

Partie II

Relativité restreinte

Introduction

Nous décrirons dans ce chapitre la théorie de la relativité restreinte, essentiellement telle qu'elle fut formulée par Einstein. Initialement, l'objet de la relativité était de réconcilier l'électromagnétisme et la cinématique : comme nous avons vu dans l'introduction à ce cours, la relativité Galiléenne n'est pas directement compatible avec l'électromagnétisme tel qu'il fut formulé par Maxwell. C'est essentiellement sur l'application à l'électromagnétisme de la relativité restreinte que portera ce chapitre. En revanche, il faut bien voir que la relativité restreinte s'applique dans un beaucoup plus large domaine. Elle joue en particulier un rôle central pour la physique des particules et la physique des accélérateurs. Elle est aussi essentielle en astronomie, beaucoup de sources de rayonnement cosmiques impliquant des déplacements à des vitesses proches de celle de la lumière.

Ce chapitre comportera deux étapes essentielles. Après un bref rappel de la relativité galiléenne, nous constaterons les difficultés que pose l'immersion de l'électromagnétisme dans cette relativité et donc dans la cinématique classique. Nous montrerons en particulier l'incompatibilité grave de l'électromagnétisme avec la loi ordinaire de composition des vitesses. Nous postulons donc un nouveau principe de relativité, imposant à toutes les lois de la physique, y compris l'électromagnétisme, d'être invariantes dans un changement de référentiel galiléen. La vitesse de la lumière devenant indépendante du référentiel, la loi de composition des vitesses et l'ensemble de la cinématique sont condamnées. Il nous faudra donc d'abord détruire la cinématique et la dynamique³ newtoniennes telles que nous les connaissons maintenant.

Il nous faudra formuler une nouvelle transformation des coordonnées et du temps décrivant les changements de référentiels, la transformation de Lorentz. Nous verrons en effet, par quelques arguments très simples, qu'un des postulats de base de la mécanique classique, l'universalité du temps et de la simultanéité, doit être abandonné. On mesure peut être assez mal aujourd'hui à quel point la démarche d'Einstein fut audacieuse, remettant en cause les postulats les plus intuitifs de la mécanique. La phase conceptuellement la plus difficile de notre travail, qui fera l'objet du premier chapitre, sera alors terminée.

Le deuxième chapitre, beaucoup plus mathématique que physique, sera consacré à l'introduction de notations tensorielles, bien adaptées à l'espace-temps à quatre dimensions de la relativité. Nous introduirons en particulier des conventions de notations très puissantes, dues à Einstein, qui permettent d'écrire de manière compacte et fiable les expressions parfois complexes auxquelles conduisent les calculs relativistes. Ces notations s'avèrent indispensables pour aborder la relativité générale, théorie géométrique de la gravitation.

Nous formulerons, au chapitre suivant, les lois de la nouvelle dynamique. Nous écrirons en particulier, dans une approche lagrangienne, le lagrangien d'une particule libre et nous en déduirons l'expression de la quantité de mouvement relativiste. Nous démontrerons en passant la formule la plus célèbre de l'histoire de la physique (nous laissons au lecteur le soin de deviner laquelle). Nous donnerons également la forme relativiste du principe fondamental de la dynamique que nous ne pourrions guère exploiter sans une forme explicite des forces, au moins de la force de Lorentz. Nous n'explorerons donc pas très en détails cette partie de la relativité qui se conclura par une brève description de la

³Rappelons que la cinématique décrit les mouvements indépendamment de leurs causes et que la dynamique permet de prévoir le mouvement si on en connaît les causes.

théorie relativiste des collisions, d'une grande importance en physique des particules.

Pour un cours centré sur l'électromagnétisme, nous consacrerons en effet l'essentiel de nos efforts au dernier chapitre de cette partie. Nous chercherons à y construire une théorie non triviale d'interaction entre particules transmise par un champ. Nous postulerons des formes simples pour le lagrangien d'interaction et pour le lagrangien décrivant ce champ et nous écrirons les équations de Lagrange correspondantes. Nous constaterons sans déplaisir que la structure de cette théorie de champ est celle de l'électromagnétisme. Nous aurons donc montré à quel point l'électromagnétisme de Maxwell s'adapte naturellement au cadre relativiste. Nous en profiterons pour examiner quelques problèmes simples d'électromagnétisme, du mouvement de particules relativistes dans des champs imposés aux bilans d'énergie–impulsion pour le champ lui même. Nous montrerons ainsi que cette approche complètement relativiste, outre son élégance, permet de dériver des lois importantes qui ne sont accessibles qu'au prix de calculs lourds en électromagnétisme “classique”.

Chapitre 1

Cinématique relativiste

La première étape est donc de comprendre les incompatibilités entre électromagnétisme et cinématique classique, et de refonder une cinématique tout à fait nouvelle. Nous allons commencer par quelques très brefs rappels de cinématique galiléenne ou newtonienne.

1.1 Rappels de relativité galiléenne

1.1.1 Transformation de Galilée

Il est très intuitif que le mouvement d'un point dépende de l'observateur. Pour utiliser un vocabulaire ferroviaire¹, le passager de train a une vitesse faible ou nulle par rapport à celle du contrôleur, alors qu'il a une vitesse élevée par rapport au garde barrière.

La notion centrale de la cinématique (classique ou relativiste) est celle de *référentiel*. Un référentiel, c'est un ensemble d'observateurs, immobiles les uns par rapport aux autres. Ces observateurs peuvent constater le passage du mobile à leur position. La connaissance de la position des observateurs concernés permet alors de déterminer la trajectoire du mobile. On peut bien sûr convenir d'un repère (cartésien, orthonormal) pour repérer ces positions au moyen de trois coordonnées. Les observateurs sont de plus munis d'horloges qui leur permettent de noter l'instant auquel le mobile passe en face d'eux, le mouvement étant 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 supposons que toutes les horloges de tous les observateurs d'un même référentiel sont synchronisées (indiquent la même valeur au même instant). 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 alors défini par les trois fonctions $x(t), y(t), z(t)$ 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}' : $x'(t'), y'(t'), z'(t')$. En mécanique classique, on admet sans restrictions l'identité des temps (à une synchronisation près) des observateurs de \mathcal{R} et de \mathcal{R}' ². Il

¹Les papiers originaux sur la relativité emploient souvent des expériences de pensée utilisant des trains et des gares, parfois même des tunnels. C'est sans doute lié au succès grandissant des transports ferroviaires au début du siècle et à leur importance sociologique. Pour céder à la tradition, nous emploierons ce genre de vocabulaire dans ce cours, bien que les effets relativistes soient complètement négligeables, même avec les trains les plus modernes.

²Cette hypothèse était déjà faite explicitement par Newton dans ses *Principia*. S'il en avait tout à fait reconnu l'importance, il n'aurait guère de doutes sur sa validité.

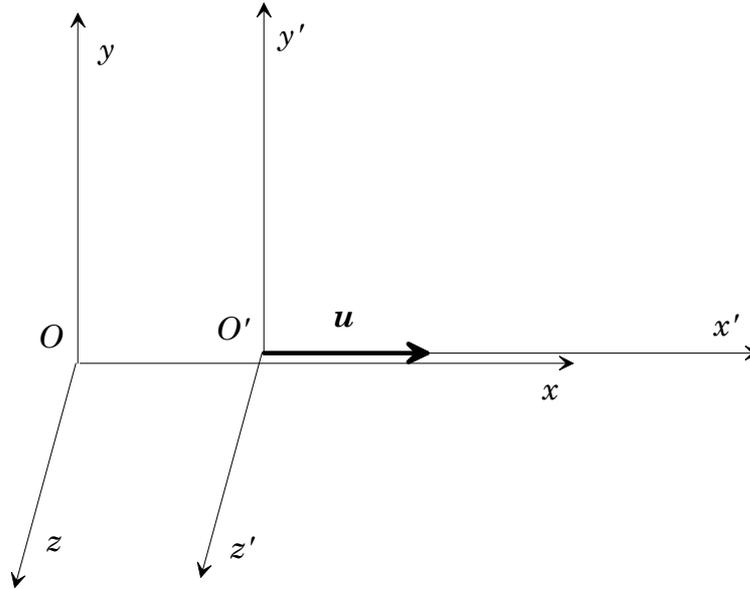


Figure 1.1: Choix des axes dans deux référentiels \mathcal{R} et \mathcal{R}' en mouvement relatif. Les axes des deux repères sont parallèles. Les axes Ox et $O'x'$, alignés avec la vitesse relative \mathbf{u} , coïncident à chaque instant.

est possible alors de donner la transformation qui fait se correspondre les mouvement 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 du tout la généralité, on peut choisir les axes dans \mathcal{R} et \mathcal{R}' de telle manière que:

- Les axes Ox et $O'x'$ coïncident à tout instant et sont parallèles à la vitesse \mathbf{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 $t = 0$.

La figure 1.1 présente la géométrie choisie. Nous l'exposons en détail parce que nous choisirons la même pour décrire les changements de référentiel en relativité restreinte.

La loi de transformation de Galilée s'écrit alors trivialement:

$$\left. \begin{aligned} x'(t) &= x(t) - ut \\ y'(t) &= y(t) \\ z'(t) &= z(t) \end{aligned} \right\} \quad (1.1)$$

C'est cette transformation, tellement triviale qu'elle est bien rarement écrite explicitement, qui sera remplacée par la transformation de Lorentz en relativité einsteinienne. Cette transformation de Galilée contient, par simple dérivation par rapport au temps, la loi de composition des vitesses:

$$\mathbf{v} = \mathbf{v}' + \mathbf{u} \quad (1.2)$$

(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, les référentiels galiléens, en mouvement de translation uniforme les uns par rapport aux autres, tels que le mouvement d'une particule libre y soit rectiligne et uniforme.

1.1.2 Les difficultés de la cinématique classique

La loi de composition des vitesses, telle que nous venons de la rappeler, est difficilement compatible avec l'électromagnétisme de Maxwell. La conséquence la plus importante et la plus nouvelle des équations de Maxwell est en effet la prédiction de l'existence d'ondes se propageant à la vitesse c . Le problème qui apparaît immédiatement est celui du référentiel dans lequel cette vitesse est définie, le seul donc dans lequel les équations de Maxwell seraient directement applicables.

Le sentiment le plus naturel, qui prédominait très largement à la fin du siècle dernier, était que les ondes électromagnétiques se propageaient dans un milieu baignant l'univers entier: *l'éther*. 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 infiniment rigide pour propager des ébranlement 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 "fluide" si particulier se trouvait ainsi doté de propriétés presque aussi extraordinaires que le calorique du siècle précédent ou, encore avant lui, le phlogistique.

Il existe aussi une difficulté philosophique grave avec l'introduction de l'éther. Les physiciens avaient mis plus de 20 siècles, entre Aristote et Copernic, pour comprendre que notre petite planète n'est pas le centre de l'univers. Le principe de relativité selon Galilée avait le mérite d'indiquer qu'aucun référentiel galiléen n'est particulièrement privilégié. L'introduction de l'éther devait briser cette "démocratie" des référentiels en introduisant un référentiel très particulier, celui de l'éther, le seul dans lequel les équations de Maxwell devaient s'appliquer. Ce genre d'argument a certainement joué un rôle essentiel pour Einstein.

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 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'éther³. Or cette expérience fut tout à fait négative (ou plutôt très positive): la vitesse de la lumière semblait indépendante du mouvement de la terre par rapport au soleil.

On pouvait, devant ce résultat négatif, adopter deux points de vue. 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 cycloïdes, 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

³Nous ne détaillerons pas ici le principe de cette expérience: cette description n'est pas indispensable pour la suite de l'exposé. Le lecteur intéressé pourra trouver une description détaillée dans pratiquement tous les manuels de relativité.

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) et donc que toute la physique était 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: "Sur l'électrodynamique des corps en mouvement"⁴. Le principe fondamental de cette nouvelle physique, le "principe de relativité" est exposé dans le prochain paragraphe.

1.2 Principe de relativité

1.2.1 Enoncé

Il existe une classe de référentiels privilégiés, en translation uniforme les uns par rapport aux autres (que nous continuerons à appeler "référentiels galiléens"), dans lesquels toutes les lois de la physique prennent la même forme.

Si toutes les lois de la physique prennent la même forme, les équations de Maxwell sont valides dans tous les référentiels et la vitesse de la lumière, c , est la même dans tous les référentiels. Nous centrerons cet exposé sur l'invariance de la vitesse de la lumière. Il serait tout aussi possible de ne pas faire jouer un rôle aussi central à l'électromagnétisme. On pourrait simplement postuler qu'il existe une vitesse limite de propagation de toutes les interactions et effectuer tous les raisonnements qui vont suivre sur cette vitesse. Il suffirait, enfin, de constater qu'expérimentalement la vitesse de la lumière dans le vide est identique à la vitesse limite, à la précision des mesures. Il n'est pas absolument exclus, en effet, bien que cela soit très peu vraisemblable, que le photon possède une très petite masse, rendant la vitesse de la lumière très légèrement inférieure à la vitesse limite qui apparaît en relativité.

Ce principe de relativité, de prime abord, semble ne rien remettre en cause d'essentiel et semble très 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 en effet que le principe de relativité a deux conséquences immédiates:

- 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).
- Deux événements qui se produisent simultanément dans un référentiel peuvent se produire à des instants différents dans un autre référentiel.

Remettre en cause des propriétés aussi intuitives de l'espace et du temps ne sera pas sans conséquences. Il est clair, en particulier, que la transformation appelée à remplacer la transformation de Galilée devra renoncer au caractère absolu du temps et mélanger les coordonnées spatiales et temporelles.

1.2.2 Deux expériences de pensée

Nous considérons donc deux référentiels en mouvement relatif, avec la géométrie décrite dans la figure 1.1. Le référentiel \mathcal{R}' sera celui du contrôleur, ou du train, pour reprendre nos analogies ferroviaires, le référentiel \mathcal{R} celui du chef de gare. Le contrôleur, situé en O' , envoie à $t' = 0$ (nous ne confondrons 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 figure 1.2). L'impulsion, réfléchi par le miroir, revient vers le contrôleur et l'atteint au bout d'un temps $T' = 2L/c$ (nous supposons, pour ce paragraphe seulement, que, pour la cinématique classique, la vitesse de l'impulsion est c dans \mathcal{R}' - l'ensemble

⁴Nous ne saurions trop recommander la lecture de cet article, ainsi que celle d'un article de revue rédigé dès 1907, qui constitue un exposé très pédagogique de la relativité (Edition de œuvres essentielles d'Einstein, Relativités I—éditions Seuil-CNRS).

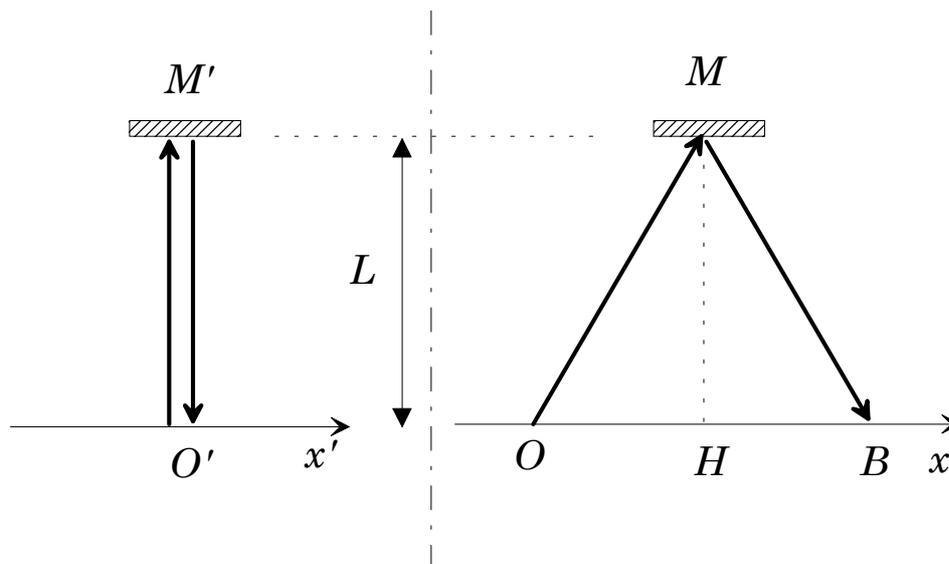


Figure 1.2: Une expérience de pensée établissant le caractère relatif du temps. Un signal lumineux est émis depuis l'observateur O' , le long de l'axe $O'z'$, vers un miroir M . Après réflexion sur ce miroir, le signal revient à l'observateur O' . La même expérience est vue, à gauche, dans le référentiel du train et, à droite, dans un référentiel immobile. Pendant l'expérience, l'observateur O' est passé de O à B .

de l'argument pouvant être transposé sans difficultés au cas, plus naturel, où on supposerait que la vitesse de l'impulsion est c dans \mathcal{R}). Notons que le contrôleur pourrait ainsi construire une horloge. Renvoyant une deuxième impulsion à l'instant précis où il reçoit la première, il établirait un phénomène périodique et donc une horloge.

Regardons maintenant cette même expérience avec l'oeil du chef de gare (partie droite de la figure 1.2). A $t' = 0$, le contrôleur est en O' et donc aussi en O . De son côté, le miroir s'est déplacé avant que l'impulsion ne l'atteigne. Il occupe donc une position M , à une certaine distance de O sur l'axe Ox . Enfin, le train continue à se déplacer pendant le retour de l'impulsion et le contrôleur occupe la position B au moment du retour. La trajectoire de l'impulsion dans \mathcal{R} est triangulaire.

Imaginons d'abord que le chef de gare ait été nommé avant 1905 et soit donc un adepte de la cinématique classique. Pour lui, la vitesse de la lumière obéit à la loi de composition. La vitesse de l'impulsion a donc une composante $\pm c$ sur Oy et une composante u sur Ox (u étant la vitesse du train). Son module est donc $\sqrt{c^2 + u^2}$. La durée du parcours OM étant L/c (on admettra dans toute la suite que la position du miroir selon y n'est pas affectée par le changement de référentiel; nous en donnerons plus tard une justification détaillée), on a $OH = uL/c$ et la longueur OM vaut $L\sqrt{c^2 + u^2}/c$. Le temps du parcours OM est donc L/c et la durée totale de l'expérience $T = 2L/c$ est identique à celle vue par le contrôleur (la distance parcourue dans \mathcal{R} est plus grande, mais le module de la vitesse est augmenté dans la même proportion). On retrouve bien, naturellement, le postulat d'universalité du temps.

Considérons maintenant un chef de gare ayant admis la validité du postulat de relativité. Pour lui, la vitesse de l'impulsion est c . Le temps de parcours OM est donc OM/c . Il en déduit $OH = uOM/c$. Comme $OM^2 = L^2 + OH^2$, on a $OM = L/\sqrt{1 - u^2/c^2}$. Il obtient donc finalement:

$$T = \gamma T' , \quad (1.3)$$

avec

$$\gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} . \quad (1.4)$$

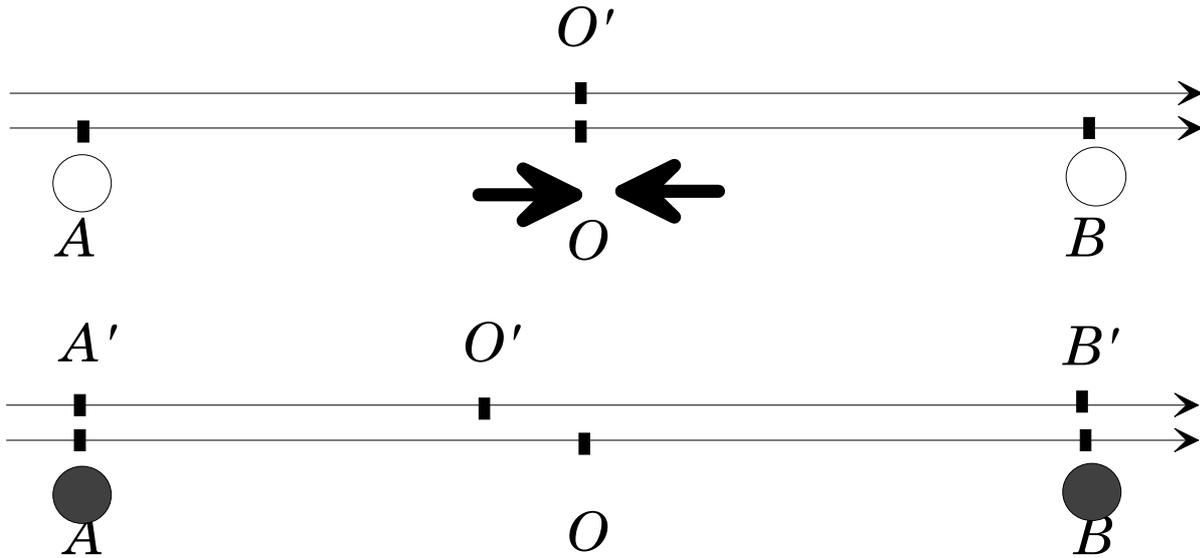


Figure 1.3: Deuxième expérience de pensée illustrant le postulat de relativité. En haut, vue de la situation au moment où le chef de gare, situé en O , et le contrôleur, en O' voient arriver simultanément les signaux lumineux émis par A et B . En bas, situation au moment où les signaux se sont allumés. O' n'est pas encore arrivé en O . Les signaux s'allument en face des observateurs A' et B' .

Le facteur γ (que nous aurons de nombreuses occasions de retrouver) est toujours supérieur à un. La durée de l'expérience mesurée par le chef de gare est plus longue que celle mesurée par le contrôleur. Si chacun construisait une horloge avec le même dispositif, celle du chef de gare battrait plus vite et avancerait par rapport à celle du contrôleur (un cauchemar pour le respect des horaires; heureusement, l'effet est petit comme on pourra le vérifier). Le postulat de relativité a pour conséquence immédiate que le temps n'est pas une notion universelle.

Cette expérience de pensée nous fournit une autre indication sur ce que sera la cinématique relativiste. Le facteur γ n'existe que si la vitesse relative des deux référentiels est plus petite que c . Si ce n'était pas le cas, l'impulsion lumineuse qui se réfléchit normalement sur le miroir dans \mathcal{R}' n'arriverait jamais à rattraper ce miroir dans \mathcal{R} , puisque sa vitesse n'est "que" c . Un événement (la réflexion) se produirait dans un référentiel et pas dans un autre, ce qui est bien sûr absurde. Deux référentiels galiléens ne peuvent donc être animés l'un par rapport à l'autre d'une vitesse supérieure (ou même égale) à c . Comme on peut attacher un référentiel galiléen à tout objet en translation uniforme, c apparaît aussi comme une vitesse limite pour tous les objets matériels.

La seconde expérience de pensée que nous allons étudier nous emmènera encore plus loin, puisqu'elle 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 . A l'instant $t = 0$, il voit ces deux signaux s'allumer simultanément. S'il sait, ou s'il mesure, que la même distance L 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 l'instant auquel ils se sont allumés (dans son échelle de temps, bien sûr)? 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.

Supposons d'abord que la cinématique classique s'applique. Le temps étant universel, les deux signaux s'allument dans \mathcal{R}' au même instant $t' = -L/c$. A cet instant O' est à une abscisse $-uL/c$ par rapport à O . Les passagers A' et B' sont alors situés respectivement en $x' = -L + uL/c$ et $x' = L + uL/c$ (voir la partie inférieure de la figure 1.3). La distance $A'O' = L(c - u)/c$ est donc inférieure à la distance $B'O' = L(c + u)/c$. En revanche, la vitesse de l'impulsion venant de A' est $c - u$ et la vitesse de l'impulsion venant de B' est $c + u$. En appliquant la procédure décrite ci-dessus, le contrôleur établira donc que les deux signaux se sont allumés au même instant. L'un était plus proche, mais la vitesse de l'autre impulsion était plus grande.

Que se passe-t-il maintenant si nous appliquons le principe de relativité? La vitesse des deux impulsions lumineuses est la même. Nous ne pourrions pas, pour le moment, déterminer la position des observateurs A' et B' (il nous faudra la transformation de Lorentz pour cela). Nous pouvons comprendre, en revanche, 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 de l'autre. Le contrôleur en déduira que le signal A s'est allumé *après* le signal B . Deux événements peuvent être vus comme simultanés ou non par des observateurs appartenant à des référentiels différents. Nous verrons bientôt qu'heureusement cet abandon de l'universalité de la simultanéité ne compromet pas la causalité.

Nous avons vu émerger, en discutant ces deux expériences de pensée, deux des notions essentielles de la relativité: l'événement et l'intervalle.

1.3 Evénements et intervalles

1.3.1 Evénements

Comme nous venons de le voir, le temps n'est plus universel et n'est plus séparable des coordonnées spatiales. Il faudra décrire les expériences en termes d'événements (très littéralement: il s'est passé quelque chose quelque part). Un événement, c'est par exemple l'allumage du signal A ou la réflexion de la lumière sur le miroir dans notre première expérience de pensée. Un événement existe indépendamment du choix du référentiel. On peut caractériser un événement, 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 produit⁵. On pourra donc complètement caractériser l'événement par quatre nombres: les trois coordonnées spatiales de l'observateur (on se munit d'un repère convenable) et le temps. On décrira donc un événement par la donnée d'un référentiel et d'un quadruplet de nombres (ct, x, y, z) (nous développerons au chapitre suivant des notations tensorielles puissantes pour traiter ces quadruplets). 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. Il y a un parallèle très fort entre la différence entre événement (indépendant du référentiel) et coordonnées spatio-temporelles et celle qui existe entre un vecteur (indépendant du repère) et ses composantes sur une base donnée.

Nous utiliserons souvent des représentations géométriques des événements. On peut en effet les représenter comme un point dans un espace à quatre dimensions. Cette représentation posant quelques problèmes techniques, on se cantonne souvent à une dimension d'espace. On représente alors un événement comme sur la figure 1.4. Pour des raisons de commodité, on porte sur l'axe vertical le produit ct . Les deux coordonnées dans cet espace ont ainsi la même dimension⁶.

⁵On supposera encore que tous les observateurs d'un même référentiel peuvent synchroniser leurs horloges. Il leur est interdit de se déplacer, mais on peut procéder de façon plus subtile. On peut, par exemple, déterminer par des moyens géométriques le milieu du segment AB joignant deux observateurs. On peut placer en ce point une source lumineuse qui s'allume à un certain instant. Si les observateurs A et B font le zéro de leurs horloges au moment où ils voient cette source s'allumer, ils auront établi leur synchronisme.

⁶Les professionnels de la relativité prennent souvent $c = 1$, ce qui simplifie énormément les écritures. A notre niveau, il est peut être imprudent de se priver d'un moyen de vérifier l'homogénéité de nos formules

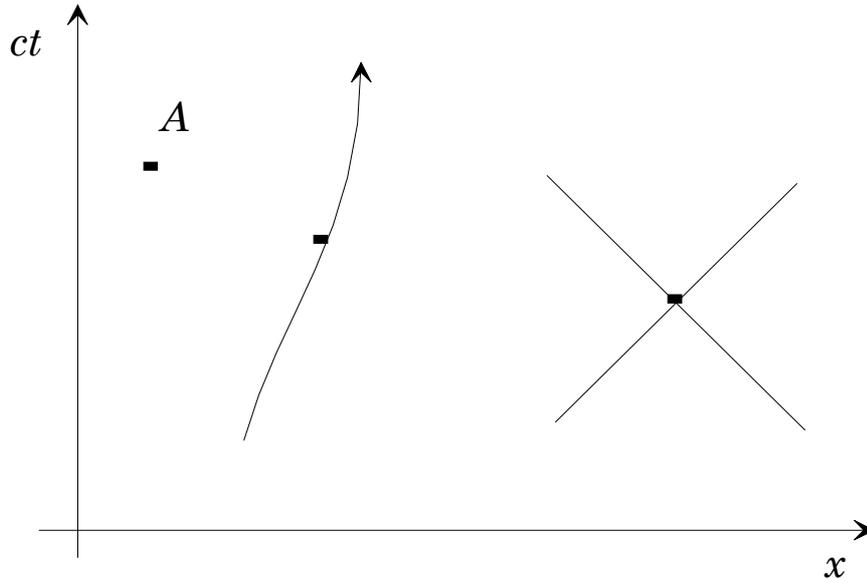


Figure 1.4: Un événement, une ligne d’univers et un cône de lumière. Un événement est représenté par un point dans un espace x, ct . Une ligne d’univers est l’ensemble des événements correspondant aux positions successives d’une particule. Le cône de lumière d’un événement est constitué des lignes d’univers d’un signal lumineux passant par cet événement.

On peut considérer le mouvement d’un point dans un référentiel comme 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 dans l’espace-temps une ligne, que nous nommerons “ligne d’univers” de la particule. Une telle ligne est représentée sur la figure 1.4.

