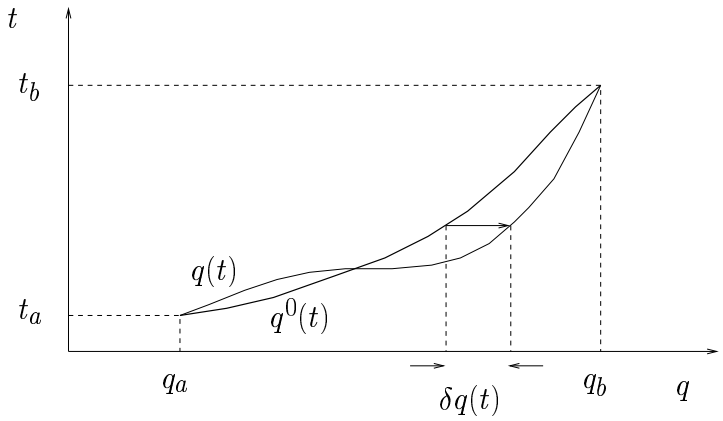
Licence Sciences de la Matière
L3 Physique - L3 Mécanique



Approche lagrangienne de la physique des milieux continus

Bertrand Berche
Groupe M,
Laboratoire de Physique des Matériaux,
Université Henri Poincaré, Nancy 1

Mécanique analytique et champs

Sommaire

Introduction 1

1 : Parois dans les milieux continus 3
2 : Application d'un principe de minimisation en électromagnétisme 11

Formalisme lagrangien 17

3 : Formulation lagrangienne de la dynamique des milieux continus 19
4 : Invariance de jauge 29

Formalisme hamiltonien 37

5 : Formulation hamiltonienne de la dynamique des milieux continus 39

Calcul tensoriel 43

6 : Eléments de calcul tensoriel 45
7 : Formulation covariante de l'électro- magnétisme et de l'électrodynamique . 59
8 : Formulation lagrangienne et relativité restreinte 69

Introduction

1

Parois dans les milieux continus

Nous allons illustrer l'application des méthodes de mécanique lagrangienne aux milieux continus par des exemples de problèmes indépendants du temps. L'accent est mis sur l'introduction de champs continus et l'expression de l'énergie totale du système en fonction du champ et de ses dérivées spatiales. Les solutions proposées ici sont obtenues par analogie avec l'approche lagrangienne pour les systèmes discrets. La généralisation à des problèmes dépendant du temps sera considérée au chapitre suivant.

Principe d'Hamilton pour les systèmes discrets

On considère un système à n degrés de liberté $n = 1, \ldots$ décrit par les coordonnées $q_n(t)$ et les vitesses $\dot{q}_n(t)$. On cherche une équation régissant l'évolution du système d'un point $A(q_n(t_a), t_a)$ à un point $B(q_n(t_b), t_b)$. La dynamique du système est entièrement déterminée par une fonction L appelée lagrangien :

$$\texttt{Lagrangien} = L(q_n(t), \dot{q}_n(t), t).$$

Le principe de moindre action d'Hamilton stipule que la trajectoire réelle du système est telle qu'une quantité $S[q_n(t)]$, appelée action, est stationnaire :

$$\begin{aligned} &\texttt{Action} = S[q_n(t)] = \int_{t_a}^{t_b} L(q_n(t), \dot{q}_n(t), t)\,\mathrm{d}t, \\ &\texttt{Principe d'Hamilton :} \quad \delta S[q_n(t)] = 0. \end{aligned} \tag{1.1}$$

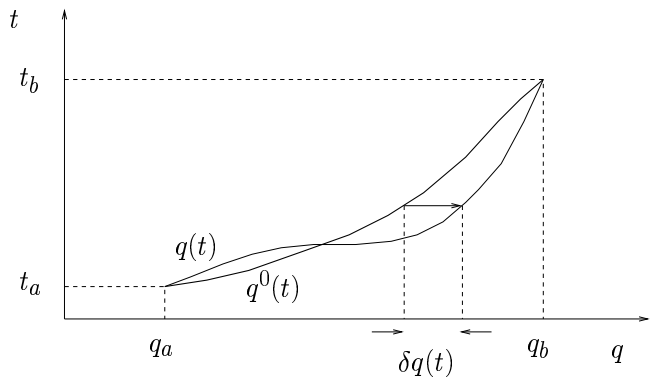
C'est une fonctionnelle [1] de $q_n(t)$ et $\dot{q}_n(t)$. Ce principe signifie que si l'on considère la trajectoire réelle notée $q_n^0(t)$ entre A et B et une autre trajectoire joignant les

