L'espace et le temps revus par Einstein

R. Lehoucq

Fleurance, 5 août 2003

Résumé

La théorie de la relativité a complètement bouleversé la physique du XX^{e} siècle. Elle est réputée difficile, mais ce sont surtout ses résultats qui sont difficiles à admettre car ils sont souvent contre-intuitifs. Pourtant ses postulats sont très simples, et les mathématiques sur lesquelles elle repose tout à fait élémentaires. Ce cours d'approfondissement va tenter d'en faire la démonstration.

1 Rappels de relativité galiléenne

1.1 Transformation de Galilée

Le référentiel est la notion centrale de la cinématique. C'est un ensemble d'observateurs, immobiles les uns par rapport aux autres, qui constatent le passage d'un mobile à leur position. Sachant où sont situés les observateurs concernés on peut déterminer la trajectoire du mobile. Il faut bien sûr convenir d'un repère (généralement cartésien, orthonormal) pour repérer ces positions au moyen de trois coordonnées. De plus, les observateurs sont munis d'horloges qui leur permettent de noter l'instant où le mobile passe en face d'eux ; le mouvement est alors complètement déterminé par la trajectoire et la loi horaire. Ces horloges peuvent être constituées de n'importe quel phénomène physique périodique, suffisamment rapide à l'échelle du mouvement pour en donner une description temporelle convenable. Nous supposerons que les horloges de tous les observateurs d'un même référentiel sont synchronisées. Cette synchronisation ne pose aucune difficulté en cinématique classique, puisque temps et espace sont complètement découplés. Il suffit, par exemple, que tous les observateurs se retrouvent en un même point pour faire le zéro de leurs horloges à un moment commun. Certes, ces précautions pour la définition du temps paraissent superfétatoires en cinématique classique. Nous verrons, en revanche, qu'elles sont très importantes en cinématique relativiste.

Un mouvement dans un référentiel $\mathcal{R}$ est donc défini par les trois fonctions $x(t)$, $y(t)$ et $z(t)$ représentant la position en fonction du temps commun des observateurs. Le même mouvement serait décrit dans un autre référentiel $\mathcal{R}'$, en mouvement par rapport à $\mathcal{R}$, par trois autres fonctions du temps commun des observateurs de $\mathcal{R}'$: $x'(t')$, $y'(t')$ et $z'(t')$. En mécanique classique, on admet sans restrictions l'identité des temps (à une synchronisation près) des observateurs de $\mathcal{R}$ et de $\mathcal{R}'$. Il est possible alors de donner la transformation qui fait se correspondre les mouvements vus dans deux référentiels différents.

Dans le cas le plus simple, où les deux référentiels sont en translation uniforme l'un par rapport à l'autre, cette transformation est la transformation dite de Galilée. Sans restreindre la généralité, on peut choisir les axes de $\mathcal{R}$ et $\mathcal{R}'$ de manière que (voir figure 1) :

- les axes Ox et $O'x'$ coïncident à tout instant et sont parallèles à la vitesse u de $\mathcal{R}'$ par rapport à $\mathcal{R}$;
- les origines O et O' sont confondues à l'instant $t = 0$;
- les axes Oy et $O'y'$, d'une part, et les axes Oz et $O'z'$, d'autre part, sont constamment parallèles et coïncident à $t = 0$.

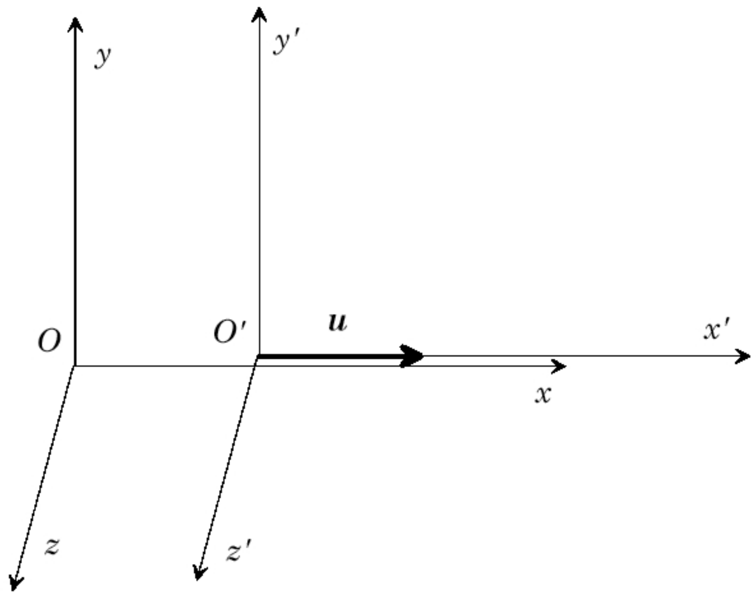

FIG. 1 – *Choix des axes des deux référentiels $\mathcal{R}$ et $\mathcal{R}'$ en mouvement relatif. Les axes des deux repères sont parallèles. Les axes Ox et $O'x'$, alignés avec la vitesse relative $\vec{u}$, coïncident à chaque instant.*

La loi de transformation de Galilée s'écrit alors :

$$\begin{cases} x'(t) = x(t) - ut \\ y'(t) = y(t) \\ z'(t) = z(t) \end{cases} \tag{1}$$

La transformation de Galilée contient, par simple dérivation par rapport au temps, la loi de composition des vitesses :

$$\vec{v} = \vec{v}' + \vec{u} \tag{2}$$

que l'on traduit aussi par «vitesse absolue égale vitesse relative plus vitesse d'entraînement». La dynamique newtonienne résulte alors du principe d'inertie de Galilée : il existe une classe de référentiels privilégiés en mouvement de translation uniforme les uns par rapport aux autres (les référentiels galiléens), tels que le mouvement d'une particule libre y soit rectiligne et uniforme.

2 Les difficultés de la cinématique classique

La loi galiléenne de composition des vitesses n'est pas compatible avec l'électromagnétisme de Maxwell (deux de ces équations ne sont pas invariantes par changement de référentiel galiléen). La conséquence la plus importante des équations de Maxwell est en effet la prédiction de l'existence d'ondes électromagnétique se propageant à une vitesse c finie. Le problème qui apparaît immédiatement est celui du référentiel dans lequel cette vitesse est définie, le seul dans lequel les équations de Maxwell seraient directement applicables.

Le sentiment le plus naturel, qui prédominait très largement à la fin du XIX[e] siècle, était que les ondes électromagnétiques se propageaient dans un milieu approprié, l'éther, baignant l'univers entier. L'analogie entre ondes électromagnétiques et ondes sonores était en effet présente à tous les esprits. Les difficultés apparaissent toutefois très vite dès qu'on examine les propriétés de cet hypothétique éther. Il doit en effet être omniprésent et extrêmement rigide pour propager des ébranlements transverses à grande vitesse. Mais il doit, dans le même temps, être impondérable et infiniment perméable au mouvement des corps matériels (puisque, par exemple, l'étude sur quelques siècles de la rotation terrestre ne révèle aucun frottement). Ce support particulier se trouvait ainsi doté de propriétés presque aussi extraordinaires que le calorique du siècle précédent ou, encore avant lui, le phlogistique.

L'introduction de l'éther présente aussi une difficulté philosophique grave. Les physiciens avaient mis plus de 20 siècles, entre Aristote et Copernic, pour se convaincre que la Terre n'est pas le centre de l'univers. Le principe d'inertie avait le mérite de ne priviligier aucun référentiel galiléen. L'introduction de l'éther devait briser cette équivalence des référentiels galiléens en introduisant un référentiel particulier, celui de l'éther, le seul dans lequel les équations de Maxwell peuvent s'appliquer.

Les dernières difficultés, les plus graves en pratique, mais qui n'ont pas forcément joué le rôle majeur qu'on leur attribue généralement dans la genèse de la relativité, sont d'ordre expérimental. Si la vitesse de la lumière est définie dans le référentiel de l'éther et si elle obéit à la loi de composition des vitesses, on doit pouvoir mesurer une variation de cette vitesse pour des mouvements assez rapides par rapport à l'éther. Le mouvement de la Terre sur son orbite autour du Soleil est suffisamment rapide (30 km/s) pour que la variation soit mesurable dans une expérience d'interférométrie optique sensible. La célèbre expérience de Michelson et Morley (1887) fut conçue dans ce but. D'une sensibilité tout à fait remarquable pour l'époque, encore honorable aujourd'hui, elle aurait dû mettre clairement en évidence le mouvement de la Terre par rapport à l'éther. Cette expérience fut catégorique : la valeur mesurée de la vitesse de la lumière semblait indépendante du mouvement de la Terre par rapport au Soleil.

Devant ce résultat, deux points de vue peuvent être adopter. Le premier était de tenter de «réparer» la théorie de l'éther. Si on ne pouvait décemment supposer que le référentiel absolu était celui de la Terre (la révolution copernicienne était passée par là), on pouvait supposer que l'éther était entraîné au voisinage des corps massifs, une analogie évidente avec l'entraînement de la couche limite en hydrodynamique. On pouvait aussi supposer, avec Lorentz, une contraction de la longueur des objets matériels dans la direction du mouvement, fondée sur une théorie électrostatique des interactions entre particules dans la matière. On pouvait supposer aussi un lien entre la vitesse de la lumière et celle de sa source (les sources utilisées par Michelson étant liées à son appareil). Si de telles modifications ad hoc de l'électromagnétisme permettaient d'expliquer le résultat négatif de l'expérience de Michelson, ils ne constituaient pas un corps théorique cohérent. Il était à craindre que de nouvelles modifications tout aussi arbitraires ne doivent être apportées au gré des résultats expérimentaux et que l'électrodynamique ne finisse, comme la théorie astronomique des épicycles, en un corps raffiné de règles arbitraires qui décrivent correctement mais ne prédisent rien.

L'autre attitude, beaucoup plus courageuse puisqu'elle conduit, comme nous le verrons, à mettre en cause des notions très fondamentales, était d'admettre que la vitesse de la lumière n'obéissait pas à la loi de composition des vitesses. Cela impliquait bien sûr que la cinématique galiléenne était erronée ou, du moins, n'était qu'une approximation valide pour des vitesses petites devant celle de la lumière. Toute la physique était donc à reconstruire (sauf, peut être, l'électrodynamique). C'est la voie que suivit Einstein avec le succès que l'on connaît et qu'il ouvrit par son célèbre article de 1905 [5]. Exposons maintenant le principe fondamental de cette nouvelle physique, le principe de relativité.

3 Principe de relativité

3.1 Enoncé

Il existe une classe de référentiels privilégiés, en translation uniforme les uns par rapport aux autres, dans lesquels toutes les lois de la physique prennent la même forme.

De prime abord, ce principe de relativité ne semble rien remettre en cause d'essentiel. Il paraît voisin du principe de relativité de la physique classique. Il n'en est rien, comme nous allons le voir en considérant deux expériences de pensée. Nous allons montrer que le principe de relativité a deux conséquences immédiates, contraires à l'intuition commune :

– Le temps ne s'écoule pas de la même façon dans deux référentiels galiléens en mouvement relatif (deux horloges en mouvement relatif bâties sur le même modèle ne battent pas au même rythme).