La ligne d’univers d’une particule qui se déplacerait à la vitesse de la lumière serait parallèle, dans notre représentation graphique, à la première ou à la deuxième bissectrice. Dans l’espace à quatre dimensions, l’ensemble des lignes d’univers partant d’un point et correspondant à un mouvement à c forme le “cône de lumière” de cet événement (voir aussi la figure 1.4). Les événements antérieurs à l’événement de référence forment le passé du cône de lumière, les autres le futur. Comme c est une vitesse limite, toutes les lignes d’univers passant par un événement donné doivent être à l’intérieur du cône de lumière. 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 est évident géométriquement que cette relation 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 dans le cas général, comme on pourra s’en persuader aisément. Si C est dans le passé du cône de lumière de B , lui-même dans le futur du cône de lumière de A , alors C n’est pas nécessairement dans le cône de lumière de A . En un mot, si A et C peuvent tous deux être la cause de B , il n’y a aucun lien de causalité a priori 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. Si la physique classique admet qu’un événement puisse être la cause d’un autre s’il lui est antérieur (admettant ainsi implicitement les actions instantanées à distance), la relativité exige que l’un des événements soit effectivement antérieur à l’autre (nous verrons dans le prochain paragraphe que la notion d’antériorité est indépendante du référentiel) mais aussi que les deux événements puissent être reliés par un signal. Nous allons maintenant pouvoir affiner beaucoup ces notions en introduisant l’intervalle.

1.3.2 Intervalle. Invariance de l'intervalle

Considérons deux événements repérés, dans un référentiel donné, par (ct_1, x_1, y_1, z_1) et (ct_2, x_2, y_2, z_2) . 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 . On a donc dans ce cas :

$$c^2(t_1 - t_2)^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 . \quad (1.5)$$

L'écriture de cette relation suggère d'introduire l'intervalle entre deux événements quelconques par :

$$s_{1,2}^2 = c^2(t_1 - t_2)^2 - (x_1 - x_2)^2 - (y_1 - y_2)^2 - (z_1 - z_2)^2 . \quad (1.6)$$

Notons que le choix du signe $+$ pour la composante temporelle de l'intervalle est tout à fait arbitraire. C'est cependant le plus répandu aujourd'hui. L'intervalle jouera 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: le fait pour deux événements d'être ou non reliés par un signal lumineux ne dépend pas de la description du mouvement. Un intervalle nul est donc un invariant dans un changement de référentiel.

Nous établirons rigoureusement, à partir de la transformation de Lorentz, le fait que l'intervalle est indépendant du référentiel (est un "invariant relativiste"), même s'il n'est pas nul. Nous allons donner ici une indication de ce fait par un raisonnement qui, bien qu'il ne soit pas tout à fait rigoureux (il fait appel à des hypothèses supplémentaires implicites), établit de manière simple l'invariance de l'intervalle. Cette invariance nous permettra, dans les prochains paragraphes, de comprendre beaucoup de propriétés de la transformation de Lorentz avant même d'en écrire la forme explicite.

Pour cela, considérons deux événements infiniment voisins. L'intervalle, lui aussi infinitésimal, entre ces événements s'écrit alors, dans un référentiel \mathcal{R} :

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 . \quad (1.7)$$

Considérons les deux mêmes événements dans un autre référentiel \mathcal{R}' . L'intervalle entre eux s'écrit

$$ds'^2 = c^2 dt'^2 - dx'^2 - dy'^2 - dz'^2 . \quad (1.8)$$

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). On doit pouvoir développer cette fonction au premier ordre pour les intervalles infinitésimaux que nous manipulons et écrire:

$$ds'^2 = a ds^2 , \quad (1.9)$$

où a est une constante ne dépendant que de la vitesse relative \mathbf{u} des deux référentiels. En fait l'isotropie de l'espace impose que a ne dépende que du module u de la vitesse \mathbf{u} . Considérons maintenant un troisième référentiel \mathcal{R}'' , en mouvement à la vitesse \mathbf{v} par rapport à \mathcal{R} et \mathbf{w} par rapport à \mathcal{R}' . L'intervalle infinitésimal dans ce référentiel, ds''^2 , est tel que:

$$ds''^2 = a(v)ds^2 = a(w)ds'^2 = a(w)a(u)ds^2 . \quad (1.10)$$

La fonction a doit donc vérifier, pour tout triplet de vitesses relatives:

$$a(w) = \frac{a(v)}{a(u)} \quad (1.11)$$

ce qui est manifestement impossible (le module de la vitesse w dépend de l'orientation relative des deux autres et pas seulement de leur module), à moins que $a = 1$. On établit ainsi l'invariance des intervalles infinitésimaux. Tout intervalle pouvant être obtenu par une intégration d'intervalles infinitésimaux entre les deux événements, on établit ainsi l'invariance d'un intervalle arbitraire.

1.3.3 Conséquences de la conservation de l'intervalle. Causalité relativiste

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

- Si $s_{1,2}^2 > 0$, 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, aussi, les deux événements sont dans le cône de lumière l'un de l'autre.
- Si $s_{1,2}^2 < 0$, la distance spatiale entre les deux événements est plus grande que la distance temporelle. Aucun signal ne peut donc avoir relié les deux événements, ce qui exclut tout lien de causalité (souvenons nous qu'aucune interaction ne peut se propager plus rapidement que la vitesse limite c). Nous dirons alors que nous avons à faire à un intervalle de genre "espace".
- Si $s_{1,2}^2 = 0$, les deux événements peuvent avoir été 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 très fortement relié à la notion de causalité. Il est donc essentiel que la nouvelle cinématique prédise l'invariance de l'intervalle, de manière que les liens de causalité entre événements soient indépendants des observateurs. La causalité classique, qui n'exigeait que des relations d'antériorité entre la cause et la conséquence exige maintenant deux conditions. D'abord, 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 en relativité). D'autre part, il faut encore que la cause précède la conséquence. 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 ce cône de lumière. Nous avons défini le futur de O comme l'ensemble des événements du cône de lumière de O de coordonnée temporelle supérieure à celle de O et nous supposons M situé dans cette partie du cône de lumière. Si, dans un changement de référentiel, M passait dans le passé de O , cela impliquerait qu'il existe un changement de référentiel pour lequel M et O soient confondus. En effet, quel que soit le changement de référentiel, M reste dans le cône de lumière de O . Par continuité, passer du futur au passé de O 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 en relativité. 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 de Lorentz.

Notons enfin, pour finir, que tous les intervalles pris sur la ligne d'univers d'une particule matérielle sont du genre temps.

1.3.4 Temps propre

Nous pouvons appliquer l'invariance de l'intervalle au problème des horloges en mouvement que nous avons déjà abordé dans notre première expérience de pensée. Nous y 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. Là encore, toutes ces notions seront beaucoup plus précises quand nous disposerons de la forme explicite de la transformation de Lorentz.

Considérons donc une particule, ou une horloge, en mouvement arbitraire 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 \mathbf{v} coïncide avec celle de la particule. Nous appellerons ce référentiel le référentiel tangent au mouvement \mathcal{R}' . 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 dt dans \mathcal{R} . Pendant ce temps, la particule se déplace de $d\mathbf{l} = \mathbf{v}dt$. Les deux événements correspondant aux deux extrémités de ce mouvement infinitésimal sont donc séparés par un intervalle

$$ds^2 = c^2 dt^2 - dl^2 = c^2 dt^2 \left(1 - \frac{v^2}{c^2} \right). \quad (1.12)$$

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 dt' . L'intervalle s'écrit donc aussi:

$$ds^2 = c^2 dt'^2. \quad (1.13)$$

Nous appellerons donc “temps propre” l'intervalle de temps s'écoulant dans \mathcal{R}' et nous le noterons $d\tau = dt'$. En rapprochant les deux expressions précédentes de l'intervalle, nous pouvons écrire

$$dt = \gamma d\tau \quad (1.14)$$

avec

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (1.15)$$

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. Le facteur de “dilatation” du temps, γ , toujours supérieur à un, a l'expression que nous avons déjà trouvée au paragraphe précédent. Il est important de constater à ce point que $d\tau$ 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). Le temps propre est donc une propriété intrinsèque de la particule. Nous nommerons une telle quantité un 4-scalaire dans le chapitre suivant.

Le raisonnement est fait ici pour des intervalles de temps infinitésimaux. On peut le 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 inférieure à la durée mesurée dans \mathcal{R} .

Cette “dilatation des temps” a plusieurs conséquences pratiques 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 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 supposons, 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 temporelle γ est très grand devant un et la durée de vie “vue” 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.

La dilatation des temps se manifeste aussi à une échelle de vitesses plus accessible, à condition de disposer d’horloges de haute précision. Le réseau d’horloges atomiques qui fixent le temps international doit périodiquement être recalé. Pour cela, on transporte physiquement d’un site à l’autre des horloges portables de haute précision. A l’arrivée du voyage, il faut corriger l’horloge mobile de la dilatation relativiste des temps⁷ que l’on pourra estimer numériquement à titre d’exercice.

Citons également le célèbre “paradoxe des jumeaux”, dû à Langevin. De deux frères jumeaux, l’un reste sur terre et l’autre vole vers Proxima du centaure, à une distance de 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. Nous verrons, quand nous aurons explicité la forme de la transformation de Lorentz, que le jumeau voyageur voit l’étoile de destination s’approcher de lui à une vitesse proche de celle de la lumière. En revanche, elle est initialement beaucoup plus proche de lui que 4 années-lumière. Le jumeau voyageur revient donc sur terre plus jeune que son frère! Le paradoxe apparent est qu’il semble y avoir une parfaite symétrie entre les deux jumeaux, incompatible avec cette différence d’âge: dans le référentiel du voyageur, le jumeau terrestre 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 du mobile. 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.

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 vitesse, 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 à proprement parler.

Nous concluons ce paragraphe par une remarque importante pour la dynamique. Si une horloge est immobile dans \mathcal{R} , le temps propre mesuré entre deux événements coïncide avec celui du référentiel. En revanche, si elle est en mouvement, le temps propre entre les deux mêmes événements est toujours inférieur au temps du référentiel (et ce quelle que soit la forme ou la loi horaire de la trajectoire). On en déduit donc que l’intégrale du temps propre entre deux événements est maximale pour une horloge immobile, une propriété qui nous sera fort utile pour formuler un principe variationnel pour la dynamique relativiste.

⁷Il faut aussi corriger un effet de “red shift” gravitationnel qui n’est descriptible que dans le cadre de la relativité générale. La fréquence de l’horloge est affectée en effet par le champ de pesanteur terrestre, légèrement diminué en vol.

1.4 Transformation de Lorentz

Après cette approche très 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, décrivant un changement de référentiel. 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.

1.4.1 Forme de la transformation de Lorentz

Le choix d'axes pour les deux repères est, encore une fois, celui illustré par la figure 1.1. Nous cherchons donc une transformation $\mathcal{L}(u)$ permettant d'exprimer les coordonnées (ct', x', y', z') d'un événement dans \mathcal{R}' en fonction de celles dans \mathcal{R} , (ct, x, y, z) . 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}' . Nous obtiendrons avec ce choix d'axes la transformation de Lorentz 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 tout à fait quelconques.

Nous imposerons d'abord à \mathcal{L} d'être une transformation linéaire, homogène. L'invariance de la physique dans une translation arbitraire de l'espace ou du temps impose cette linéarité.

L'ensemble des transformations de Lorentz, paramétrées 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} (comme nous avons abandonné le cadre de la relativité galiléenne, w n'est pas égal à $u + v$). La transformation de \mathcal{R} vers \mathcal{R}' peut s'écrire $\mathcal{L}(w)$ ou $\mathcal{L}(v)\mathcal{L}(u)$ (ce produit étant à comprendre comme la composition de deux applications linéaires et donc étant lu de droite à gauche). Le produit de deux transformations de Lorentz définit donc une application de composition interne qui possède évidemment toutes les propriétés d'une loi de groupe. Il existe un élément neutre, l'identité, correspondant au passage d'un référentiel à lui-même et donc à la vitesse 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 aurions certainement brisé le principe de relativité. La transformation inverse de $\mathcal{L}(u)$ doit donc être la transformation de Lorentz correspondant à la vitesse $-u$, qui est celle de \mathcal{R} mesurée dans \mathcal{R}' : on doit avoir $\mathcal{L}(u)^{-1} = \mathcal{L}(-u)$.

Il n'est pas évident a priori que ce groupe, que nous appellerons "groupe de Lorentz", soit commutatif. En fait, 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 réels⁸. Il en résulte immédiatement que tous ces groupes sont abéliens (ou commutatifs). Notons que cet isomorphisme indique qu'on peut, par un changement de variable adéquat, paramétrer

⁸Nous ne démontrerons pas ici cette propriété. On en trouvera une démonstration très élémentaire dans J.M. Lévy-Leblond et al., *Am. Journal of Physics*, **47**, 1045 (1980). 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.

le groupe par un paramètre additif, déterminé de façon univoque à un facteur près (à un choix d'unités près). On pourrait ainsi trouver un paramètre $\phi(u)$ tel que la composition de deux transformations de Lorentz s'écrive $\mathcal{L}(\phi(u) + \phi(v)) = \mathcal{L}(\phi(v))\mathcal{L}(\phi(u))$. 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 :

$$y' = ay + bz \quad (1.16)$$

$$z' = b'y + a'z. \quad (1.17)$$

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 :

$$y' = ay \quad z' = az. \quad (1.18)$$

Montrons maintenant que ce facteur a vaut nécessairement 1. Nous avons montré effectivement que $\mathcal{L}(u)^{-1} = \mathcal{L}(-u)$. 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 $a = 1/a$ et $a = \pm 1$. Si le choix des orientations des axes dans les deux référentiels est cohérent, on a donc 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 de x et ct . La transformation la plus générale ferait intervenir les coordonnées y et z . L'invariance par translation perpendiculaire à l'axe des x 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 :

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} a(u) & b(u) \\ e(u) & f(u) \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}. \quad (1.19)$$

On peut préciser considérablement la forme de cette transformation par un simple argument de symétrie. Considérons en effet dans le référentiel \mathcal{R}' un axe $O'X'$ confondu avec, mais d'orientation contraire à, $O'x'$. En un mot, $X' = -x'$. Considérons de même l'axe OX opposé avec l'axe Ox , avec $X = -x$. On peut considérer la transformation donnant X et ct en fonction de X' et ct' . Elle correspond au même changement de référentiel que celui 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 :

$$\begin{pmatrix} ct \\ X \end{pmatrix} = \begin{pmatrix} a & b \\ e & f \end{pmatrix} \begin{pmatrix} ct' \\ X' \end{pmatrix}, \quad (1.20)$$

qu'on peut mettre sous la forme :

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} a & -b \\ -e & f \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix}. \quad (1.21)$$

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

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \frac{1}{af - be} \begin{pmatrix} f & -b \\ -e & a \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix}. \quad (1.22)$$

De la comparaison de ces deux formules, on tire immédiatement que le déterminant de la transformation doit être égal à un:

$$af - be = 1 \quad (1.23)$$

et que

$$a = f = \gamma \quad (1.24)$$

(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 le caractère abélien du groupe de Lorentz spécial. En écrivant simplement que $\mathcal{L}(w) = \mathcal{L}(u)\mathcal{L}(v) = \mathcal{L}(v)\mathcal{L}(u)$, on trouve que:

$$\begin{aligned} \gamma(w) &= \gamma(u)\gamma(v) + e(u)b(v) \\ &= \gamma(v)\gamma(u) + e(v)b(u), \end{aligned} \quad (1.25)$$

ce qui ne peut être vérifié pour deux vitesses arbitraires que si $e(u)/b(u)$ est une constante, à moins qu'une de ces fonctions ne s'annule identiquement. Si ce rapport est une constante, un choix convenable d'unités d'espace et de temps permet d'amener sa valeur à ± 1 . Nous aurons donc à distinguer 4 cas:

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

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}. \quad (1.26)$$

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:

$$\begin{pmatrix} 1 & 0 \\ e(u) & 1 \end{pmatrix}. \quad (1.27)$$

Ecrire ensuite que le mouvement de O' s'effectue à la vitesse u , c'est à dire que $x = 0$ implique $x' = -ut'$, fixe $e(u) = -\beta = -u/c$. 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:

$$\begin{pmatrix} 1 & b(u) \\ 0 & 1 \end{pmatrix}. \quad (1.28)$$

Ce groupe, qui diffère du groupe de Galilée en ce qu'il transforme le temps et non l'espace, est le groupe de Carroll. Il est tout naturel de l'éliminer a priori.

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

$$\begin{pmatrix} \cosh \phi & -\sinh \phi \\ -\sinh \phi & \cosh \phi \end{pmatrix} \quad (1.29)$$

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 supplémentaires⁹.

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 Carroll contiennent des transformations changeant le signe de t pour tout événement. Ils sont donc exclus par de simples exigences de causalité.

Il ne nous reste 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

$$x' = -\sinh \phi(u)ct, \quad ct' = \cosh \phi(u)ct \quad (1.30)$$

et

$$\tanh \phi(u) = \frac{u}{c} = \beta. \quad (1.31)$$

soit encore

$$\sinh \phi(u) = \gamma\beta, \quad \gamma = \cosh \phi(u) = \frac{1}{\sqrt{1-\beta^2}}. \quad (1.32)$$

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, 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:

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}, \quad (1.33)$$

la transformation inverse étant évidemment donnée par

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta \\ \gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix}, \quad (1.34)$$

(il suffit de changer le signe de la vitesse relative). En termes de coordonnées et de temps, on peut aussi écrire la transformation directe sous la forme:

$$x' = \gamma(x - ut) \quad (1.35)$$

$$t' = \gamma\left(t - \frac{ux}{c^2}\right) \quad (1.36)$$

et la transformation inverse sous la forme:

$$x = \gamma(x' + ut') \quad (1.37)$$

$$t = \gamma\left(t' + \frac{ux'}{c^2}\right) \quad (1.38)$$

⁹A ce point, imposer la constance de la vitesse de la lumière ou l'invariance de l'intervalle suffirait à choisir le dernier groupe. Nous allons suivre une démarche un peu plus détaillée pour montrer que la sélection peut aller encore plus loin sans le postulat de relativité.

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 retrouve 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 $\mathcal{L}(\mathbf{u})$, où \mathbf{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 "orthochrome" contient le groupe de Lorentz combiné éventuellement avec des réflexions d'espace (nous avons exclu explicitement ces transformations dans notre discussion en imposant aux directions des axes d'être consistantes). Le déterminant de la matrice peut alors être ± 1 .

1.5 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 que la transformation de Lorentz conserve l'intervalle. Nous aurions en fait pu prendre cette hypothèse comme point de départ et construire à partir de là la transformation.

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.

1.5.1 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 :

$$\phi(w) = \phi(u) + \phi(v) . \quad (1.39)$$

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 :

$$\sinh \phi(w) = \beta(w)\gamma(w) = \gamma(u)\gamma(v)(\beta(u) + \beta(v)) \quad (1.40)$$

$$\cosh \phi(w) = \gamma(w) = \gamma(u)\gamma(v)(1 + \beta(u)\beta(v)) , \quad (1.41)$$

d'où on tire immédiatement :

$$\beta(w) = \frac{\beta(u) + \beta(v)}{1 + \beta(u)\beta(v)} , \quad (1.42)$$

ou encore

$$w = \frac{u + v}{1 + uv/c^2}, \quad (1.43)$$

Cette loi remplace la simple addition des vitesses de la relativité galiléenne. Notons là encore qu'on retrouve la loi galiléenne d'addition des vitesses 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 vitesses et de rapidité.

1.5.2 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 n'est que 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 unités réduites. Cette vitesse est celle que déterminerait le contrôleur en mesurant la durée qui s'écoule entre les deux gares avec sa propre montre. Notons enfin que célérité et vitesse sont identiques en relativité galiléenne, en raison de l'universalité du temps.

Nous sommes maintenant en possession d'une troisième "définition" de la vitesse en termes de la rapidité. Si nous savons déjà que son intérêt réside dans son caractère additif, à quelle expérience correspondrait-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 $d\tau$, la vitesse du train dans le référentiel tangent passe de 0 à dv' . L'accélération mesurée dans le train (ou plutôt dans le référentiel tangent) vaut donc $a = dv'/d\tau$. Pendant l'intervalle de temps correspondant, la vitesse dans \mathcal{R} passe de v à $v + dv$. En utilisant la loi de composition des vitesses du paragraphe précédent, nous pouvons écrire l'accroissement de vitesse dv' dans \mathcal{R}' en fonction de la nouvelle vitesse dans \mathcal{R} ($v + dv$) et de la vitesse v de \mathcal{R}' par rapport à \mathcal{R} :

$$dv' = \frac{v + dv - v}{1 - v(v + dv)/c^2}, \quad (1.44)$$

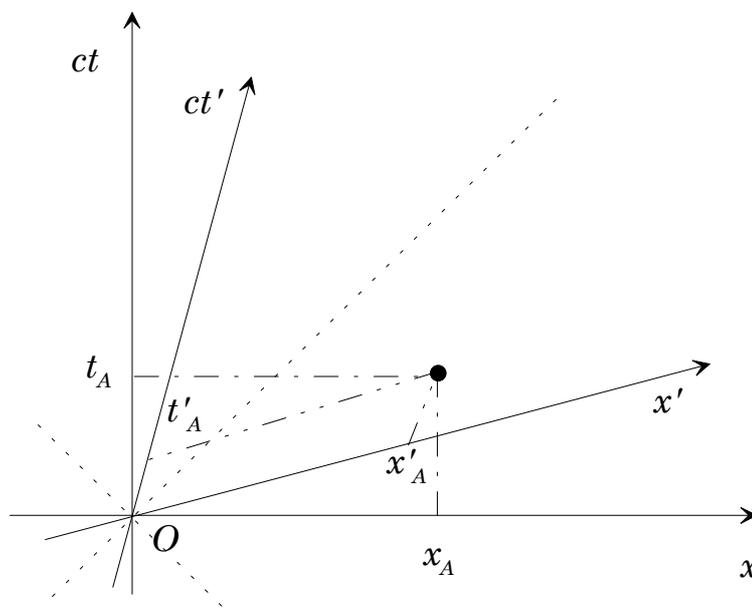


Figure 1.5: La transformation de Lorentz interprétée géométriquement comme une rotation hyperbolique dans l'espace temps. Ses directions propres coïncident avec le cône de lumière de l'origine.

soit encore

$$dv' = \frac{dv}{1 - v^2/c^2} . \quad (1.45)$$

La vitesse finale mesurée par cette méthode sera donc:

$$\int a d\tau = \int \frac{dv}{1 - v^2/c^2} = c \operatorname{arctanh} \frac{u}{c} . \quad (1.46)$$

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 (nous écrirons ce mouvement pour décrire celui de particules chargées soumises 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.

1.5.3 Géométrie de la transformation de Lorentz

Penchons nous maintenant sur l'interprétation géométrique de la transformation de Lorentz dans l'espace temps. Si une rotation ordinaire autour de l'origine est bien connue, il n'en est pas nécessairement de même pour une rotation hyperbolique. La figure 1.5 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 1.5 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:

$$\frac{1 \pm \beta}{\sqrt{1 - \beta^2}} = \frac{1}{\sqrt{1 \mp \beta}} = \exp \pm \phi . \quad (1.47)$$

Les vecteurs propres s'obtiennent sans difficulté. Ils correspondent à $x = ct$ (pour la valeur propre supérieure à un) et $x = -ct$. 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 1.5).

1.6 Conséquences de la transformation de Lorentz

1.6.1 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 pourrions en particulier préciser la valeur de la non-simultanéité dans la seconde expérience, ce que nous ne pouvons faire sans la transformation.

Pour la première expérience, il nous faut distinguer trois événements: le départ de l'impulsion de O' ($x'_1 = y'_1 = t'_1 = 0$) (nous n'écrivons 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 $x'_3 = y'_3 = 0, t'_3 = 2L/c$. La transformation de Lorentz donne les positions de ces trois événements dans \mathcal{R} :

$$x_1 = y_1 = t_1 = 0 \quad (1.48)$$

$$x_2 = \gamma u L/c, \quad y_2 = L, \quad t_2 = \gamma L/c \quad (1.49)$$

$$x_3 = 2\gamma u L/c, \quad y_3 = 0, \quad t_3 = 2\gamma L/c, \quad (1.50)$$

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} :

$$x_A = -L \quad t_A = -L/c \quad (1.51)$$

$$x_B = L \quad t_B = -L/c . \quad (1.52)$$

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

$$x'_A = -\gamma L(1 - \beta) \quad t'_A = -\gamma(1 - \beta)L/c \quad (1.53)$$

$$x'_B = \gamma L(1 + \beta) \quad t'_B = -\gamma(1 + \beta)L/c, \quad (1.54)$$

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

1.6.2 Intervalles et simultanéité

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

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 inférieur 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.

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 petit 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.

1.6.3 Loi de composition des vitesses

Nous généraliserons dans ce paragraphe 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 \mathbf{v} dans le référentiel \mathcal{R}' , entraîné à une vitesse u (selon Ox) par rapport au référentiel \mathcal{R} . En écrivant les accroissements infinitésimaux dx', dy', dz' et dt' dans \mathcal{R}' et en utilisant la transformation de Lorentz, on en déduit les accroissements correspondants dans \mathcal{R} :

$$dx = \gamma(dx' + udt') \quad (1.55)$$

$$dt = \gamma(dt' + udx'/c^2) \quad (1.56)$$

$$dy = dy' \quad (1.57)$$

$$dz = dz' . \quad (1.58)$$

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

$$v_x = \frac{v'_x + u}{1 + uv'_x/c^2} \quad (1.59)$$

$$v_y = \frac{v'_y}{\gamma(1 + uv'_x/c^2)} , \quad (1.60)$$

la transformation inverse s'obtenant trivialement en changeant le signe de u dans les expressions précédentes. On vérifiera, à titre d'exercice, que la vitesse de la lumière est bien invariante dans cette transformation.

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}' :

$$v'_x = -u \quad (1.61)$$

$$v'_y = -\frac{c}{\gamma} \quad (1.62)$$

On "voit" donc dans \mathcal{R}' la lumière de l'étoile provenir d'une direction inclinée par rapport à l'axe $O'y'$, d'un angle $\theta = \arctan \beta\gamma$. L'application de la cinématique galiléenne prévoit aussi une déviation apparente mais elle n'est, comme on le vérifiera aisément, que $\theta = \arctan \beta$. La différence entre ces deux quantités est mesurable et son observation fut une belle confirmation de la relativité restreinte.

Dans le même genre de situation, nous aurions pu nous intéresser aussi à la fréquence de la lumière reçue. Nous l'aurions trouvée modifiée, d'une quantité différente de l'effet Doppler classique (il existe en particulier un effet Doppler du second ordre pour une vitesse d'entraînement perpendiculaire à la propagation). Nous pourrions traiter ce problème plus efficacement quand nous aurons introduit, au chapitre suivant, les notations tensorielles.

1.6.4 Contraction des longueurs

Dans cette nouvelle conséquence immédiate de la transformation de Lorentz, nous allons retrouver l'hypothèse heuristique que Lorentz avait introduite pour expliquer le résultat négatif de l'expérience de Michelson.

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? Il y a deux stratégies.

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

$$x'_A = t'_A = 0 \quad (1.63)$$

$$x'_B = -L' \quad t'_B = L'/u \quad (1.64)$$

(par convention, l'extrémité O' de la règle passe d'abord devant O). Leurs coordonnées dans \mathcal{R} sont alors:

$$x_A = t_A = 0 \quad (1.65)$$

$$x_B = 0 \quad t_B = \gamma \frac{L'}{u} \left(1 - \frac{u^2}{c^2} \right) = \frac{L'}{\gamma u}. \quad (1.66)$$

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 :

$$L = \frac{L'}{\gamma} . \quad (1.67)$$

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} . Pour simplifier, nous prendrons l'instant origine. Les coordonnées dans \mathcal{R} de ces événements sont donc :

$$x_A = t_A = 0 \quad (1.68)$$

$$x_B = -L \quad t_B = 0 . \quad (1.69)$$

Leurs coordonnées dans \mathcal{R}' sont donc :

$$x'_A = t'_A = 0 \quad (1.70)$$

$$x'_B = -\gamma L \quad t'_B = \gamma u L / c^2 . \quad (1.71)$$

Comme on doit aussi avoir $x'_B = -L'$, on en déduit encore

$$L = \frac{L'}{\gamma} . \quad (1.72)$$

Les deux procédures conduisent donc, heureusement, à la même longueur. Rappelons pour finir 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 l'étoile la plus proche, 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. On pourra, à titre d'exercice, écrire complètement le paradoxe des jumeaux en termes d'événements.

Un "Paradoxe" résolu

Pour illustrer encore cette notion de contraction des longueurs, considérons un nouveau problème de trains. 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 est 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 écrivons, dans les deux référentiels, 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:

$$x_2 = L \quad t_2 = \frac{L}{u} \left(1 - \frac{1}{\gamma}\right) . \quad (1.73)$$

On peut obtenir simplement t_2 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 $c^2(1 - \gamma)/u^2\gamma^2$, négatif. 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:

$$x'_2 = L , \quad (1.74)$$

ce qui ne fait guère que vérifier la cohérence du calcul et

$$t'_2 = -\frac{L}{u} \frac{\gamma - 1}{\gamma} . \quad (1.75)$$

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. Ce n'est que la différence entre les prédictions de la relativité sur le temps et notre sens commun qui crée l'apparence paradoxale de ces situations. Ce n'est qu'avec une fréquentation assidue de ce genre de problèmes que peut se développer une intuition relativiste. Le problème est similaire, bien que moins ardu, à celui qu'on rencontre pour se forger une intuition en mécanique quantique.