[1] On reviendra sur cette notion un peu plus loin dans ce cours.

points A et B et distinctes de la trajectoire réelle d'une quantité $\delta q_n(t)$:

$$q_n(t) = q_n^0(t) + \delta q_n(t) = q_n^0(t) + \alpha\zeta_n(t),$$

alors la variation (fonctionnelle) de l'action est nulle au premier ordre en $\delta q_n(t) = \alpha\zeta_n(t)$. On peut montrer qu'en fait la trajectoire réelle correspond à un minimum pour l'action [2] .



Par généralisation de la différentielle d'une fonction de plusieurs variables (il faut sommer sur les degrés de liberté n, ce qui est implicite ici),

$$\mathrm{d}L(q_n, \dot{q}_n) = \frac{\partial L}{\partial q_n}\,\mathrm{d}q_n + \frac{\partial L}{\partial \dot{q}_n}\,\mathrm{d}\dot{q}_n,$$

la variation de S s'écrit

$$\delta S[q_n(t)] = \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial q_n}\delta q_n + \frac{\partial L}{\partial \dot{q}_n}\delta \dot{q}_n \right) \mathrm{d}t.$$

Les variables $q_n(t)$ et $\dot{q}_n(t)$ ne sont pas indépendantes, car entre t_a et t_b la connaissance de $q_n(t)$ détermine $\dot{q}_n(t)$. Pour éliminer $\delta\dot{q}_n(t) = \alpha\dot{\zeta}_n(t) = \frac{\mathrm{d}}{\mathrm{d}t}(\alpha\zeta_n(t)) = \frac{\mathrm{d}}{\mathrm{d}t}(\delta q_n(t))$, on intègre par parties le second terme :

$$\int_{t_a}^{t_b} \frac{\partial L}{\partial \dot{q}_n}\delta\dot{q}_n\,\mathrm{d}t = \left[\frac{\partial L}{\partial \dot{q}_n}\delta q_n(t)\right]_{t_a}^{t_b} - \int_{t_a}^{t_b} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_n}\right)\delta q_n\,\mathrm{d}t.$$

Le terme intégré est nul à cause des conditions aux bornes, $\delta q_n(t_a) = \delta q_n(t_b) = 0$, et il reste :

$$\delta S[q_n(t)] = \int_{t_a}^{t_b} \left[\frac{\partial L}{\partial q_n} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_n}\right)\right]\delta q_n(t)\,\mathrm{d}t.$$

[2] Corollaire : le calcul de l'action à partir d'une trajectoire différente conduit nécessairement à une borne supérieure pour S.

Comme la variation $\delta q_n(t)$ autour de la trajectoire réelle est arbitraire, pour assurer $\delta S = 0$, l'intégrant doit être nul, d'où l'on déduit les équations d'Euler – Lagrange, satisfaites par la trajectoire réelle $q_n^0(t)$:

$$\texttt{Equations d'Euler-Lagrange :} \quad \frac{\partial L}{\partial q_n} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_n}\right) = 0. \tag{1.2}$$

On obtient une équation différentielle semblable pour chaque degré de liberté.