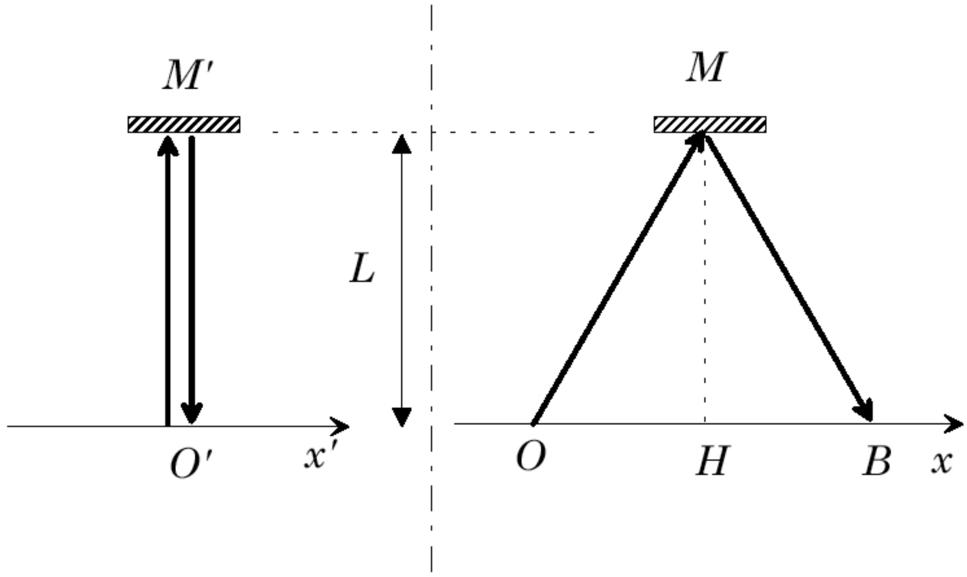

FIG. 2 – *Expérience de pensée établissant le caractère relatif du temps si l'on admet le principe de relativité. Dans le référentiel du train (à gauche), un signal lumineux est émis depuis l'observateur O', le long de l'axe $O'y'$, vers un miroir M. Après réflexion sur ce miroir, le signal revient à l'observateur O'. La même expérience est vue, à droite, dans le référentiel du quai ; dans ce cas le contrôleur du train est passé de O à B.*

– Deux événements qui se produisent simultanément dans un référentiel peuvent se produire à des instants différents quand ils sont observés depuis un autre référentiel.

Remettre en cause des propriétés apparemment aussi intuitives de l'espace et du temps ne sera pas sans conséquences. Il est clair, en particulier, qu'il faudra renoncer au caractère absolu du temps et «mélanger» les coordonnées spatiales et temporelles.

3.2 Deux expériences de pensée

Considérons donc deux référentiels en mouvement relatif, avec la géométrie décrite dans la figure 1. Le référentiel $\mathcal{R}'$ sera celui du contrôleur d'un train, le référentiel $\mathcal{R}$ celui du chef de gare qui le voit passer. Le contrôleur, situé en O', envoie à $t' = 0$ (ne confondons pas les temps dans les deux référentiels) une impulsion lumineuse de durée négligeable dans la direction y' vers un miroir situé en $y' = L$ (voir partie gauche de la figure 2). L'impulsion, réfléchie par le miroir, revient vers le contrôleur et l'atteint au bout d'un temps $T' = 2L/c$. Ainsi, le contrôleur peut construire une horloge : il lui suffit de renvoyer une deuxième impulsion à l'instant précis où il reçoit la première.

Regardons maintenant cette même expérience avec l'œil du chef de gare (partie droite de la figure 2). A $t' = 0$, le contrôleur est en O', qui coïncide avec O. Pendant que l'impulsion progresse, le miroir s'est déplacé. Lors de la réflexion, il occupe une position M, à une certaine distance de O sur l'axe Ox. Le train continue à se déplacer pendant le retour de l'impulsion et le contrôleur occupe la position B au moment de son retour. La trajectoire de l'impulsion dans $\mathcal{R}$ est donc triangulaire. Pour le chef de gare, qui admet le postulat de relativité, la vitesse de l'impulsion est aussi égale à c. Le temps de parcours de O à M est donc $T = OM/c$. Il en déduit $OH = OM\, u/c$. Comme $OM^2 = L^2 + OH^2$, on a donc $OM = L/\sqrt{1 - u^2/c^2}$ et finalement :

$$T = \gamma T' \qquad \text{avec} \qquad \gamma = \frac{1}{\sqrt{1 - u^2/c^2}} \tag{3}$$

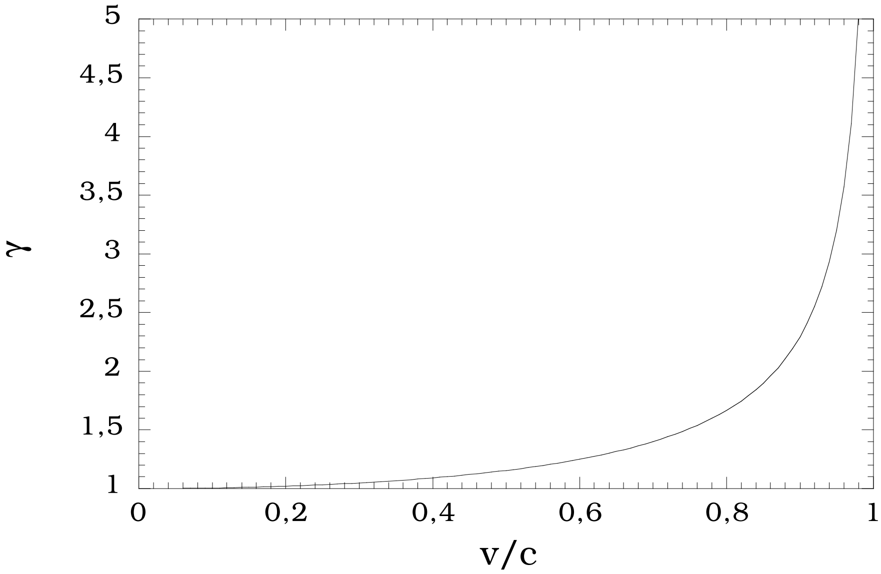

FIG. 3 – *Variation du facteur γ en fonction de la vitesse réduite.*

Le facteur γ, que nous aurons de nombreuses occasions de retrouver, est toujours supérieur à un (voir figure 3). Donc la durée de l'aller-retour mesurée par le chef de gare est toujours plus longue que celle mesurée par le contrôleur. Si chacun construisait une horloge avec le même dispositif, le chef de gare verrait la sienne battre plus vite et avancer par rapport à celle du contrôleur. Le postulat de relativité a pour conséquence immédiate que le temps n'est plus une notion universelle.

Cette expérience de pensée nous fournit une autre indication sur ce que sera la cinématique relativiste. Le facteur γ n'est défini que si u, la vitesse relative des deux référentiels, est plus petite que c, la vitesse de l'impulsion lumineuse. Si ce n'était pas le cas, l'impulsion lumineuse, qui, dans $\mathcal{R}'$, se réfléchit forcément sur le miroir, n'arriverait jamais à le rattraper dans $\mathcal{R}$. Un événement (la réflexion) se produirait dans un référentiel et pas dans un autre, ce qui semble absurde. Deux référentiels galiléens ne peuvent donc être animés l'un par rapport à l'autre d'une vitesse supérieure à c qui apparaît donc comme une vitesse limite pour tous les objets matériels.

La seconde expérience de pensée que nous allons étudier nous forcera à renoncer au caractère absolu de la simultanéité. Nous utiliserons encore les services du chef de gare et du contrôleur embarqué sur son train. Le chef de gare est situé en O, à mi-chemin de deux signaux lumineux A et B (voir figure 4, en haut). A l'instant $t = 0$, il voit ces deux signaux s'allumer simultanément. S'il connaît la distance L qui le sépare des deux signaux, il en déduira qu'ils se sont allumés simultanément à l'instant $t = -L/c$.

Au même instant $t' = t = 0$, le contrôleur, situé en O', passe devant le chef de gare. Il voit donc, à cet instant précis, les deux signaux A et B s'allumer. Comment peut-il en déduire, dans son échelle de temps, l'instant auquel ils se sont allumés ? Il lui faut d'abord déterminer où les deux signaux se sont allumés dans son référentiel. Pour cela, il peut parcourir son train et rechercher les deux voyageurs A' et B' (les observateurs) qui étaient juste en face des signaux quand ils se sont allumés. Il pourra ensuite leur demander à quel instant cet événement s'est produit ou utiliser leur position et la vitesse de la lumière pour calculer cet instant (les deux impulsions lumineuses ont la même vitesse). Nous ne pouvons pas, pour le moment, déterminer la position des observateurs A' et B'. En revanche, nous pouvons comprendre que la distance $A'O'$ est nécessairement inférieure à la distance $O'B'$. Le temps de parcours de l'impulsion provenant de A est donc inférieur à celui venant de B. Le contrôleur en déduira que le signal A s'est allumé après le signal B. Deux événements, l'allumage des signaux, peuvent être vus ou non comme simultanés par des observateurs appartenant à des référentiels différents. Nous verrons bientôt que cet abandon

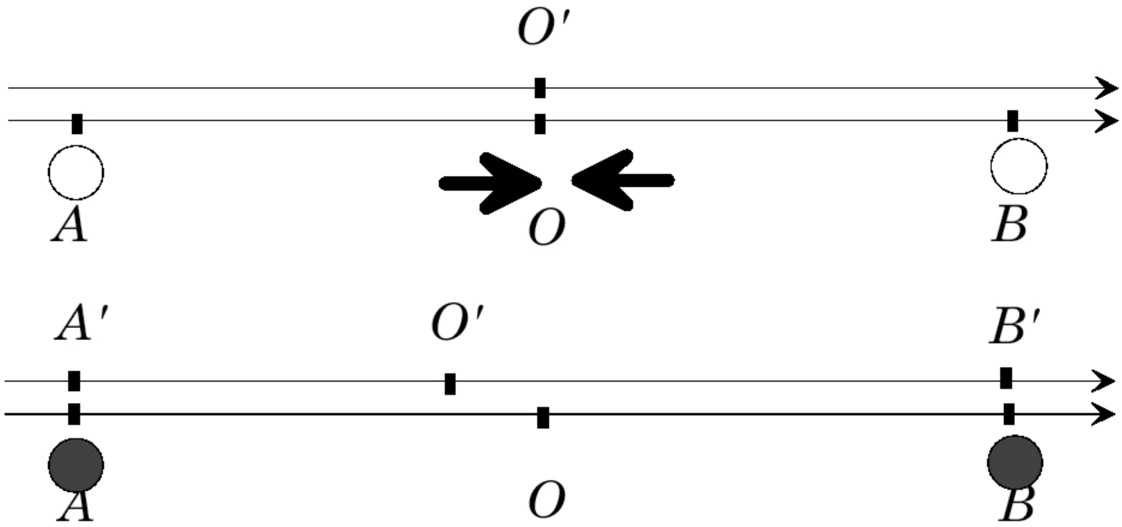

FIG. 4 – *Deuxième expérience de pensée illustrant le postulat de relativité. En haut, vue de la situation au moment où le chef de gare, situé en* O*, et le contrôleur, en* O'*, voient arriver simultanément les signaux lumineux émis par* A *et* B*. En bas, situation au moment où les signaux se sont allumés.* O' *n'est pas encore arrivé en* O*. Les signaux s'allument en face des observateurs* A' *et* B'*.*

de la simultanéité ne compromet pas la causalité.

En discutant ces deux expériences de pensée, deux notions essentielles de la relativité émergent : l'événement et l'intervalle.