Chapitre 2

Notations Quadridimensionnelles

Nous allons introduire dans ce chapitre des notions mathématiques permettant de traiter la relativité restreinte d'une façon particulièrement aisée. En nous plaçant dans un espace-temps à quatre dimensions, en le munissant d'une structure d'espace vectoriel, d'un produit scalaire et d'une norme, nous pourrons réinterpréter d'une façon très efficace les notions introduites dans le chapitre précédent. Les objets physiques, indépendants du choix du référentiel, seront alors décrits comme des vecteurs, ou, plus généralement des tenseurs (matrices) dans cet espace. La transformation de Lorentz apparaîtra alors comme un simple changement de base. Les lois de transformation de toutes les quantités physiques se déduiront alors simplement des règles de changement de base. Pour manipuler commodément ces vecteurs et ces tenseurs, nous introduirons également des notations tensorielles très puissantes dues à Einstein. Nous allons commencer par introduire la notion d'espace temps à partir de l'événement.

2.1 4-vecteur position d'un événement.

2.1.1 Coordonnées contravariantes

Un événement est complètement défini dans un référentiel par la donnée des quatre nombres (ct, x, y, z) . Nous considérerons donc un espace vectoriel à quatre dimensions (d'où le nom de 4-vecteur) muni d'une base \mathbf{e}_μ (l'indice μ variant entre 0 et 3) et nous associerons à cet événement le vecteur (ou plutôt le 4-vecteur):

$$\mathbf{R} = \sum_{\mu} x^{\mu} \mathbf{e}_{\mu} , \quad (2.1)$$

en posant

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z . \quad (2.2)$$

Nous appellerons les x^{μ} les composantes contravariantes du 4-vecteur événement. Par convention, dans toute la suite, on utilisera des indices supérieurs pour représenter ces composantes (le risque de confusion avec des exposants est pratiquement nul). Par convention aussi, nous ne préciserons pas le domaine sur lequel s'effectuent les sommations. Un indice représenté par une lettre grecque variera toujours entre 0 et 3 (0 représentant la coordonnée temporelle). Quand nous aurons besoin d'un indice ne parcourant que les valeurs 1 à 3 (et donc associé à de simples composantes spatiales), nous utiliserons un indice romain¹:

$$x^i = x^{1,2,3} . \quad (2.3)$$

L'écriture systématique des signes somme serait extrêmement fastidieuse, surtout quand nous manipulerons, avec les tenseurs, des sommes multiples. Nous adopterons donc la convention d'Einstein

¹Cette convention est très largement répandue dans les manuels récents. Signalons cependant que le Landau de théorie des champs, dont nous recommandons la lecture, utilise une convention strictement inverse.

de sommation sur les indices répétés. Quand, dans une expression, un même indice apparaît en haut et en bas, on doit sommer l'expression sur toutes les valeurs possibles de cet indice (0 à 3 pour un indice grec, 1 à 3 pour un indice romain). On écrira donc simplement:

$$\mathbf{R} = \sum_{\mu} x^{\mu} \mathbf{e}_{\mu} = x^{\mu} \mathbf{e}_{\mu} . \quad (2.4)$$

Insistons sur le fait, dont la raison apparaîtra plus clairement plus tard, que l'on ne somme implicitement sur un indice que s'il apparaît une fois en haut et une fois en bas. Le même indice apparaissant deux fois en bas ne doit pas être sommé. Une écriture comme $T_{\mu\mu}$, par exemple, représentera un élément diagonal d'un tenseur (disons à ce stade d'une matrice) et pas une somme. La trace de cette matrice s'écrirait T_{μ}^{μ} , la convention de somme s'appliquant alors à l'indice répété. Nous verrons dans la suite de cet exposé à quel point cette simple convention de sommation implicite améliore les écritures. Nous appellerons indice "libre" un indice sur lequel la règle de sommation ne s'applique pas (qui donc reste en tant qu'indice dans l'expression finale) et "muet" un indice faisant l'objet d'une sommation implicite (et n'apparaissant pas en tant que tel dans l'expression finale²). Nous respecterons, pour les indices libres, la règle de "balancement". Dans une équation, les indices libres apparaissant dans les deux membres doivent se correspondre un à un et apparaître en même position (haut ou bas). Nous comprendrons plus tard que, si ce n'était pas le cas, les objets décrits par les deux membres de l'équation ne seraient pas de même nature. Ces règles de sommation et de balancement, qui portent sur la typographie des équations, constituent en fait, en plus d'une simplification notable, un système de garde-fous rendant impossible l'écriture d'expressions absurdes. Elles jouent en relativité le rôle des notations de Dirac de la mécanique quantique qui relient de façon univoque la nature de l'objet (fonction d'onde, objet du dual, produit scalaire, opérateur) à son écriture (ket, bra, bra ket, ket bra).

Nous pouvons munir notre espace-temps d'un produit scalaire. Considérant deux 4-vecteurs associés à deux événements, $\mathbf{R} = x^{\mu} \mathbf{e}_{\mu}$ et $\mathbf{S} = y^{\nu} \mathbf{e}_{\nu}$, nous pouvons écrire le produit scalaire comme

$$\mathbf{R} \cdot \mathbf{S} = x^{\mu} y^{\nu} \mathbf{e}_{\mu} \cdot \mathbf{e}_{\nu} . \quad (2.5)$$

Nous souhaiterions bien sûr que notre produit scalaire ait une expression compatible avec celle de l'intervalle. Il faut pour cela que $\mathbf{R} \cdot \mathbf{R} = c^2 t^2 - x^2 - y^2 - z^2$. En posant:

$$g_{\mu\nu} = \mathbf{e}_{\mu} \cdot \mathbf{e}_{\nu} , \quad (2.6)$$

le produit scalaire s'écrit:

$$\mathbf{R} \cdot \mathbf{S} = g_{\mu\nu} x^{\mu} y^{\nu} \quad (2.7)$$

(où l'on voit bien la simplification apportée par la convention de sommation implicite). Nous aurons un produit scalaire convenable si le tableau de nombres $g_{\mu\nu}$ (que nous ne pouvons encore considérer comme un opérateur ou un tenseur – et que nous nommerons néanmoins le "tenseur métrique") s'écrit:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (2.8)$$

Dans ce tableau, μ est l'indice ligne et ν l'indice colonne. Cette convention sur le tenseur métrique peut aussi être vue comme une condition d'orthonormalité pour la base \mathbf{e}_{μ} . Nous ne considérerons dans la suite que des bases "orthonormées" dans ce sens. Notons également que le carré scalaire ne conduit pas à une norme définie positive (nous savons bien que le carré d'un intervalle de genre espace est négatif). Le tenseur métrique qui n'apparaît ici que comme une notation commode dans l'écriture du produit scalaire de deux événements est une des notions centrales en relativité générale, puisqu'il constitue la variable dynamique du champ de gravitation.

Cette écriture du produit scalaire peut être rendue plus compacte encore en introduisant la notion de coordonnée covariante.

²Remarquons à ce propos que le nom d'un indice muet n'a aucune importance dans l'écriture.

2.1.2 Coordonnées covariantes

Posons:

$$y_\mu = g_{\mu\nu} y^\nu . \quad (2.9)$$

Avant d'aller plus avant, commentons le fonctionnement des règles de sommation sur cette expression. L'indice répété (haut et bas) dans le membre de droite est ν . On doit donc sommer sur toutes les valeurs de cet indice. L'indice μ est un indice libre, qui apparaît sous le même nom et dans la même position (basse) dans les deux membres de l'équation. Nous appellerons "coordonnées covariantes de l'événement" les composantes y_μ .

En pratique, les règles de correspondance entre composantes contravariantes et composantes covariantes sont très simples: $y_0 = y^0$, $y_i = -y^i$. Les composantes covariantes nous permettent d'abaisser (ou d'élever pour la transformation inverse) les indices. Les écritures précédentes révèlent une règle très générale: l'abaissement ou l'élévation d'un indice spatial change le signe, alors que l'élévation ou l'abaissement d'un indice temporel ne s'accompagne pas d'un changement de signe.

Avec ces notations, le produit scalaire de deux 4-vecteurs s'écrit simplement:

$$\mathbf{R} \cdot \mathbf{S} = x^\mu y_\mu \quad (2.10)$$

mais aussi

$$\mathbf{R} \cdot \mathbf{S} = x_\mu y^\mu \quad \text{avec} \quad x_\mu = g_{\mu\nu} x^\nu . \quad (2.11)$$

Le mérite de ces notations est donc de dissimuler les conventions de signes du tenseur métrique dans la définition des coordonnées covariantes et d'obtenir un produit scalaire s'écrivant "normalement". On peut bien sûr écrire la transformation inverse, donnant les coordonnées contravariantes en fonction des coordonnées covariantes. En définissant un nouveau "tableau de nombres" $g^{\mu\nu}$ par:

$$y^\nu = g^{\nu\mu} y_\mu , \quad (2.12)$$

on peut écrire:

$$y^\nu = g^{\nu\mu} y_\mu = g^{\nu\mu} g_{\mu\sigma} y^\sigma . \quad (2.13)$$

Les $g^{\mu\nu}$ sont donc tels que:

$$g^{\nu\mu} g_{\mu\sigma} = \delta_\sigma^\nu , \quad (2.14)$$

où les symboles δ_σ^ν sont de simples symboles de Kronecker (1 si les deux indices sont égaux, 0 sinon). En termes de matrices, la matrice des $g^{\mu\nu}$ est donc inverse de la matrice des $g_{\mu\nu}$. On a donc:

$$g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \quad (2.15)$$

A titre d'exercice, montrons que les deux écritures du produit scalaire de l'équation (2.11) sont bien équivalentes:

$$x^\mu y_\mu = g^{\mu\nu} x_\nu g_{\mu\rho} y^\rho = g^{\mu\nu} g_{\mu\rho} x_\nu y^\rho = g^{\nu\mu} g_{\mu\rho} x_\nu y^\rho = \delta_\rho^\nu x_\nu y^\rho = x_\nu y^\nu , \quad (2.16)$$

la première et la dernière expression étant bien sûr équivalentes, puisque le nom d'un indice muet n'a aucune importance. Nous avons utilisé ici explicitement la symétrie du tenseur métrique par une permutation des indices.

2.1.3 Coordonnées covariantes, contravariantes et dualité

Nous avons introduit dans le paragraphe précédent les coordonnées covariantes comme une simple commodité de notation. En fait, elles ont une signification mathématique beaucoup plus profonde et plus fructueuse en termes de dualité. Nous allons donc rappeler quelques propriétés essentielles de la dualité (une notion d’algèbre linéaire qui joue aussi un rôle central en mécanique quantique).

Sur un espace vectoriel M , nous pouvons définir des *formes linéaires*. Une forme linéaire associe à tout vecteur un nombre réel (complexe dans le cas des espaces de Hilbert). On notera \tilde{R} une forme linéaire et $\tilde{R}(\mathbf{S})$ le nombre réel associé au vecteur \mathbf{S} . Une forme linéaire est, comme son nom l’indique, une fonction linéaire de son argument vectoriel. On a donc des relations telles que: $\tilde{R}(\mathbf{S} + \mathbf{T}) = \tilde{R}(\mathbf{S}) + \tilde{R}(\mathbf{T})$ (le lecteur rétablira aisément l’ensemble des règles convenables).

On peut définir sur l’ensemble des formes linéaires une addition (simple addition des images) et la multiplication par un scalaire réel. Ces deux opérations confèrent à l’ensemble des formes linéaires une structure d’espace vectoriel. Nous l’appellerons l’espace *dual* de notre espace vectoriel initial (nous noterons M^* le dual de l’espace M).

On montre que, si M est de dimension finie, le dual est de même dimension. On peut de plus, si M est muni d’un produit scalaire, définir une bijection entre l’espace et le dual. On associe à chaque vecteur \mathbf{R} la forme linéaire \tilde{R} définie par $\tilde{R}(\mathbf{S}) = \mathbf{R} \cdot \mathbf{S}$. A chaque vecteur est associée une forme linéaire et chaque forme linéaire peut s’écrire comme un produit scalaire avec un vecteur fixe. Les propriétés de linéarité du produit scalaire assurent que cette bijection est un isomorphisme entre l’espace et son dual. Cette relation très forte fait que l’on peut considérer un même objet soit comme un vecteur, soit comme une forme linéaire. En particulier, nous pourrions, dans l’espace temps à quatre dimensions, considérer au choix un événement comme un vecteur ou comme une forme linéaire. En fait, nous confondrons souvent les deux représentations en un seul objet.

Dans l’espace dual, nous pouvons choisir une base. En raison de l’isomorphisme avec l’espace d’origine, il est logique de prendre pour base les formes linéaires associées aux vecteurs de base de l’espace d’origine. En fait, pour des questions de notation, nous choisirons dans l’espace dual la base $\tilde{\mathbf{e}}^\mu$ définie par:

$$\tilde{\mathbf{e}}^\mu(\mathbf{e}_\nu) = \delta_\nu^\mu. \quad (2.17)$$

Notons que cette expression ne fait intervenir aucune sommation implicite. Notons également que les $\tilde{\mathbf{e}}^\mu$ diffèrent des formes associées aux vecteurs de base par le signe, quand l’indice μ est spatial. Nous pouvons alors former, à partir d’un 4-vecteur événement $\mathbf{R} = x^\mu \mathbf{e}_\mu$, la forme linéaire $x_\mu \tilde{\mathbf{e}}^\mu$. L’action de cette forme sur le 4-vecteur $\mathbf{S} = y^\mu \mathbf{e}_\mu$ s’écrit alors $x_\mu \tilde{\mathbf{e}}^\mu(y^\nu \mathbf{e}_\nu) = x_\mu y^\mu = \mathbf{R} \cdot \mathbf{S}$. La forme ainsi construite coïncide donc avec la forme linéaire \tilde{R} associée au vecteur \mathbf{R} . Si les composantes contravariantes sont les composantes du 4-vecteur, les composantes covariantes sont les composantes de la forme linéaire associée à ce vecteur sur la base duale. Ceci donne à ces composantes une signification mathématique beaucoup plus forte qu’une simple convention de signe. Comme nous pouvons confondre vecteur et forme linéaire en un seul objet physique, les composantes contravariantes et covariantes ne sont que deux écritures différentes d’une même quantité.

D’un point de vue “typographique”, notons que les indices sont en bas pour les bases dans l’espace d’origine, en haut pour les bases dans l’espace dual, alors que les composantes dans l’espace d’origine (contravariantes) ont des indices en haut et les composantes dans l’espace dual (covariantes) des indices en bas. On comprend bien ici l’origine mathématique des règles de sommation sur les indices. N’écrire de sommations implicites que si l’un des indices est en haut et l’autre en bas, c’est reconnaître que la seule opération légitime est l’action d’une forme linéaire sur un vecteur (ou, autrement dit, le produit scalaire).

2.1.4 Changement de référentiel, changement de base

Un changement de référentiel transforme, comme nous l’avons vu au chapitre précédent, les coordonnées contravariantes d’un événement par la transformation de Lorentz. Nous allons d’abord ex-

primer cette transformation pour les coordonnées contravariantes et en déduire les transformations des coordonnées covariantes. Nous profiterons, pour ces écritures, de nos nouvelles notations. Nous reviendrons ensuite sur les transformations des vecteurs de base pour notre espace-temps et nous montrerons que la transformation de Lorentz est un simple changement de base orthonormée.

Les transformations que nous considérerons dans ce chapitre sont tout à fait générales. Elles englobent la forme spéciale de la transformation de Lorentz mais elles contiennent aussi les rotations, réflexions d'espace... constituant le groupe de Lorentz complet. Toutefois, quand nous expliciterons la forme des transformations, nous nous limiterons à la forme spéciale. Le passage au cas général ne pose que des problèmes d'écriture.

On peut écrire la transformation de Lorentz pour les coordonnées contravariantes comme:

$$x'^{\mu} = \mathcal{L}^{\mu}_{\nu} x^{\nu} , \quad (2.18)$$

où les x^{ν} sont les composantes de l'événement dans le référentiel \mathcal{R} et les x'^{μ} les composantes contravariantes dans le nouveau référentiel \mathcal{R}' . En considérant, dans le tableau de nombres \mathcal{L}^{μ}_{ν} , l'indice μ (relatif donc au nouveau référentiel) comme un indice ligne et l'indice ν (relatif à l'ancien référentiel) comme un indice colonne, l'expression (2.18) apparaît comme un simple produit matriciel. Dans le cas de la forme spéciale de la transformation, on a simplement:

$$\mathcal{L}^{\mu}_{\nu} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \quad (2.19)$$

qu'on pourrait encore écrire en termes de la rapidité ϕ , paramètre additif du groupe de Lorentz. La transformation inverse peut a priori s'écrire

$$x^{\mu} = (\mathcal{L}^{-1})^{\mu}_{\nu} x'^{\nu} , . \quad (2.20)$$

La matrice inverse \mathcal{L}^{-1} s'obtenant simplement en changeant le signe de β dans les expressions précédentes. Nous verrons qu'en fait l'utilisation de cette matrice inverse n'est pas indispensable.

Examinons maintenant la loi de transformation des coordonnées covariantes (dans l'espace dual). Définissons un nouveau "tableau de nombres" \mathcal{L}_{μ}^{ν} par:

$$x'_{\mu} = \mathcal{L}_{\mu}^{\nu} x_{\nu} . \quad (2.21)$$

Notons tout de suite que \mathcal{L}_{μ}^{ν} est forcément différent de \mathcal{L}^{μ}_{ν} . L'ordre et la position (haute ou basse) des indices sont donc tous deux essentiels pour définir les quantités que nous manipulons. On peut déduire le lien entre \mathcal{L}_{μ}^{ν} et \mathcal{L}^{μ}_{ν} de l'invariance du produit scalaire (qui découle de l'invariance de l'intervalle). On a en effet:

$$x'^{\mu} y'_{\mu} = x^{\nu} y_{\nu} \quad \text{avec} \quad x'^{\mu} = \mathcal{L}^{\mu}_{\rho} x^{\rho} \quad \text{et} \quad y'_{\mu} = \mathcal{L}_{\mu}^{\sigma} y_{\sigma} . \quad (2.22)$$

On en déduit donc:

$$\mathcal{L}^{\mu}_{\rho} x^{\rho} \mathcal{L}_{\mu}^{\sigma} y_{\sigma} = x^{\mu} y_{\mu} = x^{\rho} y_{\sigma} \delta_{\rho}^{\sigma} . \quad (2.23)$$

Cette relation devant être vérifiée pour tout couple de vecteurs, on a:

$$\mathcal{L}^{\mu}_{\rho} \mathcal{L}_{\mu}^{\sigma} = \delta_{\rho}^{\sigma} . \quad (2.24)$$

Notons que le membre de gauche de cette équation ne décrit pas un produit de matrices. En toute rigueur, on effectue la sommation sur deux indices lignes. Dans le cas de la forme spéciale de la transformation de Lorentz, les matrices sont symétriques et la matrice de transformation des coordonnées covariantes est simplement l'inverse de la matrice de transformation des coordonnées contravariantes,

inverse que l'on obtient en changeant le signe de la vitesse relative. On aurait pu établir ce résultat en remarquant simplement que le passage des coordonnées contravariantes aux coordonnées covariantes est un simple changement de signe pour les composantes spatiales. On change donc, dans la matrice de transformation, le signe des éléments spatio-temporels. Dans la forme spéciale de la transformation de Lorentz, cela revient à changer le signe de β .

On peut préciser encore le lien entre ces deux transformations en faisant intervenir le tenseur métrique dans l'expression du produit scalaire:

$$\mathbf{R} \cdot \mathbf{S} = x^\rho g_{\rho\sigma} y^\sigma = x'^\mu g_{\mu\nu} y'^\nu . \quad (2.25)$$

Le tenseur métrique qui exprime l'orthogonalité de la base est en effet évidemment invariant dans une transformation de Lorentz. En exprimant les nouvelles coordonnées en fonction des anciennes, on a:

$$x^\rho g_{\rho\sigma} y^\sigma = \mathcal{L}^\mu{}_\rho x^\rho g_{\mu\nu} \mathcal{L}^\nu{}_\sigma y^\sigma . \quad (2.26)$$

Cette relation étant vérifiée quelque soient \mathbf{R} et \mathbf{S} , on en déduit:

$$g_{\rho\sigma} = \mathcal{L}^\mu{}_\rho g_{\mu\nu} \mathcal{L}^\nu{}_\sigma . \quad (2.27)$$

Le déterminant de la matrice représentant le tenseur métrique étant 1, on retrouve à partir de cette expression que le déterminant de la matrice représentant la transformation de Lorentz est ± 1 (rappelons qu'il n'est négatif que si la transformation fait intervenir une réflexion d'espace ou de temps). On peut mettre la dernière relation sous une forme plus parlante en multipliant les deux membres par $g^{\tau\rho}$ et en effectuant les sommations:

$$g^{\tau\rho} g_{\rho\sigma} = \delta_\sigma^\tau = g^{\tau\rho} \mathcal{L}^\mu{}_\rho g_{\mu\nu} \mathcal{L}^\nu{}_\sigma , \quad (2.28)$$

et donc:

$$(g^{\tau\rho} g_{\mu\nu} \mathcal{L}^\mu{}_\rho) \mathcal{L}^\nu{}_\sigma = \delta_\sigma^\tau . \quad (2.29)$$

En rapprochant cette équation de (2.24) en utilisant la symétrie de g , on voit que:

$$\mathcal{L}_\nu{}^\tau = g_{\nu\mu} g^{\tau\rho} \mathcal{L}^\mu{}_\rho . \quad (2.30)$$

D'un point de vue purement typographique, l'interprétation de cette relation est transparente. Nous savons, pour les coordonnées simples, que l'application d'un tenseur métrique avec les indices hauts transforme une coordonnée covariante (indice bas) en une coordonnée contravariante (indice haut). Elever les indices nécessite l'action d'un tenseur métrique à indices hauts (nous dirons complètement contravariant). De même un tenseur métrique complètement covariant (indices bas) abaisse l'indice, transformant une coordonnée contravariante en une coordonnée covariante. La relation que nous venons d'établir pour les changements de base généralise ce principe. Pour passer de $\mathcal{L}^\mu{}_\rho$, avec le premier indice en haut et le second en bas, à $\mathcal{L}_\nu{}^\tau$, il faut appliquer deux tenseurs métriques, l'un complètement contravariant qui élève l'indice colonne, l'autre complètement covariant qui abaisse l'indice ligne. Si on fait confiance à ces règles typographiques, l'établissement détaillé de (2.30) est inutile et on peut écrire cette transformation a priori. On écrira ainsi la relation inverse:

$$\mathcal{L}^\nu{}_\tau = g_{\tau\rho} g^{\nu\mu} \mathcal{L}^\mu{}_\rho . \quad (2.31)$$

Remarquons que la forme très simple du tenseur métrique rend triviales les modifications dans un abaissement ou une élévation d'indice. Comme pour les coordonnées, le changement de position d'un indice spatial change le signe alors qu'un indice temporel ne change pas le signe. Dans le passage d'une transformation à l'autre, seuls changent de signe les coefficients spatio-temporels. Les coefficients spatiaux et temporels sont inchangés. C'est bien ce qu'on observe sur la forme spéciale de la transformation de Lorentz.

On pourrait penser alors ne faire agir sur les transformations de Lorentz qu'un tenseur métrique et à définir deux nouvelles quantités:

$$\mathcal{L}^{\mu\tau} = g^{\tau\rho} \mathcal{L}^{\mu}_{\rho} \quad (2.32)$$

$$\mathcal{L}_{\mu\sigma} = g_{\mu\nu} \mathcal{L}^{\nu}_{\sigma} . \quad (2.33)$$

En fait ces nouvelles quantités décrivent, comme nous allons le voir, le passage simultané d'un référentiel à l'autre et d'un type de coordonnées à l'autre. On a en effet:

$$x'^{\mu} = \mathcal{L}^{\mu}_{\rho} x^{\rho} = \mathcal{L}^{\mu}_{\rho} g^{\rho\nu} x_{\nu} = \mathcal{L}^{\mu\nu} x_{\nu} \quad (2.34)$$

et de même:

$$x'_{\mu} = \mathcal{L}_{\mu\nu} x^{\nu} . \quad (2.35)$$

On peut enfin utiliser la relation (2.24) pour exprimer le changement de référentiel inverse sans faire intervenir la transformation \mathcal{L}^{-1} . On a en effet:

$$\mathcal{L}^{\mu}_{\rho} x'_{\mu} = \mathcal{L}^{\mu}_{\rho} \mathcal{L}_{\mu}^{\nu} x_{\nu} = \delta_{\rho}^{\nu} x_{\nu} = x_{\rho} \quad (2.36)$$

et donc

$$x_{\rho} = \mathcal{L}^{\mu}_{\rho} x'_{\mu} . \quad (2.37)$$

On peut bien sûr écrire des transformations similaires pour les coordonnées contravariantes ou toute combinaison de composantes mixtes:

$$x'_{\mu} = \mathcal{L}_{\mu}^{\nu} x_{\nu} \quad (2.38)$$

$$x'^{\mu} = \mathcal{L}^{\mu}_{\nu} x^{\nu} \quad (2.39)$$

$$x_{\mu} = \mathcal{L}^{\nu}_{\mu} x'^{\nu} \quad (2.40)$$

$$x^{\mu} = \mathcal{L}_{\nu}^{\mu} x'^{\nu} , \quad (2.41)$$

les différentes expressions de la transformation se déduisant de la forme originale par les règles d'élévation et d'abaissement des indices. Ces différentes combinaisons s'obtiennent automatiquement (ou plutôt typographiquement) en respectant les règles de balancement des indices, en attribuant le premier indice (ligne) au nouveau référentiel, le second à l'ancien et en sommant sur l'indice correspondant au référentiel de la coordonnée à transformer. L'ensemble de ces règles très strictes, que nous avons justifiées en détail, minimise le risque d'erreurs dans ces écritures, beaucoup plus efficacement qu'avec les notations standard de l'algèbre linéaire.

Nous terminerons ce paragraphe en examinant les lois de transformation des vecteurs de base de notre espace-temps. De manière évidente, la transformation de Lorentz correspond à un simple changement de base dans l'espace temps. Le même événement, le même vecteur, s'exprime par deux jeux de composantes différentes quand on le représente dans deux référentiels ou deux bases différentes. En remarquant que $x_{\nu} = \mathbf{R} \cdot \mathbf{e}_{\nu} = \tilde{R}(\mathbf{e}_{\nu})$ et que $x'^{\mu} = \mathbf{R} \cdot \mathbf{e}'_{\mu}$, où les \mathbf{e}'_{μ} sont les transformés des vecteurs de base, on peut écrire $\mathbf{R} \cdot \mathbf{e}'_{\mu} = \mathcal{L}_{\mu}^{\nu} \mathbf{R} \cdot \mathbf{e}_{\nu}$, relation qui doit être vérifiée pour tout vecteur \mathbf{R} . On en déduit donc que:

$$\mathbf{e}'_{\mu} = \mathcal{L}_{\mu}^{\nu} \mathbf{e}_{\nu} . \quad (2.42)$$

En d'autres termes, la loi de transformation des vecteurs de base est celle des composantes covariantes, inverse (au sens précisé plus haut) de la transformation des composantes contravariantes. Les appellations covariantes et contravariantes proviennent précisément de ces comportements.

On peut aussi exprimer la loi de transformation de la base duale. En écrivant $x'^{\mu} = \tilde{\epsilon}'^{\mu}(\mathbf{R})$, $x^{\mu} = \tilde{\epsilon}^{\mu}(\mathbf{R})$, on en déduit:

$$\tilde{\epsilon}'^{\mu} = \mathcal{L}_{\nu}^{\mu} \tilde{\epsilon}^{\nu} . \quad (2.43)$$

Les vecteurs de la base duale se transforment donc comme les composantes contravariantes.

2.2 Autres 4-vecteurs

2.2.1 Définition

Nous n'avons considéré jusque là que le 4-vecteur associé à la position d'un événement dans l'espace-temps. Nous allons maintenant généraliser la notion à d'autres quantités physiques. Nous appellerons 4-vecteur toute quantité physique s'exprimant par 4 coordonnées (covariantes ou contravariantes) se transformant comme le 4-vecteur position dans un changement de référentiel (c'est à dire par une transformation de Lorentz). De manière évidente, toute collection de quatre quantités physiques arbitraires ne constitue pas un 4-vecteur. Nous verrons dans les prochains paragraphes quelques exemples de 4-vecteurs.