Dans le cas d'une particule ponctuelle de masse m, d'énergie cinétique T, soumise à une énergie potentielle extérieure V, le lagrangien s'écrit :

$$\texttt{cas conservatif :}\ L = T - V$$

et l'équation d'Euler redonne le principe fondamental de la dynamique de Newton :

$$\begin{cases} T = \dfrac{1}{2}m\dot{q}^2, \quad V = U(q) \\ \dfrac{\partial L}{\partial q} = -\dfrac{\partial V}{\partial q} = F \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial L}{\partial \dot{q}}\right) = \dfrac{\mathrm{d}}{\mathrm{d}t}(m\dot{q}) = \dfrac{\mathrm{d}p}{\mathrm{d}t} \end{cases}$$

soit

$$\texttt{Equation de Newton :} \quad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \mathbf{F}.$$

Milieux diélectriques anisotropes : cristaux liquides

▷ *Energie d'anisotropie.* Un matériau anisotrope a des propriétés différentes dans les différentes directions de l'espace. C'est le cas des cristaux liquides constitués de molécules allongées orientées majoritairement, pour des raisons d'encombrement stérique, dans la direction d'un vecteur unitaire appelé directeur et noté $\mathbf{n}$. Les propriétés physiques (optiques, électriques,...) sont donc très différentes parallèlement à $\mathbf{n}$ et dans le plan perpendiculaire. Dans l'approximation linéaire, la susceptibilité est définie par un tenseur diagonal dans le repère des axes principaux. Sous un champ parallèle à $\mathbf{n}$, la réponse du milieu est plus forte que dans le plan perpendiculaire :

$$\mathbf{P}_{\parallel} = \varepsilon_0 \chi_{\parallel} \mathbf{E}_{\parallel},$$
$$\mathbf{P}_{\perp} = \varepsilon_0 \chi_{\perp} \mathbf{E}_{\perp}.$$

Pour un champ d'orientation arbitraire on écrit $\mathbf{E} = \mathbf{E}_{\parallel} + \mathbf{E}_{\perp}$, soit

$$\begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} = \varepsilon_0 \begin{pmatrix} \chi_{\perp} & 0 & 0 \\ 0 & \chi_{\perp} & 0 \\ 0 & 0 & \chi_{\parallel} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}.$$

On peut également écrire $\mathbf{E}_\perp = \mathbf{E} - (\mathbf{E}\cdot\mathbf{n})\mathbf{n}$ soit

$$\mathbf{P} = \varepsilon_0[\chi_\perp \mathbf{E} + \underbrace{(\chi_\parallel - \chi_\perp)}_{\chi_a}(\mathbf{E}\cdot\mathbf{n})\mathbf{n}].$$

Cette expression permet de définir l'énergie potentielle d'un cristal liquide de directeur $\mathbf{n}$ dans un champ $\mathbf{E}$ incliné d'un angle θ par rapport à $\mathbf{n}$:

$$U = -W = -\int_0 (\mathbf{P}\wedge\mathbf{E})\,\overrightarrow{\mathrm{d}\theta} = -\varepsilon_0\chi_a\int_0 (\mathbf{E}\cdot\mathbf{n})\mid \mathbf{n}\wedge\mathbf{E}\mid \mathrm{d}\theta = -\frac{1}{2}\varepsilon_0\chi_a E^2\sin^2\theta.$$

Cette énergie est totalement due à l'anisotropie.

▷ *Cristal liquide semi-infini.* On considère un espace semi-infini limité par un plan de verre xOz en $y=0$ et s'étendant sur $y>0$, empli d'un milieu cristal liquide. On définit la densité d'énergie potentielle d'anisotropie

$$u_a = \frac{\mathrm{d}U_a}{\mathrm{d}y} = -\frac{1}{2}\varepsilon_0\chi_a E^2\sin^2\theta$$

pour $\mathrm{d}x = \mathrm{d}z = 1$.

Si un traitement approprié (surfactant) est appliqué sur le verre, le vecteur directeur au niveau de la surface xOz est contraint à être parallèle à ce plan : $\mathbf{n}(y=0) \equiv \mathbf{u}_z$. Lorsque l'on applique un champ électrique perpendiculaire à la surface de verre, les molécules tendent à s'orienter dans le sens du champ et l'angle θ du directeur par rapport à l'axe Oz est une fonction de la coordonnée y. La surface induit alors une contrainte élastique de la forme $u_e = \frac{\mathrm{d}U_e}{\mathrm{d}y} = \frac{1}{2}k\left(\frac{\mathrm{d}\theta}{\mathrm{d}y}\right)^2$.