4 Evénements et intervalles

4.1 Evénements

Comme nous venons de le voir, le temps n'est plus universel. Il faut décrire les expériences en termes d'événements : il s'est passé quelque chose quelque part. Deux exemples d'événements : la réflexion de la lumière sur le miroir ou l'allumage du signal A. Un événement existe indépendamment du choix du référentiel. On peut le caractériser, dans un référentiel donné, par l'observateur qui était sur place (le passager A') et par l'instant, mesuré sur l'horloge de cet observateur, où l'événement s'est produit. Dans un référentiel fixé, un événement est donc caractérisé par quatre nombres, les trois coordonnées spatiales de l'observateur et le temps. Bien sûr, les coordonnées spatio-temporelles du même événement dans un autre référentiel sont différentes et l'essentiel de notre tâche sera de donner la loi de transformation qui remplace et étend la transformation de Galilée. On peut établir un parallèle entre, d'une part, événement et vecteur (indépendants du repère) et d'autre part, entre coordonnées spatio-temporelles et composantes du vecteur (dépendants du référentiel ou de la base choisis). Les événements sont souvent représentés de façon géométrique, comme des points dans un espace-temps à 4 dimensions. Pour des raisons pratiques, les représentations se cantonnent souvent à une dimension d'espace et une de temps. Pour des raisons de commodité, on porte sur l'axe vertical le produit ct ; les deux coordonnées dans cet espace ont ainsi la même unité (voir figure 5).

Le mouvement d'un point dans un référentiel est une suite d'événements (la suite des observateurs devant lesquels la particule est passée associée aux instants correspondants). Une telle suite continue d'événements forme une ligne de l'espace-temps, nommée «ligne d'univers» de la particule (voir figure 5). La ligne d'univers d'une particule qui se déplace à la vitesse de la lumière est parallèle, dans notre représentation graphique, à la première ou à la deuxième bissectrice. Dans

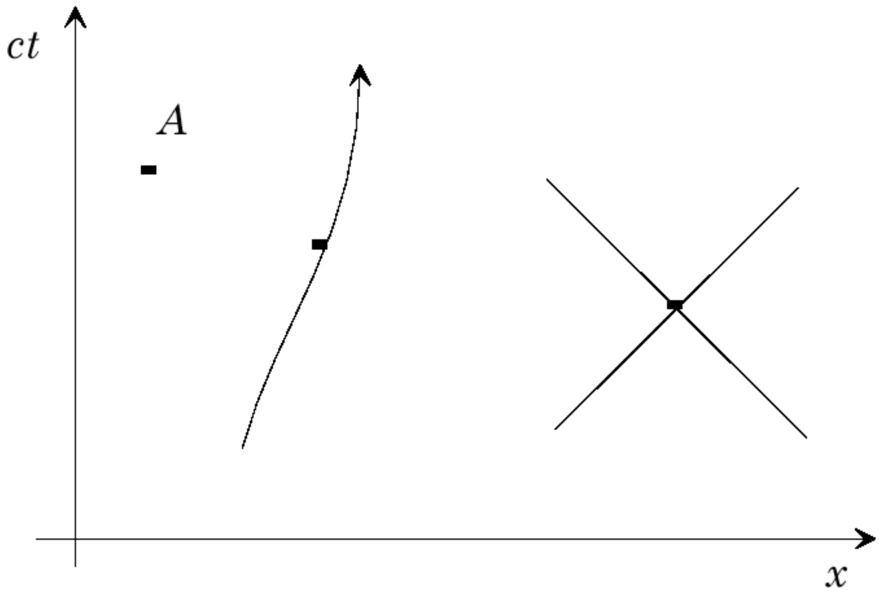

FIG. 5 – *Un événement, une ligne d'univers et un cône de lumière. Un événement est représenté par un point dans un espace* (x, ct)*. Une ligne d'univers est l'ensemble des événements correspondant aux positions successives d'une particule. Le cône de lumière d'un événement est constitué des lignes d'univers d'un signal lumineux passant par celui-ci.*

l'espace à quatre dimensions, l'ensemble des lignes d'univers partant d'un point et correspondant à un mouvement à la vitesse de la lumière forme le «cône de lumière» de cet événement (voir figure 5). Comme c est la vitesse limite, toutes les lignes d'univers passant par un événement donné doivent être à l'intérieur du cône de lumière. Les événements antérieurs à l'événement de référence et situés dans le cône de lumière sont dans le passé, les autres dans le futur. Deux événements ne pourront être reliés par un signal ou une relation causale, que s'ils sont dans le cône de lumière l'un de l'autre. Il semble évident que la relation «être dans le cône de lumière» est symétrique : si A est dans le cône de lumière de B, alors B est dans le cône de lumière de A. En revanche, cette relation n'est pas transitive, comme on pourra s'en persuader aisément. Si C est dans le passé de B, lui-même dans le futur de A, alors C n'est pas nécessairement dans le cône de lumière de A. En un mot, A et C peuvent être tous deux cause de B, sans qu'il existe un lien causal entre eux. En revanche, si C est dans le futur de B, il est nécessairement dans le cône de lumière de A : si A est la cause de B qui est lui même la cause de C, alors A peut être la cause de C. En ces termes, la version relativiste de la causalité apparaît très clairement. La physique classique admet qu'un événement est la cause d'un autre si et seulement si il lui est antérieur (admettant ainsi implicitement les actions instantanées à distance). La physique relativiste exige en plus que les deux événement puissent être reliés par un signal lumineux. Ces idées vont être affinées en introduisant la notion d'intervalle.

4.2 Intervalle entre deux événements

Considérons deux événements dont les coordonnées sont (x_1, y_1, z_1, ct_1) et (x_2, y_2, z_2, ct_2) dans un référentiel donné. Si ces deux événements sont sur le cône de lumière l'un de l'autre, ils peuvent être reliés par un signal lumineux se propageant à la vitesse c. Dans ce cas $c^2(t_1 - t_2)^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2$. Cette relation suggère d'introduire une nouvelle quantité, nommée intervalle entre deux événements, définie par :

$$s_{1,2}^2 = c^2(t_1 - t_2)^2 - (x_1 - x_2)^2 - (y_1 - y_2)^2 - (z_1 - z_2)^2$$

Notons que le choix du signe + pour la composante temporelle de l'intervalle est tout à fait arbitraire ; c'est le plus répandu aujourd'hui. L'intervalle joue le rôle d'une distance dans notre espace-temps à quatre dimensions. Sa seule propriété évidente à ce point est de s'annuler quand les deux événements sont sur le cône de lumière l'un de l'autre. Cette propriété est indépendante du référentiel :que deux événements soit reliés ou non par un signal lumineux ne dépend pas de l'état de mouvement. Un intervalle nul est donc invariant par changement de référentiel.

Plus tard, nous établirons rigoureusement l'invariance d'un intervalle quelconque par changement de référentiel. Cette propriété peut aussi être établie par un raisonnement moins rigoureux, faisant appel à des hypothèses supplémentaires implicites. Pour cela, considérons deux événements infiniment voisins. L'intervalle, lui aussi infinitésimal, entre ces événements s'écrit alors, dans le référentiel $\mathcal{R}$:

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 \tag{4}$$

Dans un autre référentiel $\mathcal{R}'$, l'intervalle entre les deux mêmes événements s'écrit

$$\mathrm{d}s'^2 = c^2\mathrm{d}t'^2 - \mathrm{d}x'^2 - \mathrm{d}y'^2 - \mathrm{d}z'^2$$

On doit pouvoir écrire l'intervalle dans le nouveau référentiel comme une fonction de celui dans $\mathcal{R}$, fonction qui s'annule avec son argument parce qu'un intervalle nul est conservé. Cette fonction se développe à l'ordre le plus bas pour les intervalles infinitésimaux que nous manipulons : $\mathrm{d}s'^2 = a\mathrm{d}s^2$ où a est une constante ne dépendant que de la vitesse relative $\vec{u}$ des deux référentiels. En fait l'isotropie de l'espace impose que a ne dépende que du module u de la vitesse relative. Considérons maintenant un troisième référentiel $\mathcal{R}''$, en mouvement à la vitesse $\vec{v}$ par rapport à $\mathcal{R}$ et $\vec{w}$ par rapport à $\mathcal{R}'$. L'intervalle infinitésimal dans ce référentiel, $\mathrm{d}s''^2$, est tel que $\mathrm{d}s''^2 = a(v)\mathrm{d}s^2 = a(w)\mathrm{d}s'^2 = a(w)a(u)\mathrm{d}s^2$. Ainsi, la fonction a doit vérifier, pour tout triplet de vitesses relatives, la relation $a(v) = a(w)a(u)$. Ceci est manifestement impossible, sauf si $a = 1$, car le module de la vitesse $\vec{v}$ dépend de l'orientation relative des deux autres vitesses et pas seulement de leur module. Ainsi les intervalles infinitésimaux sont invariants par changements de référentiel. Tout intervalle pouvant être obtenu par une intégration d'intervalles infinitésimaux, on établit ainsi l'invariance d'un intervalle arbitraire.

4.3 Conséquences de la conservation de l'intervalle

Comme l'intervalle est un invariant relativiste, son signe l'est aussi, bien sûr. Nous distinguerons donc trois types d'intervalles :

– Si $s^2_{1,2} > 0$, nous dirons que nous avons à faire à un intervalle du genre «temps». La différence entre les temps des deux événements est supérieure à la distance (en unités convenables). Cela signifie que les deux événements peuvent être reliés par un signal se propageant moins vite que la lumière et qu'il peut donc y avoir un lien de causalité entre eux. En d'autres termes, les deux événements sont dans le cône de lumière l'un de l'autre. Notons enfin que tous les intervalles pris sur la ligne d'univers d'une particule matérielle sont du genre temps.

– Si $s^2_{1,2} < 0$, la distance spatiale entre les deux événements est plus grande que la distance temporelle. Aucun signal ne peut donc relier les deux événements, ce qui exclut tout lien de causalité (souvenons nous que c est une vitesse limite). Nous dirons alors que nous avons à faire à un intervalle de genre «espace».

– Si $s^2_{1,2} = 0$, les deux événements peuvent être reliés par un signal se propageant à la vitesse de la lumière. Nous dirons alors que l'intervalle est du genre «lumière».

L'intervalle, ou du moins son signe, est fortement lié à la notion de causalité. Il est donc essentiel que la nouvelle cinématique prédise l'invariance de l'intervalle, de sorte que les liens de causalité entre événements soient indépendants des observateurs. La causalité classique n'exigeait que des relations d'antériorité entre la cause et la conséquence. La causalité relativiste exige maintenant une deuxième condition : la cause et la conséquence doivent être dans le cône de lumière l'une de l'autre pour qu'une interaction ait eu le temps de se propager entre elles. La notion d'interaction instantanée à distance, commune en mécanique classique, disparaît donc en relativité. Il est donc important que les notions de passé et de futur, à l'intérieur du cône de lumière, soient elles aussi des invariants relativistes. Pour établir cette invariance, considérons le cône de lumière de l'événement

O et un événement M dans le futur de O. Nous avons défini le futur de O comme l'ensemble des événements du cône de lumière issu de O de coordonnée temporelle supérieure à celle de O ; M est donc situé dans cette partie du cône de lumière de O. Si, dans un changement de référentiel, M pouvait se retrouver dans le passé de O, cela impliquerait qu'il existe un changement de référentiel pour lequel M et O sont confondus. En effet, quel que soit le changement de référentiel, M reste dans le cône de lumière de O. Par continuité, passer du futur de O à son passé impose qu'il existe un changement de référentiel amenant M et O à coïncidence. Mais ceci est contraire à l'invariance de l'intervalle, qui deviendrait nul dans ce changement de référentiel, alors qu'il ne l'est pas initialement. Nous en déduisons donc que les notions de passé et de futur sont des invariants relativistes, ce qui est d'une importance cruciale pour que la causalité garde un sens. Notons que cette invariance ne tient que pour deux événements situés dans le cône de lumière l'un de l'autre. Si ce n'est pas le cas et si les deux événements ne peuvent être reliés par aucun lien de causalité, l'ordre des temps peut être modifié par un changement de référentiel (c'est par exemple le cas dans l'expérience de pensée du train et des deux signaux que nous avons détaillée plus haut). Bien sûr, nous préciserons quantitativement ces notions dans le paragraphe suivant quand nous disposerons de la forme explicite de la transformation qui permet de passer d'un référentiel à un autre.