Les composantes contravariantes de tout 4-vecteur \mathbf{A} peuvent s'écrire sous la forme de trois composantes spatiales qui forment un vecteur (ce qui résulte des lois de transformation dans le sous-groupe du groupe de Lorentz qui décrit les changements de repère sans changement de référentiel) et d'une coordonnée temporelle. Nous noterons $\mathbf{A} = (a^0, \mathbf{a})$, où a^0 est sa composante temporelle et \mathbf{a} le vecteur tridimensionnel. Les composantes covariantes pourraient alors se mettre sous la forme $(a_0 = a^0, -\mathbf{a})$. A priori, ces composantes ont une dimension qui peut être différente de celles des composantes du 4-vecteur position. Chaque espèce de 4-vecteur devrait donc, en toute rigueur, exister dans un espace-temps propre. En fait, comme en dynamique classique, nous ferons évoluer tous nos 4-vecteurs, positions, vitesses et accélérations dans le même espace. Nous prendrons simplement garde à ne pas additionner des objets de natures différentes.

Les notions de produit scalaire, de composantes covariantes et contravariantes, les règles de changement de base ou de référentiel s'appliquent bien sûr aux 4-vecteurs quelle que soit leur nature.

Nous profiterons aussi de ce paragraphe pour introduire la notion de 4-scalaire. Il s'agit simplement d'une quantité physique indépendante du référentiel, telle que la vitesse limite, la constante de structure fine, la masse d'une particule ou, plus simplement, le produit scalaire de deux 4-vecteurs (le produit scalaire de vecteurs de dimensions différentes n'étant pas exclu). Le produit d'un 4-vecteur par un 4-scalaire donne bien sûr un autre 4-vecteur.

2.2.2 4-vitesse, 4-impulsion, 4-accélération

Nous chercherons ici à définir la vitesse d'une particule relativiste. Il nous faut pour cela paramétrer la ligne d'univers de la particule par un paramètre temporel. Ce paramètre doit être un 4-scalaire. Nous pourrions ainsi définir simplement la vitesse comme étant le rapport de l'accroissement du 4-vecteur position pendant un petit intervalle de temps à l'accroissement de ce paramètre temporel. Le résultat sera alors évidemment un 4-vecteur. Le temps absolu d'un référentiel ne convient pas pour paramétrer la trajectoire. Nous avons vu, en revanche, que le temps propre τ de la particule en mouvement est indépendant du choix de référentiel. Ce temps propre permet donc de définir un 4-vecteur vitesse par :

$$\mathbf{U} = \frac{d\mathbf{R}}{d\tau} . \quad (2.44)$$

Ecrivons cette quantité en fonction de la vitesse spatiale ordinaire \mathbf{v} de la particule dans un référentiel \mathcal{R} donné. Si le temps propre varie de $d\tau$, le temps dans \mathcal{R} varie de $dt = \gamma d\tau$ (dilatation des temps). La position de la particule variant de $d\mathbf{r} = \mathbf{v}dt$, on a immédiatement $d\mathbf{R}/dt = (c, \mathbf{v})$ et

$$\mathbf{U} = (c\gamma, \gamma\mathbf{v}) . \quad (2.45)$$

Remarquons que la partie spatiale de la 4-vitesse n'est autre que la célérité (vitesse calculée dans l'espace du référentiel et dans le temps propre de la particule). \mathbf{U} étant un 4-vecteur, il se transforme par la transformation de Lorentz dans un changement de référentiel. Un excellent exercice consiste à appliquer les lois de transformations décrites au paragraphe précédent pour retrouver la loi de

composition des vitesses relativistes. Notons enfin que le carré de la norme de \mathbf{U} , $U^\mu U_\mu = c^2\gamma^2 - \gamma^2 v^2 = c^2$ n'est autre que le carré de la vitesse de la lumière.

En multipliant la 4-vitesse par la masse de la particule, qui est évidemment un 4-scalaire, on obtient encore un 4-vecteur, l'impulsion de la particule:

$$\mathbf{P} = m\mathbf{U} = (p^0, \mathbf{p}) = (m\gamma c, m\gamma \mathbf{v}) . \quad (2.46)$$

Cette quantité jouera un rôle essentiel dans notre nouvelle dynamique.

La 4-vitesse \mathbf{U} est en général une fonction du temps propre τ de la particule. Nous pouvons donc la dériver encore par rapport à ce temps, pour aboutir à une définition de la 4-accélération

$$\mathbf{\Gamma} = \frac{d\mathbf{U}}{d\tau} = \gamma \frac{d\mathbf{U}}{dt} = (c\gamma\gamma', \gamma\gamma' \mathbf{v} + \gamma^2 \mathbf{a}) , \quad (2.47)$$

où $\gamma' = d\gamma/dt$ est la dérivée temporelle ordinaire du facteur γ et \mathbf{a} est l'accélération tridimensionnelle de la particule. On remarque que la 4-accélération est perpendiculaire à la 4-vitesse: $\mathbf{\Gamma}^\mu U_\mu = 0$. Si on peut établir cette relation à partir des expressions précédentes de ces 4-vecteurs, elle découle beaucoup plus simplement du fait que le module de \mathbf{U} est une constante.

2.2.3 Densité de courant

Nous allons maintenant former un 4-vecteur à partir de la densité de charges ρ et de la densité de courant \mathbf{j} . Il est clair en effet que ces deux quantités se transforment de manières profondément reliées dans un changement de référentiel. Pour comprendre comment former ce 4-vecteur, nous allons considérer le cas d'une charge q "ponctuelle", que nous modéliserons comme une densité de charge constante contenue dans un petit élément de volume dV (défini dans un référentiel \mathcal{R}). Nous écrirons donc $q = \rho dV$. La particule se déplace à la vitesse \mathbf{v} dans \mathcal{R} . La densité de courant est donc $\mathbf{j} = \rho \mathbf{v}$ en tous points de dV (elle est nulle en dehors).

Dans un autre référentiel \mathcal{R}' , toutes ces quantités sont a priori modifiées. La "particule" occupe, à un instant donné dans \mathcal{R}' , un élément de volume dV' . Seule la charge totale q , qui est une propriété intrinsèque de la particule, est un 4-scalaire. On a donc $\rho dV = \rho' dV'$. En d'autres termes, ρdV est un 4-scalaire (notons que le lien entre dV et dV' n'est pas évident, puisqu'il faut faire la transformation de Lorentz sur tous les "coins" de l'élément de volume. Nous reviendrons bientôt sur ce point). Considérons maintenant un intervalle de temps dt dans \mathcal{R} . Pendant ce temps, la particule se déplace de $dx^\mu = (cdt, d\mathbf{r})$ qui est évidemment un 4-vecteur. La quantité $\rho dV dx^\mu = \rho dt dV (dx^\mu/dt) = \rho (d\Omega/c) dx^\mu/dt$ est donc, elle aussi un 4-vecteur, si $d\Omega = c dt dV$ représente un élément d'intégration dans l'espace-temps (en revanche, dx^μ/dt n'est pas un 4-vecteur, de même que dt ou dV pris séparément ne sont pas des 4-scalaires).

Nous montrerons maintenant que $d\Omega = c dt dV$ est un 4-scalaire. D'un point de vue physique, $d\Omega$ représente un petit domaine de l'espace-temps considéré entre deux instants infiniment voisins et a la dimension d'un volume multiplié par un temps. Modélisons ce petit domaine par un cube élémentaire de dimensions dx, dy et dz (voir Fig. 2.1). Si nous choisissons les axes convenablement, nous n'aurons à considérer que la forme spéciale de la transformation de Lorentz. Les coordonnées y et z étant inchangées, nous nous ramenons à démontrer l'invariance de l'élément de surface dans le plan (x, t) . L'étendue de notre élément est la surface rectangulaire de côtés dx et dt dans le référentiel \mathcal{R} . Sa mesure est bien $dx dt$. Remarquons tout de suite que $dx dt$ n'est pas un invariant de Lorentz, comme on peut s'en convaincre aisément. Vu dans le référentiel \mathcal{R}' , le petit élément de surface prend l'aspect représenté sur la partie droite de la figure 2.1. Il s'agit d'un petit losange. La surface de ce losange peut s'exprimer facilement en écrivant les transformés de Lorentz de trois des sommets. En fait, on trouve que cette mesure, qui est manifestement différente de $dx' dt'$ est égale à $dx dt$ (ce résultat découle directement du fait que le déterminant de la transformation de Lorentz est un, ce qui confirme sa généralité). L'étendue totale dans l'espace-temps d'un petit élément d'intégration est

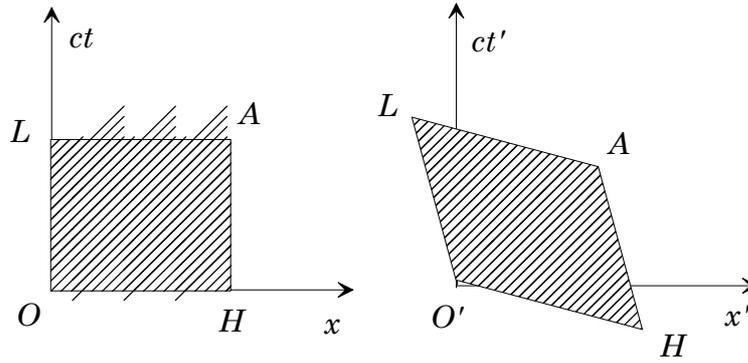


Figure 2.1: Un même élément de volume de l'espace-temps vu dans deux référentiels. La forme change, mais la surface (ou l'hypervolume dans l'espace quadridimensionnel) est constante.

donc un 4-scalaire (résultat qui nous permettra plus tard de définir convenablement des intégrales de volume dans l'espace-temps).

On déduit de tout ce qui précède que

$$\mathbf{J}^\mu = \rho \frac{dx^\mu}{dt} \quad (2.48)$$

est un 4-vecteur réunissant les densités de charge et de courant créés par une charge ponctuelle. On peut aussi écrire

$$\mathbf{J}^\mu = (c\rho, \mathbf{j}) , \quad (2.49)$$

en faisant intervenir la densité de courant tridimensionnelle. Pour une répartition quelconque de charges et de courants, résultant de la superposition des mouvements d'un grand nombre de particules ponctuelles, la quantité écrite en termes des densités totales de charges et de courants est bien sûr encore un 4-vecteur.

2.2.4 Vecteur d'onde

Considérons, dans le référentiel \mathcal{R} , une onde plane monochromatique de fréquence ω et de vecteur d'onde \mathbf{k} . Les différents champs et potentiels de cette onde sont tous proportionnels à $\exp(-i\phi)$ avec:

$$\phi = \omega t - \mathbf{k} \cdot \mathbf{r} . \quad (2.50)$$

La phase ϕ de l'onde plane se doit d'être un 4-scalaire. Sa valeur pour un événement donné représente en effet la valeur relative commune des champs et potentiels par rapport à leur maximum. Les transformations de Lorentz de ces quantités, dont nous ne préjugerons pas encore, étant sûrement linéaires, le même facteur $\exp(i\phi)$ décrit, dans un autre référentiel \mathcal{R}' , la valeur des quantités transformées par rapport à leurs nouvelles amplitudes.

En posant

$$k^\mu = (\omega/c, \mathbf{k}) , \quad (2.51)$$

on peut, au moins formellement, écrire:

$$\phi = k_\mu x^\mu , \quad (2.52)$$

où les x^μ sont les composantes contravariantes de l'événement considéré. ϕ étant un 4-scalaire et x^μ un 4-vecteur, il est évident que k^μ est, lui aussi, un 4-vecteur généralisant le vecteur d'onde dans l'espace-temps. On notera que sa norme est nulle:

$$k_\mu k^\mu = 0 \quad (2.53)$$

puisque $\omega^2 - c^2 k^2 = 0$ pour une onde plane dans le vide. Le vecteur d'onde se transforme donc simplement par une transformation de Lorentz quand on passe d'un référentiel à un autre:

$$k'^{\mu} = \mathcal{L}^{\mu}_{\nu} k^{\nu} . \quad (2.54)$$

En explicitant cette transformation pour les coordonnées spatiales et temporelle, on obtient:

$$\omega' = \gamma(\omega - uk_x) \quad (2.55)$$

$$k'_x = \gamma(k_x - \beta\omega/c) \quad (2.56)$$

$$k'_y = k_y \quad (2.57)$$

$$k'_z = k_z \quad (2.58)$$

Dans la première équation, on reconnaît le changement de fréquence d'une onde dans un changement de référentiel, c'est à dire l'effet Doppler. L'expression que nous trouvons ici ne diffère de l'expression habituelle en relativité galiléenne ($\omega' = \omega - \mathbf{k} \cdot \mathbf{u}$) que par le facteur de dilatation des temps γ (il est bien naturel que ce facteur intervienne dans la modification de la fréquence). Pour des mobiles qui ne sont pas trop rapides (ce qui est presque toujours le cas, par exemple, pour des sources atomiques), l'effet Doppler reste dominé par l'effet classique. Il est un cas cependant où l'effet relativiste domine, celui où la vitesse \mathbf{u} est strictement perpendiculaire au vecteur d'onde. En cinématique classique, la fréquence n'est pas affectée. En relativité einsteinienne, en revanche, la fréquence est multipliée par le facteur γ de dilatation des temps. La modification relative de la fréquence est donc du second ordre en β (d'où le nom d'effet Doppler du second ordre), alors que l'effet Doppler classique est du premier ordre. C'est donc un effet faible pour des vitesses ordinaires qui est cependant gênant dans des expériences de haute précision. La spectroscopie atomique la plus précise actuellement est réalisée sur des atomes d'hydrogène excités par laser. Pour s'affranchir de l'effet Doppler, on fait en sorte que la vitesse des atomes soit très précisément perpendiculaire à la direction du laser. La cause essentielle d'erreur devient alors l'effet Doppler du second ordre qu'il faut corriger très précisément. Pour cela, il est indispensable de réduire autant que possible la vitesse des atomes et de la déterminer précisément.

Les équations portant sur la partie spatiale décrivent le changement de direction de propagation. Les composantes transverses ne sont pas modifiées. La composante longitudinale change de telle manière que la relation de dispersion des ondes planes dans le vide $k' = \omega'/c$ reste vérifiée, comme on pourra s'en convaincre aisément. Notons qu'on peut ainsi retrouver de manière très rapide l'aberration relativiste des étoiles. La situation que nous considérons au chapitre précédent était en effet une onde se propageant vers l'origine le long de l'axe Oy : k_y est donc la seule composante non nulle dans le référentiel \mathcal{R} . En appliquant la transformation précédente, on trouve immédiatement:

$$k'_y = k_y = -\omega/c \quad (2.59)$$

$$k'_x = -\beta\gamma\omega/c \quad (2.60)$$

La direction de propagation fait donc, dans le référentiel \mathcal{R}' un angle $\theta = \arctan \beta\gamma$ avec l'axe Oy' . Nous retrouvons bien le résultat du chapitre précédent.

2.3 Tenseurs

Nous aurons à manipuler des quantités plus complexes que les 4-vecteurs. Dans un espace vectoriel, on peut bien sûr définir des opérateurs linéaires, représentés par des matrices dans une base convenable. En fait, nous considérerons des objets plus généraux pouvant dépendre de plus de deux indices. De tels objets décrivent des lois de composition multilinéaires entre vecteurs ou opérateurs. Les tenseurs se retrouvent en fait dans de nombreux domaines de la physique. En élasticité, par exemple, la déformation d'un solide dans une direction peut dépendre d'une contrainte appliquée dans une autre

direction. La relation linéaire correspondante est décrite par un tenseur de rang 2. En optique non linéaire, un ensemble de trois champs électriques peut créer une polarisation dans le milieu. Il faudra donc écrire une application linéaire donnant un vecteur à partir de trois autres, ce qui doit être décrit par un tenseur à quatre indices. Notons tout de suite que n'importe quel tableau de nombres n'est pas un tenseur, comme toute collection de nombres ne représente pas forcément un vecteur. Il faut, de plus, que ces quantités se transforment "bien" dans un changement de base.

2.3.1 Tenseurs contravariants

L'opération de produit tensoriel permet d'associer à l'espace vectoriel M un espace $M \otimes M$, plus grand. A tout couple de vecteurs \mathbf{R} et \mathbf{S} de M , on associe un vecteur $\mathbf{R} \otimes \mathbf{S}$ de $M \otimes M$. Une base de $M \otimes M$ est formée des 16 produits tensoriels possibles formés avec les 4 vecteurs de base de M : $\mathbf{e}_\mu \otimes \mathbf{e}_\nu$. Les composantes de $\mathbf{R} \otimes \mathbf{S}$ sur cette base sont les produits des composantes de \mathbf{R} et \mathbf{S} :

$$\mathbf{R} \otimes \mathbf{S} = x^\mu y^\nu \mathbf{e}_\mu \otimes \mathbf{e}_\nu . \quad (2.61)$$

La dimension de l'espace produit tensoriel est 16. Un produit tensoriel de deux vecteurs ne dépend que de 8 paramètres libres. Nous pouvons donc définir des objets plus généraux que les produits tensoriels de vecteurs: les tenseurs de rang 2 complètement contravariants, éléments de l'espace $M \otimes M$. Nous noterons $T^{\mu\nu}$ les composantes d'un tel tenseur sur la base $\mathbf{e}_\mu \otimes \mathbf{e}_\nu$. En fait, nous avons déjà rencontré de tels objets, par exemple avec le tenseur métrique complètement contravariant.

On peut déduire la règle de transformation d'un tenseur dans un changement de base (c'est à dire une transformation de Lorentz) de celle d'un produit tensoriel de vecteurs: $T^{\mu\nu}$ se transforme comme un produit de composantes contravariantes:

$$T'^{\mu\nu} = \mathcal{L}^\mu{}_\rho \mathcal{L}^\nu{}_\sigma T^{\rho\sigma} . \quad (2.62)$$

En appliquant les règles du paragraphe précédent, nous aurons aussi la transformation inverse:

$$T^{\mu\nu} = \mathcal{L}_\rho{}^\mu \mathcal{L}_\sigma{}^\nu T'^{\rho\sigma} . \quad (2.63)$$

Notons que ces expressions ne décrivent pas des produits de trois matrices au sens ordinaires. Dans la première, par exemple, la sommation sur σ est une sommation sur un indice colonne dans les deux termes où il apparaît. Il faut donc prendre garde, dans les calculs pratiques, d'effectuer les transpositions nécessaires si on veut utiliser les règles standard du produit matriciel.

L'opération de produit tensoriel peut être généralisée aisément à un nombre arbitraires de termes. On peut définir l'espace $M^{\otimes k}$, produit tensoriel de M k fois avec lui même. Les éléments de cet espace, de dimension 4^k , sont les tenseurs complètement contravariants de rang k et leurs composantes s'écriront $T^{\mu\nu\rho\dots\tau}$. Ces composantes se transforment par un "produit" de k transformations de Lorentz. Nous laissons au lecteur le soin de l'écrire. Un tenseur de rang n décrit une transformation multilinéaire qui, à $n - 1$ vecteurs associe un vecteur.

2.3.2 Tenseurs covariants, tenseurs mixtes

Ce que nous avons fait pour l'espace M peut être repris pour son dual M^* . On peut définir ainsi des tenseurs de rang deux, complètement covariants, dont les composantes sur la base produit tensoriel $\tilde{\mathbf{e}}^\mu \otimes \tilde{\mathbf{e}}^\nu$ s'écriront $T_{\mu\nu}$. La transformation de Lorentz de ces quantités s'écrit simplement:

$$T'_{\mu\nu} = \mathcal{L}_\mu{}^\rho \mathcal{L}_\nu{}^\sigma T_{\rho\sigma} . \quad (2.64)$$

On peut faire le produit tensoriel d'un nombre arbitraires d'espaces duaux. On peut aussi définir des objets appartenant au produit tensoriel de l'espace M par son dual M^* . On obtient alors des tenseurs mixtes de rang deux (ou plus si on utilise plusieurs M et M^*) dont les composantes s'écriront

$T^\mu{}_\nu$ (pour $M \otimes M^*$) ou $T_\mu{}^\nu$ (pour $M^* \otimes M$). Notons que ces deux écritures recouvrent a priori des objets différents. La règle de transformation d'un tel tenseur mixte est simplement:

$$T'^\mu{}_\nu = \mathcal{L}^\mu{}_\rho \mathcal{L}_\nu{}^\sigma T^\rho{}_\sigma, \quad (2.65)$$

et se généralise aisément à tout tenseur mixte de n'importe quel rang.

En fait, nous savons bien que les composantes contravariantes et les composantes covariantes recouvrent le même objet physique. Il en est de même pour les tenseurs: une quantité physique s'exprimant comme un tenseur peut être écrite à volonté comme un tenseur complètement contravariant, complètement covariant, ou mixte de façon arbitraire. Comme pour les composantes des 4-vecteurs (et cela résulte du lien entre tenseur et produits de ces composantes), le tenseur métrique $g^{\mu\nu}$ ou $g_{\mu\nu}$ peut être utilisé pour élever ou abaisser les indices. Nous pourrions écrire par exemple:

$$T^{\mu\nu} = g^{\mu\rho} g^{\nu\sigma} T_{\rho\sigma} \quad (2.66)$$

$$T^\mu{}_\nu = g^{\mu\rho} g_{\nu\sigma} T_\rho{}^\sigma \quad (2.67)$$

$$T^\mu{}_\nu = g_{\nu\rho} T^{\mu\rho}. \quad (2.68)$$

Le tenseur métrique se réduisant en relativité restreinte, à un tableau diagonal de signes, ces règles de transformation prennent une forme très simple: l'abaissement ou l'élévation d'un indice spatial change le signe de la quantité, alors que l'élévation ou l'abaissement d'un indice temporel ne change rien. On trouve ainsi, par exemple, que seuls les indices spatio-temporels changent de signe dans le passage d'un tenseur de rang 2 de la forme complètement contravariante à la forme complètement covariante.

Ces définitions des tenseurs peuvent paraître abstraites. On peut aussi voir les tenseurs de rang deux comme des applications linéaires de l'espace M dans lui-même (ou de son dual dans lui-même). Un tenseur complètement covariant de rang 2, par exemple, décrit naturellement une application linéaire de M dans M^* . L'image W d'un vecteur V s'écrit alors:

$$W_\mu = T_{\mu\nu} V^\nu. \quad (2.69)$$

On aurait des écritures similaires pour les tenseurs complètement contravariants (de M^* dans M) ou les tenseurs mixtes, décrivant la transformation entre deux objets de même nature. Nous laisserons au lecteur le soin de les écrire.

Les règles de changement de base pour les tenseurs peuvent se déduire de cette forme. Nous prendrons une transformation de Lorentz \mathcal{L} . Les transformations des composantes de W et de V s'écrivent:

$$W'_\rho = \mathcal{L}_\rho{}^\mu W_\mu \quad V^\nu = \mathcal{L}_\sigma{}^\nu V'^\sigma. \quad (2.70)$$

On peut donc écrire:

$$W'_\rho = \mathcal{L}_\rho{}^\mu W_\mu = \mathcal{L}_\rho{}^\mu T_{\mu\nu} V^\nu = \mathcal{L}_\rho{}^\mu T_{\mu\nu} \mathcal{L}_\sigma{}^\nu V'^\sigma = T'_{\rho\sigma} V'^\sigma. \quad (2.71)$$

On lit directement sur cette équation l'expression des composantes transformées du tenseur. On constatera qu'elles coïncident bien avec celles que nous avons obtenues plus haut.

2.3.3 Vocabulaire et exemples

Un tenseur de rang 2 est dit symétrique si:

$$T^{\mu\nu} = T^{\nu\mu} \quad (2.72)$$

On en déduit immédiatement $T_{\mu\nu} = T_{\nu\mu}$ et $T^\mu{}_\nu = T_\nu{}^\mu$. On pourra donc écrire la forme mixte sans préciser l'ordre des indices, comme $T^\mu{}_\nu$. Notons que la symétrie du tenseur n'implique *pas* l'égalité de $T^\mu{}_\nu$ et de $T_\mu{}^\nu$ (il n'est que d'examiner le cas des indices spatio-temporels pour s'en convaincre).

Un tenseur de rang 2 est dit antisymétrique si:

$$T^{\mu\nu} = -T^{\nu\mu} . \quad (2.73)$$

Les termes diagonaux de ce tenseur $T^{\mu\mu}$ sont évidemment nuls. On montre qu'un tenseur antisymétrique peut se mettre sous la forme:

$$T^{\mu\nu} = \begin{pmatrix} 0 & a_x & a_y & a_z \\ -a_x & 0 & -b_z & b_y \\ -a_y & b_z & 0 & -b_x \\ -a_z & -b_y & b_x & 0 \end{pmatrix} = (\mathbf{a}, \mathbf{b}) \quad (2.74)$$

où \mathbf{a} est un vecteur spatial et \mathbf{b} un pseudo-vecteur (qui se transforme en l'opposé de son symétrique dans un changement de base comprenant une réflexion d'espace). Le couple champ électrique/champ magnétique obéissant à ces conditions, nous ne nous étonnerons pas que le champ électromagnétique s'exprime comme un tenseur antisymétrique de rang 2.

Nous appellerons "trace" d'un tenseur de rang 2 la quantité $T^{\mu}_{\mu} = T_{\mu}^{\mu}$. Plus généralement, nous appellerons "contraction" d'un tenseur sur un indice une expression comme $T^{\mu}_{\mu}{}^{\rho}$. La contraction d'un tenseur de rang k sur un indice est un tenseur de rang $k - 2$ (la trace étant une contraction d'un tenseur de rang 2, elle donne un tenseur de rang 0, c'est à dire un 4-scalaire). La contraction d'un tenseur de rang trois donne, pour sa part, un tenseur de rang 1, c'est à dire un 4-vecteur.

Montrons, à titre d'exercice, que la contraction sur un indice d'un tenseur de rang trois se transforme bien comme un vecteur:

$$\begin{aligned} T^{\mu}{}_{\mu}{}^{\rho} &= \mathcal{L}^{\mu}{}_{\sigma} \mathcal{L}_{\mu}{}^{\nu} \mathcal{L}^{\rho}{}_{\tau} T^{\sigma}{}_{\nu}{}^{\tau} \\ &= \delta_{\sigma}^{\nu} \mathcal{L}^{\rho}{}_{\tau} T^{\sigma}{}_{\nu}{}^{\tau} \\ &= \mathcal{L}^{\rho}{}_{\tau} T^{\sigma}{}_{\sigma}{}^{\tau} , \end{aligned} \quad (2.75)$$

ce qui établit bien la propriété cherchée.

Comme tenseurs, nous connaissons déjà le tenseur métrique. On pourra vérifier directement qu'il est bien invariant dans une transformation de Lorentz. Il s'agit d'un tenseur symétrique. Sa forme mixte est évidemment: $g_{\nu}^{\mu} = g^{\mu\rho} g_{\nu\rho} = \delta_{\nu}^{\mu}$. Le symbole de Kronecker n'est donc que la forme mixte du tenseur métrique! La relation entre les formes contravariantes et covariantes $g^{\mu\nu} g_{\nu\rho} = \delta_{\rho}^{\mu}$ peut donc s'interpréter comme un simple abaissement d'indice.

Notons à ce point que les opérateurs \mathcal{L} de changement de base ont toutes les caractéristiques de tenseurs et se comportent normalement vis à vis de l'élévation ou de l'abaissement des indices. En toute rigueur, nous ne devons pas les considérer comme des tenseurs. On sait bien, en algèbre linéaire, que les matrices de changement de base ne sont pas à proprement parler des opérateurs. On peut aussi s'en convaincre en s'interrogeant sur le sens physique d'un changement de base pour une transformation de Lorentz.

Nous définirons finalement le tenseur de rang 4 complètement antisymétrique $\epsilon^{\mu\nu\rho\sigma}$. Parmi les 256 éléments de ce tenseur, seuls sont non nuls ceux dont les indices correspondent à une permutation de $(0, 1, 2, 3)$. Si la permutation est paire, l'élément correspondant vaut +1. Il vaut -1 si la permutation est impaire. Il n'y a donc que 24 éléments non nuls dans ϵ , 12 valant +1 et 12 -1. On a $\epsilon^{\mu\nu\rho\sigma} = -\epsilon_{\mu\nu\rho\sigma}$, puisqu'on abaisse toujours trois indices spatiaux dans cette opération. On a enfin $\epsilon^{\mu\nu\rho\sigma} \epsilon_{\mu\nu\rho\sigma} = -24$ (une somme de 256 termes qui nous fait apprécier à sa juste valeur l'élégance des notations d'Einstein).

2.4 Dérivation et analyse vectorielle

La dernière étape à franchir est de refonder, dans notre formalisme quadridimensionnel, l'analyse vectorielle. Nous commencerons par redéfinir le gradient, pour généraliser ensuite aux autres opérateurs

différentiels. Nous verrons que des lois bien connues de l'électromagnétisme prennent une forme très simple en termes de ces opérateurs. Nous définirons enfin une intégration dans l'espace-temps et généraliserons les théorèmes de Stokes/Ostrogradski.