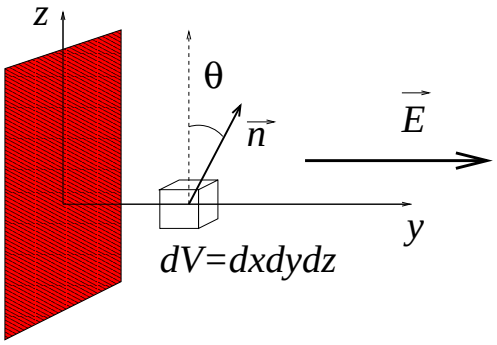


Figure 1.1 Cristal liquide semi-infini.

La densité d'énergie $u_{\mathrm{tot}}(\theta(y), \mathrm{d}\theta/\mathrm{d}y) = u_e + u_a$ est une fonction de $\theta(y)$. A l'équilibre, l'énergie totale définie par une fonctionnelle,

$$U_{\mathrm{tot}}[\theta(y)] = \int_0^\infty u_{\mathrm{tot}}(\theta(y), \mathrm{d}\theta/\mathrm{d}y)\ \mathrm{d}y$$

est minimale. Le profil $\theta(y)$ à l'équilibre est donc solution de l'équation d'Euler-Lagrange

$$\frac{\mathrm{d}}{\mathrm{d}y}\frac{\partial u_{\mathrm{tot}}}{\partial(\mathrm{d}\theta/\mathrm{d}y)} - \frac{\partial u_{\mathrm{tot}}}{\partial\theta} = 0$$

et $u_e - u_a$ est une constante. On peut justifier ces résultats par une analogie avec la mécanique analytique. On a en effet

$$U_{\text{tot}}[\theta(y)] = \int_0^\infty \left(\frac{1}{2} k \left(\frac{\mathrm{d}\theta}{\mathrm{d}y} \right)^2 - \frac{1}{2} \varepsilon_0 \chi_a E^2 \sin^2 \theta \right) \mathrm{d}y$$

à comparer à

$$S[x(t)] = \int_0^\infty \underbrace{\left(\frac{1}{2} m \left(\frac{\mathrm{d}x}{\mathrm{d}t} \right)^2 - V(x) \right)}_{L(x,\dot{x})=T-V} \mathrm{d}t$$

où la quantité conservée s'écrit

$$T + V \longrightarrow u_e - u_a = \frac{1}{2} k \left(\frac{\mathrm{d}\theta}{\mathrm{d}y} \right)^2 + \frac{1}{2} \varepsilon_0 \chi_a E^2 \sin^2 \theta.$$

Lorsque $y \to \infty$, $u_e - u_a \to \frac{1}{2} \varepsilon_0 \chi_a E^2$, de sorte que l'on obtient l'équation différentielle pour le profil d'orientation du directeur

$$\frac{\mathrm{d}\theta}{\mathrm{d}y} = \underbrace{\sqrt{\frac{\varepsilon_0 \chi_a E^2}{k}}}_{1/\xi} \cos\theta$$

qui s'intègre en

$$-\frac{y}{\xi} + \text{const} = \ln \tan \left(\frac{\pi}{4} - \frac{\theta}{2} \right).$$

La constante se détermine en $y = 0$, $\text{const} = \ln \tan \pi/4 = 0$ et la solution pour le profil d'équilibre est finalement

