Chapter 2

La théorie de la relativité restreinte

> Daß die Elekrodynamik Maxwells [...] in ihrer Anwendung auf bewegte Körper zu Asymmetrien führt, welche den Phänomenen nicht anzuhaften scheinen, ist bekannt.
>
> A. Einstein

2.1 Introduction

La théorie de la relativité restreinte fait partie intégrale de la physique moderne, s'appliquant à toutes les interactions connues dans la physique, comme par exemple les forces nucléaires, sauf à la gravité aux plus grandes échelles où il faut prendre en compte les effets de la théorie de la relativité générale. Cependant la relativité restreinte est souvent présentée comme une généralisation de la mécanique Newtonienne, ce qui est une approche tout à fait justifiée. Or, historiquement et logiquement la relativité trouve ses origines dans l'électromagnétisme de Maxwell. Nous allons suivre donc la perspective adoptée par Einstein lui-même en 1905 dans son article “Zur Elektrodynamik bewegter Körper”.

2.1.1 Maxwell et Einstein

[Schwinger, § 3.4, 10.1, Einstein 1905]

Les résultats de la section (1.6.3) suggèrent qu'une description de la physique incorporant la mécanique classique ainsi que l'électrodynamique a été achevée. Par exemple, le membre de gauche du théorème de Poynting sous la forme (1.79) fait référence à la (densité de la) quantité de mouvement combinée des champs et un système mécanique

$$\mathbf{p}_{\text{tot}} = \mathbf{p}_{\text{méca}} + \mathbf{p}_{\text{EM}} \tag{2.1}$$

Dans le cas où le système mécanique comprend une collection de particules nous avons

$$\mathbf{p}_{\text{tot}} = \frac{1}{4\pi c}\mathbf{E} \times \mathbf{B} + \sum_{n=1}^{N} \delta(\mathbf{x} - \mathbf{x}_n) m_n \mathbf{v}_n \tag{2.2}$$

la première contribution représentant le vecteur de Poynting des champs et la deuxième la quantité de mouvement de toutes les particules, chacune avec son impulsion $\mathbf{p}_n = m_n\mathbf{v}_n$. Le théorème de Poynting exprime la conservation de la quantité de mouvement totale; ni celle des particules ni celle des champs sont conservée individuellement. Or nous avons déjà remarqué, dans la section 1.6.4 que l'impulsion des champs électromagnétiques est reliée a l'énergie comme

$$\mathbf{P}_{\text{EM}} = \frac{E}{c^2}\mathbf{v} \tag{2.3}$$

Si l'on exige que l'équation de la conservation de la quantité de mouvement des champs et particules soit cohérente, il suit que

$$\mathbf{P}_{\text{méca}} = \frac{E}{c^2}\mathbf{v}_{\text{méca}} \tag{2.4}$$

pour l'impulsion d'une particule. Ce n'est évidemment pas le cas dans la mécanique Newtonienne, où l'énergie d'une particule animée d'une vitesse $\mathbf{v}$ est donnée par

$$E_{\text{Newton}} = \frac{1}{2}mv^2 \qquad \text{et donc} \qquad \mathbf{P}_{\text{Newton}} = m\mathbf{v} \neq \frac{E_{\text{Newton}}}{c^2}\mathbf{v} \tag{2.5}$$

Nous avons donc le choix entre modifier la mécanique classique d'après Newton ou de renoncer à l'interpretation des nos résultats dans le cadre de l'électrodynamique. Comme exprimé dans la citation au debut de ce chapitre, Einstein était conscient de cette situation paradoxale, ainsi que d'autres, qui nous n'abordons pas ici. Etant donné que les modifications de la mécanique sont nécessaires seulement si la vitesse des particules est proche de celle de la lumière, un domaine peu exploré à l'époque, et étant donné que les prédictions des équations de Maxwell concernant le rayonnement électromagnétique avaient été bien vérifiées par les experiences[1] Einstein développa une nouvelle théorie de la mécanique classique, la relativité restreinte. Si l'on compare l'équation (2.4) avec la forme habituelle

$$\mathbf{P}_{\text{méca}} = m\mathbf{v}_{\text{méca}}$$

on en tire la conclusion que

$$E = mc^2 \tag{2.6}$$

ce qui est déjà une indication de ce que nous allons découvrir en suivant la logique d'Einstein.

2.2 Les postulats de la relativité restreinte

La théorie de la relativité restreinte repose sur deux postulats (Einstein, 1905):

1. *Principe de la relativité:* Les lois de la nature sont les mêmes dans tous les référentiels inertiels.

2. *Constance de la vitesse de la lumière:* La vitesse de la lumière a la même valeur dans tous référentiels inertiels

Le premier postulat introduit la notion d'un référentiel inertiel. Celui-ci est défini comme suit. Dans un référentiel inertiel, K, tout corps, sur lequel ne s'exerce aucune force[2], est en mouvement de translation rectiligne, ou au repos. Par conséquent sa vitesse est constante et aucune force n'agit sur le corps. Ainsi, un deuxième référentiel inertiel, K', ne peut être lié au premier que par translation avec une vitesse $\mathbf{v}$ constante.

[1] Principalement par Heinrich Hertz (1857- 1894) et Gugielmo Marconi (1874 - 1937).

[2] ou la résultante de toutes les forces est nulle.

Tandis que le premier postulat est presque identique au principe de la relativité d'après Galilée, le deuxième postulat est plus radical et, de plus, ne semble pas cohérent avec l'intuition habituelle. Effectivement il suit qu'un éclair de lumière émis par une source en mouvement avec vitesse v ne se propage pas à la vitesse $c+v$, ce qui serait le cas pour tout expérience impliquant des objets de notre vie quotidienne. Ce postulat a pour conséquence les célèbres effets de la relativité restreinte, notamment la notion que le temps n'est pas absolu, tel qu'exprimé dans la dilatation du temps, et l'ambiguïté de la simultanéité de deux événements. La raison pour laquelle ces effets ne semblent pas être présents dans notre expérience quotidienne est que les vitesses à l'échelle humaine sont très petites devant celle de la lumière, l'échelle naturelle de la relativité restreinte.

2.3 Les transformations de Lorentz

En relativité restreinte on a souvent recours à la notion d'*événement*. Ceci est caractérisé par le lieu, $\mathbf{x} = (x, y, z)$, et le temps, t, où il se produit. Il s'avérera utile de noter les coordonnées spatio-temporelles où s'est produit un événement par les quatre symboles

$$(ct, x, y, z) := (x^0, x^1, x^2, x^3)\,. \tag{2.7}$$

La raison pour l'inclusion du facteur de c dans cette définition sera éclaircie bientôt, mais la presence d'une vitesse suit immédiatement d'une analyse dimensionnelle.

2.3.1 L'intervalle d'espace-temps

Exprimons le principe de la constance de la vitesse de la lumière dans un language plus précis. Soient deux référentiels inertiels K et K' et $\mathbf{v}$ leur vitesse relative, et soient (ct, x, y, z) et (ct', x', y', z') les coordonnées dans les deux systèmes. Faisons coïncider les deux axes x et x' et supposons que les axes y, z soient parallèles aux axes y', z'.

Nous supposons maintenant qu'un premier événement ait lieu: un éclair de lumière apparaît à l'instant t_1 à la position (x_1, y_1, z_1). L'arrivée de cet éclair défini un deuxième événement, au point (x_2, y_2, z_2) à l'instant t_2. Il est claire que

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - c^2(t_1 - t_2)^2 = 0\,. \tag{2.8}$$

La propagation de ce signal, du point de vue du deuxième référentiel conduit à l'expression

$$(x_1' - x_2')^2 + (y_1' - y_2')^2 + (z_1' - z_2')^2 - c^2(t_1' - t_2')^2 = 0\,, \tag{2.9}$$

où les définitions des valeurs des coordonnées des deux événements dans K' sont analogues à celles dans K. Seule la constante c, la vitesse de la lumière, n'est pas modifiée, en vertu du deuxième postulat d'Einstein. La quantité apparaissant dans les deux dernières équations est appelée *l'intervalle* ('s') d'espace-temps entre deux événements,

$$s_{12}^2 = -c^2(t_1 - t_2)^2 + (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2\,. \tag{2.10}$$

En vertu du deuxième postulat de la relativité restreinte un intervalle entre deux événements vaut zéro dans tous référentiels inertiels s'il vaut zéro dans un seul référentiel. L'intervalle entre deux événements qui sont infiniment proches l'un de l'autre, s'écrit

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2\,. \tag{2.11}$$

En général l'intervalle n'est pas forcement nul entre deux événements, par exemple si l'on considère la propagation d'une particule animée d'une vitesse inférieure à celle de la lumière. On distingue trois cas

1. Intervalle de genre lumière: $s^2 = 0$, comme il était le cas pour l'éclair de lumière ci-dessus.

2. Intervalle de genre temps: $s^2 < 0$, comme il était le cas pour (2.10) si $\Delta x = 0$.

3. Intervalle de genre espace: $s^2 > 0$, comme il était le cas pour (2.10) si $\Delta t = 0$.

Nous voyons que la notion d'intervalle relativiste nous a conduit à une définition d'un type de distance, donnée formellement par le carré d'un nombre (2.10), qui peut devenir négatif. Les mathématiques qui se cachent derrière sont celles de la géométrie de l'espace de Minkowski, où plus généralement celles de la géométrie pseudo-Euclidienne, mais nous n'allons pas aborder ce sujet dans ce cours.