4.4 Temps propre

Nous pouvons appliquer l'invariance de l'intervalle au problème des horloges en mouvement déjà abordé dans notre première expérience de pensée. Nous avons vu que la période d'une horloge (l'aller et retour d'un signal lumineux) n'était pas la même pour le contrôleur et le chef de gare. Nous allons établir ce résultat de manière plus générale en utilisant l'invariance de l'intervalle.

Considérons donc une particule, ou une horloge, en mouvement par rapport à un référentiel $\mathcal{R}$. Si ce mouvement est accéléré, il n'existe pas de référentiel galiléen dans lequel la particule soit au repos à tout instant. En revanche, on peut considérer à chaque instant le référentiel galiléen dont la vitesse $\vec{v}$ coïncide avec celle de la particule. Ce référentiel le référentiel $\mathcal{R}'$ est dit «tangent» au mouvement. A l'instant considéré, on peut faire en sorte que la particule soit située à l'origine O' de $\mathcal{R}'$, avec une vitesse nulle.

Considérons maintenant un intervalle de temps infinitésimal $\mathrm{d}t$ dans $\mathcal{R}$ durant lequel la particule se déplace de $\mathrm{d}\vec{l} = \vec{v}\,\mathrm{d}t$. Les deux événements correspondant aux deux extrémités de ce mouvement infinitésimal sont donc séparés par un intervalle $\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \mathrm{d}l^2 = c^2\mathrm{d}t^2(1 - v^2/c^2)$. Considérons maintenant les deux mêmes événements dans $\mathcal{R}'$, le référentiel tangent. Dans ce référentiel, la vitesse de la particule est nulle. Son déplacement est donc nul au premier ordre en $\mathrm{d}t'$. L'intervalle s'écrit donc $\mathrm{d}s^2 = c^2\mathrm{d}t'^2$. Nous appellerons «temps propre» l'intervalle de temps s'écoulant dans $\mathcal{R}'$ et nous le noterons $\mathrm{d}\tau = \mathrm{d}t'$. En rapprochant les deux expressions précédentes de l'intervalle, nous pouvons écrire $\mathrm{d}t = \gamma\mathrm{d}\tau$, en accord avec l'équation (3). Nous retrouvons, de façon plus générale, que l'intervalle de temps mesuré dans le référentiel tangent ou le référentiel propre dans le cas d'un mouvement uniforme, est plus court que l'intervalle mesuré dans un autre référentiel. A ce point, il est important de constater que $\mathrm{d}\tau$ est une quantité indépendante de l'observateur. Tout observateur, indépendamment de son état de mouvement par rapport à la particule, peut calculer un intervalle de temps propre à partir d'un intervalle de temps dans son référentiel, et en déduire, par intégration, le temps propre de la particule entre deux événements servant de référence. Le résultat obtenu sera le même pour tous les observateurs (si ils choisissent les mêmes événements de référence).

Ce raisonnement, appliqué à des intervalles de temps infinitésimaux, peut être généraliser à des intervalles arbitraires. A chaque instant, on peut définir un référentiel tangent pour la particule. On peut alors déterminer le «temps propre» de la particule, τ, en intégrant les intervalles de temps propre infinitésimaux. En intégrant aussi la relation entre temps propre et intervalle de temps dans $\mathcal{R}$, on montrera que la durée propre est toujours inférieur à la durée mesurée dans $\mathcal{R}$.

Cette «dilatation des temps» a plusieurs conséquences pratiques et mesurables qui ont apporté des confirmations éclatantes au principe de relativité. Considérons d'abord une particule instable de durée de vie moyenne T. Dans quel référentiel doit-on utiliser cette durée de vie? Elle n'a bien sûr de signification que dans le référentiel propre de la particule. C'est en effet une «horloge» interne à la particule qui déclenche sa désintégration. Dans le référentiel du laboratoire, $\mathcal{R}$, la durée de vie moyenne de la particule sera alors γT (nous supposerons, pour fixer les idées, que la particule est en mouvement rectiligne uniforme – γ est donc une constante). Si la vitesse de la particule est très proche de celle de la lumière, le facteur de dilatation temporel est très grand et la durée de vie mesurée dans le référentiel du laboratoire est très grande par rapport à la durée de vie intrinsèque. C'est cet effet qui permet d'observer, dans les chambres à bulles ou à fils, les traces de particules à durée de vie très courte. C'est aussi la dilatation temporelle qui permet d'expliquer le parcours des muons dans l'atmosphère.

Les particules cosmiques qui traversent l'atmosphère sont, entre autre, constituées de mésons μ ou muons, particules de charge $\pm e$ et de masse 207 fois supérieure à celle de l'électron. Les muons de désintègrent exponentiellement avec une durée de vie moyenne T_p, mesurée en laboratoire et égale à 2,2 μs. En 1941, B. Rossi et D. Hall ont comparé les flux de muons, dont la quantité de mouvement était de 500 MeV$/c$, à des altitudes différentes. Entre les altitudes $z_1 = 3\,240$ m et $z_2 = 1\,616$ m ils ont mesuré un rapport de flux $n_2/n_1 = 0,594$. En écrivant la loi de décroissance des muons dans le référentiel terrestre, on peut montrer que leur durée de vie moyenne observée est égale à $T = -(z_1 - z_2)/v \ln(n_2/n_1)$. La comparaison de T et T_p donne : $T/T_p \approx 1/\sqrt{1 - v^2/c^2}$.

Citons également le célèbre «paradoxe des jumeaux», dû au physicien Paul Langevin. De deux frères jumeaux, l'un reste sur Terre et l'autre vole vers α du Centaure, distante d'à peu près 4 années-lumière, avec une vitesse constante, proche de celle de la lumière. A peine arrivé, le jumeau voyageur fait demi-tour et revient sur Terre à la même vitesse. A l'arrivée, le temps écoulé pour le jumeau terrestre est de huit ans (4 ans pour l'aller, autant pour le retour). En revanche, pour le voyageur, le temps écoulé n'est que de $8/\gamma$ ans, beaucoup plus court. Le jumeau voyageur revient donc sur Terre plus jeune que son frère ! Le paradoxe est lié à l'apparente symétrie entre les deux jumeaux, incompatible avec cette différence d'âge : dans le référentiel du voyageur, c'est le jumeau terrestre qui s'éloigne et se rapproche à grande vitesse.

La «solution» de ce paradoxe apparent est que le référentiel du jumeau voyageur n'est pas un référentiel galiléen. Le temps propre tel que nous l'avons défini n'est pas le temps mesuré dans un référentiel donné. C'est une accumulation de temps infinitésimaux tous calculés dans des référentiels galiléens différents, les référentiels tangents au mouvement accéléré de la fusée. Considérer la situation du point de vue du jumeau voyageur reviendrait à définir un temps pour un référentiel bien défini (celui du jumeau voyageur), accéléré. Ceci n'est pas possible dans le cadre de la relativité restreinte. En relativité générale, le temps est affecté par la gravitation ou de manière équivalente par l'accélération. On peut alors effectivement définir un temps pour le jumeau voyageur et retrouver rigoureusement la dissymétrie entre les deux jumeaux. Ce paradoxe des jumeaux peut aussi être traité dans un espace torique, par exemple, dans lequel il n'est nul besoin de faire un demi-tour pour revenir à son point de départ. Ce cas est traité dans l'article [7].

Cette brève étude du paradoxe des jumeaux introduit naturellement la notion de célérité. Si on désire voyager loin, ce qui importe c'est le temps propre utilisé (celui dans lequel on vieillit) et la distance parcourue dans le référentiel immobile. On peut définir alors une quantité, que nous nommerons célérité, en termes du temps estimé dans un référentiel et de l'espace estimé dans un autre. De façon évidente, la célérité est le produit de la vitesse ordinaire par le facteur γ de dilatation du temps. Elle peut donc être très supérieure à la vitesse de la lumière, sans que la causalité relativiste ne soit violée puisque la célérité n'est pas une vitesse définie dans un référentiel unique.

5 Transformation de Lorentz

Après une approche qualitative, qui nous a permis de comprendre certaines caractéristiques essentielles de la nouvelle cinématique, il nous reste à donner la forme explicite de la transformation de Lorentz, qui permet de passer d'un référentiel à un autre. Nous allons, en fait, oublier pour un temps tout ce que nous avons appris dans les paragraphes précédents et essayer de construire toutes les transformations obéissant à un certain nombre de symétries fondamentales, telles que l'isotropie de l'espace ou l'invariance par translation dans le temps. Nous verrons qu'il n'y a en fait que quatre formes possibles pour une telle transformation. Deux d'entre elles sont inacceptables parce qu'elles conduiraient à abandonner le principe de causalité. Les deux dernières sont la transformation de Galilée, que nous rejetterons également car elle n'obéit pas au principe de relativité, et enfin la transformation de Lorentz. Au cours de cette recherche, nous verrons apparaître certaines propriétés essentielles de la transformation de Lorentz que nous discuterons dans le dernier paragraphe de cette section.

5.1 Forme de la transformation de Lorentz

Le choix d'axes pour les deux repères est, encore une fois, celui illustré par la figure 1. Nous cherchons donc une transformation $L(u)$ permettant d'exprimer les coordonnées (x', y', z', ct') d'un événement dans $\mathcal{R}'$ en fonction de celles dans $\mathcal{R}$, (x, y, z, ct). Rappelons que u est la projection algébrique de la vitesse de $\mathcal{R}'$ par rapport à $\mathcal{R}$ sur l'axe du mouvement. Notons tout de suite qu'avec nos conventions l'événement $(0, 0, 0, 0)$ dans $\mathcal{R}$ se transforme en l'événement origine $(0, 0, 0, 0)$ dans $\mathcal{R}'$. Avec ce choix d'axes nous obtiendrons la transformation de Lorentz dite «spéciale». Une simple combinaison avec les rotations et symétries nous permettra ensuite d'obtenir le groupe de Lorentz complet, dont le groupe spécial est un sous-groupe, décrivant des changements de référentiels quelconques.

Nous imposerons d'abord à L d'être une transformation linéaire et homogène. L'invariance de la physique dans une translation arbitraire de l'espace ou du temps impose cette linéarité. Ensuite, l'ensemble des transformations de Lorentz, paramétré par la vitesse relative u, doit former un groupe. Considérons en effet trois référentiels : $\mathcal{R}$, $\mathcal{R}'$ en mouvement à la vitesse u par rapport à $\mathcal{R}$, et $\mathcal{R}''$ en mouvement à la vitesse v par rapport à $\mathcal{R}'$ et w par rapport à $\mathcal{R}$ (w n'est bien sûr pas égal à $u+v$ puisque nous avons abandonné le cadre de la relativité galiléenne). La transformation de $\mathcal{R}$ vers $\mathcal{R}''$ peut s'écrire $L(w)$ ou $L(v)L(u)$ (ce produit étant à comprendre comme la composition de deux applications linéaires il doit être lu de droite à gauche). Le produit de deux transformations de Lorentz définit donc une application de composition interne qui possède toutes les propriétés d'une loi de groupe. Il existe bien un élément neutre, l'identité $L(0)$, correspondant au passage d'un référentiel à lui-même c'est-à-dire par une vitesse relative nulle. Chaque élément possède un inverse. Il doit en être ainsi, pour qu'à tout événement dans $\mathcal{R}$ corresponde un seul jeu de coordonnées dans $\mathcal{R}'$. La transformation inverse est celle qui donne les coordonnées dans $\mathcal{R}$ en fonction de celles dans $\mathcal{R}'$. La vitesse de l'origine O dans $\mathcal{R}'$ doit bien sûr être $-u$. Si la vitesse de O par rapport à O' n'était pas opposée à la vitesse de O' par rapport à O, nous briserions le principe de relativité. La transformation inverse de $L(u)$ doit ainsi être la transformation de Lorentz correspondant à la vitesse $-u$, qui est celle de $\mathcal{R}$ mesurée dans $\mathcal{R}'$, donc $L(u)^{-1} = L(-u)$.