2.4.1 Dérivation

Nous pouvons définir, pour une fonction du 4-vecteur position d'un événement, la dérivation par rapport aux coordonnées contravariantes de l'événement:

$$\partial_\mu = \frac{\partial}{\partial x^\mu} . \quad (2.76)$$

Il est évident que les 4 quantités ∂_μ forment les quatre composantes covariantes d'un opérateur différentiel vectoriel qui généralise la notion de "nabla" à notre espace à quatre dimensions. Leur covariance est manifeste si nous examinons l'accroissement infinitésimal d'une fonction scalaire des x^μ pour un accroissement dx^μ :

$$df = \partial_\mu f(x^\nu) dx^\mu . \quad (2.77)$$

df étant un scalaire et dx^μ un vecteur contravariant, ∂_μ est un "vecteur" covariant. Il se transforme donc comme tel dans une transformation de Lorentz:

$$\partial'_\mu = \mathcal{L}_\mu{}^\nu \partial_\nu , \quad (2.78)$$

où les ∂' représentent les dérivées par rapport aux nouvelles coordonnées contravariantes. Cette loi de transformation peut aussi s'établir péniblement à partir des lois de transformation des composantes.

On peut aussi définir la dérivation par rapport aux coordonnées covariantes:

$$\partial^\mu = \frac{\partial}{\partial x_\mu} . \quad (2.79)$$

Ces opérateurs différentiels forment évidemment les composantes contravariantes d'un vecteur (comme on peut s'en convaincre en écrivant l'accroissement d'une fonction scalaire des coordonnées covariantes). On a de plus

$$\partial^\mu = g^{\mu\nu} \partial_\nu , \quad (2.80)$$

généralisation aux opérateurs différentiels des lois d'élévation ou d'abaissement des indices.

2.4.2 Analyse vectorielle

On peut définir à partir des ∂_μ des analogues des opérateurs de l'analyse vectorielle à trois dimensions. Si f est une fonction scalaire, $\partial_\mu f$, généralisant le gradient, peut s'écrire

$$\partial_\mu f = \left(\frac{1}{c} \frac{\partial f}{\partial t}, \nabla f \right) \quad (2.81)$$

et

$$\partial^\mu f = \left(\frac{1}{c} \frac{\partial f}{\partial t}, -\nabla f \right) \quad (2.82)$$

Si on considère un champ de 4-vecteurs $A^\mu(x^\nu) = (a^0, \mathbf{a})$, on peut définir sa divergence comme:

$$\partial^\mu A_\mu = \partial_\mu A^\mu = \frac{1}{c} \frac{\partial a^0}{\partial t} + \nabla \cdot \mathbf{a} , \quad (2.83)$$

qui n'est pas sans nous rappeler l'équation de conservation de la charge ou la jauge de Lorentz. L'analogie du rotationnel sera le tenseur de rang 2 complètement antisymétrique:

$$\partial^\mu A^\nu - \partial^\nu A^\mu . \quad (2.84)$$

On pourra aussi écrire le rotationnel sous sa forme complètement covariante:

$$\partial_\mu A_\nu - \partial_\nu A_\mu . \quad (2.85)$$

Le “laplacien” de l’espace temps est la norme du vecteur ∂^μ :

$$\partial_\mu \partial^\mu = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta = \square \quad (2.86)$$

et n’est pas autre chose que le d’alembertien \square .

Nous pouvons ainsi réécrire de façon extrêmement simple quelques lois de l’électromagnétisme classique. L’équation de conservation de la charge s’écrit simplement $\partial_\mu J^\mu = 0$, en utilisant le 4–vecteur courant.

Les équations de Poisson pour les potentiels peuvent se regrouper en $\partial_\mu \partial^\mu \mathcal{A}^\nu = \mu_0 J^\nu$, à condition de regrouper les potentiels scalaire et vecteur en un 4–vecteur $\mathcal{A}^\mu = (V/c, \mathbf{A})$. L’équation de Poisson prouve immédiatement que cette quantité est un 4–vecteur. On pourrait alors en déduire les lois relativistes de transformation des potentiels dans un changement de référentiel et, par dérivation, les lois de transformation des champs. Nous établirons plus simplement ces résultats dans les prochains chapitres.

La condition de jauge de Lorentz s’écrit tout simplement $\partial_\mu \mathcal{A}^\mu = 0$. L’ensemble de ces relations prouve que l’électromagnétisme se coule de manière très naturelle dans le cadre mathématique de la cinématique relativiste.

2.4.3 Intégration

On peut définir une intégrale de volume dans l’espace–temps pour n’importe quel type de quantité par

$$\int d\Omega , \quad (2.87)$$

où $d\Omega = c dt dx dy dz$ est l’élément d’intégration dans l’espace temps dont nous avons déjà démontré le caractère scalaire.

Une surface dans l’espace à trois dimensions est une variété à trois dimensions. On peut définir une intégrale sur ces surfaces (un flux) à condition de définir un 4–vecteur élément de surface dS^μ .

Un élément de surface est un petit objet à trois dimensions. On peut le considérer comme sous-tendu par trois 4–vecteurs dx^μ , dy^μ et dz^μ (comme un petit élément de surface à deux dimensions est sous-tendu par deux vecteurs infinitésimaux). dS^μ doit être orthogonal à tout vecteur de l’élément et sa longueur doit être une mesure du “volume” de l’élément de surface. On peut obtenir dS^μ par une procédure d’orthogonalisation standard. On forme d’abord le tenseur de rang 3 $dS^{\mu\nu\rho}$ tel que:

$$dS^{\mu\nu\rho} = \begin{vmatrix} dx^\mu & dy^\mu & dz^\mu \\ dx^\nu & dy^\nu & dz^\nu \\ dx^\rho & dy^\rho & dz^\rho \end{vmatrix} \quad (2.88)$$

L’élément de surface cherché s’obtient alors en contractant ce tenseur avec le tenseur de rang 4 complètement antisymétrique:

$$dS^\sigma = -\frac{1}{6} \epsilon^{\sigma\mu\nu\rho} dS_{\mu\nu\rho} , \quad (2.89)$$

dont on montre qu’il a toutes les propriétés requises. Le lecteur sceptique aura avantage à écrire explicitement le vecteur élément de surface pour un petit élément de volume purement spatial (qui est bien un élément de surface dans l’espace–temps). Il trouvera un 4–vecteur dont seule la composante temporelle est non nulle. Sa mesure est précisément le volume spatial de l’élément considéré.

On peut établir pour les intégrales de surface un théorème généralisant le théorème de gauss:

$$\int_S A^\mu dS_\mu = \int_V \partial_\mu A^\mu d\Omega , \quad (2.90)$$

où V est un volume dans l'espace-temps et S sa surface frontière. L'intégrale de la divergence étendue à tout l'espace est donc égale au flux sur la "sphère de l'infini"³. Celui-ci est en général nul pour des champs physiques.

On peut aussi définir une intégration sur des variétés à deux dimensions (que nous n'appellerons pas surfaces). L'élément d'intégration est un tenseur antisymétrique de rang 2 formé sur les vecteurs dx^μ et dy^ν sous-tendant l'élément d'intégration:

$$df^{\mu\nu} = dx^\mu dy^\nu - dx^\nu dy^\mu . \quad (2.91)$$

On peut enfin définir une intégrale curviligne sur une ligne d'univers. Le théorème de Stokes relie l'intégrale sur une variété à deux dimensions à l'intégrale sur son contour:

$$\int A_\mu dx^\mu = \int df^{\mu\nu} (\partial_\mu A_\nu - \partial_\nu A_\mu) . \quad (2.92)$$

³Cette sphère de l'infini est assez peu intuitive. Elle est formée de tous les points de l'espace à $t = -\infty$ et de tout l'espace à nouveau à $t = \infty$. Entre les deux, elle est formée à tout instant de tous les points à l'infini dans l'espace (la "sphère de l'infini" habituelle de l'analyse vectorielle).

Chapitre 3

Dynamique relativiste

Après avoir jeté les bases d'une nouvelle cinématique, il nous faut bien sûr établir la nouvelle dynamique. Nous commencerons par postuler une forme très simple pour l'action d'une particule libre, dont nous déduirons les équations de Lagrange (qui donnent trivialement un mouvement rectiligne uniforme). Nous en déduirons, de manière plus intéressante, l'expression relativiste de l'impulsion de la particule. Nous définirons alors la notion de force pour une particule en interaction. Nous n'irons pas beaucoup plus loin dans ce chapitre. Pour utiliser la notion de force, il faut en effet la relier aux causes du mouvement. C'est ce que nous ferons au prochain chapitre dans le cas de l'électromagnétisme, en "redécouvrant" l'expression de la force de Lorentz. Notons à ce point que toutes les forces phénoménologiques utilisées en mécanique classique (frottements, tensions..) n'ont aucun sens en relativité (elles ne peuvent se transformer correctement dans un changement de référentiel). Notons également que la force de gravitation ne peut être traitée correctement que dans le cadre de la relativité générale. Nous concluons ce chapitre par quelques brèves remarques sur le vaste problème des collisions de particules relativistes.

3.1 Particule Libre

Nous essayons ici de définir une action ou un lagrangien pour une particule relativiste libre. La condition d'extrémalité de l'action devrait alors nous donner la ligne d'univers suivie par cette particule.

Pour que toutes les quantités que nous manipulons se comportent bien dans un changement de référentiel (on dit souvent qu'elles sont manifestement covariantes), il faut que l'action et le lagrangien soient tous les deux des 4-scalaires. Le temps sur lequel on intègre le lagrangien doit donc être lui aussi un 4-scalaire. De manière évidente, seul le temps propre τ de la particule convient. Enfin, les bornes de l'intégration doivent être deux événements bornant la portion de ligne d'univers cherchée (et jouant le rôle de conditions aux limites dans les équations de Lagrange). On écrira donc:

$$S = \int_a^b L d\tau , \quad (3.1)$$

ou encore:

$$S = \frac{1}{c} \int_a^b L ds , \quad (3.2)$$

où ds est un intervalle infinitésimal sur la ligne d'univers: $ds = cd\tau$.

Pour L , le choix le plus simple est une constante $-\alpha$, relative à la particule. Nous choisissons un signe moins et une constante positive. Ce choix conduira à un minimum pour l'action¹. On peut

¹On peut rendre l'intégrale du temps propre aussi petite que l'on veut en imaginant des mouvements à de très grandes vitesses entre a et b . En revanche, l'intégrale du temps propre est maximale si la particule se déplace à vitesse constante entre a et b . L'action, pour avoir un minimum, doit donc être opposée à l'intégrale du temps propre.

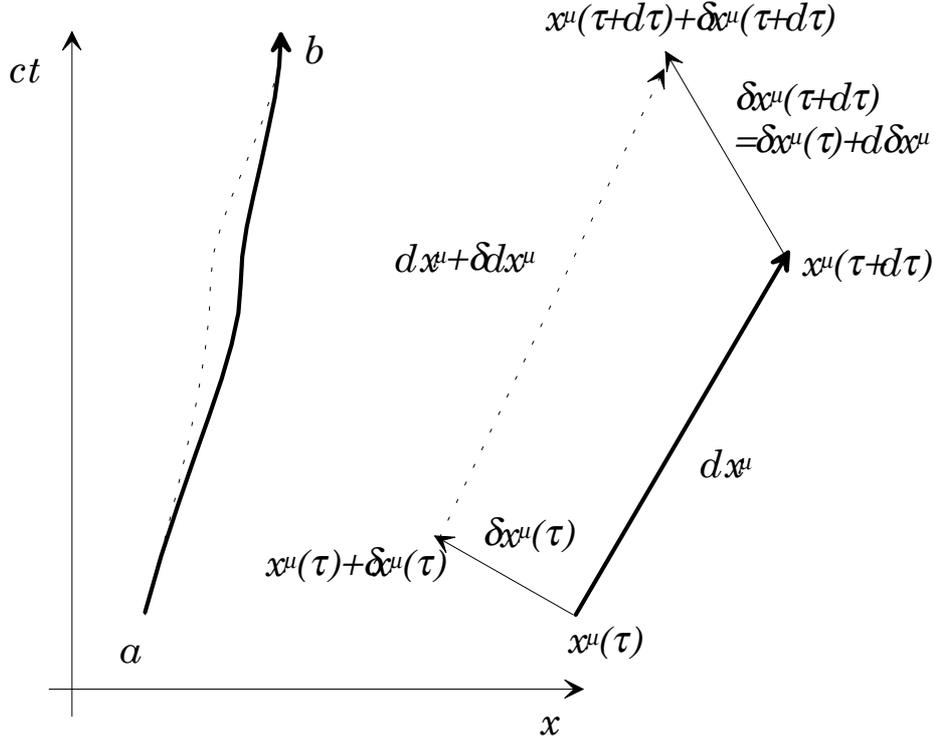


Figure 3.1: Deux lignes d'univers joignant deux événements. L'une est la trajectoire de référence (en traits pleins), l'autre une trajectoire infiniment proche (en pointillés). Les deux trajectoires coïncident pour les événements initial et final. La partie de droite schématise des éléments infinitésimaux correspondants des deux trajectoires.

identifier aisément cette constante en imposant à notre action de coïncider avec l'action classique si la vitesse de la particule est petite devant c . En remarquant que $d\tau = dt/\gamma$, on écrira:

$$S = -\alpha \int \sqrt{1 - v^2/c^2} dt = \int (-\alpha + \alpha v^2/2c^2) dt . \quad (3.3)$$

Le terme constant ne joue aucun rôle dans l'action classique. Pour que le terme en v^2 coïncide avec l'énergie cinétique, il faut que $\alpha = mc^2$. Le lagrangien d'une particule libre est donc une simple constante:

$$L = -mc^2 \quad (3.4)$$

et l'action s'écrit simplement:

$$S = -mc^2 \int d\tau = -mc \int ds . \quad (3.5)$$

Nous allons maintenant établir, à partir de cette expression de l'action, les équations de Lagrange. Ce raisonnement est très proche de celui utilisé dans la première partie de ce cours pour établir les équations de Lagrange d'un système quelconque. Nous considérerons donc deux lignes d'univers légèrement différentes entre les événements a et b qui nous serviront de "conditions aux limites" (voir figure 3.1). Nous ne pouvons plus en effet spécifier, sans perdre l'invariance relativiste, de positions à un instant initial et à un instant final. L'une de ces trajectoires sera la trajectoire effectivement suivie, qui réalise un extremum pour l'action. Elle est paramétrée par le temps propre τ de la particule. Les événements de cette ligne d'univers peuvent donc s'écrire $x^\mu(\tau)$.

L'autre ligne d'univers s'écarte de la trajectoire de référence par des quantités infinitésimales. Les coordonnées spatio-temporelles de chaque événement sur cette ligne d'univers pourront s'écrire $x^\mu(\tau) + \delta x^\mu(\tau)$, où $\delta x^\mu(\tau)$ est une quantité infinitésimale. Notons qu'avec ce choix d'écriture, la

trajectoire variée et la trajectoire de référence sont paramétrées par le même temps propre τ . C'est en effet une condition importante pour pouvoir écrire les intégrales d'action avec le même élément différentiel. Notons que le temps propre est utilisé comme un simple paramètre. τ n'est pas le temps propre sur la trajectoire variée (c'est précisément à la détermination de celui-ci que nous allons nous consacrer).

Considérons un intervalle infinitésimal de temps propre $d\tau$ (voir partie droite de la figure 3.1). Sur la trajectoire de référence, on passe de l'événement x^μ à $x^\mu + dx^\mu$ avec $dx^\mu = (dx^\mu/d\tau)d\tau$. Sur la ligne d'univers variée, en revanche, on passe de l'événement $x^\mu + \delta x^\mu$ à $x^\mu + \delta x^\mu + dx^\mu + d\delta x^\mu$ avec $d\delta x^\mu = (d\delta x^\mu/d\tau)d\tau = \delta(dx^\mu/d\tau)d\tau = \delta dx^\mu$, l'accroissement de la vitesse étant manifestement la dérivée temporelle de l'accroissement. Dans toute la suite du raisonnement, nous traiterons les accroissements δ comme des infiniments petits d'ordre supérieur par rapport aux éléments d de trajectoire. L'action sur la trajectoire de référence est donc

$$S = -mc \int_a^b ds = -mc \int_a^b \sqrt{dx_\mu dx^\mu} . \quad (3.6)$$

L'action sur la trajectoire variée s'écrit:

$$S + \delta S = -mc \int_a^b \sqrt{(dx_\mu + \delta dx_\mu)(dx^\mu + \delta dx^\mu)} . \quad (3.7)$$

En développant cette dernière expression au premier ordre non nul dans les petits accroissements, on obtient l'accroissement de l'action

$$\begin{aligned} S + \delta S &= -mc \int_a^b \left[1 + \frac{dx_\mu \delta dx^\mu}{ds^2} \right] ds \\ \delta S &= -mc \int_a^b \frac{dx_\mu}{ds} \delta dx^\mu \\ &= -m \int_a^b U_\mu \delta dx^\mu \\ &= , -m \int_a^b U_\mu \frac{d\delta x^\mu}{d\tau} d\tau \end{aligned} \quad (3.8)$$

où nous faisons intervenir la 4-vitesse $U_\mu = cd x_\mu/ds$ de la particule.

Pour obtenir les équations de Lagrange, nous pouvons transformer cette expression par une intégration par parties:

$$\delta S = -m \int_a^b U_\mu \delta dx^\mu = -m [U_\mu \delta x^\mu]_a^b + m \int_a^b \frac{dU_\mu}{d\tau} \delta x^\mu d\tau . \quad (3.9)$$

Le terme tout intégré est manifestement nul puisque les deux lignes d'univers coïncident sur les événements limites. La nullité de δS ne peut être assurée pour tous les accroissements que si:

$$\frac{dU^\mu}{d\tau} = 0 , \quad (3.10)$$

ce qui constitue l'équation de Lagrange, nécessairement triviale, du mouvement de la particule libre.

En fait, le principal intérêt de ce calcul ne réside pas dans cette équation que nous aurions pu obtenir en invoquant l'invariance galiléenne. Nous avons d'abord pris la mesure des difficultés mises en jeu dans un calcul variationnel en relativité (en particulier, nous avons du prendre garde de ne manipuler que des quantités 4-vectorielles ou scalaires). Ensuite, nous avons établi, en passant, l'équation (3.9) qui va nous permettre de déterminer les variations de l'action dans une variation du point d'arrivée, et donc d'établir la forme relativiste de l'impulsion.

3.2 Energie–impulsion

Nous avons introduit naturellement l'énergie et l'impulsion en dynamique classique comme les dérivées temporelles et spatiales de l'action par rapport au point d'arrivée. La contrepartie relativiste en est manifestement le 4–gradient de l'action par rapport à l'événement b . Nous poserons donc :

$$P_\mu = -\partial_\mu S , \quad (3.11)$$

les dérivées s'entendant par rapport à la position de b .

Pour estimer P_μ , nous considérerons deux lignes d'univers effectivement suivies par la particule, l'une (ligne d'univers de référence) connectant a et b , l'autre (variée) a et $b+db$. Le calcul de la variation de l'action suit alors celui effectué au paragraphe précédent. Les deux trajectoires sont paramétrées par le temps propre de la trajectoire de référence. La seule différence est que l'accroissement ne s'annule pas pour l'événement b : $\delta x^\mu(b) = db$. L'équation (3.9) est donc encore correcte. Comme les deux trajectoires sont des trajectoires effectivement suivies, elles correspondent à des vitesses constantes et le terme intégral est identiquement nul. En revanche, le terme tout intégré est non nul et vaut $U_\mu \delta x^\mu(b)$. On en déduit par simple comparaison que :

$$P_\mu = mU_\mu \quad P^\mu = mU^\mu , \quad (3.12)$$

un résultat pour le moins attendu.

Penchons nous maintenant sur la signification physique des composantes de P^μ . On a bien sûr $P^\mu = -(1/c)dS/dt, \nabla S$. Une simple comparaison avec la mécanique classique nous indique que la composante temporelle de P^μ est l'énergie mécanique \mathcal{E} de la particule, alors que les composantes spatiales représentent la quantité de mouvement :

$$P^\mu = (\mathcal{E}/c, \mathbf{p}) . \quad (3.13)$$

L'énergie et la quantité de mouvement apparaissent donc comme les composantes temporelles et spatiales d'un 4–vecteur. Elles doivent donc se transformer ensemble dans un changement de référentiel. Nous pouvons de plus identifier ces quantités à partir de l'expression de la 4–vitesse :

$$\mathcal{E} = m\gamma c^2 \quad (3.14)$$

$$\mathbf{p} = m\gamma \mathbf{v} \quad (3.15)$$

Commentons tout d'abord l'expression de l'énergie mécanique. Pour des mouvements à vitesse faible devant c , on peut développer le facteur γ . On obtient alors :

$$\mathcal{E} = mc^2 + \frac{1}{2}mv^2 . \quad (3.16)$$

A une constante près, nous retrouvons donc bien l'énergie cinétique galiléenne.

Si les énergies sont toujours définies à une constante additive près en mécanique classique, il n'en est pas de même en relativité. On ne saurait en effet ajouter une constante arbitraire à la composante temporelle d'un 4–vecteur en conservant une forme correcte pour les lois de changement de référentiel. La constante additive qui, ci-dessus, représente l'énergie mécanique d'une particule au repos :

$$\mathcal{E}_0 = mc^2 , \quad (3.17)$$

doit donc avoir une signification physique².

Si une masse au repos possède de l'énergie, cela indique qu'on peut, au moins sur un plan de pur bilan d'énergie, transformer une certaine quantité de masse en énergie ou une certaine quantité

²Remarquons que nous venons d'établir ce qui est sans doute la formule de physique la plus célèbre.

d'énergie en masse. Les applications les plus spectaculaires de ces conversions sont les créations et annihilations de particules élémentaires. Quand deux antiparticules s'annihilent en émettant de l'énergie sous forme de rayonnement électromagnétique, on a une conversion totale d'une certaine quantité de masse en énergie. Le phénomène réciproque, la création d'une paire particule/antiparticule à partir d'un rayonnement suffisamment énergétique, correspond à la conversion d'énergie en masse. Pour créer, par exemple, une paire électron/positron, il faudra au moins fournir une énergie $2mc^2$, soit 2 fois 511 keV. Cette notion de seuil de réaction joue un rôle essentiel pour l'analyse des collisions de particules. Notons à ce point que, si nous avons montré que de tels processus sont envisageables du point de vue du simple bilan énergétique, nous n'avons pas montré qu'ils pouvaient se produire. Il faut ajouter à ces conditions de bilan les règles de sélection qui indiquent quelles particules on peut effectivement créer dans une situation donnée. Ces règles ne peuvent s'obtenir que par une approche quantique qui est hors de propos ici. En dépit des conséquences très importantes de cette énergie de repos pour la physique des particules, nous n'évoquerons pas ce problème beaucoup plus avant ici, et nous renverrons le lecteur intéressé aux nombreux manuels qui couvrent ce sujet.

La quantité de mouvement, de son côté, s'écrit donc $m\gamma\mathbf{v}$. Au premier ordre en v/c , elle coïncide donc bien avec la quantité de mouvement galiléenne ordinaire. En revanche, pour des vitesses proches de celles de la lumière, la quantité de mouvement croît beaucoup plus vite que la vitesse (elle est en fait simplement proportionnelle à la célérité). Comme l'accroissement de la quantité de mouvement est manifestement relié à la force qui s'exerce sur la particule (nous anticipons quelque peu sur la suite de cet exposé), on retrouve qu'il est impossible d'accélérer une particule matérielle jusqu'à la vitesse de la lumière, puisqu'il faudrait lui communiquer une quantité de mouvement infinie. Notons enfin que, du point de vue de la quantité de mouvement, et dans une large mesure du point de vue de l'ensemble de la dynamique, "tout se passe comme si" la masse de la particule dépendait de la vitesse comme le facteur γ . Cette approche, largement développée dans des ouvrages de vulgarisation, n'est bien sûr pas correcte. L'invariance relativiste impose que la masse de la particule soit un 4-scalaire, indépendant du référentiel.

Notons, pour terminer ce paragraphe, un lien utile entre énergie et quantité de mouvement. Nous savons en effet que le module de la 4-vitesse est égal à c . On en déduit $P_\mu P^\mu = m^2 c^2$ et donc

$$p^2 c^2 + m^2 c^4 = \mathcal{E}^2 . \quad (3.18)$$

Examinons le cas d'une particule ultra-relativiste, dont la vitesse est proche de celle de la lumière. L'énergie \mathcal{E} est alors très supérieure à l'énergie de repos \mathcal{E}_0 , et on peut négliger le terme $m^2 c^4$ dans l'expression précédente. On a alors simplement $\mathcal{E} = pc$ et on peut pratiquement, à une constante dimensionnelle près, confondre l'énergie et l'impulsion de la particule.

Un exemple extrême de particule ultra-relativiste est une particule de masse nulle, comme le photon ou le neutrino (si la masse du premier est très certainement nulle, un petit doute théorique et expérimental subsiste pour le second). Si la vitesse de propagation de cette particule était inférieure à c , le facteur γ étant fini, l'énergie de la particule serait nulle. Pour que de telles particules aient une énergie non nulle, il faut que leur vitesse de propagation soit la vitesse limite de la relativité (qu'on peut donc confondre à ce point seulement, en toute rigueur, avec la vitesse de la lumière). Par un simple passage à la limite, on voit donc que l'énergie et l'impulsion de la particule sont liées par $\mathcal{E} = pc$. Dans le cas particulièrement important du photon, l'énergie est $\mathcal{E} = h\nu$, où h est la constante de Planck et l'impulsion est donc, en module, $h\nu/c$. Si on y ajoute que l'impulsion doit avoir la direction du vecteur d'onde (nous verrons pourquoi à la fin de cette partie), on en déduit que l'impulsion spatiale du photon peut s'écrire $\mathbf{p} = \hbar\mathbf{k}$.

3.3 Particule soumise à une force

Cette section ne peut être à ce niveau qu'embryonnaire. Si la particule est soumise à des interactions qui modifient son mouvement, on pourra écrire l'équation du mouvement sous la forme d'une

équation de Lagrange à condition de connaître le lagrangien décrivant l'interaction. Le seul cas où nous pourrions effectuer cette démarche est celui de l'électromagnétisme que nous traiterons dans le prochain chapitre. En fait, aucune autre interaction ne peut être incluse correctement dans le cadre de la relativité restreinte à notre niveau. Tout d'abord, toutes les interactions phénoménologiques (contacts, ressorts..) n'ont plus de sens en relativité, ainsi que la notion de solide. Elles contiennent en effet toutes plus ou moins la notion d'interaction instantanée. Il est possible de formuler une hydrodynamique relativiste, par exemple, mais au prix de grands efforts. Même la gravitation ne peut être en toute rigueur coulée dans ce cadre. Elle n'est décrite convenablement que dans celui de la relativité générale.

En fait, nous ne pouvons ici qu'intuiter que la dérivée par rapport au temps propre de l'impulsion, $G^\mu = dP^\mu/d\tau$, jouera le rôle d'une force, qu'on pourra exprimer simplement, dans le cas de l'électromagnétisme, en fonction de la vitesse de la particule et des champs (cette force ne sera autre que la force de Lorentz). Le module de l'impulsion étant constant, la force est nécessairement perpendiculaire à l'impulsion: $G_\mu P^\mu = 0$. En termes de l'énergie et de la quantité de mouvement spatiale, on pourra écrire:

$$G^\mu = \left(\frac{\gamma}{c} \frac{d\mathcal{E}}{dt}, \gamma \frac{d\mathbf{p}}{dt} \right). \quad (3.19)$$

On pourra poser $\mathbf{f} = d\mathbf{p}/dt$ et définir ainsi la force spatiale, dérivée par rapport au temps ordinaire de la quantité de mouvement. L'orthogonalité de G_μ et de P^μ est simplement équivalente, comme on s'en convaincra aisément, au fait que $\mathcal{E}^2 - c^2 p^2$ est une constante.

3.4 Conservation de l'énergie–impulsion. Application aux collisions

On peut montrer, comme en mécanique classique, que l'invariance dans une translation globale dans l'espace–temps implique que tout système isolé possède un 4–vecteur énergie impulsion qui est conservé. Pour un ensemble de particules matérielles sans interaction, l'impulsion globale du système est simplement la somme des impulsions individuelles. Pour des particules en interaction (en particulier électromagnétique), la situation est moins simple. L'interaction, qui ne peut être instantanée, doit en effet être véhiculée par un champ. Ce champ, possédant une énergie, doit aussi posséder une quantité de mouvement et entrer dans le bilan de l'impulsion globale. L'analyse détaillée de la conservation de l'impulsion dans une telle situation est alors difficile.

La conservation de l'impulsion permet néanmoins d'obtenir des renseignements précis sur les collisions de particules relativistes. Dans une telle collision, on considère en effet un état initial où les deux particules sont très éloignées et n'interagissent pratiquement pas. Après avoir interagi, les deux particules s'éloignent à nouveau l'une de l'autre et on considère un état final où, à nouveau, l'interaction est négligeable. Même si on renonce à examiner ce qui se passe pendant l'interaction, on pourra écrire que l'impulsion globale initiale, qui est la somme des impulsions de particules incidentes, est égale à la somme des impulsions finales. On regroupe ainsi dans une même équation liant deux 4–vecteurs les lois de conservation de l'énergie et de l'impulsion qu'on écrit séparément dans une analyse classique de la collision³.