L'intervalle (2.10) s'écrit de façon plus commode si on défini l'objet x^μ, ($\mu = 0, 1, 2, 3$), avec

$$x^0 = ct\,, \qquad x^1 = x\,, \qquad x^2 = y\,, \qquad x^3 = z \tag{2.12}$$

et la matrice

$$[g]_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.13}$$

A l'aide de ces deux outils, l'intervalle s'écrit

$$s_{12}^2 = \sum_{\mu,\nu=0}^{3} \Delta x^\mu g_{\mu\nu} \Delta x^\nu = \Delta x^\mu g_{\mu\nu} \Delta x^\nu\,, \tag{2.14}$$

où le dernier membre de droit utilise la convention de sommation d'Einstein. Aussi nous avons indroduit

$$\Delta x^\mu = (c\Delta t, \Delta x, \Delta y, \Delta z)\,. \tag{2.15}$$

La matrice $g_{\mu\nu}$ s'appelle le tenseur métrique de l'espace Minkowski, et nous en parlerons plus dans la suite. Remarquons que nous avons fait un choix de ce qu'on appelle la signature de l'espace-temps de Minkowski, notre choix étant $(-+++)$, c'est-à-dire que la diagonale de la métrique est donnée par $(-1, +1, +1, +1)$. Un choix alternatif, également cohérent aurait été la signature $(+---)$.

2.3.2 La forme des transformations de Lorentz

Relativité d'après Galiée

Nous cherchons maintenant les transformations qui nous amènent du référentiel K au référentiel K'. Il s'agit de trouver la formule qui produit les coordonnées $(x')^\mu$ d'un événement dans le référentiel K', étant donnée que les coordonnées du même événement soient x^μ dans le référentiel K.

Dans la mécanique classique d'après Newton et Galilée on postule que le temps soit absolu, autrement dit que $t = t'$. Si la vitesse relative entre K et K' est donnée par le vecteur $\mathbf{v}$, on constate que

$$\begin{aligned} t' &= t \\ \mathbf{x}' &= \mathbf{x} - \mathbf{v}t\,, \end{aligned}$$

la formule habituelle pour une transformation Galiléenne. Pourtant cette transformation ne laisse pas invariant l'intervalle (2.10) et donc viole le deuxième postulat d'Einstein. Il est donc nécessaire de trouver la forme des transformations qui satisfait aux postulats d'Einstein et qui, dans la limite de petites vitesses se réduit a celle de Galilée. Ce dernier point découle en considérant l'invariance de l'intervalle dans la limite où toutes les vitesses du problème sont petites devant la vitesse de la lumière $v \ll c$. C'est donc la limite $c \to \infty$ de l'invariance de l'intervalle, dans laquelle on tire la conclusion que $t = t'$.

Démonstration de la transformation de Lorentz

D'après le premier postulat d'Einstein la transformation Λ des coordonnées

$$\Lambda : (ct, x, y, z) \longrightarrow (ct', x', y', z') \tag{2.16}$$

entre K et K' ne peut pas dépendre du point dans l'espace temps où l'on se trouve, ce qui est satisfait si et seulement si les applications apparaissant dans (2.16) sont linéaires. Considérons en premier lieu le cas où le référentiel K' se déplace par rapport à K le long de l'axes des x. Les relations (2.8) et (2.9) mènent à l'équation

$$c^2t^2 - x^2 = c^2(t')^2 - (x')^2\,. \tag{2.17}$$

La transformation linéaire satisfaisant cette relation est donnée par

$$\begin{aligned} ct &= ct' \cosh\eta + x' \sinh\eta \\ x &= ct' \sinh\eta + x' \cosh\eta \end{aligned}$$

où η est une grandeur quantifiant la transformation Λ. Afin d'obtenir une interprétation physique du paramètre η nous considérons le mouvement de l'origine des coordonnées du référentiel K' dans le référentiel K. Dans ce cas, $x' = 0$, et nous déduisons

$$\frac{x}{ct} = \tanh \eta \,. \tag{2.18}$$

Puisque nous suivons l'origine des coordonnées du système K' dans le référentiel K, le rapport $\frac{x}{t} = V$ est la vitesse relative du système K' par rapport à K et donc

$$\tanh \eta = \frac{V}{c} := \beta \,. \tag{2.19}$$

Le paramètre η est appèlé la rapidité. On peut donc réécrire

$$\begin{aligned} \cosh \eta &= \frac{1}{\sqrt{1-\beta^2}} := \gamma \\ \sinh \eta &= \frac{\beta}{\sqrt{1-\beta^2}} := \gamma\beta \,. \end{aligned}$$

Avec cette identification la transformation complète reliant le référentiel K au référentiel K' s'écrit

$$\begin{aligned} x^0 &= \gamma\left(x'^0 + \beta x'^1\right) \\ x^1 &= \gamma\left(x'^1 + \beta x'^0\right) \\ x^2 &= x'^2 \\ x^3 &= x'^3 \end{aligned} \tag{2.20}$$

tandis que la transformation reliant K' à K s'écrit

$$\begin{aligned} x'^0 &= \gamma\left(x^0 - \beta x^1\right) \\ x'^1 &= \gamma\left(x^1 - \beta x^0\right) \\ x'^2 &= x^2 \\ x'^3 &= x^3 \end{aligned} \tag{2.21}$$

On vérifie facilement que la transformation Galiéenne résulte de (2.21) dans la limite $c \to \infty$ comme nous l'avons requis. Si les référentiels K et K' se déplacent à la vitesse $\mathbf{v}$, pas forcement parallèle à l'axe des x, la transformation prend la forme

$$\begin{aligned} x'^0 &= \gamma\left(x^0 - \boldsymbol{\beta}\cdot\mathbf{x}\right) \\ \mathbf{x}' &= \mathbf{x} + \frac{\gamma-1}{\beta^2}(\boldsymbol{\beta}\cdot\mathbf{x})\boldsymbol{\beta} - \gamma\boldsymbol{\beta}x^0\,. \end{aligned} \tag{2.22}$$

Il est sous-entendu que l'orientation des axes dans K et K' soit la même qu'auparavant. La forme de cette transformation est démontrée dans la feuille d'exercise.

Il est souvent commode d'écrire la transformation de Lorentz sous forme d'une relation matricielle. Alors, la transformation Λ, s'écrit

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{2.23}$$

Soit Λ la matrice avec composantes $\Lambda^\mu{}_\nu x^\nu$ données par (2.23). A l'aide de la définition (2.12) de x^μ la transformation entre K et K' s'écrit

$$(x')^\mu = \sum_{\nu=0}^{3} \Lambda^\mu{}_\nu x^\nu := \Lambda^\mu{}_\nu x^\nu\,, \tag{2.24}$$

où nous avons utilisé la convention de sommation d'Einstein. L'équation précédente est un exemple d'un concept plus général, la transformation d'un quadrivecteur sous transformation de Lorentz. Avant de traiter ce sujet systématiquement, réexaminons la forme de la transformation en terme de la rapidité η. L'équation (2.23) s'écrit alors

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{2.25}$$

ce qui ressemble très fortement à une rotation d'un vecteur tridimensionel, $\mathbf{x} \in \mathbb{R}^3$

donnée par la matrice

$$R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.26}$$

appartenant au groupe de matrices 3×3 orthogonales, à savoir $R \in O(3)$. Remarquons que le produit vectoriel de $\mathbb{R}^3$, a la signature $(+++)$, dite Euclidienne. Or une transformation de Lorentz peut être considérée comme une généralisation d'une rotation ordinaire à un espace quadridimensionel avec signature $(-+++)$. A cause de cette signature, chaque fois qu'on tourne une cordonnée de genre $+$ vers la direction d'une cordonnée de genre $-$ (et vice versa) on remplace les fonctions trigonométriques avec les fonction hyperboliques. La rapidité joue le rôle de l'angle hyperbolique de cette rotation. L'objet x^μ joue le rôle d'un vecteur dans cet espace grace à son comportement sous la transformation (2.23) qui ressemble à celui d'un vecteur de $\mathbb{R}^3$. Tout objet satisfaisant à la règle de transformation (2.23) est dénommé un quadrivecteur. L'ensemble des relations que nous avons commencé à explorer définie ce qu'on appelle la géométrie de l'espace de Minkowski, dénommé $\mathbb{R}^{1,3}$, et nous passons maintenant à une discussion plus approfondie de cet être mathématique.