Ce groupe, que nous appellerons «groupe de Lorentz», est-il commutatif ? Il est possible de montrer que tout groupe paramétré par un paramètre unique, à condition que ce paramétrage soit suffisamment continu et dérivable, est isomorphe au groupe additif des réels [1]. Il en résulte immédiatement que tous ces groupes sont commutatifs. Notons que cet isomorphisme indique qu'on peut, par un changement de variable adéquat, paramétrer le groupe par un paramètre additif, déterminé de façon univoque à un facteur près (ce qu'un physicien traduirait par «à un choix d'unités près»). On pourrait ainsi trouver un paramétrage $\phi(u)$ tel que la composition de deux transformations de Lorentz s'écrive $L(\phi(v))L(\phi(u)) = L(\phi(u) + \phi(v))$. Nous verrons par la suite que ce paramétrage additif a une signification physique très claire.

Penchons nous d'abord sur les lois de transformation des coordonnées y et z. Comme les axes Ox et $O'x'$ coïncident à tout instant, le fait que $y = z = 0$ implique que $y' = z' = 0$ pour toutes les valeurs de x et t. Ces deux dernières coordonnées ne peuvent donc intervenir dans les lois de transformation de y et z, qui se résument donc à :

$$\left\{ \begin{array}{ccccc} y' & = & ay & + & bz \\ z' & = & b'y & + & a'z \end{array} \right.$$

Les axes Oy et $O'y'$ doivent coïncider à $t = 0$. Si b et b' n'étaient pas nuls, l'axe $O'y'$ correspondrait à des valeurs simultanément non nulles de y et z et ne pourrait donc coïncider avec Oy. La transformation se réduit donc à un simple facteur d'échelle sur y et z. L'isotropie de l'espace impose de plus que les facteurs affectant y et z soient identiques. On a donc simplement $y' = ay$ et $z' = az$.

1. Nous ne démontrerons pas ici cette propriété. On en trouvera une démonstration très élémentaire dans [6]. On peut donner des exemples simples de cette propriété. Le groupe multiplicatif des réels (paramétré par la valeur de l'élément) admet une représentation additive évidente qui n'est autre que le logarithme népérien. Le groupe des rotations autour d'un point, paramétré par l'angle de rotation, est directement paramétré sous forme additive.

Montrons maintenant que ce facteur a vaut nécessairement 1. Nous avons montré effectivement que $L(u)^{-1} = L(-u)$. La transformation inverse est donc décrite par le facteur $a(-u)$, mais aussi par le facteur $1/a(u)$. L'isotropie de l'espace impose de plus que le facteur a ne dépende pas de l'orientation de la vitesse par rapport à l'axe Oy. On a donc $a = 1/a$ soit $a = \pm 1$. Si le choix des orientations des axes dans les deux référentiels est cohérent, on a finalement $a = 1$. Nous avons montré que la transformation de Lorentz laisse invariantes les coordonnées perpendiculaires à la vitesse relative.

Intéressons nous maintenant à la transformation des coordonnées x et ct. L'invariance par translation perpendiculaire à l'axe Ox impose évidemment que y et z n'interviennent pas dans la loi de transformation de x. De même, à x et ct donnés, le temps ct' ne doit pas dépendre de y ou z. Finalement, on peut exprimer la loi de transformation la plus générale par une relation matricielle 2×2 :

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} a(u) & b(u) \\ e(u) & f(u) \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}$$

On peut considérablement préciser la forme de cette transformation par un simple argument de symétrie. Considérons dans le référentiel $\mathcal{R}'$ un axe $O'X'$ confondu avec $O'x'$ mais d'orientation opposée. En un mot, $X' = -x'$. De même, considérons dans $\mathcal{R}$ l'axe OX de directions opposée à celui de l'axe Ox, avec $X = -x$. La transformation donnant X et ct en fonction de X' et ct' est identique au changement de référentiel que nous étudions. En effet, dans les deux cas, la vitesse du nouveau référentiel selon l'axe des x ou X est u. La vitesse de O est en effet $-u$ sur $O'x'$ et donc u sur $O'X'$. On en déduit que:

$$\begin{pmatrix} ct \\ X \end{pmatrix} = \begin{pmatrix} a & b \\ e & f \end{pmatrix} \begin{pmatrix} ct' \\ X' \end{pmatrix}$$

qu'on peut mettre sous la forme:

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} a & -b \\ -e & f \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix}$$

Mais cette dernière relation est aussi la transformée inverse de la transformation cherchée :

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \frac{1}{af - be} \begin{pmatrix} f & -b \\ -e & a \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix}$$

De la comparaison de ces deux formules, on tire immédiatement que $af - be = 1$ mais aussi que $a = f \equiv \gamma$ (nous changeons un peu les notations pour évoluer vers la forme standard de la transformation de Lorentz).

Nous pouvons préciser encore la forme de la transformation en utilisant la commutativité du groupe de Lorentz spécial. En écrivant simplement que $L(w) = L(u)L(v) = L(v)L(u)$, on trouve les relations :

$$\begin{aligned} \gamma(w) &= \gamma(u)\gamma(v) + e(u)b(v) \\ &= \gamma(v)\gamma(u) + e(v)b(u) \end{aligned}$$

qui ne peuvent être vérifiées pour deux vitesses arbitraires que si $e(u)/b(u)$ est une constante ou si l'une de ces fonctions s'annule identiquement. Dans le premier cas, un choix convenable des unités d'espace et de temps permet de prendre la constante égale à ± 1. Nous avons donc 4 cas à distinguer :

– $e(u) = -b(u)$. La matrice est donc antisymétrique et ses deux coefficients vérifient $\gamma^2 + b^2 = 1$. On peut donc poser $\gamma = \cos\theta$ et $b = \sin\theta$ et la matrice de transformation s'écrit :

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

C'est une simple rotation autour de l'origine dans l'espace-temps. La représentation additive de ce groupe est l'angle de rotation θ.

– $b(u) = 0$. La valeur du déterminant impose alors $\gamma = 1$ et la matrice de la transformation s'écrit :

$$\begin{pmatrix} 1 & 0 \\ e(u) & 1 \end{pmatrix}$$

Ecrire ensuite que le mouvement de O' s'effectue à la vitesse u, c'est-à-dire que $x = 0$ implique $x' = -ut'$, fixe $e(u) = -u/c \equiv -\beta$. Ce groupe est simplement celui de Galilée dont le paramètre additif est la vitesse u ou la vitesse réduite $\beta = u/c$.

– $e(u) = 0$. Là encore, on doit avoir $\gamma = 1$ et la matrice s'écrit :

$$\begin{pmatrix} 1 & b(u) \\ 0 & 1 \end{pmatrix}$$

Ce groupe diffère du groupe de Galilée en ce qu'il transforme le temps et non l'espace ; on le nomme groupe de Caroll.

– $e(u) = b(u)$. On a alors $\gamma^2 - b^2 = 1$ et on peut poser $\gamma = \cosh\phi$ et $b(u) = -\sinh\phi$ (la raison de ce choix de signe apparaîtra clairement plus tard). La matrice de transformation:

$$\begin{pmatrix} \cosh\phi & -\sinh\phi \\ -\sinh\phi & \cosh\phi \end{pmatrix}$$

est alors simplement celle d'une rotation hyperbolique (rotation autour de l'origine d'un angle imaginaire pur). La représentation additive de ce groupe (que nous appellerons «groupe de Lorentz», en faisant fi du suspense) est simplement l'angle de rotation ϕ, dont nous donnerons dans un moment l'interprétation physique.

Nous n'avons utilisé jusque là que des arguments très généraux de symétrie et de réciprocité. Il est déjà très remarquable que nous n'ayons plus le choix qu'entre quatre groupes, chacun paramétré par un seul paramètre additif. Pour choisir la forme correcte, nous pouvons employer deux arguments supplémentaires. Si nous voulons que la causalité soit une notion indépendante du référentiel, il faut au moins exiger de la transformation qu'elle préserve l'ordre temporel des événements (au moins pour certains couples d'événements, ceux qui sont dans le cône de lumière l'un de l'autre). Il doit donc exister des classes d'événements pour lesquelles le signe de t ne doit changer dans aucune transformation. Il est facile de voir que le groupe des rotations spatio-temporelles et le groupe de Caroll contiennent des transformations changeant le signe de t pour tout événement. Ils sont donc exclus pour préserver la causalité. Il ne nous reste alors que le groupe de Galilée et le groupe de Lorentz. Le premier, qui conduit à la composition des vitesses au sens ordinaire, n'est pas acceptable. La transformation cherchée doit appartenir donc au groupe de Lorentz (à celui des rotations hyperboliques). Il ne nous reste plus qu'à préciser la valeur du paramètre additif $\phi(u)$, que nous appellerons «rapidité». Il suffit pour cela d'écrire que O est animé, dans $\mathcal{R}'$, d'un mouvement uniforme à la vitesse $-u$ ($x' = -ut'$). On déduit immédiatement de la transformation que $x' = \sinh\phi\, ct$ et $ct' = \cosh\phi\, ct$. D'où $\tanh\phi = u/c = \beta$, soit encore $\cosh\phi = \gamma$ et $\sinh\phi = \gamma\beta$.

Notons que le facteur γ peut prendre des valeurs arbitrairement grandes. La rapidité peut donc aussi être arbitrairement grande. Si la vitesse u est limitée par la vitesse de la lumière c, il n'en est pas de même pour le paramètre «naturel» du groupe de Lorentz. Avec ces valeurs, nous avons complètement déterminé la transformation de Lorentz, qui peut s'écrire :

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \tag{5}$$

la transformation inverse étant évidemment donnée par (il suffit de changer le signe de la vitesse relative)

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta \\ \gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix} \tag{6}$$

A toutes les relations précédentes, il convient bien sûr d'ajouter l'invariance des coordonnées transverses y et z. Notons immédiatement que, si on ne retient dans la transformation de Lorentz que les termes du premier ordre en u/c, on retombe sur la transformation de Galilée. La cinématique classique apparaît donc bien comme une limite de la cinématique relativiste pour des vitesses d'entraînement faibles devant celle de la lumière.

Notons que nous avons fait, dans tout ce paragraphe, un choix d'axes bien particulier. Comme il nous a conduit à une forme univoque de la transformation de Lorentz, nous n'avons pas restreint la généralité. On peut avoir cependant à composer des transformations de Lorentz correspondant à des directions de vitesses différentes. La transformation de Lorentz s'écrira alors $L(\vec{u})$, où $\vec{u}$ est le vecteur vitesse de $\mathcal{R}'$ par rapport à $\mathcal{R}$, de direction arbitraire. Pour écrire ce genre de changement de référentiel, il convient de composer la transformation que nous venons d'écrire avec les rotations arbitraires d'espace, avec les réflexions d'espace et même, éventuellement, avec les réflexions du temps. On obtient ainsi le «groupe de Lorentz complet», qui décrit tous les changements de référentiels. On distingue parfois, à l'intérieur du groupe complet , plusieurs sous-groupes :

- Le groupe de Lorentz «propre» comprenant la transformation de Lorentz combinée avec les rotations spatiales. Sauf cas spécial, il suffit à décrire un changement de référentiel avec une direction de vitesse arbitraire. Le déterminant de la matrice correspondante est 1.
- Le groupe de Lorentz «orthochrone» contient le groupe de Lorentz éventuellement combiné avec des réflexions d'espace (nous avions exclu explicitement ces transformations dans notre discussion en imposant aux directions des axes d'être consistantes). Le déterminant de la matrice peut alors être égale à ± 1.