Pour illustrer très brièvement l'étude des collisions relativistes, nous traiterons deux exemples: le calcul d'un seuil de réaction et le cas de l'effet Compton. Ces deux exemples à eux seuls sont loin de

³Si les particules sont chargées, un champ électromagnétique est associé à chaque particule avant et après la collision. Ce champ doit entrer en toute rigueur dans le bilan d'impulsion. En fait, quand on mesure la masse ou l'impulsion d'une particule chargée, on ne le fait jamais indépendamment du champ qui accompagne cette particule. La masse que nous mettons dans l'impulsion est donc une masse "habillée", qui englobe effectivement la contribution du champ à l'inertie de la particule. Nous n'aurons donc effectivement pas à tenir compte des champs des particules libres dans notre bilan d'impulsion. Notons qu'avec une particule chargée ponctuelle, la différence entre la masse "nue" et la masse "habillée" de la particule est évidemment infinie. Ce n'est qu'au prix d'un calcul complètement quantique mené avec les techniques de renormalisation qu'on peut définir proprement la masse "habillée" de la particule.

couvrir toutes les applications de la conservation de la 4-impulsion à la physique des particules, mais ils permettent d'illustrer de façon simple des techniques et des concepts importants.

3.4.1 Seuil de réaction

La plupart des collisions réalisées dans les accélérateurs visent à produire de nouvelles particules. Il est clair que l'énergie cinétique incidente doit être suffisante pour que la réaction soit énergétiquement possible. Le seuil de réaction est précisément l'énergie cinétique à fournir aux particules en collision pour que la réaction soit énergétiquement possible. Ceci ne signifie pas nécessairement que la réaction se produise effectivement. Des règles de sélection quantiques ou une section efficace insuffisante peuvent faire que les produits ne soient jamais observés.

Le calcul du seuil est complexe pour les réactions complexes. Nous ne considérerons donc que l'une des plus simples des réactions de création de particules: la création d'une paire proton/antiproton par collision de deux protons. L'équation de la réaction s'écrit:



L'énergie minimale pour que cette réaction puisse se produire correspond à une situation où toutes les particules finales seraient au repos. Elle est donc définie dans un référentiel où la quantité de mouvement totale est nulle. Pour tout système de particules matérielles, il existe un tel référentiel, appelé "référentiel du centre de masse" \mathcal{R}_{CM} . Il est l'analogie du référentiel barycentrique en mécanique classique. Dans \mathcal{R}_{CM} , l'énergie minimale est donc $4mc^2$ (les particules et les antiparticules ont même masse m). Dans ce référentiel, les deux protons incidents ont initialement des énergies égales et des quantités de mouvement opposées. L'énergie de chacun doit donc être $2mc^2$ et son énergie cinétique mc^2 . Le seuil de réaction, dans le référentiel barycentrique, est donc de mc^2 pour chaque particule, un résultat assez intuitif.

Le problème est que l'énergie à fournir effectivement est celle définie dans le référentiel du laboratoire, qui ne coïncide pas nécessairement avec le référentiel du centre de masse. Ce n'est que dans le cas des anneaux de collision (LEP pour les collisions électron/électron, par exemple), que les deux projectiles sont de quantités de mouvement opposées et que le référentiel du laboratoire est aussi celui du centre de masse.

Dans beaucoup d'expériences, plus anciennes, un seul des protons est en mouvement (le projectile) et l'autre (la cible) est immobile. On peut estimer le seuil de réaction dans ce cas. La première chose à faire est de déterminer les éléments de \mathcal{R}_{CM} par rapport au référentiel du laboratoire \mathcal{R} . Nous noterons \mathcal{E} et p l'énergie et la quantité de mouvement du projectile. Tous les mouvements s'effectuant sur l'axe projectile/cible, nous n'écrirons que des quantités en projection sur cet axe. Il est clair aussi que la vitesse de \mathcal{R}_{CM} par rapport à \mathcal{R} est selon cet axe. Le 4-vecteur énergie-impulsion total avant la collision s'écrit donc $(\mathcal{E}/c + mc, p)$. De manière évidente la quantité de mouvement totale n'est pas nulle.

Nous cherchons donc une transformation de Lorentz (paramètres β et γ) telle que la nouvelle impulsion p' soit nulle. Il faut donc choisir:

$$\beta = \frac{pc}{\mathcal{E} + mc^2} , \quad (3.21)$$

ce qui définit le mouvement du référentiel du centre de masse. On peut alors écrire l'énergie totale disponible dans \mathcal{R}_{CM} en utilisant la même transformation de Lorentz:

$$\mathcal{E}' = \gamma(\mathcal{E} + mc^2 - \beta pc) . \quad (3.22)$$

En y portant l'expression de β et en notant que $\mathcal{E}^2 - p^2c^2 = m^2c^4$, on a simplement:

$$\mathcal{E}' = 2\gamma mc^2 . \quad (3.23)$$

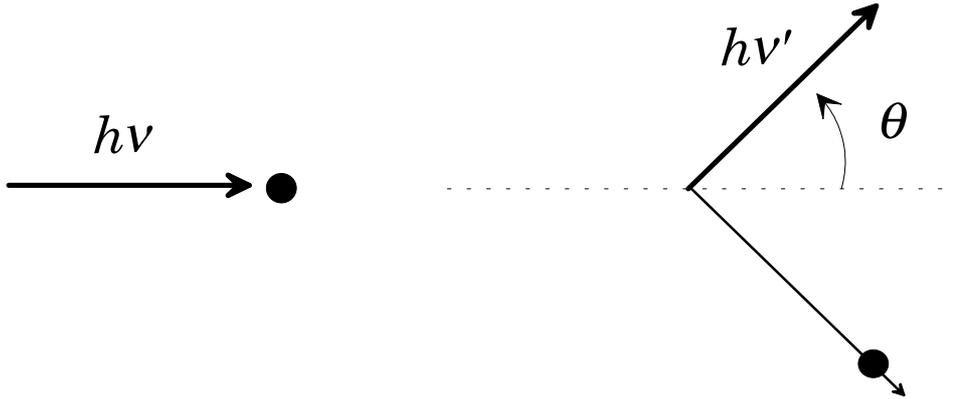


Figure 3.2: Diffusion d'un photon par un électron initialement au repos. Partie gauche: état initial. Partie droite: état final. Le photon est diffusé avec une fréquence modifiée et l'électron est éjecté.

On peut enfin exprimer γ en fonction de β puis de \mathcal{E} et p :

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{\sqrt{\mathcal{E} + mc^2}}{\sqrt{2}\sqrt{mc^2}}. \quad (3.24)$$

L'énergie disponible dans \mathcal{R}_{CM} est donc simplement:

$$\mathcal{E}' = \sqrt{2mc^2(\mathcal{E} + mc^2)}. \quad (3.25)$$

Le seuil de réaction s'obtient alors en écrivant que \mathcal{E}' doit être plus grand que $4mc^2$. Après quelques manipulations, on voit que le seuil s'écrit:

$$\mathcal{E} > 7mc^2, \quad (3.26)$$

l'énergie cinétique du projectile devant être d'au moins $6mc^2$. On doit comparer ce résultat à celui obtenu quand le référentiel du laboratoire est aussi celui du centre de masse. Au lieu d'une énergie cinétique par particule de mc^2 , on a besoin d'une énergie 6 fois plus grande. Avec une puissance donnée d'accélérateur, il est donc, de beaucoup, préférable de travailler avec des collisions entre deux particules en mouvement. Ceci explique que pratiquement toutes les grandes machines sont maintenant des anneaux de collision. Le prix à payer est bien sûr que la densité du faisceau est très petite devant celle d'une cible solide: le taux de collisions est beaucoup moins grand que ce qu'on obtient en envoyant un faisceau unique sur une cible fixe.

3.4.2 Effet Compton

Notre second exemple sera l'effet Compton, la diffusion d'un photon de haute énergie par un électron initialement immobile (ou lié à un atome: les énergie mises en jeu dans les expériences sont telles, comme nous le verrons, que l'énergie de liaison d'un électron dans un atome est négligeable par rapport à l'énergie finale de l'électron diffusé). Les paramètres importants du problème sont représentés sur la figure 3.2.

Sans restreindre la généralité, le photon de fréquence ν et d'énergie $h\nu$ est incident dans la direction de l'axe Ox . Après la collision, l'électron et le photon sont diffusés dans des directions qui dépendent des détails de l'interaction. Toujours sans restreindre la généralité, nous pouvons supposer que les trajectoires finales de l'électron et du photon sont dans le plan Oxy . L'électron emportant de l'énergie cinétique, le photon perd nécessairement de l'énergie. Sa nouvelle fréquence sera notée ν' . Nous chercherons seulement ici à calculer la nouvelle fréquence du photon en fonction de son angle de

diffusion θ . C'est en effet la seule quantité accessible dans les expériences: l'électron est rapidement diffusé et son énergie amortie par le milieu où se produit la diffusion. De façon très remarquable, la simple application de la conservation de l'impulsion relativiste permet d'explicitier ce calcul⁴.

Les 4-impulsions du photon avant et après la collision seront notées \mathbf{P} et \mathbf{P}' . On notera de même \mathbf{Q} et \mathbf{Q}' les impulsions initiale et finale de l'électron. En nous souvenant que la quantité de mouvement spatiale de l'électron est $h\nu/c$, nous pourrions écrire:

$$\mathbf{P} = (h\nu/c, h\nu/c, 0) , \quad (3.27)$$

$$\mathbf{P}' = (h\nu'/c, h\nu'/c \cos\theta, h\nu'/c \sin\theta) \quad (3.28)$$

$$\mathbf{Q} = (mc, 0, 0) \quad (3.29)$$

$$\mathbf{Q}' = (\mathcal{E}/c, q_x, q_y) . \quad (3.30)$$

Dans les équations précédentes, nous n'avons écrit que les composantes spatiales dans le plan Oxy . \mathcal{E} et q_x, q_y représentent l'énergie et la quantité de mouvement de l'électron diffusé. Pour calculer ν' en fonction de θ , nous allons chercher à éliminer ces inconnues. La loi de conservation de l'énergie-impulsion s'écrit:

$$\mathbf{P} + \mathbf{Q} = \mathbf{P}' + \mathbf{Q}' . \quad (3.31)$$

On en déduit immédiatement:

$$(\mathbf{P} - \mathbf{P}')^2 = (\mathbf{Q} - \mathbf{Q}')^2 \quad (3.32)$$

$$\mathbf{P}^2 + \mathbf{P}'^2 - 2\mathbf{P} \cdot \mathbf{P}' = \mathbf{Q}^2 + \mathbf{Q}'^2 - 2\mathbf{Q} \cdot \mathbf{Q}' . \quad (3.33)$$

Le photon étant une particule de masse nulle, $\mathbf{P}^2 = \mathbf{P}'^2 = 0$. De plus $\mathbf{Q}^2 = \mathbf{Q}'^2 = m^2c^2$ où m est la masse de l'électron. Enfin, dans le produit scalaire $\mathbf{Q} \cdot \mathbf{Q}'$, le produit des composantes spatiales est nul et il ne reste que le produit des composantes temporelles: $\mathbf{Q} \cdot \mathbf{Q}' = -2m\mathcal{E}$. On en déduit:

$$\frac{h^2\nu\nu'}{c^2}(1 - \cos\theta) = m(\mathcal{E} - mc^2) . \quad (3.34)$$

La simple conservation de l'énergie assure alors que:

$$h\nu + mc^2 = h\nu' + \mathcal{E} , \quad (3.35)$$

d'où on déduit finalement

$$\lambda' - \lambda = \lambda_c(1 - \cos\theta) , \quad (3.36)$$

$\lambda = c/\nu$ et $\lambda' = c/\nu'$ étant les longueurs d'ondes associées au photon avant et après la collision. La longueur λ_c , appelée longueur d'onde de Compton de l'électron, vaut:

$$\lambda_c = \frac{h}{mc} = 2.10^{-11} \text{ m} . \quad (3.37)$$

La très petite valeur numérique de cette quantité fait que l'effet n'est notable que pour des photons incidents de courte longueur d'onde, c'est à dire de très haute énergie. Pour des photons visibles, par exemple, la modification de fréquence due au "recul" de l'électron est tout à fait négligeable. Quand nous étudierons la diffusion de rayonnement par un atome nous négligerons complètement cet effet.

⁴Notons que la mise en évidence de l'effet Compton et l'accord quantitatif avec la loi que nous allons démontrer a constitué, sans doute, une des premières preuves indiscutables de l'existence du photon. L'effet photoélectrique, s'il est parfaitement explicable en termes de photons, peut en effet être décrit dans une théorie semi-classique couplant un champ électromagnétique classique à un détecteur quantifié.

Chapitre 4

Electromagnétisme relativiste

Nous désirons traiter maintenant de façon relativiste l'interaction entre particules chargées par l'intermédiaire d'un champ électromagnétique. Nous pourrions bien sûr supposer connues les équations de Maxwell et montrer directement qu'elles s'insèrent sans difficultés dans un cadre relativiste. Nous allons en fait procéder d'une façon beaucoup moins directe, mais plus profonde. Nous allons chercher à écrire l'interaction relativiste non triviale entre des particules par l'intermédiaire d'un champ de vecteurs. En d'autres termes, nous allons tenter de construire, dans le cadre de la relativité, la théorie de champ la plus simple qui ne soit pas triviale. Nous postulons pour cela la forme des actions associées aux particules et au champ. Nous verrons que ces formes sont en effet tout à fait naturelles. Nous déduirons alors de ces actions des équations de Lagrange qui décrivent la dynamique des particules couplées au champ et la dynamique du champ couplé aux particules. Nous constaterons enfin que ces équations ont la forme des équations de Maxwell. Nous aurons donc pu déduire les équations de Maxwell d'une approche lagrangienne très générale.

Nous aurons montré également que l'électromagnétisme est la plus simple des théories de champ vectorielles dans le cadre de la relativité. Bien sûr, nous ne nous contenterons pas dans ce chapitre de réécrire des équations bien connues. L'écriture en termes de quantités explicitement covariantes nous fournira des résultats nouveaux, difficiles ou impossibles à établir dans le cadre de l'électromagnétisme classique. Nous établirons ainsi la transformation des champs électromagnétiques dans un changement de référentiel, nous isolerons des invariants scalaires formés à partir des champs, dont nous montrerons qu'ils ont une signification physique importante. Nous pourrions enfin établir des bilans d'énergie-impulsion pour le champ qui nous permettront de jeter une lumière nouvelle sur des phénomènes bien connus, comme la pression de radiation ou même la force de Coulomb électrostatique.

Notre système sera donc constitué d'un ensemble de particules en interaction avec un champ représenté par un champ de 4-vecteurs. L'action totale pour ce système peut a priori se décomposer sous la forme:

$$S = S_{\text{Particules Libres}} + S_{\text{Champ Libre}} + S_{\text{Interaction}} , \quad (4.1)$$

où $S_{\text{Particules Libres}}$ représente l'action des particules en l'absence de champ (une simple collection de particules libres si on néglige toute autre forme d'interaction), $S_{\text{Champ Libre}}$ représente l'action décrivant le champ seul, en l'absence de toutes particules. Enfin, $S_{\text{Interaction}}$ représente l'interaction entre particules et champ: d'une part les particules sont la source du champ, d'autre part la présence du champ modifie la trajectoire des particules.

Pour aborder ce problème, nous allons procéder en deux temps, comme on le fait souvent dans les exposés élémentaires d'électromagnétisme. Nous allons d'abord considérer une particule unique en présence d'un champ imposé. Nous supposons donc qu'un grand ensemble de particules crée un champ qui agit sur une particule test. Si cette particule "test" est suffisamment petite, elle ne modifie pas notablement le champ ni la dynamique des particules "sources". Les seules variables dynamiques dans ce cas sont donc celles de la particule libre, les valeurs du champ étant des quantités imposées.

Nous pourrions alors écrire aisément les équations de Lagrange, dont nous verrons qu'elles redonnent l'expression attendue pour la force de Lorentz. Nous verrons en effet que le seul champ de vecteurs introduit (qui n'est autre que la version relativiste du potentiel vecteur) intervient dans la force sous la forme de son rotationnel, c'est à dire d'un tenseur de rang deux antisymétrique. Ce tenseur pouvant être exprimé à partir de deux champs de vecteurs, nous retrouverons que l'électromagnétisme est une théorie à deux champs. En écrivant les propriétés du rotationnel relativiste, nous obtiendrons des relations entre ces deux champs qui se trouveront coïncider avec le groupe des équations de Maxwell homogènes. Notons enfin que nous pourrions dès ce point établir la forme des changements de référentiels pour les champs.

Dans un deuxième temps, nous considérerons un champ en interaction avec des particules dont la dynamique est imposée. Nous représenterons les mouvements de ces particules par des densités de charges et de courants et nous postulerons une forme simple pour $S_{\text{Champ Libre}}$. Nous obtiendrons alors, comme équations de Lagrange, les équations de Maxwell faisant intervenir les sources. Nous aurons alors terminé notre programme: en combinant les équations décrivant la dynamique des particules dans un champ imposé et les équations décrivant la dynamique du champ sous l'action de courants imposés, on peut, au moins en principe, résoudre tout problème d'électromagnétisme. La dernière partie de ce chapitre sera alors consacrée à l'exploitation de ces résultats. Nous y établirons en particulier les bilans d'énergie-impulsion pour le champ.

4.1 Particule libre dans un champ imposé

4.1.1 Equations de Lagrange

Nous considérons donc ici une particule de masse m plongée dans un champ imposé. L'action décrivant la particule libre s'écrit simplement

$$S_{\text{Particules Libres}} = -mc \int_a^b ds , \quad (4.2)$$

où a et b sont deux événements décrivant les conditions aux limites imposées à la particule. En l'absence de champ, la ligne d'univers de la particule serait simplement la droite joignant a et b .

Nous postulerons que le champ peut être représenté par un champ unique de 4-vecteurs que nous noterons $\mathcal{A}^\mu(x^\nu) = (V/c, \mathbf{A})$. Pour des raisons qui apparaîtront évidentes plus tard, nous nommerons le champ \mathcal{A} "potentiel". L'interaction entre le champ et la particule doit être représentée par l'intégrale d'une quantité scalaire sur la ligne d'univers entre les événements limites a et b . La quantité la plus simple non triviale que nous puissions former est donc:

$$S_{\text{Interaction}} = -q \int_a^b \mathcal{A}_\mu dx^\mu , \quad (4.3)$$

où q est une quantité scalaire représentant l'intensité du couplage de la particule au champ que nous nommerons simplement "charge". Dans l'expression de l'action d'interaction, il faut comprendre, comme au chapitre précédent, que la ligne d'univers de la particule est paramétrée par son temps propre τ et que dx^μ est en fait égal à $(dx^\mu/d\tau)d\tau$.

Nous pouvons tout de suite nous rassurer sur la pertinence de ce lagrangien d'interaction. En écrivant que $(dx^\mu/d\tau)d\tau = U^\mu d\tau = U^\mu dt/\gamma$, en particulierisant pour un instant un référentiel \mathcal{R} et en développant le produit scalaire, on met l'action d'interaction sous la forme:

$$S_{\text{Interaction}} = -q \int_a^b (V - \mathbf{v} \cdot \mathbf{A}) dt , \quad (4.4)$$

où \mathbf{v} est la vitesse tridimensionnelle de la particule dans \mathcal{R} . On retrouve bien là la forme du lagrangien d'interaction avec une particule chargée obtenu dans le premier chapitre de la première partie de ce

cours, à condition bien sûr d'assimiler la composante temporelle du potentiel au potentiel scalaire de l'électromagnétisme et ses composantes spatiales au potentiel vecteur. C'est à la justification détaillée de cette assimilation que nous allons procéder maintenant.

Comme dans le chapitre précédent (on se reportera en particulier à la figure 3.1), on considère entre a et b la trajectoire effectivement suivie (trajectoire de référence) et une trajectoire variée de façon infinitésimale. Les deux trajectoires sont paramétrées par le temps propre de la trajectoire de référence et on se reportera au chapitre précédent pour une description des éléments différentiels importants.

La variation de l'action de la particule libre s'obtient comme au chapitre précédent. En écrivant la variation de l'intervalle élémentaire et en se livrant à la traditionnelle intégration par parties, on trouve:

$$\delta S_{\text{Particule Libre}} = \int_a^b \frac{dP^\mu}{d\tau} \delta x^\mu d\tau . \quad (4.5)$$

Calculons maintenant la variation de l'action d'interaction:

$$\begin{aligned} \delta S_{\text{Interaction}} &= -q\delta \int_a^b \mathcal{A}_\mu dx^\mu \\ &= -q \int_a^b [\delta \mathcal{A}_\mu] dx^\mu - q \int_a^b \mathcal{A}_\mu \delta dx^\mu , \end{aligned} \quad (4.6)$$

où le premier terme représente la variation due au fait que la trajectoire variée n'échantillonne pas le potentiel aux mêmes points que la trajectoire de référence. Le deuxième terme représente pour sa part la variation de l'action due à la modification de géométrie de la trajectoire à potentiel constant.

Traisons d'abord le deuxième terme. On a en fait

$$\begin{aligned} -q \int_a^b \mathcal{A}_\mu \delta dx^\mu &= -q \int_a^b \mathcal{A}_\mu \frac{d\delta x^\mu}{d\tau} d\tau \\ &= [-q \mathcal{A}_\mu \delta x^\mu]_a^b + q \int_a^b \left(\frac{d\mathcal{A}_\mu}{d\tau} d\tau \right) \delta x^\mu . \end{aligned} \quad (4.7)$$

Le terme tout intégré de l'intégration par parties est identiquement nul, puisque les trajectoires coïncident aux extrémités. La parenthèse dans l'intégrale restante représente l'accroissement $d\mathcal{A}_\mu$ du potentiel quand on passe d'une extrémité à l'autre d'un élément différentiel de la trajectoire de référence. Le potentiel étant une fonction de l'événement auquel il est estimé, nous pourrions écrire:

$$d\mathcal{A}_\mu = \partial_\nu \mathcal{A}_\mu dx^\nu . \quad (4.8)$$

Le deuxième terme s'écrit donc:

$$q \int_a^b \partial_\nu \mathcal{A}_\mu dx^\nu \delta x^\mu . \quad (4.9)$$

Revenons maintenant au premier terme. La variation $\delta \mathcal{A}_\mu$ du potentiel quand on passe de la trajectoire de référence à la trajectoire variée s'écrit $\delta \mathcal{A}_\mu = \partial_\nu \mathcal{A}_\mu \delta x^\nu$. Le terme à intégrer, après une permutation sans conséquences des indices muets, s'écrit donc $\partial_\mu \mathcal{A}_\nu \delta x^\mu dx^\nu$. La variation de l'action d'interaction peut donc finalement s'écrire:

$$\delta S_{\text{Interaction}} = q \int_a^b [\partial_\nu \mathcal{A}_\mu - \partial_\mu \mathcal{A}_\nu] \delta x^\mu dx^\nu . \quad (4.10)$$

En remarquant finalement que, sur la trajectoire de référence, $dx^\nu = (dx^\nu/d\tau)d\tau = U^\nu d\tau$, où U^ν est la 4-vitesse, on peut écrire la variation totale de l'action sous la forme:

$$\delta S = \int_a^b \left[\frac{dP^\mu}{d\tau} - G_\mu \right] \delta x^\mu d\tau , \quad (4.11)$$

avec

$$G_\mu = q [\partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu] U^\nu . \quad (4.12)$$

Les équations de Lagrange s'obtiennent alors immédiatement. La variation de l'action ne peut s'annuler au premier ordre dans les écarts entre les trajectoires que si tous les coefficients des δx^μ sont identiquement nuls. Les équations de mouvement s'écrivent donc:

$$\frac{dP_\mu}{d\tau} = G_\mu , \quad (4.13)$$

la quantité G_μ n'étant autre que la 4-force que nous avons pressentie au chapitre précédent. Notons que cette équation 4-vectorielle contient aussi bien la variation de l'énergie de la particule que celle de sa quantité de mouvement.

La force s'exprime en fonction du rotationnel du potentiel, qui est un tenseur de rang deux, antisymétrique, écrit ici sous sa forme complètement covariante. Nous appellerons tenseur champ ce rotationnel et nous poserons:

$$F_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu . \quad (4.14)$$

La force à laquelle est soumise la particule s'écrit alors simplement:

$$G_\mu = q F_{\mu\nu} U^\nu \quad (4.15)$$

et n'est pas autre chose que la contraction du tenseur champ avec la vitesse de la particule.

4.1.2 Tenseur champ électromagnétique

$F_{\mu\nu}$ est par définition un tenseur antisymétrique de rang 2, le 4-rotationnel du potentiel ($V/c, \mathbf{A}$). Il ne dépend donc que de six coordonnées indépendantes. Les trois coordonnées spatio-temporelles forment les composantes d'un vecteur spatial, alors que les trois coordonnées purement spatiales forment les composantes d'un pseudo-vecteur.

On peut écrire les composantes spatio-temporelles sous la forme:

$$\begin{aligned} F_{0i} &= \partial_0 \mathcal{A}_i - \partial_i \mathcal{A}_0 \\ &= \frac{1}{c} \left(-\frac{\partial \mathcal{A}_i}{\partial t} - \partial_i V \right) \\ &= \frac{E_i}{c} , \end{aligned} \quad (4.16)$$

en posant

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t} . \quad (4.17)$$

Nous nommerons bien sûr "champ électrique" le vrai vecteur spatial ainsi défini.

Nous pouvons de même mettre les composantes purement spatiales du tenseur champ sous la forme:

$$F_{12} = -B_z \quad (4.18)$$

$$F_{13} = B_y \quad (4.19)$$

$$F_{23} = -B_x , \quad (4.20)$$

en introduisant le pseudo-vecteur "champ magnétique"

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (4.21)$$

De manière toute naturelle, notre théorie de champ décrite par un 4-vecteur potentiel s'exprime en fonction de deux champs et ressemble de plus en plus à l'électromagnétisme de Maxwell.

On peut écrire F sous forme matricielle:

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix}, \quad (4.22)$$

ou encore, sous forme complètement contravariante:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}. \quad (4.23)$$

L'invariance de jauge est contenue dans la définition même du tenseur champ électromagnétique en fonction du potentiel. L'équation du mouvement de la particule (qui a seule un sens physique non ambigu) est inchangée si nous ajoutons au 4-potential un 4-gradient arbitraire (il est aisé de vérifier que le 4-rotationnel d'un champ de 4-gradient est identiquement nul). La transformation "de jauge"

$$\mathcal{A}_\mu \longrightarrow \mathcal{A}_\mu + \partial_\mu \Phi, \quad (4.24)$$

où Φ est un champ de 4-scalaires arbitraire, laisse invariante la dynamique. En exprimant cette transformation en termes des composantes spatiales et temporelles du potentiel, nous retrouvons la forme standard de la transformation de jauge:

$$\begin{aligned} \mathbf{A} &\longrightarrow \mathbf{A} - \nabla \Phi \\ V &\longrightarrow V + \frac{\partial \Phi}{\partial t}. \end{aligned} \quad (4.25)$$

Pour lever l'ambiguïté sur le potentiel, nous pouvons imposer une condition de jauge supplémentaire. Pour respecter l'invariance relativiste, cette condition de jauge se doit d'être manifestement covariante. La plus naturelle, la "jauge de Lorentz", est d'imposer la nullité de la 4-divergence du potentiel:

$$\partial_\mu \mathcal{A}^\mu = 0, \quad (4.26)$$

qui s'écrit en termes des composantes spatiales et temporelles:

$$\frac{1}{c^2} \frac{dV}{dt} + \nabla \cdot \mathbf{A} = 0. \quad (4.27)$$

Notons que la jauge de l'électrostatique, ou jauge de Coulomb, $\nabla \cdot \mathbf{A} = 0$, brise la covariance. Si elle peut être employée sans restriction dans un référentiel donné, elle est à proscrire quand s'imposent des changements de référentiel.

4.1.3 Force de Lorentz

Pour nous rapprocher encore de l'électromagnétisme sous sa forme classique, nous allons exprimer la force tridimensionnelle subie par la particule chargée dans un référentiel donné en fonction des champs électriques et magnétiques. En fonction du temps t dans le référentiel \mathcal{R} , l'équation de la dynamique s'écrit:

$$\gamma \frac{dP_\mu}{dt} = \gamma \left(\frac{d\mathcal{E}/c}{dt}, -\frac{d\mathbf{p}}{dt} \right) = q F_{\mu\nu} U^\nu, \quad (4.28)$$

où \mathcal{E} est l'énergie totale de la particule et \mathbf{p} sa quantité de mouvement tridimensionnelle. En développant le dernier terme (qui peut s'écrire simplement comme un produit matriciel) et en isolant les composantes temporelles et spatiales, on trouve:

$$\frac{d\mathcal{E}}{dt} = q \mathbf{E} \cdot \mathbf{v}, \quad (4.29)$$

où \mathbf{v} est la vitesse de la particule dans \mathcal{R} et

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) . \quad (4.30)$$

On retrouve ainsi, dans un référentiel donné, la forme standard de la force de Lorentz. On trouve aussi que la variation d'énergie est entièrement due au champ électrique. Bien sûr, dans ces équations, \mathcal{E} et \mathbf{p} sont des quantités relativistes ($\mathcal{E} = m\gamma c^2$ et $\mathbf{p} = m\gamma\mathbf{v}$). La dynamique de la particule est donc en général différente de la dynamique classique.