2.3.3 La géométrie de l'espace de Minkowski

L'espace de Minkowski est définie par l'ensemble des points $\{(x^0, x^1, x^2, x^3) \,:\, x^\mu \in \mathbb{R}\}$ doté d'un produit scalaire

$$\sum_{\mu=0}^{3} x^\mu y_\mu = \sum_{\mu,\nu=0}^{3} x^\mu g_{\mu\nu} y^\mu = -x^0y^0 + x^1y^1 + x^2y^2 + x^3y^3 \tag{2.27}$$

L'ensemble $x^\mu, (\mu = 0,1,2,3)$ comprend les composantes dites *contravariantes* du quadrivecteur x^μ tandis que l'ensemble $x_\mu, (\mu = 0,1,2,3)$ comprend les composantes dites *covariantes*, définies par

$$x_\mu = \sum_{\nu=0}^{3} g_{\mu\nu} x^\nu := g_{\mu\nu} x^\nu \,. \tag{2.28}$$

et de même pour tout autre quadrivecteur. La dernière égalité introduit la convention de sommation d'Einstein dans l'espace de Minkowski. Comme auparavant on somme sur chaque paire d'indices répétés, avec la condition supplémentaire qu'un indice de chaque paire doit toujours être de genre covariant et l'autre de genre contravariant. On appelle deux indices répétés, un indice du haut et l'autre du bas, des indices *muets*.

Soient x_1^μ, x_2^μ les coordonnées de deux événements et soit y^μ leur difference, $y^\mu = x_1^\mu - x_2^\mu$. Soit Λ une transformation de Lorentz d'un référentiel à l'autre

$$(y')^\mu = \Lambda^\mu{}_\nu y^\nu \,. \tag{2.29}$$

Alors l'invariance de l'intervalle

$$y^\mu g_{\mu\nu} y^\nu = (y')^\mu g_{\mu\nu} (y')^\nu \tag{2.30}$$

conduit à l'équation matricielle

$$\Lambda^T g \Lambda = g \qquad \longleftrightarrow \qquad \Lambda^\mu{}_\alpha g_{\mu\nu} \Lambda^\nu{}_\beta = g_{\alpha\beta} \tag{2.31}$$

Cette relation peut être consideré comme la définition d'une transformation de Lorentz, à l'instar d'une rotation en $\mathbb{R}^3$, sous laquelle le produit scalaire Euclidien tridimensionel reste invariant. Formellement, l'ensemble des matrices

$$\left\{ \Lambda \in \mathbb{R}^{4\times 4} \,:\, \Lambda^T g \Lambda = g \right\} := O(1,3) \tag{2.32}$$

définie le groupe de Lorentz. La notation $O(1,3)$ désigne le fait que ces transformations laissent inchangé un produit scalaire de signature $(-+++)$, tandis que le groupe des rotations de $\mathbb{R}^3$ laisse inchangé un produit scalaire de signature $(+++)$ et alors est dénommé $O(3)$.

On trouve qu'il y a deux classes des solutions de (2.32). Soit $R \in O(3)$ une matrice 3×3 orthogonale, c'est-à-dire $R^T R = \mathbf{1}$. La première classe s'écrit schématiquement

$$\Lambda = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & R & \\ 0 & & & \end{array}\right) \qquad \text{“rotation”} \tag{2.33}$$

Donc le groupe du Lorentz contient comme sous-ensemble le groupe des rotation de $\mathbb{R}^3$ agissant sur la partie spatiale d'un quadrivecteur quelconque. Les transformations de Lorentz de la deuxième classe sont appelées “boosts” et ils impliquent toujours une rotation entre une direction spatiale et le temps, comme par exemple

$$\Lambda = \left(\begin{array}{cc|cc} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \qquad \text{“boost”} \tag{2.34}$$

et de même pour des boosts selon les directions y et z au lieu de x. Ces sont des transformations appropriés au change d'un référentiel à l'autre que nous avons démontrées ci-dessus. On peut démontrer que chaque matrice Λ qui satisfait à (2.32) avec $\det(\Lambda) = +1$ peut être composée d'un boost et d'une rotation, ce qui caractérise le groupe propre de Lorentz.

Cône de lumière

Considérons, de nouveau, les différents genres d'intervalle. Soit $s^2 = -c^2t^2 + \mathbf{x}^2$ l'intervalle entre deux événements dans le référentiel K. Lorsque l'intervalle est de genre temps, $s^2 < 0$ il existe toujours un référentiel K', joignable par une transformation de Lorentz ou les événements ont lieu en un même point

$$s^2 = -c^2t^2 + \mathbf{x}^2 = -c^2(t')^2 < 0 \qquad \text{(genre temps)} \tag{2.35}$$

Si les deux événements concernent le même corps, leur intervalle est forcement de genre temps.

Lorsque l'intervalle et de genre espace, $s^2 > 0$, il existe toujours un référentiel dans

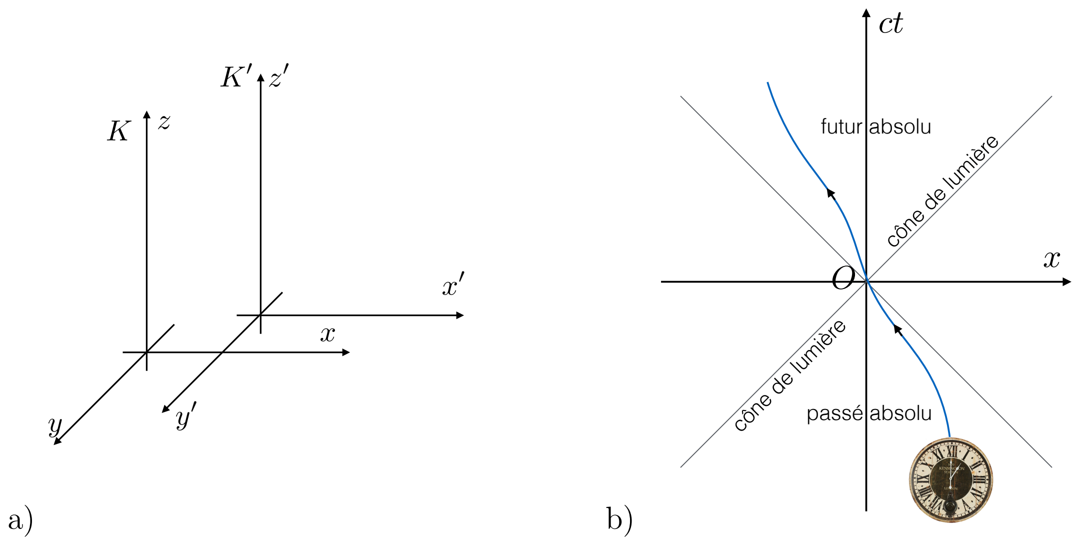

Figure 2.1: a) Deux référentiels inertiels, K et K' avec vitesse relative $\mathbf{v}$. b) Cône de lumiére d'un événement O, choisie comme origine du système des coordonnées. Tout autre événement P à l'intérieur du cône est séparé de O par un intervalle de genre temps, alors il n'existe aucun référentiel dans lequel O et P aient lieu au même temps. Tout événement Q à l'extérieur du cône est séparé de O par un intervalle de genre espace, alors il n'existe aucun référentiel dans lequel O et Q aient lieu au même point. En bleu: ligne d'univers d'une horloge animée d'un mouvement arbitraire $\mathbf{x}(t)$.

lequel les deux événements ont lieu simultanément,

$$s^2 = -c^2t^2 + \mathbf{x}^2 = \mathbf{x}'^2 > 0 \qquad \text{(genre espace)} \tag{2.36}$$

Soit O un événement quelconque, et choisissons-le comme origine du système des coordonnées. Le bord entre l'ensemble des événements séparé de O par un intervalle de genre temps et l'ensemble des événements séparé de O par un intervalle de genre espace est dénommé le cône *de lumière* (voir Figure 2.1b). Le dernier peut de façon équivalente être défini comme l'ensemble des événements séparés de O par un intervalle de genre lumière, autrement dit, l'ensemble des points tracés par un faisceau quelconque de lumière emits depuis O.

La vitesse de la lumière étant la limite pour tout corps et tout signal se propageant dans l'espace temps, seuls les événements à l'intérieur du cône de la lumière peuvent interagir avec O. Ceci est la base de la *causalité* d'après Einstein, définissant le *futur absolu* et le *passé absolu*. Le complement du futur et du passé absolus est *l'éloignement absolu*, voire l'ensemble des événements qui ne peuvent pas influencer O et qui eux-mêmes ne peuvent pas être influencés par O.

Temps propre

Comme nous l'avons mentionné auparavant, les deux postulats d'Einstein impliquent que le temps s'écoule différemment dans deux différents référentiels. Supposons qu'on observe, depuis un référentiel inertiel, une horloge animée d'un mouvement arbitraire $\mathbf{x}(t)$ (voir Figure 2.1). Pendant le temps dt l'horloge parcourt la distance $|d\mathbf{x}| = \sqrt{dx^2 + dy^2 + dz^2}$, ce qui correspond à l'intervalle

$$ds^2 = -c^2 dt^2 + d\mathbf{x}^2 = -c^2 d\tau^2 \,. \tag{2.37}$$

Le dernier membre de droit fait recours au temps τ, défini comme le temps mesuré dans un référentiel lié à l'horloge dans lequel celle-ci est instantanément au repos. Autrement dit, τ est le temps indiqué par l'horloge animée du mouvement, aussi appelé le *temps propre*. On trouve facilement que

$$dt = \gamma(\mathbf{v}) d\tau \tag{2.38}$$

où $\mathbf{v}(t)$ est la vitesse de l'horloge. Puisque $c d\tau = ds$, avec ds étant l'intervalle invariant, le temps propre est indépendant du référentiel. Il est alors commode de parametriser la ligne d'univers d'un objet à l'aide du temps propre, $X^\mu(\tau)$. Une ligne d'univers est indiqué en bleu sur la Figure 2.1b. Par intégration on trouve

$$t_2 - t_1 = \int_{\tau_1}^{\tau_2} \gamma d\tau \geq \tau_2 - \tau_2 \,. \tag{2.39}$$

On constate que le temps indiqué par une horloge en mouvement s'écoule plus lentement que celui indiqué par une horloge au repos ce qui est bien vérifié par des nombreuses expériences, dont le temps de désintégration des particules instables aux accélérateurs (p.e. au CERN).