6 Propriétés de la transformation de Lorentz

Avant d'appliquer la transformation de Lorentz à des situations physiques, nous allons nous pencher plus en détails sur certaines de ses propriétés. La première, qui découle directement de l'expression de la transformation comme une rotation hyperbolique, est qu'elle conserve l'intervalle. Nous aurions en fait pu prendre cette hypothèse comme point de départ et construire la transformation à partir de là. Un point essentiel de ce paragraphe sera de comprendre qu'il y a trois quantités fondamentales différentes décrivant la vitesse d'un référentiel par rapport à un autre, correspondant à trois situations expérimentales différentes pour déterminer cette vitesse.

6.1 Composition des transformations

Revenons à la situation déjà évoquée de trois référentiels en mouvement relatif. La loi de composition des rapidités nous permet d'écrire de manière évidente: $\phi(w) = \phi(u) + \phi(v)$. Nous pouvons en tirer facilement la loi de composition des vitesses relativistes, sous une forme simplifiée correspondant à des vitesses qui sont toutes colinéaires (nous généraliserons au paragraphe suivant), en écrivant w en fonction de u et v. Il suffit pour cela de prendre le cosh et le sinh de l'équation précédente. On obtient alors :

$$\begin{array}{lclcl} \sinh\phi(w) & = & \beta(w)\gamma(w) & = & \gamma(u)\gamma(v)(\beta(u)+\beta(v)) \\ \cosh\phi(w) & = & \gamma(w) & = & \gamma(v)\gamma(u)(1+\beta(u)\beta(v)) \end{array}$$

d'où on tire immédiatement:

$$\beta(w) = \frac{\beta(u)+\beta(v)}{1+\beta(u)\beta(v)} \quad \text{ou} \quad w = \frac{u+v}{1+uv/c^2} \tag{7}$$

Cette loi remplace la simple addition des vitesses de la relativité galiléenne. Notons là encore qu'on retrouve la loi galiléenne d'addition des vitesse pour la composition de vitesses toutes deux petites devant la vitesse de la lumière. Remarquons également que cette loi prédit correctement l'invariance de la vitesse de la lumière : si $\beta(u) = 1$ ou si $\beta(v) = 1$, on trouve immédiatement $\beta(w) = 1$. Cela montre aussi qu'on ne peut dépasser la vitesse de la lumière en courant dans le couloir d'un train se déplaçant à une vitesse proche de c. Encore une fois, si l'accumulation de vitesses ne peut conduire à une vitesse supérieure à celle de la lumière, les rapidités s'accumulent sans limite. Il est d'ailleurs fructueux à ce point de comparer ces notions de vitesse, de célérité (vue au paragraphe 4.4) et de rapidité.

6.2 Vitesse, célérité et rapidité

Nous sommes maintenant en possession de trois quantités différentes décrivant le mouvement d'un référentiel par rapport à un autre. Il est temps d'examiner les différences entre ces quantités et de préciser leur sens physique. La première définition de la vitesse relative de $\mathcal{R}'$ (le référentiel du contrôleur, pour reprendre le vocabulaire ferroviaire) par rapport à $\mathcal{R}$ (le chef de gare) est la vitesse u, vitesse de l'origine O' dans $\mathcal{R}$. Imaginons, pour bien insister sur le sens physique de ces quantités, que le contrôleur désire déterminer sa vitesse. La première méthode est de repérer, sur les horloges de deux gares successives, ses temps de passage. La liste des tarifs, imprimée dans $\mathcal{R}$, lui donnant la distance (mesurée dans $\mathcal{R}$) entre ces gares, il en déduira sa vitesse. Cette vitesse, mesurée dans l'espace de $\mathcal{R}$ avec le temps de $\mathcal{R}$, ou dans l'espace de $\mathcal{R}'$ avec le temps de $\mathcal{R}'$, est bien entendu la vitesse u.

L'autre définition que nous avons déjà rencontrée est celle de la célérité. Rappelons que c'est la définition qui intéresse le voyageur interstellaire, puisqu'elle mesure la distance parcourue dans le référentiel fixe par unité de temps du mobile. Nous avons vu que la célérité s'exprimait par γu (ou $\gamma\beta = \sinh\phi$ en unité de la vitesse de la lumière). Cette vitesse est celle que déterminerait le contrôleur en mesurant la durée qui s'écoule entre les deux gares avec sa propre montre, connaissant les distances dans $\mathcal{R}$. Notons enfin que célérité et vitesse sont identiques en relativité galiléenne, en raison de l'universalité du temps.

Nous sommes aussi en possession d'une troisième définition de la vitesse en termes de la rapidité ϕ. Si nous savons déjà que son intérêt réside dans son caractère additif (elle sert à paramétrer le groupe de Lorentz), à quelle expérience correspond-elle ? Le contrôleur dispose d'une troisième méthode pour déterminer sa vitesse, même si les vitres sont occultées. Supposons qu'il soit initialement immobile dans la première gare mais qu'il dispose d'un accéléromètre (un simple fil à plomb lui suffirait). Cet accéléromètre mesure la variation de la vitesse du train par unité de temps du train (ce temps est un temps propre). En intégrant les indications de l'accéléromètre sur toute la phase d'accélération, le contrôleur pourra déterminer sa vitesse finale.

On peut traiter ce cas très simple de cinématique d'un mouvement accéléré sans recourir à la relativité générale. Le référentiel du train n'est plus galiléen. Nous considérerons en revanche comme référentiel $\mathcal{R}'$ le référentiel tangent au mouvement pour une valeur du temps propre τ (obtenu, rappelons-le, par intégration des temps propres dans les référentiels tangents successifs). A un instant donné, la vitesse de ce référentiel par rapport à $\mathcal{R}$ est v, qui passe de 0 à u pendant la phase d'accélération. Pendant un intervalle de temps infinitésimal $\mathrm{d}\tau$, la vitesse du train dans le référentiel tangent passe de 0 à $\mathrm{d}v'$. L'accélération mesurée dans le train (ou plutôt dans le référentiel tangent) vaut donc $a = \mathrm{d}v'/\mathrm{d}\tau$. Pendant l'intervalle de temps correspondant, la vitesse dans $\mathcal{R}$ passe de v à $v+\mathrm{d}v$. En utilisant la loi de composition des vitesses du paragraphe précédent, nous pouvons écrire l'accroissement de vitesse $\mathrm{d}v'$dans $\mathcal{R}'$ en fonction de la nouvelle vitesse $(v+\mathrm{d}v)$ dans $\mathcal{R}$ et de la vitesse v de $\mathcal{R}'$ par rapport à $\mathcal{R}$:

$$\mathrm{d}v' = \frac{v + \mathrm{d}v - v}{1 - v(v + \mathrm{d}v)/c^2}$$

soit encore

$$\mathrm{d}v' = \frac{\mathrm{d}v}{1 - v^2/c^2}$$

La vitesse finale mesurée par cette méthode sera donc:

$$\int a\mathrm{d}\tau = \int \mathrm{d}v' = \int \frac{\mathrm{d}v}{1 - v^2/c^2} = c\tanh^{-1}\frac{u}{c} = c\phi$$

Elle coïncide donc avec la définition de la rapidité, qui prend ainsi un sens physique très fort. Cette définition en termes d'accélération accumulée nous fait comprendre pourquoi la rapidité n'est pas bornée. On peut en effet avoir un mouvement indéfiniment accéléré en relativité restreinte (comme celui d'une particule chargée soumise à une force constante dans un champ électrique uniforme). La vitesse tend asymptotiquement vers c, l'accélération habituelle (mesurée dans le référentiel fixe) tend vers zéro mais l'accélération mesurée comme ci-dessus demeure constante et la rapidité s'accumule. Notons pour finir qu'en relativité galiléenne la rapidité et la vitesse coïncident aussi.

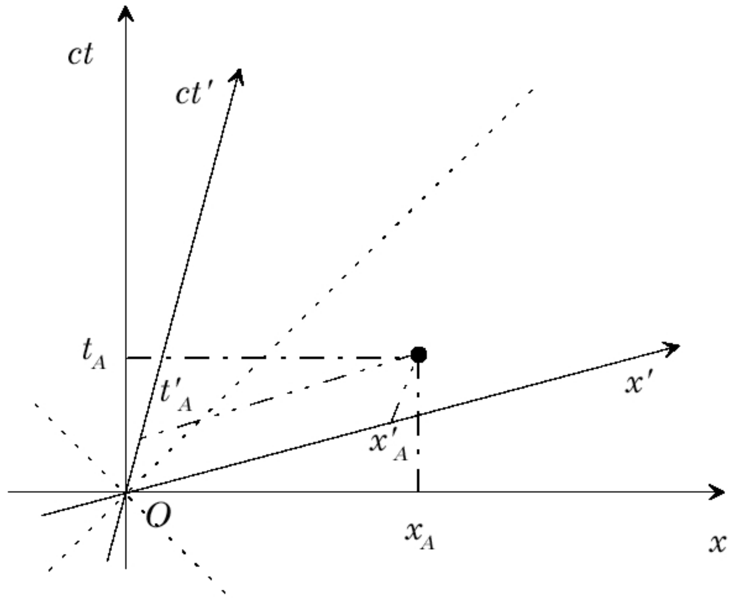

FIG. 6 – *La transformation de Lorentz interprétée géométriquement comme une rotation hyperbolique dans l'espace-temps. Ses directions propres coïncident avec le cône de lumière de l'origine.*

6.3 Géométrie de la transformation de Lorentz

Penchons nous maintenant sur l'interprétation géométrique de la transformation de Lorentz dans l'espace-temps. Si la rotation ordinaire autour de l'origine est bien connue, il n'en est pas nécessairement de même pour une rotation hyperbolique. La figure 6 illustre la géométrie de cette transformation. Au contraire d'une rotation, elle ne conserve pas l'angle entre les axes. Pour $u > 0$, les deux nouveaux axes sont à l'intérieur du premier quadrant. On a illustré sur la figure 6 un événement (par un point), ainsi que ses nouvelles et ses anciennes coordonnées, obtenues par projection sur les axes correspondants.

On peut préciser encore notre interprétation de cette transformation, en considérant ses valeurs propres et vecteurs propres. L'équation caractéristique s'écrit $\lambda^2 - 2\gamma\lambda + 1 = 0$. Elle admet donc deux valeurs propres réelles inverses l'une de l'autre $\exp \pm\phi$. Les vecteurs propres s'obtiennent sans difficulté. Ils correspondent à $x = ct$ (pour la valeur propre supérieure à un) et $x = -ct$ (pour l'autre valeur propre). Les directions propres de la transformation de Lorentz ne sont autres que celles du cône de lumière du point O (directions que nous avons également représentées sur la figure 6 en pointillé).