Comme la force de Lorentz est la première force que nous ayons explicitée dans le cadre relativiste, nous allons, à titre d'application et d'exemple, étudier en détail le mouvement d'une particule chargée dans un champ magnétique ou électrique uniforme, statique. Nous verrons ainsi comment la nature relativiste du mouvement modifie la dynamique et nous pourrons jeter un regard nouveau sur la notion de vitesse limite.

Champ magnétique uniforme

Nous nous placerons, dans ce paragraphe et le suivant, dans un référentiel particulier \mathcal{R} et nous abandonnerons donc la covariance manifeste. Nous considérerons le mouvement d'une particule dans un champ magnétique \mathbf{B} uniforme et constant.

Le champ magnétique ne modifiant pas l'énergie totale de la particule, le facteur γ est une constante et l'équation du mouvement s'écrit simplement:

$$\gamma m \frac{d\mathbf{v}}{dt} = q\mathbf{v} \times \mathbf{B} . \quad (4.31)$$

Elle est donc la même que dans le cas non relativiste, avec la simple substitution de la masse m par γm . En particulier, la trajectoire est une hélice admettant le champ magnétique pour axe, avec un rayon $R = m\gamma v/qB$ où v est le module de la vitesse perpendiculaire au champ magnétique. la pulsation du mouvement circulaire uniforme dans le plan perpendiculaire à \mathbf{B} (pulsation cyclotron) étant:

$$\omega_c = \frac{qB}{m\gamma} . \quad (4.32)$$

S'il n'y a pas de différence qualitative entre le mouvement relativiste et le mouvement classique, le facteur γ induit néanmoins des complications techniques dans les applications. Dans de nombreux types d'accélérateurs, un champ magnétique est utilisé pour confiner les particules au voisinage d'une trajectoire circulaire. Le facteur γ fait que le rayon de ces trajectoires est, pour des particules ultra-relativistes, beaucoup plus grand que ce que prédit la mécanique classique. L'encombrement de ce type de dispositif est en partie dû à cet effet¹. De plus, la fréquence des champs accélérateurs, qui doit être adaptée à la fréquence cyclotron, doit être ajustée pendant toute la phase d'accélération pour tenir compte de la variation de ce facteur relativiste.

Notons que cette "contraction" relativiste de la fréquence cyclotron peut être mise en évidence même pour des électrons de très basse énergie. Dans une très spectaculaire série d'expériences, Hans Dehmelt et ses collaborateurs (Université de Seattle) ont étudié des électrons confinés dans un piège constitué d'un champ magnétique et d'un champ quadripolaire électrique (piège de Penning). Ils ont ainsi mesuré avec une précision remarquable, sur un électron unique, le célèbre "facteur gyro-magnétique anormal", qui constitue un test sévère de l'électrodynamique quantique. Une des étapes de l'expérience est d'exciter, par un champ radiofréquence convenable, le mouvement cyclotron de l'électron. Dehmelt a pu observer que la fréquence de résonance cyclotron se déplace avec l'énergie de l'électron, conformément à la loi relativiste. Les énergies mises en jeu n'étant que d'une fraction d'électron-volt, on pourra juger de la sensibilité de l'expérience.

¹Un autre effet important limite la compacité des accélérateurs: plus une particule est accélérée, plus elle perd d'énergie par rayonnement. Nous discuterons de ce "rayonnement de freinage" dans un prochain chapitre. Pour des particules légères (électrons), cet effet est la principale limitation au rayon des accélérateurs. Notons aussi que les accélérateurs linéaires échappent à ces deux types de limitations.

Champ électrique uniforme

Nous considérerons maintenant le mouvement d'une particule dans un champ électrique uniforme et constant. Pour simplifier l'algèbre, sans trop restreindre la physique, nous prendrons comme condition initiale une particule au repos. Le champ électrique étant par convention orienté le long de l'axe Ox , il est évident que le mouvement s'effectue le long de cet axe. L'équation de la dynamique, projetée sur Ox , s'écrit alors:

$$m \frac{d\gamma \dot{x}}{dt} = qE . \quad (4.33)$$

Dans ce cas, bien sûr, l'énergie de la particule et donc le facteur γ ne sont pas des constantes. On tire de cette équation immédiatement:

$$\gamma \dot{x} = \frac{qEt}{m} , \quad (4.34)$$

la valeur initiale de cette quantité étant nulle par convention. Nous poserons, pour alléger les notations, $V = qEt/m$ (notons que V serait la vitesse de la particule si nous ne tenions pas compte des corrections relativistes au mouvement). On déduit alors de ce qui précède:

$$\dot{x} = \frac{V}{\sqrt{1 + V^2/c^2}} , \quad (4.35)$$

qui s'intègre aisément en

$$x = \frac{mc^2}{qE} \left[\sqrt{1 + \left(\frac{qEt}{mc} \right)^2} - 1 \right] , \quad (4.36)$$

à condition de prendre $x = 0$ comme condition initiale.

Pour des temps suffisamment petits, la vitesse de la particule est faible et on peut développer l'expression précédente au premier ordre en qEt/mc . On trouve alors

$$x = \frac{1}{2m} qEt^2 , \quad (4.37)$$

mouvement uniformément accéléré de la dynamique classique. Aux temps longs, en revanche, x tend simplement vers ct : la vitesse de la particule tend vers la vitesse de la lumière, comme nous pouvions nous y attendre. On pourra montrer, à titre d'exercice, que la rapidité de la particule continue, pour sa part, à croître indéfiniment. La généralisation de ce calcul à trois dimensions ne pose aucune autre difficulté qu'algébrique.

4.1.4 Changements de référentiels pour le champ

La formulation explicitement relativiste du tenseur champ électromagnétique nous permet d'écrire sans difficultés la loi de transformation des champs dans un changement de référentiel. Nous aurons en effet:

$$F'^{\mu\nu}(x'^\alpha = \mathcal{L}^\alpha_\beta x^\beta) = \mathcal{L}^\mu_\rho \mathcal{L}^\nu_\sigma F^{\rho\sigma}(x^\beta) , \quad (4.38)$$

où les quantités primées sont relatives au nouveau référentiel. Le champ étant un champ de tenseur, il est une fonction de l'événement auquel il est estimé. Il faut bien sûr estimer le champ dans les deux référentiels pris au même événement et donc à des coordonnées spatio-temporelles qui se déduisent les unes des autres dans une transformation de Lorentz. Dans l'expression précédente, \mathcal{L} peut décrire n'importe quel élément du groupe de Lorentz le plus général.

Nous précisons maintenant les nouvelles valeurs du champ pour une transformation spéciale de Lorentz avec les conventions habituelles pour l'orientation des axes. Le calcul ne présente aucune difficulté de principe. Il faut toutefois prendre garde que le produit de "tenseurs" au second membre ne peut être calculé directement comme un produit de leurs trois représentations matricielles. Les deux

derniers termes correspondent en effet à une sommation sur deux indices colonne. On peut mettre le calcul sous la forme de produits matriciels standard en effectuant les transpositions nécessaires. Après quelques lignes d'algèbre, on arrive aux lois suivantes pour les transformations des composantes des champs électriques et magnétiques:

$$E'_x = E_x \quad (4.39)$$

$$E'_y = \gamma(E_y - c\beta B_z) \quad (4.40)$$

$$E'_z = \gamma(E_z + c\beta B_y) \quad (4.41)$$

$$B'_x = B_x \quad (4.42)$$

$$B'_y = \gamma\left(B_y + \beta \frac{E_z}{c}\right) \quad (4.43)$$

$$B'_z = \gamma\left(B_z - \beta \frac{E_y}{c}\right). \quad (4.44)$$

La transformation inverse s'obtient trivialement en changeant le signe de β^2 . Si, jusqu' alors, nous n'avions fait que retrouver les caractéristiques essentielles de l'électromagnétisme (il n'est peut-être plus utile de cacher que notre théorie de champ est bien l'électromagnétisme), nous obtenons ici, grâce à la formulation manifestement covariante, un résultat nouveau et fort important.

On peut obtenir une approximation galiléenne à la loi de transformation des champs en ne gardant que l'ordre le plus bas en u/c dans les équations précédentes. On obtient alors, sous forme vectorielle:

$$\mathbf{E}' = \mathbf{E} - \mathbf{B} \times \mathbf{u} \quad (4.45)$$

$$\mathbf{B}' = \mathbf{B} + \frac{\mathbf{E} \times \mathbf{u}}{c^2}. \quad (4.46)$$

4.1.5 Invariants du champ électromagnétique

On peut se poser le problème de former des quantités 4-scalaires à partir du tenseur champ. De telles quantités seront en effet conservées dans un changement de référentiel et constitueront des invariants du champ électromagnétique, fort utiles. Là encore, nous allons ajouter des résultats nouveaux à l'électromagnétisme standard. Nous ne chercherons pas systématiquement tous les invariants possibles. En fait, il n'en existe que deux qui présentent un intérêt physique³.

Formons d'abord la quantité:

$$F_{\mu\nu}F^{\mu\nu}. \quad (4.47)$$

Cette quantité est manifestement un 4-scalaire et donc un invariant du champ. Ecrivons-la en termes des champs électriques et magnétiques pour en comprendre la signification physique. Interviennent des composantes spatio-temporelles et des composantes spatiales. La contribution des composantes spatio-temporelles est manifestement $2F_{0i}F^{0i}$ (en effet les deux termes se déduisant l'un de l'autre par permutation des indices sont manifestement égaux en raison de l'antisymétrie de F) ou encore $-2E^2/c^2$. De même, les composantes spatiales font intervenir le carré scalaire du champ magnétique et on a enfin:

$$F_{\mu\nu}F^{\mu\nu} = \frac{2}{c^2}(c^2B^2 - E^2). \quad (4.48)$$

La quantité $c^2B^2 - E^2$ est donc un invariant du champ.

Donnons dès maintenant une application de cet invariant. Dans un référentiel \mathcal{R} , considérons une onde électromagnétique plane. Les modules du champ électrique et du champ magnétique sont reliés

²On notera que ces expressions ne sont pas invariantes par échange de y et z . Ceci n'est en rien contradictoire avec la symétrie du problème. Echanger ces axes revient à changer l'orientation de l'espace et donc le signe de \mathbf{B} qui est un pseudo-vecteur.

³On montrera en particulier que le déterminant de $F^{\mu\nu}$ est proportionnel au carré de notre second invariant

par $E = cB$. L'invariant du champ considéré est donc nul. Dans un autre référentiel, sans préjuger de la structure du champ⁴, on aura aussi $E' = cB'$.

Le deuxième invariant que nous allons former s'écrit :

$$\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} , \quad (4.49)$$

où ϵ est le tenseur complètement antisymétrique de rang 4. Rappelons rapidement que les éléments de ce tenseur valent +1 si les quatre indices sont une permutation paire de (0, 1, 2, 3), -1 si ils constituent une permutation impaire, 0 dans tous les autres cas.

Il n'y a donc dans cet invariant que 24 termes non nuls. En fait, ces termes sont égaux 8 à 8. Si nous considérons une permutation μ, ν, ρ, σ donnée, nous obtenons en effet un terme identique en échangeant μ et ν et/ou ρ et σ (le tenseur antisymétrique et le tenseur champ concernés changent tous les deux de signe). Il y a donc au total quatre permutations de ce type et 4 termes identiques dans le développement de notre invariant. De plus, nous obtenons un terme identique en échangeant la première paire (μ, ν) et la seconde (ρ, σ). En effet les termes en tenseur champ ne changent pas. Cette permutation des deux paires peut s'effectuer avec quatre permutations des indices. Elle ne change pas non plus la valeur du tenseur antisymétrique. Ensuite, les opérations de permutations à l'intérieur des paires ainsi permutées peuvent être effectuées sans changer la valeur. Nous introduisons ainsi quatre nouvelles permutations des indices donnant la même valeur. Au total, il y a donc 8 termes identiques. Comme nous n'avons manifestement que 24 termes non nuls, il n'y a que trois termes différents, correspondant à un ensemble d'indices (par exemple 0,1,2,3) et aux deux manières d'échanger un terme de la première paire et un terme de la deuxième (dans ce cas 2,1,0,3 et 3,1,2,0). Le calcul de ces trois termes est alors trivial à partir de l'expression de F (ϵ valant +1 pour la première permutation et -1 pour les deux autres). On trouve finalement :

$$\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} = -8\mathbf{E} \cdot \mathbf{B}/c , \quad (4.50)$$

un résultat particulièrement simple.

Le produit scalaire des champs électriques et magnétiques est donc invariant dans un changement de référentiel (bien sûr, cette invariance pourrait être établie, de manière assez pénible, directement à partir des lois de transformation). Donnons tout de suite une application de cette propriété. Dans un référentiel, considérons une onde plane. \mathbf{E} et \mathbf{B} sont alors perpendiculaires et leurs modules sont dans un rapport c . Dans un autre référentiel, ils sont donc encore perpendiculaires avec des modules dans un rapport c .

4.1.6 Premier groupe d'équations de Maxwell

Il nous reste à tirer parti du fait que le tenseur champ est le rotationnel du potentiel. A trois dimensions, cette propriété impliquerait la nullité de sa divergence. Nous allons maintenant établir la propriété correspondante à quatre dimensions. En écrivant cette propriété en termes des champs électriques et magnétiques, nous établirons des relations différentielles entre eux qui ne seront autres que les deux équations de Maxwell homogènes. Nous avons donc

$$F_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu . \quad (4.51)$$

On en déduit immédiatement :

$$\begin{aligned} \partial_\rho F_{\mu\nu} &= \partial_\rho \partial_\mu \mathcal{A}_\nu - \partial_\rho \partial_\nu \mathcal{A}_\mu \\ \partial_\nu F_{\rho\mu} &= \partial_\nu \partial_\rho \mathcal{A}_\mu - \partial_\nu \partial_\mu \mathcal{A}_\rho \\ \partial_\mu F_{\nu\rho} &= \partial_\mu \partial_\nu \mathcal{A}_\rho - \partial_\mu \partial_\rho \mathcal{A}_\nu \end{aligned} \quad (4.52)$$

⁴Qui se trouve être également une onde plane.

En remarquant que les dérivées secondes croisées du potentiel sont égales, et en faisant la somme de ces trois équations, on obtient immédiatement:

$$\partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} + \partial_\mu F_{\nu\rho} = 0 . \quad (4.53)$$

Cette équation est une conséquence directe du fait que le champ dérive d'un potentiel (c'est d'ailleurs une condition nécessaire et suffisante).

Pour en comprendre la signification physique, écrivons cette équation en termes des champs électriques et magnétiques. Remarquons d'abord que, s'il y a a priori 64 équations possibles, seules 4 ne sont pas triviales. Si les trois indices sont identiques, tous les F sont nuls et l'équation est un truisme. Si deux indices sont égaux (par exemple $\rho = \mu$), l'équation se réduit à $\partial_\mu(F_{\mu\nu} + F_{\nu\mu}) = 0$, une tautologie en raison de l'antisymétrie de F . L'équation n'est non triviale que si les trois indices sont différents et il ne reste donc que quatre équations indépendantes.

La première correspond aux indices 1,2,3. Elle s'écrit:

$$\partial_1 F_{23} + \partial_3 F_{12} + \partial_2 F_{31} = 0 , \quad (4.54)$$

soit encore

$$\nabla \cdot \mathbf{B} = 0 . \quad (4.55)$$

On montrera de même que les trois autres équations peuvent se résumer, sous forme vectorielle, par:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} . \quad (4.56)$$

Nous retrouvons donc ainsi les équations de Maxwell homogènes, qui sont équivalentes à l'existence d'un potentiel scalaire et d'un potentiel vecteur.

4.2 Champ en fonction des sources

Nous allons maintenant établir les équations qui relient le tenseur champ à ses sources, c'est à dire au mouvement des particules chargées. Dans le paragraphe précédent, nous nous intéressions au mouvement d'une particule unique. Les variables dynamiques du problème étaient donc la position et l'impulsion de la particule, situation habituelle en mécanique analytique. Dans tout ce chapitre, conformément à notre programme initial, nous supposons imposées les dynamiques des particules (c'est-à-dire le courant) et nous ne nous intéresserons qu'à la dynamique du champ. Les variables dynamiques sont donc les valeurs du potentiel ou des champs en tous points de l'espace et à chaque instant. Il nous faudra donc adapter nos techniques variationnelles pour des variables dynamiques continues. En particulier, nous n'écrirons plus l'action en termes de lagrangien mais d'une densité de lagrangien que nous intégrerons sur tout l'espace et sur le temps pour obtenir l'action. Il nous faudra aussi réécrire l'action d'interaction comme l'intégrale sur tout l'espace d'une densité de lagrangien qui devra faire intervenir le 4-vecteur courant au lieu des positions et vitesses individuelles des particules.

Le fait que nous traitons de plusieurs particules pose une difficulté technique immédiate. Pour obtenir les équations du mouvement d'une particule unique, nous avons intégré le lagrangien entre deux événements limites relatifs à cette particule. Nous ne pouvons définir de façon aussi simple les bornes d'intégration si nous considérons plusieurs particules qui ne partagent pas le même temps propre. Pour éviter toute difficulté ou le recours à un formalisme complexe nous éluderons le problème en nous plaçant, pour un temps, dans un référentiel donné \mathcal{R} . Dans ce référentiel, le temps est bien défini, et nous pourrions intégrer la densité de lagrangien entre deux instants de référence. Nous n'aurons à manipuler que des intégrales d'espace et de temps sous forme habituelle. En abandonnant ainsi la covariance manifeste, nous risquons bien sûr d'obtenir des équations de Lagrange qui ne seraient pas des invariants relativistes. Nous verrons qu'il n'en sera heureusement rien: les équations que nous obtiendrons dans un référentiel donné s'écriront en termes de quantités explicitement covariantes, et seront donc valables dans n'importe quel référentiel.

4.2.1 Interaction champ–courant

Nous allons, dans ce premier paragraphe, réécrire l'action d'interaction en termes du courant macroscopique. Nous considérons un ensemble de particules chargées ponctuelles, dont le mouvement est imposé, que nous indiquerons par un indice α entre parenthèses, pour éviter toute confusion entre cet indice qui numérote simplement les particules et un indice relativiste en position covariante.

Le 4–vecteur courant J^μ peut donc s'écrire: $J^\mu = (c\rho, \mathbf{j})$, avec:

$$\rho = \sum_{\alpha} q_{(\alpha)} \delta(\mathbf{r} - \mathbf{r}_{(\alpha)}) \quad (4.57)$$

$$\mathbf{j} = \sum_{\alpha} q_{(\alpha)} \mathbf{v}_{(\alpha)} \delta(\mathbf{r} - \mathbf{r}_{(\alpha)}) , \quad (4.58)$$

où $q_{(\alpha)}$, $\mathbf{r}_{(\alpha)}$ et $\mathbf{v}_{(\alpha)}$ sont respectivement la charge, la position et la vitesse de la particule α . En généralisant l'action d'interaction introduite au paragraphe précédent, on écrira:

$$S_{\text{Interaction}} = - \sum_{\alpha} q_{(\alpha)} \int_{t_a}^{t_b} \mathcal{A}_{\mu}(x^{\nu}_{(\alpha)}) U^{\mu} d\tau_{(\alpha)} , \quad (4.59)$$

où $x^{\nu}_{(\alpha)}$ est la position de la particule α . t_a et t_b sont les instants dans \mathcal{R} où nous spécifierons les conditions initiales imposées au champ. $\tau_{(\alpha)}$ est le temps propre de la particule α , qui peut être paramétré lui même par le temps t du référentiel \mathcal{R} dans lequel nous nous sommes placés. En raison de la “dilatation des temps”, $d\tau_{(\alpha)} = dt/\gamma_{(\alpha)}$ où $\gamma_{(\alpha)}$ est le facteur de dilatation relativiste calculé à chaque instant avec la vitesse de la particule α . En substituant cette expression dans l'action d'interaction et en explicitant enfin les composantes temporelles et spatiales, on trouve:

$$S_{\text{Interaction}} = - \int_{t_a}^{t_b} dt \left[\sum_{\alpha} \mathcal{A}_{\mu}(x_{(\alpha)}^{\nu}) (q_{(\alpha)} c, q_{(\alpha)} \mathbf{v}_{(\alpha)}) \right] . \quad (4.60)$$

Notons que l'équation précédente, qui mélange notations d'Einstein et séparation des parties temporelles et spatiales constitue un abus de notations manifeste. Dans cette expression, le potentiel est à évaluer à l'endroit où se trouve la particule α . Pour mettre l'expression précédente sous la forme de l'intégrale d'une densité de lagrangien, on peut écrire:

$$\mathcal{A}_{\mu}(x_{(\alpha)}^{\nu}) = \int dV \mathcal{A}_{\mu}(ct, \mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_{(\alpha)}(t)) , \quad (4.61)$$

l'intégrale portant sur tout l'espace. On a alors:

$$S_{\text{Interaction}} = - \int_{t_a}^{t_b} dt \int dV \left[\sum_{\alpha} (q_{(\alpha)} c, q_{(\alpha)} \mathbf{v}_{(\alpha)}) \delta(\mathbf{r} - \mathbf{r}_{(\alpha)}(t)) \right] \mathcal{A}_{\mu}(ct, \mathbf{r}) \quad (4.62)$$

On reconnaît, entre les crochets, l'expression du 4–vecteur courant. On a donc finalement:

$$S_{\text{Interaction}} = \int_{t_a}^{t_b} dt \int dV \mathcal{L}_{\text{Interaction}} \quad (4.63)$$

où la densité de lagrangien d'interaction s'exprime par:

$$\mathcal{L}_{\text{Interaction}} = -\mathcal{A}_{\mu} J^{\mu} . \quad (4.64)$$

Bien que nous ayons établi cette expression dans un référentiel donné, elle est manifestement un 4–scalaire et est donc correcte dans tous les référentiels.

4.2.2 Lagrangien du champ

Il nous faut maintenant postuler l'expression de la densité de lagrangien pour le champ libre. Nous allons choisir bien sûr une quantité qui soit manifestement un 4-scalaire. Il faudra de plus qu'elle soit invariante de jauge et donc qu'elle ne s'exprime en définitive qu'en fonction du tenseur champ et non du potentiel. Enfin, il faudra que ce soit une quantité quadratique dans le champ. Une densité de lagrangien est en effet homogène à une densité d'énergie qui doit être une fonction quadratique des variables dynamiques. Les deux invariants scalaires du champ que nous avons construits au paragraphe précédent remplissent tous deux ces conditions. Le second, équivalent au produit scalaire des champs électriques et magnétiques, ne semble pas convenir. On ne voit pas, en particulier, comment il pourrait décrire de façon convenable une situation purement électrostatique. Seule l'autre invariant est convenable et nous postulerons que la densité de lagrangien pour le champ libre peut s'écrire:

$$\mathcal{L}_{\text{Champ Libre}} = -\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu}, \quad (4.65)$$

où μ_0 est a priori une constante dimensionnelle telle que la densité de lagrangien ait la dimension d'une densité d'énergie. L'action totale intégrée entre les instants t_a et t_b s'écrit donc:

$$S = S_{\text{Particules Libres}} + \int_{t_a}^{t_b} dt \int dV \left[-\mathcal{A}_\mu J^\mu - \frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} \right]. \quad (4.66)$$

Dans cette expression, l'action des particules libres est une simple constante, puisque leur dynamique est imposée.

4.2.3 Equations de Lagrange

Pour établir les équations de Lagrange, nous allons considérer, entre les deux instants de référence t_a et t_b , deux histoires possibles du champ. L'une, qui sera la "trajectoire effectivement suivie", autrement dit la solution des équations de Lagrange, correspondra, à chaque instant, au potentiel \mathcal{A}_μ . L'autre, infinitésimalement différente, correspondra au potentiel $\mathcal{A}_\mu + \delta\mathcal{A}_\mu$. Pour assurer que le champ vrai et le champ varié obéissent aux mêmes conditions aux limites, nous imposerons à l'accroissement infinitésimal $\delta\mathcal{A}$ de s'annuler, en tous points de l'espace, en t_a et en t_b . Nous allons ensuite exprimer la variation de l'action due à cette variation du potentiel en tous points de l'espace à chaque instant. En exprimant que cette variation est nulle au premier ordre dans l'accroissement, quel que soit celui-ci, nous obtiendrons une relation qui devra être vérifiée par le champ en tous points, à tout instant. Ce raisonnement généralise de façon évidente à un ensemble continu de variables dynamiques celui que nous avons déjà utilisé fréquemment pour un nombre fini de degrés de liberté.

La trajectoire des particules étant imposée, le courant ne doit pas être varié et la variation de l'action totale, s'écrit donc:

$$\delta S = - \int_{t_a}^{t_b} dt \int dV \left[\delta(\mathcal{A}_\mu) J^\mu + \frac{1}{4\mu_0} \delta(F_{\mu\nu} F^{\mu\nu}) \right]. \quad (4.67)$$

On a de manière évidente:

$$\delta(F_{\mu\nu} F^{\mu\nu}) = 2(\delta F_{\mu\nu}) F^{\mu\nu}. \quad (4.68)$$

En exprimant ensuite $\delta F_{\mu\nu}$ en termes du potentiel, on a:

$$\delta S = - \int_{t_a}^{t_b} dt \int dV \left[\delta(\mathcal{A}_\mu) J^\mu + \frac{1}{2\mu_0} F^{\mu\nu} \partial_\mu \delta\mathcal{A}_\nu - \frac{1}{2\mu_0} F^{\mu\nu} \partial_\nu \delta\mathcal{A}_\mu \right]. \quad (4.69)$$

En permutant les deux indices muets du terme central, et en utilisant l'antisymétrie de F , on constate que les deux derniers termes dans l'intégrale sont égaux, et que donc:

$$\delta S = - \int_{t_a}^{t_b} dt \int dV \left[\delta(\mathcal{A}_\mu) J^\mu - \frac{1}{\mu_0} F^{\mu\nu} \partial_\nu \delta\mathcal{A}_\mu \right]. \quad (4.70)$$

Considérons le deuxième terme dans l'intégrale. A un facteur $1/c$ près, il s'agit de l'intégrale d'espace temps d'une quantité scalaire. Le volume d'intégration \mathcal{V} est l'ensemble de l'espace à trois dimensions pris entre les instants t_a et t_b . En utilisant le théorème d'Ostrogradski pour les intégrales quadridimensionnelles, on peut réaliser sur ce terme une intégration par parties. En posant $d\Omega = c dt dV$ on a en effet:

$$\int_{\mathcal{V}} d\Omega F^{\mu\nu} \partial_\nu \delta \mathcal{A}_\mu = \int_{\mathcal{S}} F^{\mu\nu} \delta \mathcal{A}_\mu dS_\nu - \int_{\mathcal{V}} d\Omega (\partial_\nu F^{\mu\nu}) \delta \mathcal{A}_\mu , \quad (4.71)$$

la surface \mathcal{S} , à trois dimensions, étant la frontière du volume \mathcal{V} et dS_ν l'élément différentiel de cette surface. Cette surface est constitué de l'ensemble de l'espace pris à l'instant initial t_a , de la sphère de l'infini (une sphère de rayon R , dont on prend la limite pour $R \rightarrow \infty$) à tous les instants entre t_a et t_b et à nouveau de tout l'espace à l'instant final t_b . L'accroissement du potentiel est nul aux instants limites en tous points de l'espace. Il est nul aussi, ainsi que tous les champs physiques, en tous points de la sphère de l'infini à chaque instant. L'intégrale de surface dans l'expression précédente est donc identiquement nulle. Notons que nous retrouvons ici, sous une forme un peu plus complexe, l'intégration par parties à laquelle nous devons toujours procéder pour établir les équations de Lagrange.

L'accroissement de l'action s'écrit donc finalement:

$$\delta S = - \int_{t_a}^{t_b} dt \int dV \left[J^\mu + \frac{1}{\mu_0} \partial_\nu F^{\mu\nu} \right] \delta \mathcal{A}_\mu . \quad (4.72)$$

Elle ne peut être nulle quel que soit l'accroissement du potentiel que si le champ entre crochets est identiquement nul. Les équations de Lagrange déterminant le champ en fonction des sources s'écrivent donc simplement:

$$\partial_\nu F^{\mu\nu} = -\mu_0 J^\mu , \quad (4.73)$$

ou encore

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\nu . \quad (4.74)$$

en exploitant les propriétés de symétrie du tenseur champ.

Dans tout ce raisonnement, nous avons abandonné la covariance manifeste en nous plaçant dans un référentiel donné. En revanche, les équations obtenues ne font intervenir que des quantités covariantes. Elles sont donc très générales, et valables dans tout référentiel galiléen.