Quadrivitesse, quantité de mouvement

Soit $X^\mu(\tau)$ la ligne d'univers d'une particule. On définit la quadrivitesse

$$U^\mu(\tau) := \frac{dX^\mu}{d\tau} = \gamma \begin{pmatrix} c \\ \mathbf{v} \end{pmatrix} \tag{2.40}$$

et, à l'aide de la quadrivitesse, la quantité de mouvement relativiste

$$P^\mu = mU^\mu(\tau) = \gamma \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix}, \tag{2.41}$$

où $E = \gamma mc^2$ et $\mathbf{p} = \gamma m\mathbf{v}$. Or nous avons retrouve la relation entre l'énergie et la quantité de mouvement d'une particule exigé par la cohérence de la nouvelle mécanique (d'après Einstein) et l'électrodynamique.

La raison pour laquelle on définit la quadrivitesse comme la dérivée par rapport au temps propre est la covariance relativiste: la grandeur $d\tau$ est un scalaire sous Lorentz, et donc la quadrivitesse hérite sa nature quadrivectorielle de X^μ sans interférence par la dérivée, ce qui n'aurait pas été le cas si l'on avait défini la quadrivitesse come une dérivée par rapport au t dans un référentiel quelconque. C'est là l'importance du temps propre dans la relativité restreinte.

Addition des vitesses

Une conséquence du deuxième postulat d'Einstein est que l'addition des vitesses Galiléenne ne peut pas être correcte. Nous cherchons ici la formule qui la remplaçe dans la relativité restreinte. Dans ce but considérons nos deux référentiels d'inertie, K et K', avec l'orientation des axes habituelle (Figure 2.1a), mais maintenant avec une vitesse relative égale à $\mathbf{v} = -v\hat{\mathbf{x}}$. La raison pour le choix du signe est que du point de vue d'un observateur dans K', l'origine de K se déplace vers x positifs, ce qui donnera lieu à l'addition (et non pas à la soustraction) des vitesses. Evidemment, ceci n'est qu'un choix commode.

La quadrivitesse étant un quadrivecteur contravriant, nous considérons donc la transformation (2.21) avec $\beta \to -\beta$. D'après la définition, nous avons

$$
\begin{aligned}
(U)^\mu &= \gamma(u) \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} \\
(U')^\mu &= \gamma(u') \begin{pmatrix} c \\ \mathbf{u}' \end{pmatrix}
\end{aligned} \tag{2.42}
$$

où $\mathbf{u}$ et $\mathbf{u}'$ sont également dirigés selon l'axe des x, $\mathbf{u} = u\hat{\mathbf{x}}$, $\mathbf{u}' = u'\hat{\mathbf{x}}$. Nous appliquons la transformation (2.20) avec $\beta(v) \to -\beta(v)$, conscient que cela met en jeu aussi le troisième facteur de gamma $\gamma(v)$. On obtient deux équations non-triviales, à savoir

$$
\begin{aligned}
\gamma(u')c &= \gamma(v)\gamma(u)\left(c + \frac{vu}{c}\right) \\
\gamma(u')u' &= \gamma(v)\gamma(u)\,(u+v)
\end{aligned} \tag{2.43}
$$

On en déduit alors la formule pour l'addition des vitesses

$$
u' = \frac{u+v}{1+\frac{uv}{c^2}}\,. \tag{2.44}
$$

Il reste apparemment une équation non-triviale, mais on vérifie (exercise) qu'elle est satisfaite identiquement, si l'on utilise (2.44). Nous remarquons que la vitesse résultante reste toujours inférieure ou égale à celle de la lumière, et si par exemple $u = c$, alors aussi $u' = c$. C'est une manifestation explicite des postulats d'Einstein, autrement dit, de la constance de la vitesse de lumière dans tous référentiels d'inertie et de la limite supérieure c pour toute vitesse. Si la vitesse $\mathbf{v}$ n'est pas parallèle à $\mathbf{u}$ on obtient par décomposition

$$
\begin{aligned}
u'_\parallel &= \frac{u_\parallel + v}{1+\frac{\mathbf{v}\cdot\mathbf{u}}{c^2}} \\
\mathbf{u}'_\perp &= \frac{\mathbf{u}_\perp}{\gamma(v)\left(1+\frac{\mathbf{v}\cdot\mathbf{u}}{c^2}\right)}
\end{aligned} \tag{2.45}
$$

Si toutes les vitesses sont petites devant c on retrouve facilement la formule Galiléenne,

$$
\mathbf{u}' = \mathbf{u} + \mathbf{v}\,, \qquad\qquad (|u|, |v| \ll c)\,. \tag{2.46}
$$

Quadrivecteurs, tenseurs de rang $r > 1$

Comme nous avons vu ci-dessus, on appelle un ensemble V^μ de quatre grandeurs, se transformant lorsqu'on change d'un référentiel à l'autre comme

$$V^\mu \longrightarrow (V')^\mu = \Lambda^\mu{}_\nu V^\nu \tag{2.47}$$

un quadrivecteur. Plus précisément, V^μ est un quadrivecteur contravariant, tandis que V_μ est un quadrivecteur covariant. La quantité

$$V^2 = V^\mu V_\mu = V^\mu g_{\mu\nu} V^\nu = (V')^\mu g_{\mu\nu} (V')^\nu \tag{2.48}$$

est une invariante de Lorentz, d'après la définition (2.31) du groupe de Lorentz. De (2.47) en tire la conclusion

$$(V')^\mu (\Lambda^{-1})^\alpha{}_\mu = V^\alpha \qquad \text{où} \qquad (\Lambda^{-1})^\alpha{}_\mu \Lambda^\mu{}_\beta = \delta^\alpha_\beta \,. \tag{2.49}$$

avec le symbol δ^α_β de Kronecker

$$\delta^\alpha_\beta = \begin{cases} 1\,, & \alpha = \beta \\ 0\,, & \alpha \neq \beta \end{cases} \tag{2.50}$$

Alors, si l'on remplace (2.49) dans l'équation (2.47) on en déduit que

$$(V')_\mu = V_\nu (\Lambda^{-1})^\nu{}_\alpha \,, \tag{2.51}$$

les composantes du vecteur V^μ se transforment selon l'inverse $(\Lambda^{-1})^T$ si les composantes du vecteur V_μ se transforment selon Λ. C'est là la raison pour le nom *contravriant* versus *covariant*. On peut démontrer (exercise) que l'inverse s'écrit de façon plus commode comme suit:

$$\Lambda_\mu{}^\alpha = (\Lambda^{-1})^\alpha{}_\mu \qquad \text{donc} \qquad \Lambda_\mu{}^\alpha \Lambda^\mu{}_\beta = \delta^\alpha_\beta \tag{2.52}$$

Un quadrivecteur auquel nous aurons souvent recours dans la suite est le quadrivecteur de derivées, dénommé ∂_μ. Il est définie comme suit

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial (ct)}, \nabla \right) \tag{2.53}$$

L'indice de ∂_μ dénomme que celui-ci est un vecteur covariant, ce qui suit d'une application de la transformation de Lorentz du quadrivecteur de la position x^μ au dénominateur (exercise). Il existe aussi la version contravariante, à savoir

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = \begin{pmatrix} -\frac{\partial}{\partial(ct)} \\ \nabla \end{pmatrix} . \tag{2.54}$$

Nous allons éclairer leur connexion ci-dessous, ainsi que la connexion entre tout autre quadrivecteur covariant et sa version contravariante.

Indices du haut, indices du bas

Commençons avec la remarque que l'inverse de la métrique $g_{\mu\nu}$, (2.27)

$$g^{\mu\alpha} g_{\alpha\nu} = \delta^\mu_\nu$$

a la même forme matricielle, $g^{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$. Notons que les indices de l'inverse de la métrique sont en haut. A l'aide de la métrique et son inverse on peut transformer un quadrivecteur contravariant en un quadrivecteur covariant et vice versa

$$V_\mu = g_{\mu\nu} V^\nu \qquad \longleftrightarrow \qquad V^\mu = g^{\mu\nu} V_\nu \tag{2.55}$$

Par exemple, pour le quadrivecteur vitesse U^μ on trouve

$$U^\mu = \gamma \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} \qquad \longleftrightarrow \qquad U_\mu = \gamma(-c, \mathbf{u}) \tag{2.56}$$

Et, pareillement, pour la dérivée $\partial_\mu \leftrightarrow \partial^\mu$, comme énoncé ci-avant.