$$\theta_0(y) = \frac{\pi}{2} - 2 \operatorname{Arctan} \mathrm{e}^{-y/\xi}.$$

L'énergie totale vaut ainsi au minimum de la fonctionnelle

$$\begin{aligned}
\Delta U_{\text{tot}} &= U_{\text{tot}}[\theta_0(y)] - U_{\text{tot}}[\theta = \pi/2] \\
&= \int_0^\infty \mathrm{d}y \left(\frac{1}{2} \left(\frac{\mathrm{d}\theta_0}{\mathrm{d}y} \right)^2 - \frac{1}{2} \varepsilon_0 \chi_a E^2 \sin^2 \theta_0 \right) - \int_0^\infty \mathrm{d}y \left(-\frac{1}{2} \varepsilon_0 \chi_a E^2 \right) \\
&= \int_0^\infty \Bigg(\underbrace{\frac{1}{2} \left(\frac{\mathrm{d}\theta_0}{\mathrm{d}y} \right)^2}_{\frac{1}{2} k \frac{\varepsilon_0 \chi_a E^2}{k} \cos^2 \theta_0(y)} + \frac{1}{2} \varepsilon_0 \chi_a E^2 \cos^2 \theta_0 \Bigg) \\
&= \int_0^\infty \underbrace{\mathrm{d}y}_{\xi \frac{\mathrm{d}\theta}{\cos\theta_0}} \varepsilon_0 \chi_a E^2 \cos^2 \theta_0(y) \\
&= \varepsilon_0 \chi_a E^2 \xi \int_0^{\pi/2} \mathrm{d}\theta_0 \cos\theta_0 \\
&= \varepsilon_0 \chi_a E^2 \xi
\end{aligned}$$

On définit également la polarisation totale suivant Oz : $\mathcal{P}_z = \int_0^\infty (\mathbf{P}\cdot\mathbf{u}_z)\ \mathrm{d}y$ pour estimer l'effet de désorientation dû à la surface.

$$\begin{aligned}\Delta\mathcal{P}_z &= \vec{\mathcal{P}}[\theta_0(y)]\cdot\mathbf{u}_z - \vec{\mathcal{P}}[\theta=\pi/2]\cdot\mathbf{u}_z \\ &= \varepsilon_0\int_0^\infty \mathrm{d}y\chi_a(\mathbf{E}\cdot\mathbf{n})\mathbf{n}\cdot\mathbf{u}_z \\ &= \varepsilon_0\chi_a E\xi.\end{aligned}$$

Système ferromagnétique inhomogène, parois de Bloch

▷ *Fonctionnelle d'énergie.* Un système ferromagnétique fini (au dessous de sa température de transition) à l'équilibre a avantage à créer des domaines magnétiques séparés par des parois de Bloch pour minimiser son énergie magnétique (mesurée par l'extension des lignes de champ), malgré le coût en énergie interfaciale dû à la présence des parois.

Au niveau des parois de Bloch, il y a retournement progressif des spins. On cherche ici à déterminer le profil de l'orientation des moments magnétiques au niveau d'une paroi de Bloch à 180^o. Pour cela, on minimise l'énergie de la paroi. On considère donc un matériau ferromagnétique ($T < T_c$) inhomogène, ce qui implique de tenir compte de l'énergie d'anisotropie (qui dépend de l'orientation des moments). On se limite à un problème unidimensionnel en supposant qu'il y a invariance par translation parallèlement à la paroi (axes Oy et Oz) et on traite les $\boldsymbol{\mu}_i$ comme des vecteurs du plan yOz :

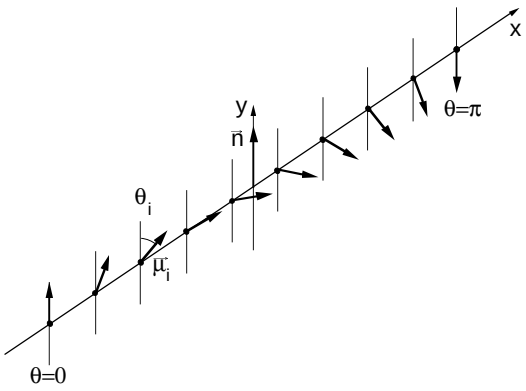


Figure 1.2 Schéma du retournement des spins au niveau d'une paroi de Bloch.

En limitant l'interaction d'échange aux premiers voisins dans la direction x, l'énergie s'écrit, pour une rangée de moments magnétiques et par unité de surface a^2 dans le plan yOz :

$$U = \frac{1}{a^2}\left[-J\sum_i \boldsymbol{\mu}_i\boldsymbol{\mu}_{i+1} - K\sum_i(\mathbf{n}\cdot\boldsymbol{\mu}_i)^2\right].$$

En notant μ le module des moments magnétiques et en introduisant les angles θ, on obtient

$$U\{\theta_i\} = -\frac{J\mu^2}{a^2}\sum_i\cos(\theta_{i+1}-\theta_i) - \frac{K\mu^2}{a^2}\sum_i\cos^2\theta_i.$$

