and J. Kaplan pour leurs utiles critiques, discussions, et suggestions et, surtout, les nombreux étudiants des divers cours d'introduction à la relativité restreinte pour leur réticence justifiée à accepter les dérivations habituelles de la transformation de Lorentz, qui a

mené aux considérations présentes.

1. Voir les exemples récents et non exhaustifs suivants : A. L. Harvey, Am. J. Phys. **36**, 901 (1968) ; A. B. Kaiser, *ibid.* **37**, 216 (1969) ; J. Rekveld, *ibid.* **37**, 717 (1969) ; L. Parker and G. M. Schmieg, *ibid.* **38**, 218 (1970) ; L. Parker *ibid.* **39**, 223 (1971) ; J. R. Shepanski and R. Simons, *ibid.* **40**, 486 (1972).

2. A. Einstein, Ann. Phys. (Leipz.) **17**, 891 (1905) ; réédité traduit en anglais dans *The Principle of Relativity* (Dover, New York).

3. La relativité générale *pourrait* être interprétée comme une théorie relativiste particulière d'un champ de tenseurs auto-couplé, qui se trouve jouer le rôle d'une véritable métrique, masquant celle présumée de Minkowski : W. Thirring, Ann. Phys. N.Y. **16**, 96 (1961) ; voir aussi R. V. Sexl, Fortschr. Phys. **15**, 269 (1967).

4. E. P. Wigner, Phys. Today **17** (3), 34 (1964).

5. E. P. Wigner, Ann. Math. **40**, 149 (1939).

6. C'est ici le lieu pour montrer combien la terminologie peut elle-même être mystifiante. Pourquoi appeler c la vitesse de la lumière quand il s'agit de la vitesse de *tous* les objets de masse nulle, par exemple, la vitesse des neutrinos ? Par ailleurs, si les photons, les neutrinos, et les gravitons avaient une très faible masse, mais non nulle (Réf. 7), c ne serait pas la vitesse d'un objet réel. Je ne suggère bien sûr pas un changement de vocabulaire de la physique qui s'est instauré par des années de pratique ; l'abus de langage n'est pas seulement inévitable, il est probablement nécessaire. Au moins, reconnaissons-le comme tel.

7. A. S. Goldhaber and M. M. Nieto [Rev. Mod. Phys. **43**, 277 (1971)] ont examiné les limites présentes sur la masse du photon. Dans le cas d'un photon avec une masse non nulle, les auteurs présentent une dérivation plus sophistiquée mais plutôt artificielle de la relativité restreinte, par des considérations toujours basées sur la (alors variable) vitesse de la lumière (see p. 294 de leur article).

8. H. Bacry and J. M. Levy-Leblond, J. Math. Phys. **9**, 1605 (1968), et d'autres références à l'intérieur.

9. Je n'inclue *pas* dans ces travaux le très beau théorème de E. C. Zeeman [J. Math. Phys. **5**, 490 (1964)], qui dérive les transformations de Lorentz à partir d'un principe de causalité. En effet, son point de départ est l'existence d'un cône de causalité, ce qui implique *ab initio* l'existence d'une vitesse limite.

10. NDT On a $x = [C^{ste} + K(a)t]/H(a)$, et comme $v = dx/dt$, on a bien $v = K(a)/H(a)$.

11. Cette annulation de K , impliquant le repos absolu comme seul mouvement inertiel, caractérise la cinématique singulière décrite par le « groupe statique » ou « groupe de Carroll » étudié en Réf. 8.

12. NDT v ne pouvant être que le rapport de deux fonctions de v , on a forcément $a = v$. A noter que $d^2x/dt^2 = 0$ car H et K ne sont pas des fonctions du temps.

13. NDT On pose $H(a) = \gamma(v)$.

14. NDT v n'est pas un vecteur car nous sommes en une dimension d'espace.

15. Il est peut-être nécessaire de souligner ici que, dans l'état actuel de nos connaissances, la symétrie de l'espace-temps par réflexion d'espace ou de temps n'a rien à voir avec la parité ou l'invariance par renversement du temps des interactions physiques. La violation de la parité dans les interactions faibles et l'échec de l'invariance par renversement du temps dans celles super-faibles ont lieu dans un espace-temps relativiste ayant une parfaite symétrie par réflexion de l'espace ou du temps. Il serait tentant de trouver les origines de ces cassures de loi d'invariance d'une dynamique spécifique dans la distorsion de l'espace-temps lui-même. Toutefois, de telles tentatives ont jusqu'à présent échouées.

16. Le fait que la même loi de composition des vitesses apparaisse comme dans le cas Lorentzien peut se comprendre en utilisant un paramètre de « rapidité » φ , tel que $\rho v = \tanh \varphi$, pour réécrire (17) sous la forme

$$x' = \exp(-\varphi)(\cosh \varphi x - \sinh \varphi t),$$

$$t' = \exp(-\varphi)(\cosh \varphi t - \sinh \varphi x).$$

Cette transformation apparaît alors comme une transformation de Lorentz combinée à une dilatation globale dépendante de la vitesse (et non invariante par réflexion).