Quelques définitions:

On appelle l'objet $T^{\mu_1 \cdots \mu_p}$ se transformant selon

$$T^{\mu_1 \cdots \mu_p} \longrightarrow (T)'^{\mu_1 \cdots \mu_p} = \Lambda^{\mu_1}{}_{\nu_1} \cdots \Lambda^{\mu_p}{}_{\nu_p} T^{\nu_1 \cdots \nu_p} \tag{2.57}$$

un *tenseur contravariant* de rang p. Remarquons qu'un quadrivecteur V^μ est donc un tenseur contravariant de rang 1. Un tenseur contravariant de rang plus haut se transforme comme une chaine de tels vecteurs, chaque indice se transformant individuellement comme V^μ. Il est donc naturel d'introduire un tenseur covariant de rang q, dénommé $T_{\mu_1 \ldots \mu_q}$, se transformant selon

$$T_{\mu_1 \ldots \mu_q} \longrightarrow (T)'_{\mu_1 \ldots \mu_p} = \Lambda_{\mu_1}{}^{\nu_1} \cdots \Lambda_{\mu_p}{}^{\nu_p} T_{\nu_1 \cdots \nu_p} \tag{2.58}$$

avec, comme nous l'avons démontré auparavant,

$$\Lambda_\alpha{}^\mu \Lambda^\alpha{}_\nu = \delta^\nu_\mu \,.$$

Finalement la combinaison de ces notions conduit à la définition d'un tenseur de rang contravariant p et de rang covariant q, dénommé $T^{\mu_1 \cdot \mu_p}{}_{\nu_1 \cdot \nu_q}$, se transformant selon

$$T^{\mu_1 \cdots \mu_p}{}_{\nu_1 \cdots \nu_q} \longrightarrow (T')^{\mu_1 \cdots \mu_p}{}_{\nu_1 \cdots \nu_q} = \Lambda^{\mu_1}{}_{\alpha_1} \cdots \Lambda^{\mu_p}{}_{\alpha_p} \Lambda_{\mu_1}{}^{\beta_1} \cdots \Lambda_{\mu_q}{}^{\beta_q} T^{\alpha_1 \cdots \alpha_p}{}_{\beta_1 \cdots \beta_q} \,. \tag{2.59}$$

A l'instar d'un quadrivecteur on peut monter et descendre un indice quelconque d'un tenseur de rang (p, q) pour en produire un tenseur du rang $(p + 1, q - 1)$ ou $(p - 1, q + 1)$. Par exemple

$$g_{\nu\mu} T^{\mu_1 \cdots \mu \cdots \mu_p}{}_{\nu_1 \cdots \nu_q} = T^{\mu_1 \cdots}{}_{\nu}{}^{\cdots \mu_p}{}_{\nu_1 \cdots \nu_q} \,. \tag{2.60}$$

Au premier regard on pourrait avoir l'impression que ses définitions soient assez compliqués, mais il ne s'agit là que d'une généralisation de vecteur à un objet avec une règle de transformation répétée plusieurs fois. Toutefois dans ce cours nous n'utilisons pas de tenseurs de rang supérieur à deux.

Résumé

Rassemblons quelques remarques utiles qui suivent de nos définitions jusqu'ici.

1. **Maxwell → Einstein**: l'électrodynamique de Maxwell nous a conduit à une réflexion sur la transformation entre deux référentiels K et K', ce qui nous a

amène à la forme des transfomations de Lorentz. Le principe de l'invariance de l'intervalle peut être exprimé, de façon géométrique, comme l'invariance d'un produit scalaire (2.27) dans l'espace de Minkowski. Ainsi on définit géometriquement le groupe de Lorentz comme la condition d'invariabilité de ce produit scalaire.

2. **Tenseurs**: Les objets naturels dans la relativité restreinte sont ceux qui se transforment de façon bien définie sous une transformation de Lorentz. Un tel être est représenté par des tenseurs de rang (p, q), comme par exemple les coordonnées d'un événement, qui sont représentées par un quadrivecteur contravariant, x^μ, c'est-à-dire, par un tenseur de rang $(0, 1)$.

3. **Covariance I**: L'avantage de cette formulation plus abstraite de la relativité restreinte est qu'il est immédiatement claire comment un objet se transforme sous Lorentz selon le nombre et la position des indices. On parle donc d'une formulation covariante (= se variant selon des règles simples et connues) de la physique. En particulier, chaque quantité avec tous indices contractés ("muets"), reste inchangé, et par consequent est une quantité scalaire, l'exemple le plus courant étant le carré d'un vecteur $V^2 = V^\mu V_\mu$.

4. **Covariance II**: Les tenseurs aussi servent à formuler facilement des équations qui prennent la même forme dans tous référentiels liés par des transformations de Lorentz, c'est-à-dire des équations qui satisfont aux postulats d'Einstein. Toute équation qui comprend du même nombre et type d'indices de chaque coté est automatiquement une équation cohérente dans la relativité restreinte. D'ici nous pouvons écrire une théorie qui respecte la relativité restreinte en suivant deux règles simples.

   (a) Utiliser que des quantité tensorielles

   (b) Le rang tensoriel de chaque expression dans une équation doit être le même

Nous allons formuler l'électromagnétisme de façon covariante dans la suite de ce chaptire, ce qui est la forme la plus naturelle de la théorie, vu qu'elle implique la relativité restreinte logiquement.

2.3.4 L'électrodynamique covariante

Nous avons commencé ce cours avec la définition de la charge et du courant, les sources des champs $\mathbf{E}$ et $\mathbf{B}$. Dans le but d'écrire les équations de l'électrodynamique sous forme covariante, il s'avérera utile de commencer au même endroit et donc de trouver un objet covariant qui comprend la densité de charge et du courant.

Le quadrivecteur courant et le quadripotentiel

Définisson le *quadrivecteur courant*

$$j^\mu = \begin{pmatrix} c\rho \\ \mathbf{j} \end{pmatrix} . \tag{2.61}$$

A l'aide de cette grandeur, l'équation de conservation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0$$

s'écrit

$$\partial_\mu j^\mu = 0 \, . \tag{2.62}$$

Le membre de droite étant un scalaire, et ∂_μ étant un quadrivecteur covariant, pour que cette équation soit cohérente, il suit que j^μ est un quadrivecteur contravariant. Un autre approche démontre la propriété quadrivectorielle de j^μ à partir de l'invariance de la charge[3] (Voir, par exemple, Jackson §11.9).

Posons aussi le quadripotentiel

$$A^\mu = \begin{pmatrix} \Phi \\ \mathbf{A} \end{pmatrix} \tag{2.63}$$

ou, avec l'indice en bas

$$A_\mu = (-\Phi, \mathbf{A}) \, . \tag{2.64}$$

[3] mais non pas la *densité* de charge.

Sous une transformation de jauge le quadripotentiel se transforme

$$A_\mu \longrightarrow (A')_\mu = A_\mu + \partial_\mu \lambda\,. \tag{2.65}$$

Le terme $\partial_\mu \lambda$ du membre de droite est un quadrivecteur covariant. Pour que la formule de la transformation de jauge soit cohérente il suit alors que le quadripotentiel A_μ aussi est un quadrivecteur covariant (et alors A^μ est un quadrivecteur contravariant). A l'aide du quadripotentiel la condition de jauge de Lorenz s'écrit

$$\partial_\mu A^\mu = 0 \tag{2.66}$$

Le membre de gauche est un scalaire sous Lorentz, comme le membre du droit.

Le tenseur de Faraday

Vu que les champs $\mathbf{E}$ et $\mathbf{B}$ se transforment comme vecteurs dans $\mathbb{R}^3$ sous rotation, il est nécessaire de trouver un object généralisant ceux-ci à des grandeurs covariantes sous Lorentz. Nous avons réussi d'en achever pour le potentiel vectoriel $\mathbf{A}$, à l'aide d'une quatrième quantité, le potentiel scalaire Φ. Or, pour les champs $\mathbf{E}$ et $\mathbf{B}$, on ne trouve nul part dans la théorie des grandeurs qui pourraient servir comme quatrième composantes des champs. La solution se présente en regardant la définition des champs en terme du potentiel A^μ (qui, lui-même, est un quadrivecteur),

$$\begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A} \\ \mathbf{E} &= -\nabla\Phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\,. \end{aligned}$$

Par conséquent

$$\begin{aligned} B_x &= \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} = \partial^2 A^3 - \partial^3 A^2 \\ E_x &= -\frac{1}{c}\frac{\partial A_x}{\partial t} - \frac{\partial \Phi}{\partial x} = \partial^0 A^2 - \partial^2 A^0 \end{aligned} \tag{2.67}$$

où les derniers membres de droite utilisent la dérivée contravariante $\partial^\mu = \left(-\frac{\partial}{\partial(ct)}, \nabla\right)$. Etant donné que le quadripotentiel A^μ est un tenseur de rang $(1,0)$ et de même pour la dérivée ∂^μ, il suit que

$$F^{\alpha\beta} = \partial^\alpha A^\beta - \partial^\beta A^\alpha \tag{2.68}$$

est un tenseur de rang deux, plus précisément de rang $(2,0)$, appelé le *tenseur de Faraday*. Cet tenseur est antisymmétrique, c'est-à-dire

$$F^{\alpha\beta} = -F^{\beta\alpha} \qquad \text{équation matricielle:} \qquad F = -F^T\,. \tag{2.69}$$

Sous forme matricielle, le tenseur de Faraday s'écrit

$$F^{\alpha\beta} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix} \tag{2.70}$$

tandis que le même tenseur avec deux indices en bas s'écrit

$$F_{\alpha\beta} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix} \tag{2.71}$$

ce qui suit du tenseur de Faraday (avec les deux indices en haut) par le remplacement $\mathbf{E} \to -\mathbf{E}$.