Dans la limite continue, si l'énergie d'anisotropie est faible devant l'énergie d'échange, la variation des θ_i est lente, ce qui permet de développer

$$\frac{1}{a^2}\cos(\theta_{i+1}-\theta_i) \sim \frac{1}{a^2} - \frac{1}{a^2}\frac{(\theta_{i+1}-\theta_i)^2}{2}.$$

Or

$$\frac{\theta_{i+1}-\theta_i}{a} \simeq \frac{\mathrm{d}\theta}{\mathrm{d}x}$$

soit, en remplaçant $\sum_i$ par $\int \frac{\mathrm{d}x}{a}$ et en omettant la constante :

$$U[\theta(x)] = \int \frac{\mathrm{d}x}{a}\left[\frac{1}{2}J\mu^2\left(\frac{\mathrm{d}\theta}{\mathrm{d}x}\right)^2 - \frac{K\mu^2}{a^2}\cos^2\theta(x)\right]. \tag{1.3}$$

▷ *Profil d'aimantation.* Posons $A = \frac{J\mu^2}{a} > 0$ et $B = \frac{K\mu^2}{a^3} > 0$. On doit minimiser la fonctionnelle

$$U[\theta(x)] = \int \mathrm{d}x\left[\frac{1}{2}A\left(\frac{\mathrm{d}\theta}{\mathrm{d}x}\right)^2 - B\cos^2\theta(x)\right]$$

analogue à l'action d'une particule de masse $m \to A$ se déplaçant au cours du temps $t \to x$ le long d'une coordonnée spatiale $\to \theta$ dans un potentiel $V = B\cos^2\theta$. On en déduit que l'énergie totale de la particule est conservée,

$$T + V = \text{const.} = \frac{1}{2}A\left(\frac{\mathrm{d}\theta}{\mathrm{d}x}\right)^2 + B\cos^2\theta(x)$$

où la constante est donnée par $x \to -\infty$ où $T + V = B$, soit

$$\frac{1}{2}A\left(\frac{\mathrm{d}\theta}{\mathrm{d}x}\right)^2 = B(1-\cos^2\theta(x))$$

ou encore

$$\frac{\mathrm{d}\theta}{\mathrm{d}x} = \sqrt{\frac{2B}{A}}\sin\theta(x),$$

avec $\sqrt{\frac{2B}{A}} = \sqrt{\frac{2K}{Ja^2}} = 1/\xi$ (on prend la détermination positive, car $\frac{\mathrm{d}\theta}{\mathrm{d}x}$ augmente en même temps que $\sin\theta$.) La constante ξ donne la dimension caractéristique de la paroi. On a donc

$$\frac{\mathrm{d}\theta}{\sin\theta} = \frac{\mathrm{d}x}{\xi}$$

que l'on intègre en

$$\ln \tan \frac{\theta(x)}{2} = \frac{x}{\xi}$$

$$\theta(x) = 2 \arctan \left[\exp \left(\frac{x}{\xi} \right) \right] \tag{1.4}$$

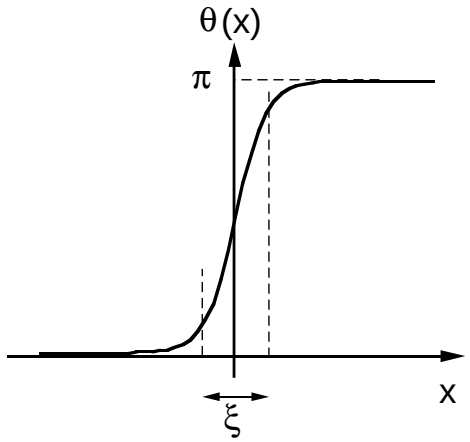


Figure 1.3 Profil de retournement des spins au niveau d'une paroi de Bloch.

Connaissant l'expression de $\theta(x)$, on pourrait ensuite calculer l'énergie de la paroi.