17. NDT Démonstration du passage de l'équation (10) à l'équation (20) :

$$x' = \gamma(v)(x - vt),$$

$$t' = \gamma(v)[\lambda(v)t - \mu(v)x],$$

$$\begin{aligned}
\frac{x' + \gamma(v)vt}{\gamma(v)} &= x \\
\frac{t' + \gamma(v)\mu(v)x}{\gamma(v)\lambda(v)} &= t \\
\frac{x'}{\gamma(v)} + v \left[\frac{t'}{\gamma(v)\lambda(v)} + \frac{\mu(v)}{\lambda(v)}x \right] &= x \\
\frac{t'}{\gamma(v)\lambda(v)} + \frac{\mu(v)}{\lambda(v)\gamma(v)} [x' + \gamma(v)vt] &= t \\
\frac{1}{\gamma(v)} \left[1 - v \frac{\mu(v)}{\lambda(v)} \right]^{-1} \left[x' + \frac{v}{\lambda(v)}t' \right] &= x \\
\frac{1}{\gamma(v)} \left[1 - v \frac{\mu(v)}{\lambda(v)} \right]^{-1} \left[\frac{t'}{\lambda(v)} + \frac{\mu(v)}{\lambda(v)}x' \right] &= t
\end{aligned}$$

qui donne directement l'équation (20).

18. NDT Détail du passage de l'équation (22) à l'équation (24) :

$$\lambda[-v/\lambda(v)] = 1/\lambda(v).$$

En utilisant la définition de $\zeta(v)$ on a :

$$\lambda[-\zeta(v)] = 1/\lambda(v).$$

En utilisant de nouveau la définition de $\zeta(v)$:

$$\begin{aligned}
v\lambda[-\zeta(v)] &= \zeta(v). \\
\frac{-\zeta(v)}{\lambda[-\zeta(v)]} &= -v \\
\zeta[-\zeta(v)] &= -v.
\end{aligned}$$

19. NDT On part de trois axiomes de la géométrie ordinaire :

$$\begin{aligned}
\zeta^{-1}(v) &= \text{sym}_{/\zeta(v)=v} \zeta(v) \\
\zeta(-v) &= \text{sym}_{/\zeta} \zeta(v) \\
-\zeta(v) &= \text{sym}_{/v} \zeta(v)
\end{aligned}$$

Avec les deux premiers axiomes :

$$\begin{aligned}
\zeta^{-1}(-v) &= \text{sym}_{/\zeta} \zeta^{-1}(v) \\
&= \text{sym}_{/\zeta} \text{sym}_{/\zeta=v} \zeta(v)
\end{aligned}$$

Puis avec l'équation (25) on remplace le premier membre de l'égalité $\zeta^{-1}(-v)$ par $-\zeta(v)$:

$$-\zeta(v) = \text{sym}_{/\zeta} \text{sym}_{/\zeta=v} \zeta(v)$$

Avec le troisième axiome on écrit :

$$\text{sym}_{/v} \zeta(v) = \text{sym}_{/\zeta} \text{sym}_{/\zeta=v} \zeta(v)$$

Enfin, en symétrisant de part et d'autre de l'égalité par rapport à l'axe v :

$$\text{sym}_{/v} \text{sym}_{/v} \zeta(v) = \text{sym}_{/v} \text{sym}_{/\zeta} \text{sym}_{/\zeta=v} \zeta(v)$$

On peut vérifier facilement en prenant un point sur un graph, que les trois symétries successives sont équivalentes à une symétrie par rapport à la droite $\zeta = -v$.

$$\text{sym}_{\zeta=-v} \zeta(v) = \zeta(v).$$

-
20. NDT $d\zeta/dv = [\lambda(v) - vd\lambda(v)/dv]/\lambda^2(v)$. Pour $v = 0$, $\lambda(0) = 1$ et l'on trouve bien 1.
21. NDT Car si $\lambda(v)$ est paire alors $\zeta(v)$ est impaire.
22. Ce n'est *pas* une preuve mathématique complètement rigoureuse de ce point, qui demande un degré de sophistication loin au-delà du niveau pédagogique adopté ici [J. Morgestern and M. Zerner (communication privée)].
23. NDT On a $\mu(v_1)/v_1 = \mu(v_2)/v_2$. Deux fonctions de variables différentes, ici v_1 et v_2 , ne peuvent être égales que si elles sont constantes.
24. C'est l'occasion de souligner la différence, pas toujours assez claire dans certaines discussions à propos de tachyons, entre la vitesse relative de deux référentiels, paramètre de la transformation de Lorentz, qui est nécessairement inférieure à c , et la vitesse de « quelque chose » dans un référentiel donné, qui serait plus grande que c pour les tachyons et *serait* plus grande que c pour les ombres, comme le spot d'un faisceau d'électrons sur l'écran d'un oscilloscope, par exemple.
25. V. N. Streltov, Dubna JINR Rapports P2-4461, P2-4462, P2-5523, P2-5823, P2-6208 (non publiés).
26. Voir Eqs. (41) et (42).
27. En des termes plus mathématiques, la causalité telle qu'exprimée ici demande la non compacité du groupe des transformations inertielles (voir Réf. 8.). Le groupe considéré ici (contre-exemple 4) est clairement isomorphe au groupe compact des rotations en deux dimensions, tandis que le groupe à deux dimension de Lorentz est non compact.