Remarque: La raison derrière le change du signe du champ **E** en passant de $F^{\alpha\beta}$ à $F_{\alpha\beta}$ est qu'il correspond aux composantes de $F^{\alpha\beta}$ avec un seul indice temporel ("0"),

$$\mathbf{E}^i = F^{0i}\,,$$

en lien avec la signature $(-+++)$ de la métrique. Il est donc nécessaire de faire attention chaque fois qu'on monte ou descend un indice temporel. C'est la raison pour laquelle j'ai choisi la signature $(-+++)$ et non pas la signature $(+---)$ dans laquelle tous les trois indices spatiaux pourraient changer leur signe.

Le tenseur dual de F

Une autre grandeur utile est le tenseur dual $\mathcal{F}^{\alpha\beta}$ de $F^{\alpha\beta}$. Introduisions d'abord le tenseur de rang 4 totalement antisymmétrique

$$\varepsilon^{\alpha\beta\gamma\delta} = \begin{cases} +1 & (\alpha\beta\gamma\delta) \text{ est une permutation paire de } (0123) \\ -1 & (\alpha\beta\gamma\delta) \text{ est une permutation impaire de } (0123) \\ 0 & \text{si deux indices sont égaux} \end{cases} \tag{2.72}$$

Par exemple $\varepsilon^{1023} = -1$, $\varepsilon^{2301} = +1$, mais $\varepsilon_{0123} = -1$ (exercise). Etudions les propriétés de cet objet naturel dans l'espace Minkowski. Soit Λ une transformation de Lorentz, satisfaisant $\det(\Lambda) = +1$. D'ailleurs, on dénomme l'ensemble de telles transformations le groupe de Lorentz propre[4], $L_+^\uparrow$.

D'après les règles de transformation des tenseur de rang 4,

$$(\varepsilon')^{\alpha\beta\gamma\delta} = \Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \Lambda^\gamma{}_\rho \Lambda^\delta{}_\sigma \varepsilon^{\mu\nu\rho\sigma} \,. \tag{2.73}$$

De cette expression on tire la conclusion que le tenseur transformé hérite l'antisymétrique totale du tenseur original. Mais tout tenseur totalement antisymétrique dans $\mathbb{R}^{1,3}$ est forcement proportionnel à $\varepsilon^{\alpha\beta\gamma\delta}$ et il reste que de trouver la constante de proportionnalité. Pour ceci, considérons

$$(\varepsilon')^{0123} = \Lambda^0{}_\mu \Lambda^1{}_\nu \Lambda^2{}_\rho \Lambda^3{}_\sigma \varepsilon^{\mu\nu\rho\sigma} = \det(\Lambda) = +1 \tag{2.74}$$

Par conséquent le tenseur $(\varepsilon')^{\alpha\beta\gamma\delta}$ reste invariant sous Lorentz (de $L_+^\uparrow$, strictement dit). Muni de ce tenseur invariant de Lorentz, nous définissions le dual

$$\mathcal{F}^{\alpha\beta} = \frac{1}{2}\varepsilon^{\alpha\beta\mu\nu} F_{\mu\nu} \tag{2.75}$$

Par exemple les composantes tempo-spatiales sont données par le champ magnétique

$$\mathcal{F}^{0i} = \frac{1}{2}\varepsilon^{0ijk} F_{jk} = \epsilon^{ijk}\partial_j A_k = B^i \tag{2.76}$$

Le deuxième epsilon dans cette expression est le symbol de Levi-Civita que nous

[4] Il y a des transformation de Lorentz avec $\det(\Lambda) = -1$. Celles-ci sont liés aux transformations propres par inversion spatiales (et du temps), mais nous n'avon besoin que du groupe propre dans ce cours.

avons rencontré au debut de ce cours. Le tenseur dual a la representation matricielle

$$\mathcal{F}^{\alpha\beta} = \begin{pmatrix} 0 & B_x & B_y & B_z \\ -B_x & 0 & -E_z & E_y \\ -B_y & E_z & 0 & -E_x \\ -B_z & -E_y & E_x & 0 \end{pmatrix} \tag{2.77}$$

On en déduit aisément que le tenseur dual peut être obtenu à partir du tenseur du Faraday en remplaçant

$$\left.\begin{array}{l} \mathbf{E} \longrightarrow \mathbf{B} \\ \mathbf{B} \longrightarrow -\mathbf{E} \end{array}\right\} \qquad \text{(dualité électromagnétique)} \tag{2.78}$$

Transformations de E et B

Muni d'un tenseur de rang deux, comprenant des champs électriques et magnétiques, F, on peut en déduire les transformations de Lorentz des champs sachant la transformation du tenseur lui-même. La dernière, comme nous avons mentionnée ci-avant suit aisément de la structure des indices du tenseur. Soit Λ une transformation de Lorentz reliant le référentiel K au référentiel K'. Alors

$$F^{\alpha\beta} \longrightarrow (F')^{\alpha\beta} = \Lambda^{\alpha}{}_{\mu}\Lambda^{\beta}{}_{\nu}F^{\mu\nu} \tag{2.79}$$

Cette équation s'écrit $F' = \Lambda F \Lambda^T$ sous forme matricielle qui peut être commode pour le calcul explicit. De même[5]

$$\mathcal{F}^{\alpha\beta} \longrightarrow (F')^{\alpha\beta} = \Lambda^{\alpha}{}_{\mu}\Lambda^{\beta}{}_{\nu}\mathcal{F}^{\mu\nu} \tag{2.80}$$

Les transformations de Lorentz comprennent comme sousensemble les rotations spatiales, écrit sous la forme (2.33) avec une matrice tridimensionelle de rotation R pour les composantes spatiales. Pour une telle transformation de Lorentz on trouve

$$(E')^i = (F')^{0i} = \Lambda^0{}_0\Lambda^i{}_jF^{0j} = R^i{}_jE^j\,. \tag{2.81}$$

En vertu de la loi de transformation de $\mathcal{F}$ ci-avant et le fait que dans $\mathcal{F}$ le champ

[5] On déduit cette loi de transformation de celle de $F^{\mu\nu}$ en sachant que le tenseur $\varepsilon^{\alpha\beta\gamma\delta}$ reste invariant.

B prend la place de **E**, il suit que le champ magnétique se transforme de manière pareille sous Lorentz. On constate que

$$\begin{aligned}\mathbf{E} \longrightarrow \mathbf{E}' &= R\mathbf{E} \qquad \text{(rotation)} \\ \mathbf{B} \longrightarrow \mathbf{B}' &= R\mathbf{B} \qquad \text{(rotation)} \end{aligned} \tag{2.82}$$

Pourtant les transformations de Lorentz comprennent aussi les boosts, les rotations hyperboliques de l'espace de Minkowski. Soit Λ un boost donné par la matrice (2.34). Etudions alors quelques composantes du tenseur de Faraday donnant lieu aux transformations de **E** et **B**.

$$\begin{aligned}E'_x = F'^{01} &= \Lambda^0{}_\mu \Lambda^1{}_\nu F^{\mu\nu} \\ &= \Lambda^0{}_0 \Lambda^1{}_1 F^{01} + \Lambda^0{}_1 \Lambda^1{}_0 F^{10} \\ &= \gamma^2(1-\beta^2)E_x = E_x \end{aligned} \tag{2.83}$$

La composante parallèle à la direction du boost reste inchangé. Cependant,

$$\begin{aligned}E'_y = F'^{02} &= \Lambda^0{}_\mu \Lambda^2{}_\nu F^{\mu\nu} \\ &= \Lambda^0{}_0 \Lambda^2{}_2 F^{02} + \Lambda^0{}_1 \Lambda^2{}_2 F^{12} \\ &= \gamma(E_x - \beta B_z)\,. \end{aligned} \tag{2.84}$$

et, de façon pareille pour

$$\begin{aligned}B'_z = F'^{12} &= \Lambda^1{}_\mu \Lambda^2{}_\nu F^{\mu\nu} \\ &= \Lambda^1{}_0 \Lambda^2{}_2 F^{02} + \Lambda^1{}_1 \Lambda^2{}_2 F^{12} \\ &= \gamma(B_z - \beta E_y)\,. \end{aligned} \tag{2.85}$$

Ce qui semblait être qu'un champ électrique dans le référentiel K implique aussi un champ magnétique du point de vue du référentiel K' et de même pour un champ purement magnétique dans K, qui se melange avec un champ électrique du point de vue du référentiel K'. Si l'on répète l'exercise ci-dessus pour les autres composantes, ou bien par le calcul matricielle $\Lambda F \Lambda^T$ on trouve (exercise)

$$\begin{aligned} E'_x &= E_x & B'_x &= B_x \\ E'_y &= \gamma(E_y - \beta B_z) & B'_y &= \gamma(B_y + \beta E_z) \\ E'_z &= \gamma(E_z + \beta B_y) & B'_z &= \gamma\left(B_z - \beta E_y\right) \end{aligned} \tag{2.86}$$

En comparant la forme de ces transformations on s'aperçoit que la transformation de **B** suit à partir de la forme de la transformation de **E** avec le remplacement (2.78) et c'est ne pas par hasard: Etant donné que F et $\mathcal{F}$ subissent la même loi de transformation sous Lorentz, ce fait est une simple conséquence de la dualité électromagnétique. Cette méthode est souvent utile, car il suffit de démontrer une certain propriété soit pour **E**, soit pour **B**, quel que soit la démonstration la plus simple, et puis on déduit aisément la propriété analogue pour **B** ou **E**. En fait, nous avons déjà utilisé cette astuce, quand nous avons déduit la transformation sous rotation du champ **B**. Je laisse comme exercise de trouver l'équation (2.82) à l'aide d'un calcul direct.