2

Application d'un principe de minimisation en électromagnétisme

A côté de la présentation traditionnelle de l'électromagnétisme, basée sur les équations de Maxwell, il existe une formulation variationnelle que nous élaborerons dans ce cours. Dans ce chapitre introductif, nous verrons comment l'application pratique d'un principe de minimisation peut être réalisée dans un exemple emprunté à l'électrostatique.

Equations de Maxwell dans le vide

Dans le vide (c'est-à-dire dans une région de l'espace vide, mais en présence de charges et de courants), les équations de Maxwell s'écrivent :

$$\begin{aligned} \mathbf{rot}\,\mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} & \operatorname{div}\mathbf{E} &= \frac{\rho}{\varepsilon_0} \\ \mathbf{rot}\,\mathbf{B} &= \mu_0\left(\mathbf{j} + \varepsilon_0\frac{\partial \mathbf{E}}{\partial t}\right) & \operatorname{div}\mathbf{B} &= 0 \end{aligned} \tag{2.1}$$

En régime permanent, c'est-à-dire indépendant du temps, les deux premières équations deviennent $\mathbf{rot}\ \mathbf{E} = \mathbf{0}$, soit $\mathbf{E} = -\vec{\nabla}\phi$, et $\operatorname{div}\mathbf{E} = \frac{\rho}{\epsilon_0}$ qui conduit au théorème de Gauss. C'est le régime de l'électrostatique. Le régime permanent n'impose pas l'immobilité des charges, celles-ci peuvent être mobiles et donner lieu à des courants. Les deux dernières équations de Maxwell donnent $\mathbf{rot}\,\mathbf{B} = \mu_0\mathbf{j}$ qui conduit au théorème d'Ampère, et $\operatorname{div}\mathbf{B} = 0$, soit $\mathbf{B} = \mathbf{rot}\,\mathbf{A}$. C'est le régime de la magnétostatique.

L'électromagnétisme est une théorie classique des champs. Ces champs sont les coordonnées du quadri–potentiel, c'est-à-dire le potentiel scalaire ϕ et les composantes du potentiel vecteur $\mathbf{A}$. Ils déterminent les champs $\mathbf{E}$ et $\mathbf{B}$ par

$$\mathbf{E} = -\vec{\nabla}\phi(\mathbf{r},t) - \frac{\partial \mathbf{A}(\mathbf{r},t)}{\partial t} \quad \text{et} \quad \mathbf{B} = \mathbf{rot}\,\mathbf{A}. \tag{2.2}$$

Comme ces relations ne sont pas univoques, on impose une condition supplémentaire : la jauge de Lorenz

$$\operatorname{div}\mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0, \quad \frac{1}{c^2} = \varepsilon_0\mu_0. \tag{2.3}$$

La formulation habituelle de l'électromagnétisme consiste à postuler les équations de Maxwell pour en déduire les équations de propagation des potentiels :

$$\begin{aligned} \Delta\phi - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} &= -\frac{\rho}{\varepsilon_0} \\ \vec{\Delta}\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} &= -\mu_0\mathbf{j}. \end{aligned} \tag{2.4}$$

Le condensateur cylindrique : minimisation empirique

Le calcul conventionnel de la capacité du condensateur cylindrique utilise par exemple l'énergie électrostatique U emmagasinée entre les armatures, $\frac{1}{2}CV^2$, si V est la tension, *i.e.* la différence de potentiel entre les armatures. Celle-ci se calcule aussi par la circulation du champ électrique, lui-même obtenu simplement grâce au théorème de Gauss : les cylindres sont supposés infinis pour éliminer les effets de bords.

$$a < \rho < b \qquad E(\rho) \times 2\pi\rho h = \frac{Q}{\varepsilon_0}$$

$$u = \frac{1}{2}\varepsilon_0 E^2 = \frac{Q^2}{8\pi^2\varepsilon_0\rho^2h^2}$$

$$U = h\int_a^b u 2\pi\rho\,\mathrm{d}\rho = \frac{1}{2}Q^2\frac{\ln b/a}{2\pi\varepsilon_0 h}$$