7 Conséquences de la transformation de Lorentz

7.1 Retour sur nos expériences de pensée

Nous reviendrons brièvement, dans ce paragraphe, sur les deux expériences de pensée que nous avons discutées au début de ce chapitre. Nous allons les décrire en termes d'événements et utiliser la transformation de Lorentz complète pour exprimer les changements de référentiels. Nous pourrons en particulier préciser la valeur de la non-simultanéité dans la seconde expérience, ce que nous ne pouvions faire sans la transformation. Pour la première expérience, il nous faut distinguer trois événements : le départ de l'impulsion de O' ($x'_1 = y'_1 = 0$, $t'_1 = 0$, nous n'écrirons jamais la coordonnée z qui ne joue aucun rôle) ; la réflexion sur le miroir de coordonnées ($x'_2 = 0, y'_2 = L, t'_2 = L/c$), et le retour de l'impulsion en O' de coordonnées ($x'_3 = 0, y'_3 = 0, t'_3 = 2L/c$). La transformation de Lorentz donne les positions de ces trois événements dans $\mathcal{R}$:

$$\left\{\begin{array}{lll} x_1 = 0 & y_1 = 0 & t_1 = 0 \\ x_2 = \gamma L u/c & y_2 = L & t_2 = \gamma L/c \\ x_3 = 2\gamma L u/c & y_3 = 0 & t_3 = 2\gamma L/c \end{array}\right.$$

qui nous redonne immédiatement la dilatation des temps.

Pour la seconde expérience, nous n'avons à considérer que l'allumage des feux. La réception des signaux coïncide en effet avec l'événement origine. On a , dans $\mathcal{R}$:

$$\left\{\begin{array}{ll} x_A = -L & t_A = -L/c \\ x_B = L & t_B = -L/c \end{array}\right.$$

On en déduit immédiatement, dans $\mathcal{R}'$,

$$\left\{\begin{array}{ll} x'_A = -\gamma(1-\beta)L & t'_A = -\gamma(1-\beta)L/c \\ x'_B = \gamma(1+\beta)L & t'_B = -\gamma(1+\beta)L/c \end{array}\right.$$

ce qui montre clairement que les instants d'allumage dans $\mathcal{R}'$ différent de $2\gamma\beta L/c$.

7.2 Intervalles et simultanéité

Nous allons établir deux propriétés qui nous permettront de revenir sur la notion de causalité relativiste.

Propriété n°1 : *Si deux événements sont séparés par un intervalle de genre temps, il existe un référentiel dans lequel ils se produisent au même endroit.*

Considérons donc deux événements dans un référentiel quelconque $\mathcal{R}$. On peut toujours choisir l'un comme origine et faire en sorte, par un choix d'axes, que l'autre se produise sur l'axe Ox. Leurs coordonnées sont alors $(0, 0)$ et (ct, x). On a $|ct| > |x|$ puisque leur intervalle est du genre temps. Prenons un nouveau référentiel $\mathcal{R}'$. Dans ce référentiel, les deux événements se produisent au même endroit si $x' = \gamma(x - ut) = 0$. Il suffit pour cela que $u = x/t$, qui est bien inférieure à c. Notons que le carré de l'intervalle est alors simplement le carré du temps propre, du temps qui s'écoule entre les deux événements dans le référentiel où ils se produisent au même point.

Si les deux événements se produisent au même endroit dans ce référentiel, il peuvent être liés par un lien de causalité. Une autre manière de voir cette propriété est d'imaginer un signal se propageant d'un événement à l'autre. Comme l'intervalle est du genre temps, ce signal se propage moins vite que la lumière. On peut donc lui associer un référentiel qui n'est, d'ailleurs, autre que $\mathcal{R}'$. On peut noter enfin que l'instant t' auquel se produit le second événement dans $\mathcal{R}'$ est positif si et seulement si t est positif. Si deux événements se produisent au même point dans un référentiel, l'ordre temporel des événements n'est modifié par aucune transformation de Lorentz. En un mot, la notion de causalité est complètement préservée par les changements de référentiels.

Propriété n°2 : *Si deux événements sont séparés par un intervalle du genre espace, il existe un référentiel où ils se produisent simultanément.*

Nous emploierons les mêmes notations. Cette fois, $|ct| < |x|$. Dans $\mathcal{R}'$ les deux événements se produisent au même instant si $t' = \gamma(t - ux/c^2) = 0$, c'est à dire si $u = c^2t/x$, qui est, là encore, plus petite que c. Le fait que les événements se produisent au même instant à des endroits différents prouve qu'ils ne peuvent être liés par un lien de causalité, la relativité n'admettant pas d'action instantanée à distance. On peut voir rapidement que, dans ce cas, le signe de t dépend de la transformation. Comme il n'y a pas de relation de causalité relativiste entre ces événements, leur ordre temporel peut dépendre du référentiel. Notons enfin que le carré de l'intervalle est dans ce cas l'opposé de la distance entre les deux événements dans le référentiel où ils se produisent au même instant. Il s'agit d'une longueur propre, telle que nous la définirons plus précisément dans un instant.

7.3 Contraction des longueurs

Le problème que nous abordons ici est celui de la définition de la longueur d'un objet en mouvement. Pour fixer les idées, nous considérons une règle rigide, de longueur L', immobile dans le référentiel mobile $\mathcal{R}'$, confondue avec l'axe $O'x'$ et ayant une extrémité en O'. Comment des observateurs de $\mathcal{R}$ peuvent-ils déterminer la longueur de cette règle? Deux stratégies sont possibles.

L'observateur O peut d'abord déterminer la vitesse u de la règle, par exemple par vélocimétrie Doppler. Il peut alors mesurer la durée Δt pendant laquelle la règle défile devant lui. Il en déduira alors sa longueur $L = u\Delta t$. On peut aussi prendre une photographie instantanée de la règle. En fait, on peut repérer, à un instant t donné, les observateurs de $\mathcal{R}$ qui sont en face des extrémités de la règle et mesurer leur distance. A titre d'exercice, nous allons examiner ces deux procédures et montrer qu'elles fournissent la même longueur.

Dans la première méthode, les deux événements A et B à considérer sont le passage en O des deux extrémités de la règle. Les coordonnées de ces deux événements dans $\mathcal{R}'$ sont $x'_A = 0$ et $t'_A = 0$ d'une part, $x'_B = -L'$ et $t'_B = L/u$ d'autre part. (par convention, l'extrémité O' de la règle passe d'abord devant O). Leurs coordonnées dans $\mathcal{R}$ sont respectivement $x_A = 0$ et $t_A = 0$, $x_B = 0$ et $t_B = \gamma L'/u(1 - v^2/c^2) = L'/\gamma u$. On retrouve évidemment que ces deux événements se produisent en O. La durée de passage Δt est égale à t_B, et la longueur L de la règle ainsi mesurée est $L = L'/\gamma$.

Pour la deuxième méthode, on considère les deux événements A et B représentant les extrémités de la règle à un instant donné dans $\mathcal{R}$, par exemple l'instant origine pour nous simplifier la tâche. Les coordonnées dans $\mathcal{R}$ de ces événements sont ainsi $x_A = 0$ et $t_A = 0$, $x_B = -L$ et $t_B = 0$. Leurs coordonnées dans $\mathcal{R}'$ sont donc $x'_A = 0$ et $t'_A = 0$, $x'_B = -\gamma L$ et $t'_B = \gamma Lu/c^2$. Comme on doit aussi avoir $x'_B = -L'$, on en déduit encore que $L = L'/\gamma$. Les deux procédures conduisent donc, heureusement, à la même longueur. Rappelons enfin que la longueur de la règle n'apparaît pas modifiée quand elle est perpendiculaire à Ox.

Une règle en mouvement dans la direction de sa longueur apparaît donc plus courte que dans un référentiel où elle est au repos. Cette contraction des longueurs est, dans la première méthode, une conséquence directe de la dilatation des temps. Dans la deuxième approche, elle est une conséquence de la non invariance de la simultanéité. Si on regarde les extrémités à un même instant dans $\mathcal{R}$, on les voit à deux instants différents dans $\mathcal{R}'$. Qui dit différence sur les temps, dit légère différence sur les positions.

Une application immédiate de la contraction des longueurs est la définition de la célérité. Reprenons le jumeau voyageur du paradoxe de Langevin. Dans le référentiel fixe, il met 4 ans à atteindre α du Centaure, à une vitesse proche de celle de la lumière. Dans son référentiel propre, il ne met que $4/\gamma$ ans. Cependant, la vitesse de l'étoile par rapport à lui est égale (en module) à sa vitesse par rapport à la Terre et donc proche de c. En revanche, la distance de l'étoile n'est plus que de $4/\gamma$ années-lumière.

7.4 Un paradoxe résolu

Pour illustrer encore cette notion de contraction des longueurs, considérons un nouveau problème. Un train de longueur L (référentiel propre $\mathcal{R}'$) entre dans un tunnel (immobile) de longueur L exactement égale. $\mathcal{R}$ est le référentiel du tunnel. Train et tunnel sont alignés avec Ox et $O'x'$. A $t = t' = 0$, l'arrière du train passe juste dans l'entrée du tunnel, située en O (ou O'). Que voient le contrôleur et le chef de gare en mission d'inspection dans le tunnel ? Pour le chef de gare, la longueur du train en mouvement est inférieure à celle du tunnel et la locomotive sort du tunnel un peu après que le dernier wagon n'y ait pénétré. Pour le contrôleur, en revanche, c'est le tunnel qui est un peu plus court que le train et la locomotive sort du tunnel avant que le dernier wagon n'y entre. La solution de ce paradoxe apparent réside bien sûr dans la non universalité de la simultanéité. Avant et après ne sont pas des absolus pour des points situés à des endroits différents (et qui ne sont pas dans le cône de lumière l'un de l'autre).

Pour nous en convaincre, nous écrirons, dans chaque référentiel, les coordonnées des deux événements importants. L'un représente l'entrée du dernier wagon dans le tunnel et coïncide avec l'événement origine dans les deux référentiels. L'autre événement est la sortie de la locomotive, dont les coordonnées dans $\mathcal{R}$ sont $x_2 = L$ et $t_2 = L/u(1-1/\gamma)$. On obtient t_2 simplement en disant que la longueur du train dans $\mathcal{R}$ est L/γ. A l'instant 0, la locomotive est donc à une distance $L(1 - 1/\gamma)$ de la sortie et elle parcourt cette distance à la vitesse u. La sortie de la locomotive dans $\mathcal{R}$ se produit donc bien après l'entrée du dernier wagon. On peut vérifier, par un calcul élémentaire, que le carré de l'intervalle entre les événements 1 et 2 est égal à $c^2(1 - \gamma)/u^2\gamma^2 < 0$;

l'intervalle entre les événements est toujours du genre espace. La notion de passé et de futur pour ces deux événements n'est donc pas nécessairement invariante. On peut alors écrire les coordonnées correspondantes dans $\mathcal{R}'$. Après un calcul sans difficulté, on trouve $x_2' = L$, ce qui ne fait guère que vérifier la cohérence du calcul, et $t_2' = -L/u(1-1/\gamma)$. L'événement 2 (sortie de la locomotive) s'est donc produit, dans $\mathcal{R}'$, avant que l'arrière du train n'entre dans le tunnel, comme nous nous y attendions. Si il y a dans ce problème un paradoxe, il ne concerne pas la validité et la cohérence de la transformation de Lorentz.