2.3.5 Application: Champs d'une charge ponctuelle animée d'un mouvement rectiligne

Utilisons les résultats ci-avant pour trouver le champ électrique d'une charge q ponctuelle en mouvement rectiligne. Soit K' un référentiel se déplaçant avec une vitesse $\mathbf{v} = -v\hat{\mathbf{x}}$ par rapport au référentiel K. Soit q, la charge au repos à l'origine de K, avec un champ purement électrique, donné par l'expression isotrope

$$\mathbf{E}(\mathbf{x}) = \frac{q}{|\mathbf{x}|^2}\hat{\mathbf{x}} = \frac{q}{|\mathbf{x}|^3}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{2.87}$$

Cette charge se déplace avec vitesse $\mathbf{v}$ dans le référentiel K'. Nous avons déjà développé les transformations de Lorentz qui s'appliquent à cette situation: les coordonnées se transforment selon (2.20), tandis que les champs se transforment selon (2.86).

$$\mathbf{E}(\mathbf{x}) \longrightarrow \mathbf{E}'(\mathbf{x}') \tag{2.88}$$

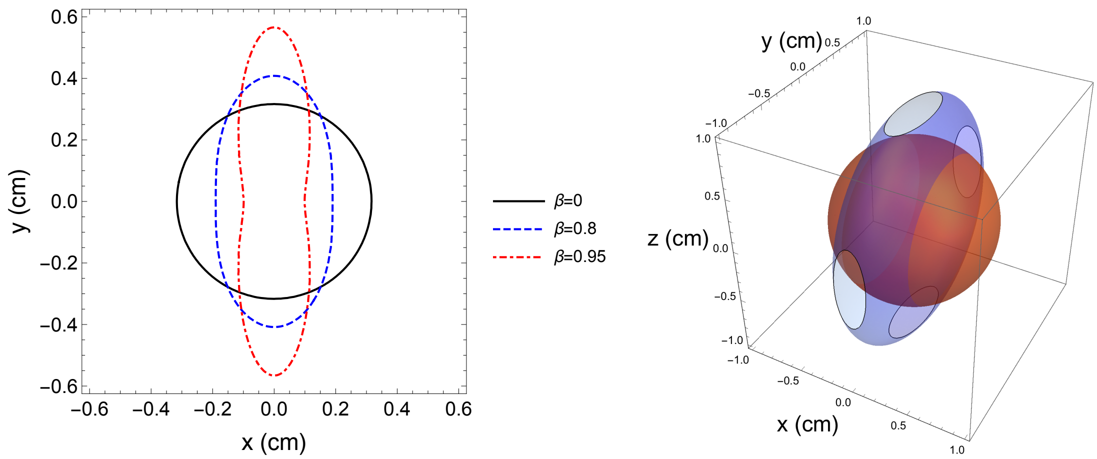

Figure 2.2: Gauche: champ électrique d'une charge $q = 1$ statC animée d'un mouvement rectiligne le long de l'axe des x avec $\beta = 0.5, 0.8, 0.95$. Dessinés sont les surfaces $|\mathbf{E}| = 10\,(\mathrm{statV/cm})$ pour les trois valeurs de β ci-avant. Droit: champ de la même charge en 3D. Dessinés sont les surfaces $|\mathbf{E}| = 2\,(\mathrm{statV/cm})$ pour $\beta = 0$ (orange) et $\beta = 0.9$ (bleu).

Le champ (2.87) donne lieu à l'expression

$$\mathbf{E}'(\mathbf{x}) = \frac{q}{(x^2+y^2+z^2)^{3/2}} \begin{pmatrix} x \\ \gamma y \\ \gamma z \end{pmatrix} \tag{2.89}$$

dans le référentiel K'. Or, une observatrice dans K' mesure les distances avec ses propres coordonnées $\mathbf{x}'$ et donc il nous reste à écrire le champ en fonction des coordonnées de K', à savoir

$$\mathbf{E}'(\mathbf{x}') = \frac{q}{[\gamma^2(x'-vt')^2+(y')^2+(z')^2]^{3/2}} \begin{pmatrix} \gamma(x'-vt') \\ \gamma y' \\ \gamma z' \end{pmatrix} \tag{2.90}$$

Le champ dans le référentiel K' est émis depuis la position de la charge $\mathbf{x}'(t') = vt'$, ce qui n'étonnerai personne. Cependent il n'est plus isotrope à cause du préfracteur. Regardons le dénominateur au temps $t' = 0$

$$\begin{aligned}\gamma^2(x')^2 + (y')^2 + (z')^2 &= (\gamma^2 - 1)(x')^2 + |\mathbf{x}'|^2 \\ &= (\beta^2\gamma^2\cos^2\vartheta' + 1)|\mathbf{x}'|^2 \\ &= (1 - \beta^2\sin^2\vartheta')\gamma^2|\mathbf{x}'|^2 \end{aligned} \tag{2.91}$$

Ici nous avons défini l'angle ϑ' entre le vecteur $\mathbf{x}'$ et l'axe des x'. A l'aide ce cette identitée le champ électrique s'écrit (au temps $t' = 0$)

$$\mathbf{E}'(\mathbf{x}') = \frac{1}{\gamma^2(1 - \beta^2\sin^2\vartheta')^{3/2}}\frac{q}{|\mathbf{x}'|^2}\hat{\mathbf{x}}' \tag{2.92}$$

C'est évidemment le champ isotrope Coulombien d'une charge ponctuelle multiplié par un facteur relativiste anisotrope. Le module s'écrit

$$|\mathbf{E}'(\mathbf{x}')| = \frac{1}{\gamma^2(1 - \beta^2\sin^2\vartheta')^{3/2}}\frac{q}{|\mathbf{x}'|^2} \tag{2.93}$$

et est dessiné dans la Figure 2.2.

Equations de Maxwell sous forme covariante

Nous souhaitons trouver la forme covariante des équations de Maxwell. Parce que celles-ci sont de deuxième ordre en A^μ on s'attend qu'ils impliquent la dérivée ∂_μ du tenseur de Faraday, car $F \sim \partial A$. Enonçons tout d'abord les équations sous forme covariante, et après démontrerons qu'ils donnent lieu aux équations dans leur forme habituelle. Les voici:

$$\partial_\alpha F^{\alpha\beta} = -\frac{4\pi}{c}j^\mu \tag{2.94}$$

$$\partial_\alpha \mathcal{F}^{\alpha\beta} = 0\,. \tag{2.95}$$

Les équations sont ici présentées sous forme covariante, ce qui n'implique pas qu'elles soient invariantes, en fait elles ne le sont pas. Autrement dit, tandis que la forme des équations reste la même quelque soit le référentiel d'inertie, l'interpretation en termes des champs **E** et **B** diffère d'un référentiel à l'autre. Effectivement les transformations de Lorentz servent à lier des différents phénomènes physiques comme étant déterminés par les mêmes équations. C'est là la signification d'une théorie covariante, autrement dit, une théorie qui satisfait aux postulats d'Einstein.

Les composantes spatio-temporelles des équations impliquant le tenseur de Faraday s'écrivent

$$\begin{aligned} \partial_i F^{i0} &= -\frac{4\pi}{c} j^0 \qquad \Rightarrow \qquad \nabla \cdot \mathbf{E} = 4\pi\rho \\ \partial_\alpha F^{\alpha i} &= -\frac{4\pi}{c} j^i \qquad \Rightarrow \qquad \frac{1}{c}\frac{\partial}{\partial t}\mathbf{E} - \nabla \times \mathbf{B} = -\frac{4\pi}{c}\mathbf{j} \end{aligned} \tag{2.96}$$

Les dernières expressions sont évidemment équivalentes aux équations inhomogènes de Maxwell dans leur forme habituelle. Pour les équations homogènes on a

$$\begin{aligned} \partial_i \mathcal{F}^{i0} &= 0 \qquad \Rightarrow \qquad \nabla \cdot \mathbf{E} = 0 \\ \partial_\alpha \mathcal{F}^{\alpha i} &= 0 \qquad \Rightarrow \qquad \frac{1}{c}\frac{\partial}{\partial t}\mathbf{B} + \nabla \times \mathbf{E} = 0 \end{aligned} \tag{2.97}$$

Notons, de nouveau, que les équations impliquant le tenseur dual, sont égaux à ceux des équations analogues impliquant le tenseur de Faraday, en remplaçant $\mathbf{E} \to \mathbf{B}$, et $\mathbf{B} \to -\mathbf{E}$, si l'on neglige les sources[6].