Nous pouvons maintenant écrire simplement ces équations en termes du potentiel. En reportant l'expression de F , nous avons:

$$\partial_\mu \partial^\mu \mathcal{A}^\nu - \partial_\mu \partial^\nu \mathcal{A}^\mu = \mu_0 J^\nu . \quad (4.75)$$

Si nous imposons au potentiel vecteur d'obéir à la condition de Jauge de Lorentz $\partial_\mu \mathcal{A}^\mu = 0$, le deuxième terme s'annule (on permutera les dérivées partielles pour le constater). L'équation aux potentiels s'écrit alors:

$$\partial_\mu \partial^\mu \mathcal{A}^\nu = \mu_0 J^\nu . \quad (4.76)$$

L'opérateur différentiel n'est autre que le carré de la norme du gradient: c'est le 4-laplacien ou encore le d'alembertien.

4.2.4 Equations de Maxwell

Pour mettre ces équations sous une forme plus familière, nous allons les exprimer en termes des champs électriques et magnétiques. La partie temporelle de cette équation vectorielle s'écrit en effet:

$$\partial_\nu F^{0\nu} = -\mu_0 c \rho , \quad (4.77)$$

ou encore

$$\nabla \cdot \mathbf{E} = \mu_0 c^2 \rho = \rho / \epsilon_0 , \quad (4.78)$$

en posant évidemment $\epsilon_0\mu_0c^2 = 1$.

Les trois composantes spatiales se mettent de leur côté évidemment sous la forme:

$$\nabla \times \mathbf{B} = \mu_0 \left[\mathbf{j} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right]. \quad (4.79)$$

Les équations aux potentiels, pourvu que ceux-ci obéissent à la Jauge de Lorentz, se mettent sous la forme:

$$\square \mathbf{A} = \mu_0 \mathbf{j} \quad (4.80)$$

$$\square V = \rho/\epsilon_0, \quad (4.81)$$

où \square est l'opérateur d'alembertien $(1/c^2)\partial^2/\partial t^2 - \Delta$, Δ étant le laplacien.

Les équations de Lagrange décrivant le champ en fonction des charges sont donc bien les équations de Maxwell faisant intervenir les courants. Si on y ajoute les équations de Maxwell homogènes et les équations de Lagrange pour la dynamique des particules qui sont équivalentes à la force de Lorentz on peut déterminer complètement, au moins en principe, la dynamique couplée des particules et du champ. L'électromagnétisme de Maxwell, formulé en termes explicitement covariants, n'est donc que la théorie de champ la plus simple qui soit décrite par un champ de vecteurs. La structure à deux champs de la théorie de Maxwell est une conséquence immédiate de la nature antisymétrique du tenseur exprimant le rotationnel du potentiel. En fait, l'électromagnétisme de Maxwell, sous sa forme standard, est déjà écrit en termes explicitement relativistes ce qui le rend bien sûr incompatible avec la cinématique classique. La découverte des équations de Maxwell aurait été une conséquence immédiate de celle de la relativité. Mais il fallait bien les difficultés soulevées par l'électromagnétisme de Maxwell pour qu'on songe à mettre en doute la mécanique Newtonienne, parfaitement vérifiée par ailleurs.

4.3 Energie–impulsion du champ

Le champ électromagnétique doit posséder une densité d'énergie. Energie et quantité de mouvement n'étant que deux aspects du même 4–vecteur impulsion, il doit aussi exister une densité de quantité de mouvement pour le champ électromagnétique. Enfin, le champ électromagnétique obéissant, comme les particules matérielles, aux grandes propriétés de symétrie par translation dans l'espace ou dans le temps, ces densités d'énergie et d'impulsion, associées à celles des particules, doivent obéir à des lois de conservation. Nous allons, dans ce paragraphe, définir les densités d'énergie et d'impulsion et établir leurs équations bilan, c'est-à-dire étudier leur propagation. Nous pourrions partir de la formulation lagrangienne de l'électromagnétisme et appliquer les grandes lois de symétrie pour établir la forme de ces densités, comme nous l'avons fait pour une particule libre. Nous nous contenterons en fait de postuler la forme d'un tenseur rassemblant ces quantités, d'établir ses propriétés et d'identifier les différents termes. Cette approche est beaucoup moins satisfaisante qu'une approche à partir des premiers principes, mais elle est beaucoup plus compacte.

Pour ce qui est du bilan d'énergie du champ électromagnétique, nous allons bien sûr retrouver des résultats bien connus sur la densité d'énergie du champ électromagnétique et sa propagation décrite par le vecteur de Poynting, dont le flux décrit le transport d'énergie à travers une surface. Nous ne ferons que rappeler très brièvement les propriétés essentielles de ces quantités. Pour la quantité de mouvement, nous obtiendrons en revanche des résultats nouveaux. Nous établirons la forme d'une densité (bien sûr vectorielle) de quantité de mouvement à trois dimensions. Le transfert d'impulsion à travers une surface s'écrira comme le flux d'une quantité tensorielle de rang 2.

4.3.1 Tenseur énergie–impulsion

Posons

$$\theta^{\alpha\beta} = \frac{1}{\mu_0} \left[g^{\alpha\mu} F_{\mu\lambda} F^{\lambda\beta} + \frac{1}{4} g^{\alpha\beta} F_{\mu\lambda} F^{\mu\lambda} \right]. \quad (4.82)$$

Nous nommerons, pour des raisons qui apparaîtront dans un instant, ce tenseur de rang 2, écrit ici sous sa forme doublement contravariante, le tenseur d'énergie impulsion du champ ("stress tensor" dans la littérature anglo-saxonne). Il s'agit d'un tenseur symétrique. La symétrie du deuxième terme du crochet est manifeste. Pour vérifier celle du premier terme, nous écrirons:

$$\begin{aligned} g^{\beta\mu} F_{\mu\lambda} F^{\lambda\alpha} &= F^{\lambda\alpha} F^{\beta}_{\lambda} = F_{\lambda}^{\alpha} F^{\beta\lambda} \\ &= F^{\alpha}_{\lambda} F^{\lambda\beta} = g^{\alpha\mu} F_{\mu\lambda} F^{\lambda\beta} . \end{aligned} \quad (4.83)$$

On a donc bien $\theta^{\alpha\beta} = \theta^{\beta\alpha}$.

Pour nous convaincre de la possible utilité de ce tenseur, nous allons l'écrire en fonction des champs électrique et magnétique. Le deuxième terme est le produit du tenseur métrique, diagonal, par l'invariant $2(B^2 - E^2/c^2)$, densité de lagrangien du champ électromagnétique. Le premier terme fait intervenir $F_{\mu\lambda} F^{\lambda\beta}$ qui se calcule comme un produit matriciel ordinaire. L'action du tenseur métrique est de changer le signe de toutes les lignes ayant un indice spatial dans la représentation matricielle obtenue. En regroupant avec le second terme et après quelques manipulations élémentaires, on peut écrire le tenseur énergie-impulsion sous la forme:

$$\theta^{\alpha\beta} = \begin{pmatrix} u & \Pi_x/c & \Pi_y/c & \Pi_z/c \\ \Pi_x/c & & & \\ \Pi_y/c & & (\mathcal{T}) & \\ \Pi_z/c & & & \end{pmatrix}, \quad (4.84)$$

où

$$u = \frac{\epsilon_0 E^2}{2} + \frac{B^2}{2\mu_0} \quad (4.85)$$

et

$$\mathbf{\Pi} = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0}. \quad (4.86)$$

On reconnaîtra bien sûr ici la densité d'énergie électromagnétique et le vecteur de Poynting.

\mathcal{T} est un tenseur purement spatial de rang 2, que nous nommerons "tenseur de Maxwell". L'expression de ses composantes est:

$$\mathcal{T}_{ij} = \epsilon_0 \left[\frac{E^2}{2} \delta_{ij} - E_i E_j \right] + \frac{1}{\mu_0} \left[\frac{B^2}{2} \delta_{ij} - B_i B_j \right], \quad (4.87)$$

où les δ_{ij} sont simplement les symboles de Kronecker. Notons ici que nous faisons une entorse sérieuse à nos conventions de notations. Quand il nous arrivera de manipuler une quantité, vectorielle ou tensorielle qui soit uniquement relative à l'espace ordinaire à trois dimensions, nous placerons tous les indices en position basse, en appliquant donc avec précautions la règle de sommation sur les indices répétés. Notons enfin que pour un tenseur spatial les composantes complètement contravariantes et complètement covariantes coïncident. Nous allons interpréter plus tard la signification physique de ce tenseur.

4.3.2 Lois de conservation. Interprétation

Si θ est associé à la propagation de l'énergie et de la quantité de mouvement, il doit vérifier des équations locales de conservation qui font intervenir sa divergence (on se souviendra de l'équation locale de bilan d'énergie qui fait intervenir la dérivée temporelle de u et la divergence de $\mathbf{\Pi}$, qui sont rassemblées dans la 4-divergence de la première ligne de θ). Nous allons donc calculer le 4-vecteur

$$\partial_{\alpha} \theta^{\alpha\beta} \quad (4.88)$$

En remarquant que les tenseurs métriques ne se dérivent pas, et que leur seule action est d'élever l'indice des dérivations, on met ce terme sous la forme:

$$\partial_\alpha \theta^{\alpha\beta} = \frac{1}{\mu_0} \left[\partial^\mu (F_{\mu\lambda} F^{\lambda\beta}) + \frac{1}{4} \partial^\beta (F_{\mu\lambda} F^{\mu\lambda}) \right]. \quad (4.89)$$

Le dernier terme peut s'écrire:

$$(\partial^\beta F_{\mu\lambda}) F^{\mu\lambda} + F_{\mu\lambda} \partial^\beta F^{\mu\lambda}. \quad (4.90)$$

Ces deux termes sont manifestement égaux (il suffit d'élever et d'abaisser les mêmes indices pour passer de l'un à l'autre). En développant également le premier terme de la divergence, on a donc:

$$\partial_\alpha \theta^{\alpha\beta} = \frac{1}{\mu_0} \left[(\partial^\mu F_{\mu\lambda}) F^{\lambda\beta} + F_{\mu\lambda} (\partial^\mu F^{\lambda\beta}) + \frac{1}{2} F_{\mu\lambda} (\partial^\beta F^{\mu\lambda}) \right]. \quad (4.91)$$

Le premier terme dans le crochet peut se transformer en utilisant les équations de Maxwell $\partial^\mu F_{\mu\lambda} = \mu_0 J_\lambda$. Après une transformation triviale, on a donc:

$$\partial_\alpha \theta^{\alpha\beta} - J_\lambda F^{\lambda\beta} = \frac{1}{2\mu_0} F_{\mu\lambda} \left[\partial^\mu F^{\lambda\beta} + \partial^\mu F^{\lambda\beta} + \partial^\beta F^{\mu\lambda} \right], \quad (4.92)$$

où nous avons artificiellement séparé un terme en deux. Les deux derniers termes du crochet peuvent s'écrire $-\partial^\lambda F^{\beta\mu} = \partial^\lambda F^{\mu\beta}$ en utilisant les équations de Maxwell homogènes. On a finalement:

$$\partial_\alpha \theta^{\alpha\beta} - J_\lambda F^{\lambda\beta} = \frac{1}{2\mu_0} F_{\mu\lambda} \left[\partial^\mu F^{\lambda\beta} + \partial^\lambda F^{\mu\beta} \right]. \quad (4.93)$$

Le second membre de cette équation est la contraction d'un tenseur antisymétrique en μ, λ avec un tenseur de rang 3, symétrique en μ, λ . Cette contraction est manifestement nulle. En écrivant le tenseur de rang 3 $K^{\mu\lambda\beta}$, on a en effet $F_{\mu\lambda} K^{\mu\lambda\beta} = F_{\lambda\mu} K^{\lambda\mu\beta}$, puisque les noms des indices muets sont indifférents. De plus, en raison des propriétés de symétrie des tenseurs, ces deux termes égaux sont opposés. Ils sont donc bien nuls. Finalement, on a

$$\partial_\alpha \theta^{\alpha\beta} = -F^{\beta\lambda} J_\lambda, \quad (4.94)$$

une équation 4-vectorielle. Pour en interpréter la signification physique, nous allons en écrire séparément la composante temporelle et les composantes spatiales et exprimer ces quantités en fonction des densités d'énergie, vecteur de Poynting et tenseur de Maxwell.

Composante temporelle: conservation de l'énergie

La composante temporelle $\beta = 0$ s'écrit simplement:

$$\partial_\alpha \theta^{\alpha 0} = -F^{0\lambda} J_\lambda. \quad (4.95)$$

Il est facile de vérifier que $F^{0\lambda} J_\lambda = \mathbf{j} \cdot \mathbf{E}/c$. En regroupant les termes, on trouve l'équation scalaire:

$$\mathbf{j} \cdot \mathbf{E} + \frac{\partial u}{\partial t} + \nabla \cdot \mathbf{\Pi} = 0. \quad (4.96)$$

Nous retrouvons ici l'équation de conservation de l'énergie électromagnétique que l'on établit dans les cours élémentaires à partir des équations de Maxwell. Rappelons que $\mathbf{j} \cdot \mathbf{E}$ est la densité de puissance cédée par le champ à la matière⁵. On peut écrire une équation bilan globale en intégrant l'équation précédente sur un volume \mathcal{V} bordé par une surface fermée \mathcal{S} . On a alors:

$$\int_{\mathcal{V}} \mathbf{j} \cdot \mathbf{E} d\tau + \frac{dU}{dt} + \int_{\mathcal{S}} \mathbf{\Pi} \cdot d\mathbf{S} = 0, \quad (4.97)$$

où $U = \int_{\mathcal{V}} u$ est l'énergie électromagnétique totale dans le volume \mathcal{V} . On trouve donc que la dérivée de l'énergie électromagnétique par rapport au temps est égale à l'opposé de la puissance totale cédée à la matière plus le flux entrant du vecteur de Poynting à travers \mathcal{S} . Le vecteur de Poynting décrit donc bien le transport d'énergie électromagnétique.

⁵Cette expression est, comme nous l'avons vu plus haut, correcte même si le mouvement des particules est relativiste.

Composantes spatiales: conservation de la quantité de mouvement

En remarquant que $F^{i\lambda}J_\lambda = -[\rho\mathbf{E}_i + (\mathbf{j} \times \mathbf{B})_i] = -\mathbf{f}_i$ n'est autre, au signe près, que la composante i de la densité de force de Lorentz \mathbf{f} , on met les composantes spatiales de notre équation de conservation sous la forme:

$$\frac{d\mathbf{g}}{dt} + \mathbf{f} + \nabla \cdot \mathcal{T} = 0, \quad (4.98)$$

où la notation $\nabla \cdot \mathcal{T}$ désigne le champ de vecteurs $\partial_i \mathcal{T}_{ij}$. Nous avons posé ici

$$\mathbf{g} = \frac{\mathbf{\Pi}}{c^2}. \quad (4.99)$$

Pour dégager plus précisément le sens physique de cette équation, nous l'intégrerons, comme dans le cas de l'énergie, sur un volume \mathcal{V} bordé par une surface fermée \mathcal{S} . On obtient alors:

$$\frac{d\mathbf{P}_m}{dt} + \frac{d\mathbf{G}}{dt} + \int_{\mathcal{S}} \mathcal{T} \cdot d\mathbf{S} = 0, \quad (4.100)$$

où la notation "produit scalaire" $\mathcal{T} \cdot d\mathbf{S}$ désigne la contraction du tenseur \mathcal{T} avec le vecteur élément de surface. Nous avons posé:

$$\frac{d\mathbf{P}_m}{dt} = \int_{\mathcal{V}} \mathbf{f}. \quad (4.101)$$

Ce terme représente donc la variation temporelle de la quantité de mouvement de toutes les particules matérielles contenues dans \mathcal{V} . Il est alors facile d'interpréter le terme $d\mathbf{G}/dt$ avec

$$\mathbf{G} = \int_{\mathcal{V}} \mathbf{g} \quad (4.102)$$

comme la variation dans le temps de la quantité de mouvement totale du rayonnement électromagnétique. Le champ \mathbf{g} représente alors simplement la densité locale de quantité de mouvement du champ. Le bilan de quantité de mouvement apparaît alors de manière transparente si nous interprétons le dernier terme de l'équation (4.100) comme la quantité de mouvement sortant par unité de temps du volume \mathcal{V} . Cette quantité de mouvement, vectorielle, apparaît bien comme le flux à travers \mathcal{S} d'une quantité tensorielle.

Nous venons bien d'établir ici le bilan de quantité de mouvement pour l'ensemble du champ et des particules chargées. Notons que, comme dans le cas du bilan d'énergie, on peut établir cette équation bilan à partir des équations de Maxwell et de la force de Lorentz. On pourra s'en convaincre aisément à titre d'exercice. Il suffit d'écrire la densité de force de Lorentz \mathbf{f} en remplaçant ρ et \mathbf{j} par leurs expressions en termes de \mathbf{E} et \mathbf{B} extraites des équations de Maxwell. Des manipulations algébriques peu agréables permettent alors de mettre le résultat sous la forme de la somme d'une dérivée partielle par rapport au temps et d'une divergence de quantité tensorielle. Une simple identification redonne alors les équations précédentes.

Notons que la quantité de mouvement du champ électromagnétique est simplement, à un facteur dimensionnel près, le vecteur de Poynting qui décrit le déplacement de l'énergie électromagnétique. Nous pouvons, qualitativement, comprendre ce résultat en termes de photons. Nous avons déjà évoqué le fait que le rayonnement se comporte tout autant comme le phénomène ondulatoire que nous traitons ici que comme un flux de particules de masse nulle, d'énergie $h\nu$, où h est la constante de Planck et ν la fréquence du rayonnement que nous supposons pour un moment monochromatique. Le vecteur de Poynting décrit alors le flux de photons. Chaque photon se déplaçant à la vitesse c , la densité numérique de photons N devra être de l'ordre de $(1/h\nu)\Pi/c$ (nous préciserons cette discussion très qualitative au prochain paragraphe dans le cas très simple de l'onde plane). L'impulsion de chaque photon (de masse nulle) étant $h\nu/c$, la densité d'impulsion est $Nh\nu/c = \Pi/c^2$. On retrouve bien l'expression de g .

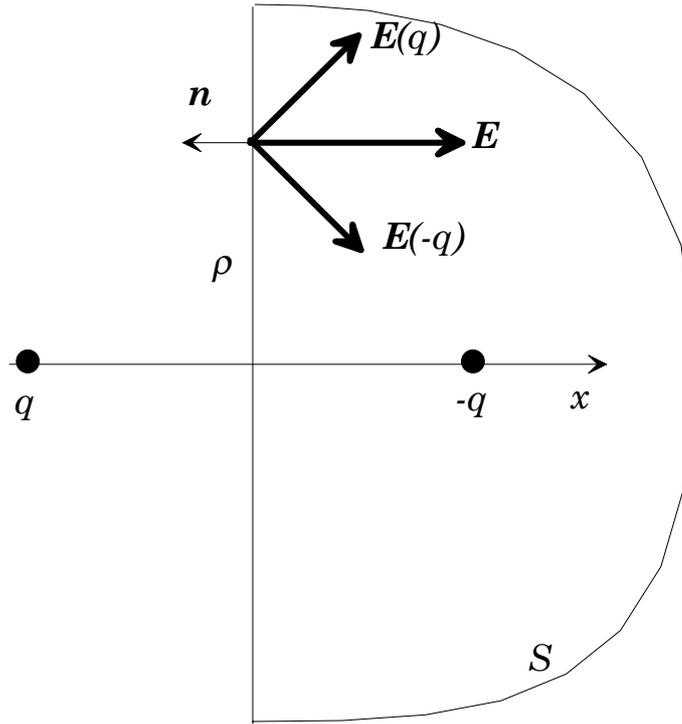


Figure 4.1: La force de Coulomb entre deux charges ponctuelles par le bilan de quantité de mouvement. Deux charges opposées sont disposées symétriquement par rapport au plan zOy . Le champ résultant est, sur ce plan, parallèle à Ox . On établit le bilan d'impulsion dans le demi-espace $x > 0$, limité par la surface S .

4.3.3 Applications

Un problème d'électrostatique

Nous allons appliquer les bilans d'énergie–quantité de mouvement à quelques situations simples. Nous allons d'abord montrer que l'équation bilan de quantité de mouvement contient simplement la force de Coulomb. Ce calcul montre que ces équations bilans peuvent présenter un intérêt en dehors des phénomènes purement propagatifs. La figure 4.1 illustre notre problème. Deux charges q et $-q$ sont situées sur l'axe Ox en $-a$ et a respectivement. Il n'échappera à personne que le calcul direct de la force de Coulomb entre ces particules ne présente aucune difficulté. Nous allons aborder ce problème par une voie un peu plus difficile. Ecrivons l'équation bilan de quantité de mouvement pour le volume \mathcal{V} correspondant au demi-espace $x > 0$. Dans tout ce volume, le champ magnétique est nul. La densité de quantité de mouvement du champ est donc nulle. La variation de la quantité de mouvement de la matière, $d\mathbf{P}_m/dt$, est simplement celle de la charge $-q$ et coïncide avec la force de Coulomb \mathbf{F} subie par cette charge. Notons que le système de charges en l'absence d'autres forces n'est manifestement pas en équilibre. On peut donc écrire la force de Coulomb sous la forme:

$$\mathbf{F} = - \int_S \mathcal{T} \cdot d\mathbf{S} . \quad (4.103)$$

La surface \mathcal{S} bordant le volume \mathcal{V} est simplement constituée du plan $x = 0$ et d'une demi sphère à l'infini. Le champ total étant à grande distance celui d'un dipôle, il décroît avec la distance R comme $1/R^3$. Le tenseur de Maxwell, proportionnel au carré du champ, décroît comme $1/R^6$. L'intégrale sur la demi sphère de l'infini est donc nulle. Sur le plan médiateur, le champ électrique est dirigé selon \mathbf{u}_x : $\mathbf{E} = E(\rho)\mathbf{u}_x$. Son module à une distance ρ de l'axe vaut $E(\rho) = (q/2\pi\epsilon_0)a/(a^2 + \rho^2)^{3/2}$. Il est facile alors de montrer que, $d\mathbf{S}$ étant orienté selon x , la seule composante du tenseur de Maxwell qui

joue un rôle est $\mathcal{T}_{xx} = -\epsilon_0 E(\rho)^2/2$. Le flux de \mathcal{T} sur \mathcal{S} se calcule alors par une intégration triviale en coordonnées polaires. On en déduit la force de Coulomb, orientée selon $-x$, dont le module a bien sûr la valeur requise. Cet exercice facile d'électrostatique nous fait comprendre l'intérêt des équations bilans d'impulsion, même dans un cas où il n'y a pas d'effets propagatifs.

Cas de l'onde plane monochromatique

Nous allons maintenant écrire les bilans d'énergie et de quantité de mouvement dans le cas beaucoup plus important de l'onde plane. Nous allons en particulier pouvoir préciser nos interprétations en termes de photons.

Nous considérons donc une onde plane monochromatique, polarisée linéairement. Elle se propage selon Oz , le champ électrique est aligné selon Ox . La pulsation de l'onde est ω . Les champs électriques et magnétiques s'écrivent donc:

$$\mathbf{E} = E_0 \mathbf{u}_x \cos(kz - \omega t) \quad (4.104)$$

$$\mathbf{B} = \frac{E_0}{c} \mathbf{u}_y \cos(kz - \omega t) \quad (4.105)$$

avec $\omega = ck$ Le vecteur de Poynting est alors dirigé selon u_z (l'énergie se propage manifestement dans cette direction) et vaut:

$$\mathbf{\Pi} = \epsilon_0 c E_0^2 \cos^2(kz - \omega t) \mathbf{u}_z . \quad (4.106)$$

On ne s'intéresse souvent qu'à la moyenne temporelle de ce vecteur (les oscillations à la fréquence 2ω n'étant pas détectables, le plus souvent, dans des mesures énergétiques). On a bien sûr:

$$\overline{\mathbf{\Pi}} = \frac{\epsilon_0 E_0^2 c}{2} \mathbf{u}_z . \quad (4.107)$$

La valeur moyenne de la densité d'impulsion est donc:

$$\overline{\mathbf{g}} = \frac{\epsilon_0 E_0^2}{2c} \mathbf{u}_z \quad (4.108)$$

La densité d'énergie électrique est égale à la densité d'énergie magnétique. La densité d'énergie électromagnétique instantanée s'écrit donc:

$$u = \epsilon_0 E_0^2 \cos^2(kz - \omega t) , \quad (4.109)$$

et sa valeur moyenne vaut:

$$\overline{u} = \frac{\epsilon_0 E_0^2}{2} . \quad (4.110)$$

On remarque immédiatement que

$$\overline{u} = \frac{\overline{\mathbf{\Pi}}}{c} \quad (4.111)$$

$$\overline{\mathbf{g}} = \frac{\overline{u}}{c} . \quad (4.112)$$

Nous pouvons comprendre quantitativement ces relations importantes en termes de densité numérique de photons. Si nous avons N photons par unité de volume, la quantité moyenne d'énergie traversant par unité de temps une surface d'aire S perpendiculaire à l'axe de propagation sera égale d'une part à $\overline{\mathbf{\Pi}}S$ et d'autre part à $\overline{u}cS$ (en un laps de temps dt un "volume" d'onde égal à $cdtS$ "traverse" la surface). On obtient ainsi la première relation: $\overline{u} = \overline{\mathbf{\Pi}}/c$. La densité numérique N de photons est $N = \overline{u}/\hbar\omega$. La densité d'impulsion $N\hbar\omega/c$ est donc bien \overline{u}/c .

Exprimons enfin le tenseur de Maxwell. Les écritures ne présentent aucune difficulté. On trouve que seule la composante \mathcal{T}_{zz} est non nulle et vaut précisément u (en valeur instantanée et donc aussi en valeur moyenne). L'onde transmet donc dans la direction z une quantité de mouvement elle aussi alignée dans la direction z . Précisons ce bilan d'impulsion en considérant le cas d'un diaphragme parfaitement absorbant d'aire S perpendiculaire à l'axe Oz . Le seul effet de ce diaphragme est d'annuler l'onde incidente immédiatement derrière lui (à plus grande distance, la diffraction par les bords du cylindre jouera un rôle et le champ électromagnétique ne sera pas exactement nul dans l'ombre géométrique du disque).

Considérons donc un volume \mathcal{V} limité par une surface \mathcal{S} entourant immédiatement le diaphragme. En première approximation, on pourra considérer que le champ est nul sur la "face arrière" de \mathcal{S} et que le champ sur la face "avant" est celui de l'onde plane non perturbée (c'est sans doute une approximation correcte si le diaphragme est parfaitement absorbant. C'est grossièrement faux s'il est réfléchissant. Nous laissons au lecteur le soin d'examiner ce dernier cas).

Ecrivons le bilan de quantité de mouvement pour le volume \mathcal{V} . Comme le volume de \mathcal{V} , qui entoure exactement le diaphragme est négligeable, la variation de la quantité de mouvement du champ est nulle. La variation de la quantité de mouvement de la matière doit être égale à la force \mathbf{F} subie par le diaphragme. Cette force est donc égale à l'opposé du flux sortant du tenseur de Maxwell. Le flux est nul partout sauf sur la surface avant du diaphragme. Il vaut alors simplement $-uS$. La force subie par le diaphragme est donc finalement

$$\mathbf{F} = uS\mathbf{u}_z . \quad (4.113)$$

L'onde électromagnétique exerce donc sur le diaphragme une pression p (la force est proportionnelle à la surface). Cette pression p est égale à la densité d'énergie du champ:

$$p = u \quad (4.114)$$

(cette relation tient entre valeurs instantanées, mais n'a d'intérêt qu'entre valeurs moyennes). En quelque sorte, nous écrivons là une équation d'état pour le rayonnement électromagnétique. On pourra se convaincre, là encore, que cette pression s'interprète en termes de collisions inélastiques des photons incidents avec le diaphragme.

Cette pression de radiation a de nombreuses manifestations. Notons tout d'abord qu'elle est en général assez faible. Si nous considérons, par exemple, un faisceau laser transportant 1W sur une surface de 1 mm², le vecteur de Poynting moyen vaut évidemment 10⁶ W/mm², la densité d'énergie et donc la pression valent 3.10⁻³ Pa. Même si la puissance est importante, la densité d'énergie est petite, parce que la vitesse de propagation est grande. Si faible soit elle, la pression de radiation permet de faire léviter des particules suffisamment petites dans un faisceau laser. Le poids d'une particule de rayon r est en effet proportionnel à r^3 alors que la force de pression de radiation varie comme r^2 . Pour r suffisamment petit, la pression de radiation l'emporte. En appliquant ce raisonnement à la pression de radiation du rayonnement solaire, on trouve que des particules suffisamment petites doivent être éjectées du système solaire. La taille limite se trouve être indépendante de la distance R au soleil. Pression de radiation et force de gravitation varient en effet toutes deux comme $1/R^2$. Avec l'ordre de grandeur de la puissance du rayonnement solaire (1.5 kW/m²) et de l'accélération de pesanteur solaire au niveau de la terre (10⁻² g), on trouve que des particules de rayon inférieur à 0.1 μm sont éjectées. Notons que, dans un modèle réaliste, il faudrait tenir compte aussi des collisions avec les particules chargées constituant le "vent solaire".