7.5 Effet Doppler-Fizeau

On appelle effet Doppler-Fizeau la variation de la fréquence d'un signal lumineux lorsque la source et le récepteur sont en mouvement relatif. Cet effet a été étudié par le physicien autrichien C. Doppler sur les ondes acoustiques, puis étendu aux ondes lumineuses par Fizeau. Considérons deux référentiels $\mathcal{R}$ et $\mathcal{R}'$, l'un lié à la source O et l'autre au récepteur O'[2]. La source émet une onde lumineuse monochromatique plane, de fréquence ν, se dirigeant dans la direction qui fait, dans le plan Oxy, l'angle θ avec Ox. L'onde lumineuse est définie par sa pulsation $\omega = 2\pi\nu$ et son nombre d'onde $\vec{k}$. La phase de l'onde, égale à $\omega t - \vec{k}\cdot\vec{r}$, est un invariant par changement de référentiel. Pour s'en convaincre, il suffit de se rappeler que les phénomènes d'interférence de deux ondes planes, qui sont déterminés par la différence de phase, sont invariants par ce changement. Ainsi, les formules de l'effet Doppler-Fizeau découle de la transformation de Lorentz entre le référentiel $\mathcal{R}$ de la source et celui $\mathcal{R}'$ du récepteur :

$$\left\{ \begin{array}{lcl} k_x & = & \gamma(k_x' + \beta\omega'/c) \\ k_y & = & k_y' \\ k_z & = & k_z' \\ \omega/c & = & \gamma(\omega'/c + \beta k_x') \end{array} \right. \tag{8}$$

Puisque $k_x' = k'\cos\theta' = (2\pi\nu'/c)\cos\theta'$, cette dernière équation donne $\nu = \gamma(\nu' + \beta\nu'\cos\theta')$ soit la formule usuelle de l'effet Doppler-Fizeau :

$$\nu' = \frac{\nu}{\gamma(1+\beta\cos\theta')} \tag{9}$$

Quand l'angle que fait la direction du mouvement par rapport à l'axe du mouvement est nul ($\theta' = 0$) on parle d'effet Doppler-Fizeau longitudinal. Dans ce cas, la formule précédente se réduit à

$$\nu' = \nu \left(\frac{1-\beta}{1+\beta}\right)^{1/2}$$

Quand la vitesse relative est assez petite, on peut faire un développement à l'ordre le plus bas en β et montrer que $\Delta\lambda/\lambda \approx \beta$ où λ est la longueur d'onde associée à l'onde lumineuse et $\Delta\lambda = \lambda' - \lambda$. Si le récepteur s'éloigne de la source ($\beta > 0$), la fréquence reçue est plus faible que la fréquence émise alors que si la source se rapproche ($\beta < 0$), cette fréquence est plus grande. Cet effet est utilisé en astrophysique pour déterminer la vitesse radiale des étoiles et des galaxies.

Quand la direction de réception est perpendiculaire à l'axe du mouvement ($\theta' = \pi/2$) on observe l'effet Doppler-Fizeau transversal et $\nu' = \nu\sqrt{1-\beta^2}$. Cet effet est plus difficile à observer car il est en général masqué par l'effet longitudinal.

7.6 Loi de composition des vitesses

Dans ce paragraphe, nous allons généraliser la loi de composition des transformations de Lorentz qui nous a déjà permis, au paragraphe précédent de traiter le cas de deux vitesses colinéaires. Nous considérerons ici une situation physique légèrement différente. Un mobile est en mouvement avec une vitesse v dans le référentiel $\mathcal{R}'$, entraîné à une vitesse u (selon Ox) par rapport au référentiel

2. Dans le cas des ondes acoustiques, s'introduite naturellement un troisième référentiel, celui lié au milieu matériel, car, contrairement aux ondes électromagnétiques, ce milieu est nécessaire à la propagation des ondes mécaniques.

$\mathcal{R}$. En écrivant les accroissements infinitésimaux $\mathrm{d}x'$, $\mathrm{d}y'$, $\mathrm{d}z'$ et $\mathrm{d}t'$ dans $\mathcal{R}'$ et en utilisant la transformation de Lorentz, on en déduit les accroissements correspondants dans $\mathcal{R}$:

$$\left\{\begin{array}{lcl} \mathrm{d}t & = & \gamma(\mathrm{d}t' + u\mathrm{d}x'/c^2) \\ \mathrm{d}x & = & \gamma(\mathrm{d}x' + u\mathrm{d}t') \\ \mathrm{d}y & = & \mathrm{d}y' \\ \mathrm{d}z & = & \mathrm{d}z' \end{array}\right.$$

On peut alors calculer sans difficultés les vitesses dans $\mathcal{R}$. On obtient par exemple :

$$v_x = \frac{v'_x + u}{1 + uv'_x/c^2} \qquad v_y = \frac{v'_y}{\gamma(1 + uv'_x/c^2)}$$

la transformation inverse s'obtenant trivialement en changeant le signe de u dans les expressions précédentes. On vérifie que la vitesse de la lumière est bien invariante dans cette transformation.

7.7 Aberration des étoiles

Nous étudierons ici une application immédiate de la loi de composition des vitesses. La lumière provenant d'une étoile apparaît déviée par la composition de sa vitesse avec celle du référentiel terrestre par rapport au système solaire. La position apparente d'une étoile dans le ciel dépend donc dans une petite mesure de la position de la Terre sur son orbite (chaque étoile semble décrire une petite ellipse annuelle autour de sa position moyenne). Pour simplifier la géométrie, nous considérerons le référentiel $\mathcal{R}$ comme celui lié au système solaire. La lumière de l'étoile arrive parallèlement à l'axe Oy. La vitesse de cette lumière est donc $v_y = -c$. Le référentiel $\mathcal{R}'$ est celui de la Terre, entraîné à la vitesse u selon Ox. De la loi de composition des vitesses, on déduit les composantes de la vitesse de la lumière de cette étoile dans le référentiel $\mathcal{R}'$: $v'_x = -u$ et $v'_y = -c/\gamma$.

Dans $\mathcal{R}'$ la lumière de l'étoile semble donc provenir d'une direction inclinée par rapport à l'axe $O'y'$, d'un angle $\theta = \arctan\gamma\beta$. L'application de la cinématique galiléenne prévoit aussi une déviation apparente mais elle n'est, comme on le vérifiera aisément, que $\theta = \arctan\beta$. On doit à l'astronome James Bradley la découverte, en 1728, du phénomène d'aberration stellaire grâce à ses observations de l'étoile γ du Dragon. L'angle mesuré par Bradley, 20,5 secondes d'arc, lui permit de déduire que la lumière se déplaçait 10 000 fois plus vite que la Terre sur son orbite.

Le phénomène d'aberration ne serait pas sans poser de problème à un voyageur interstellaire. L'aspect du ciel perçu à bord du vaisseau serait sensiblement différent de celui vu au repos (voir figure 7).

8 Conclusion

Nous venons de décrire les principaux aspects de la cinématique relativiste. Ses résultats sont souvent peu intuitifs, voire paradoxaux. Cette situation résulte la différence entre les prédictions de la relativité (sur le temps notamment) et notre sens commun. Ce n'est qu'avec une fréquentation assidue de ces problèmes que peut se développer une intuition relativiste. La difficulté est similaire, bien que moins ardue, à celle que l'on rencontre pour se forger une intuition en mécanique quantique.

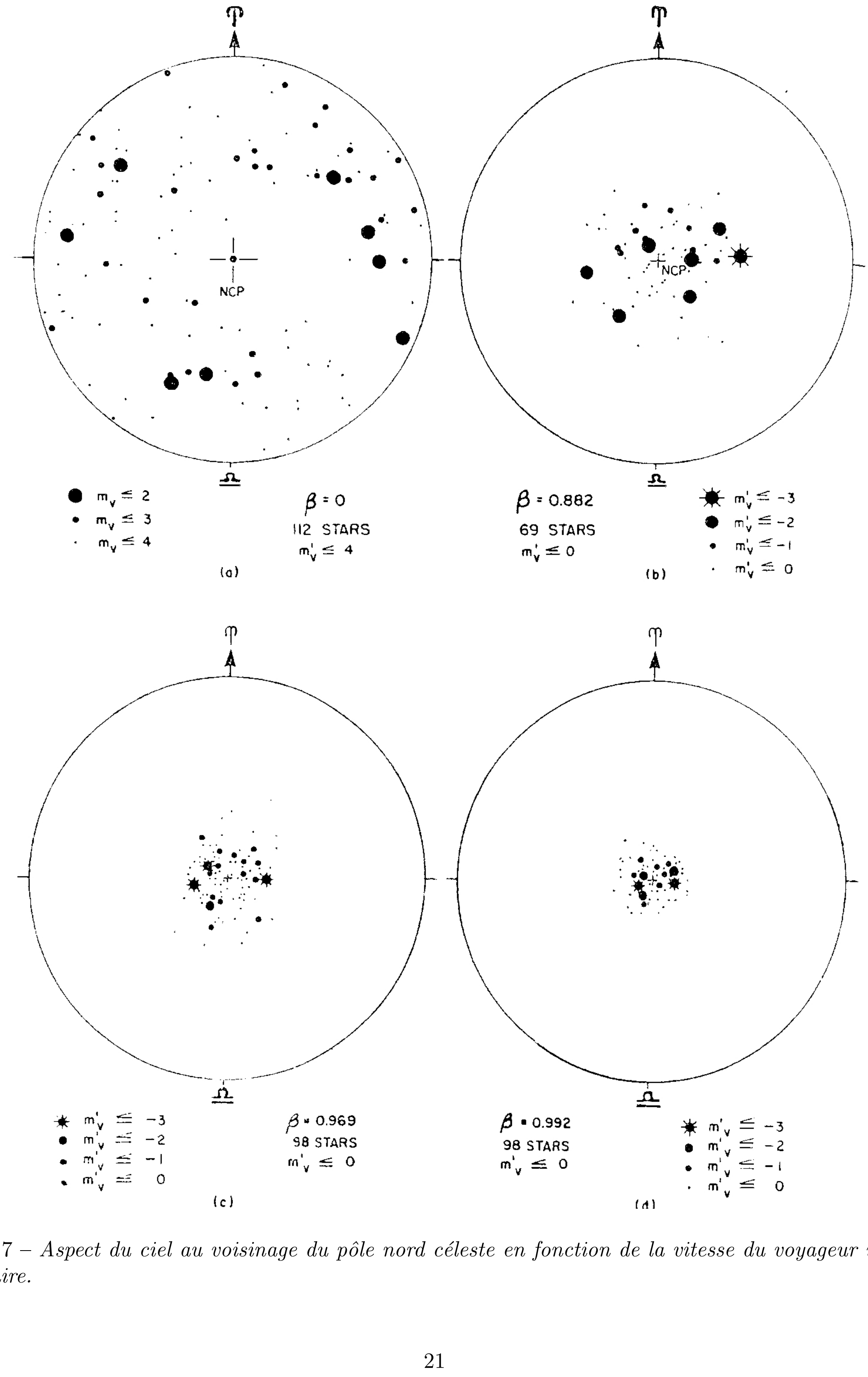

FIG. 7 – *Aspect du ciel au voisinage du pôle nord céleste en fonction de la vitesse du voyageur interstellaire.*

Références

[1] Stéphane Durand, *La relativité animée*, Editions Belin–Pour la Science.

[2] J.-Ph. Perez, *Relativité, fondements et applications*, Dunod, 1999

[3] N. Hulin, M. Hulin, L. Mousselin, *Relativité restreinte*, Dunod, 1998.

[4] G. Gamow, *M. Tompkins*, Dunod, 1992.

[5] Albert Einstein, *Sur l'électrodynamique des corps en mouvement*, 1905 ; publié dans l'ouvrage Editions de œuvres essentielles d'Einstein, Relativité I, editions Seuil–CNRS.

[6] J.-M. Lévy-Leblond et J.-P. Provost, *Additivity, rapidity, relativity*, American Journal of Physics 47, 1045 (1980).

[7] J.-P. Uzan, J.-P. Luminet, R. Lehoucq and P. Peter, *Twin Paradox and Space Topology*, European Journal of Physics 23, 277-284 (2002).

[8] Site ouèbe présentant des images simulées d'un voyage relativiste :
`http://www.fourmilab.ch/cship/cship.html`