Les équations homogènes s'écrivent de manière équivalente sous la forme

$$\partial_\alpha F_{\beta\gamma} + \partial_\gamma F_{\alpha\beta} + \partial_\beta F_{\gamma\alpha} = 0 \tag{2.98}$$

ce qu'on appèle souvent l'identité de Bianchi. Cette expression s'écoule de la définition du tenseur dual, soit

$$0 = \partial_\alpha \mathcal{F}^{\alpha\beta} = \frac{1}{2}\varepsilon^{\alpha\beta\mu\nu}\partial_\alpha F_{\mu\nu}\,. \tag{2.99}$$

Le dernier membre du droite implique que la partie totalement antisymétrique de $\partial_\alpha F_{\mu\nu}$ s'annule, ce-qui est exactement le contenu de l'identité de Bianchi (2.98).

Point final de cette section, la conservation du courant. En vertu de sa définition

[6] Ceci suggère que les équations seraient totalement symétrique si l'on introduisait des sources magnétiques au membre du droit de (2.95). Pourtant des telles sources ne semblent pas exister (voir série d'exercices 3).

(2.68) le tenseur de Faraday satisfait à

$$\partial_\alpha \partial_\beta F^{\alpha\beta} = 0 \qquad \Leftrightarrow \qquad \partial_\alpha j^\alpha = 0 \tag{2.100}$$

où la deuxième relation, la conservation de charge sous forme relativiste, est exigée par la cohérence de l'equation (2.94).

2.3.6 Equations de mouvement covariantes

Rappelons qu'une charge animée d'une vitesse $\mathbf{v} = c\boldsymbol{\beta}$ et émergée dans des champs $\mathbf{E}$ et $\mathbf{B}$ subit une force

$$\mathbf{F} = \frac{d\mathbf{P}}{dt} = q\left(\mathbf{E} + \boldsymbol{\beta} \times \mathbf{B}\right)$$

Il s'agit d'une équation impliquante les vecteurs de l'impulsion $\mathbf{P}$ et des champs $\mathbf{E}$ et $\mathbf{B}$. Sachant que ceux-ci forment des composantes spatiales des différents quadrivecteurs, nous aimerions écrire une équation tensorielle qui reproduit la ci-dessus dans un système d'inertie. Pourtant, le dérivée par rapport au temps t ne permet pas de passer à une expression covariante, car elle se transforme sous Lorentz de manière non-tensorielle. La solution, bien sûr est d'écrire

$$\frac{d\mathbf{P}}{d\tau} = \frac{q}{c}\left(\gamma c\mathbf{E} + \gamma \mathbf{v} \times \mathbf{B}\right) \tag{2.101}$$

à l'aide du temps propre. Rappelons les définitions de quadrivitesse et quadri-impulsion

$$U^\mu = \gamma \begin{pmatrix} c \\ \mathbf{v} \end{pmatrix}, \qquad P^\mu = mU^\mu$$

Evidemment, la force de Lorentz s'écrit

$$\frac{dP^i}{d\tau} = \frac{q}{c} F^{i\nu} U_\nu \qquad (i = 1, 2, 3)$$

ce qui a l'unique extension quadrivectorielle

$$\frac{dP^\mu}{d\tau} = \frac{q}{c} F^{\mu\nu} U_\nu\,. \tag{2.102}$$

Cette covariante loi de force contient une équation de plus, la composante temporelle. Nous en tirons la conclusion que

$$\frac{dP^0}{d\tau} = \frac{q}{c}\mathbf{E} \cdot \gamma\mathbf{v} \tag{2.103}$$

En fait nous avons déjà rencontré cette équation, ce n'est rien d'autre que l'équation du travail effectué par la charge (1.60), ècrit sous la forme

$$\frac{dE}{dt} = \frac{q}{c}\mathbf{E} \cdot \mathbf{v}\,. \tag{2.104}$$

Donc la loi de force relativiste incorpore la conservation de l'énergie ainsi que la conservation de la quantité de mouvement en une seule équation. Ceci pose la question si toutes les lois de conservation (théorème de Poynting, etc.) n'ont pas une forme relativiste nette. En fait, la réponse est oui, et nous passons maintenant à sa description.

2.3.7 Tenseur énergie-impulsion et équations covariantes de l'électrodynamique

Nous avons vu que l'équation de conservation de la charge a une forme naturelle du point de vue de la relativité restreinte. Effectivement toute équation de conservation impliquant une densité quelconque $\zeta(\mathbf{x})$ et un courant vectoriel $\mathbf{c}$ s'écrit de manière relativiste sous forme

$$\frac{1}{c}\frac{d\zeta(\mathbf{x})}{dt} + \nabla \cdot \mathbf{c} = 0 \qquad \leftrightarrow \qquad \partial_\mu c^\mu = 0 \qquad \text{avec} \qquad c^\mu = \begin{pmatrix} \zeta \\ \mathbf{c} \end{pmatrix}$$

Bien sûr il est nécessaire de démontrer que c^μ se transforme sous Lorentz de la façon d'un quadrivecteur pour que cette équation soit cohérente et ceci n'est pas toujours le cas. Dans le premier chapitre nous avons démontré la conservation d'enérgie (voir (1.63) pour se rappeler des définitions des quantités apparaissantes)

$$\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{S} = -\mathbf{j} \cdot \mathbf{E} = -\frac{dE_{\text{méca}}}{dt}$$

et de la quantité de mouvement

$$\nabla \cdot \left(\mathbf{1}\frac{E^2 + B^2}{8\pi} - \frac{\mathbf{EE} + \mathbf{BB}}{4\pi} \right) + \frac{\partial}{\partial t}\left(\frac{\mathbf{E} \times \mathbf{B}}{4\pi c} \right) = -\frac{d\mathbf{p}_{\text{meca}}}{dt} \, .$$

La quantité de mouvement étant une grandeur vectorielle, son équation de conservation utilise forcement une quantité tensorielle, le produit dyadique. En composantes,

$$\frac{1}{c^2}\frac{\partial S^j}{\partial t} + \nabla_i T^{ij} = -\frac{dp^j_{\text{méca}}}{dt} \tag{2.105}$$

avec le tenseur des contraintes

$$T_{ij} = \delta_{ij} u - \frac{E_i E_j + B_i B_j}{4\pi} \tag{2.106}$$

Comme il était le cas pour la densité de charge et le vecteur de courant, il s'avérera utile de fusionner la densité d'énergie, le vecteur de Poynting et le tenseur des contraintes de Maxwell en un seul objet, le *tenseur d'énergie-impulsion*. Celui-ci est une matrice (4×4) avec composantes

$$\Theta_{\alpha\beta} = \left(\begin{array}{c|c} u & -S_j/c \\ \hline -S_i/c & T_{ij} \end{array} \right) \, . \tag{2.107}$$

De manière explicite[7] cet objet s'écrit

$$\Theta_{\alpha\beta} = \frac{1}{4\pi}\left(F_{\alpha\mu} F_\beta{}^\mu - \frac{1}{4} g_{\alpha\beta} F_{\mu\nu} F^{\mu\nu} \right) \, . \tag{2.108}$$

On peut maintenant, à l'aide du tenseur d'énergie-impulsion, écrire la totalité des lois de conservation sous la forme

$$\partial_\alpha \Theta^{\alpha\beta} = -\frac{1}{c} F^{\beta\alpha} j_\alpha \tag{2.109}$$

ce qui est vraiment une manière habile et élégante de presenter le contenue de la section 1.6.3 en une seule équation. Le membre de droite a les composantes (au

[7] Montrer comme exercise (voir série 6) que $\Theta_{\alpha\beta}$ se réduit à la définition ci-dessus.

facteur $1/c$ près)

$$\begin{aligned} -F^{0\alpha}j_\alpha &= -F^{0i}j_i = -\mathbf{E}\cdot\mathbf{j} \\ -F^{i\alpha}j_\alpha &= -F^{i0}c\rho - F^{ik}j_k = -E^i c\rho - \epsilon^{ikl}j_k B_l \\ &= -c\left(\rho\mathbf{E} + \frac{\mathbf{j}}{c}\times\mathbf{B}\right)^i = -c\mathbf{f}^i_{\text{Lorentz}} \end{aligned} \tag{2.110}$$

En portant ses expression dans (2.109) on s'aperçoit que cette équation conduit aux lois habituelles de la conservation d'énergie (composante temporelle) et la conservation de la quantité de mouvement (composante spatiale).

Nous avons donc réussi à exprimer l'électrodynamique sous une forme manifestement relativiste. Tout-ce-que nous avons vu dans cette première partie du cours, ainsi que lors de la première année, est résumé dans les trois équations

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= -\frac{4\pi}{c}j^\nu && \text{(équations de Maxwell)} \\ \partial_\mu \Theta^{\mu\nu} &= -F^{\nu\mu}j_\nu && \text{(lois de conservation)} \\ \frac{dP^\mu}{d\tau} &= \frac{q}{c}F^{\mu\nu}U_\nu && \text{(équations de mouvement)} \end{aligned} \tag{2.111}$$

bien entendu que la définition du tenseur de Faraday est donné par

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{2.112}$$

ce qui implique les équations homogènes.