avec

$$Q = CV, \quad U = \frac{1}{2}\frac{Q^2}{C}$$

d'où

$$C_0 = \frac{2\pi\varepsilon_0 h}{\ln\frac{b}{a}}.$$

Pour trouver cette valeur exacte de C_0, il est nécessaire de trouver la fonction potentiel (ou champ électrique), ce qui n'est possible que pour des géométries simples. Supposons que l'on ne connaisse pas cette fonction ϕ^0 et que l'on prenne une fonction d'essai $\phi^{(\lambda)}(\rho)$. On peut calculer la densité $u^{(\lambda)} = \frac{1}{2}\varepsilon_0 \left(\vec{\nabla}\phi^{(\lambda)}\right)^2$ et intégrer sur l'espace compris entre les armatures pour trouver l'énergie $U^{(\lambda)}$ associée. Or, en électrostatique et en l'absence de charges, $\frac{1}{2}\varepsilon_0 \left(\vec{\nabla}\phi\right)^2$ est précisément la densité lagrangienne et son intégrale sur l'espace, U, doit donc être minimale lorsque la fonction d'essai coïncide avec le potentiel réel. On sait donc qu'en déterminant en fonction de λ le minimum de $U^{(\lambda)}$, on obtient la meilleure approximation du potentiel et donc de la capacité réelle, compatible avec la famille de fonctions d'essais choisie. Ce procédé nous fournit nécessairement une borne supérieure pour la capacité, évidemment d'autant plus proche de la réalité que le potentiel d'essai est réaliste. Au voisinage du minimum l'erreur sur la capacité est du second ordre lorsque celle qui est commise sur le potentiel est du premier ordre. Considérons une fonction d'essai en loi de puissance :

$$\phi^{(\lambda)}(\rho) = \alpha\rho^\lambda + \beta.$$

α et β sont des constantes déterminées par les conditions aux limites $\phi^{(\lambda)}(\rho = a) = \phi^0(a) = V$, $\phi^{(\lambda)}(\rho = b) = \phi^0(b) = 0$, soit

$$\alpha = \frac{V}{a^\lambda - b^\lambda}, \qquad \beta = -\frac{b^\lambda}{a^\lambda - b^\lambda}V$$

d'où $\phi^{(\lambda)}(\rho) = \frac{V}{a^\lambda - b^\lambda}\left[\rho^\lambda - b^\lambda\right]$ et $\left(\frac{\mathrm{d}\phi^{(\lambda)}}{\mathrm{d}\rho}\right)^2 = \left(\frac{\lambda V}{a^\lambda - b^\lambda}\rho^{\lambda-1}\right)^2$. On obtient l'énergie paramétrée par λ,

$$U^{(\lambda)} = 2\pi\varepsilon_0 V^2 \times \frac{1}{4}\lambda\frac{b^\lambda + a^\lambda}{b^\lambda - a^\lambda}h$$

(à comparer au résultat réel connu $U_0 = 2\pi\varepsilon_0 V^2 \frac{1}{2}\frac{1}{\ln\frac{b}{a}}h$!) d'où l'on déduit la capacité associée à la fonction d'essai choisie :

$$\frac{C^{(\lambda)}}{2\pi\varepsilon_0 h} = \frac{1}{2}\lambda\frac{b^\lambda + a^\lambda}{b^\lambda - a^\lambda}$$

que l'on peut évaluer numériquement pour plusieurs valeurs de λ et comparer à titre d'exemple à la valeur exacte :

$\frac{b}{a}$	$\frac{C_0}{2\pi\varepsilon_0 h}$	$C^{(\lambda)}/2\pi\varepsilon_0 h$				
		$\lambda = 1$	$\lambda = 2$	$\lambda = 0.5$	$\lambda = 0.1$	$\lambda = -1$
1.1	10.4921	10.5000	10.5238	10.4940	10.4921	10.5000
1.5	2.4663	2.5000	2.6000	2.4747	2.4666	2.5000
2.	1.4427	1.5000	1.6667	1.4571	1.4433	1.5000
4.	0.7213	0.8333	1.1333	0.7500	0.7225	0.8333

Cette illustration est une application "expérimentale" ou "numérique" d'un principe de minimisation. On peut bien entendu fonder de manière un peu plus rigoureuse cette méthode variationnelle.

Principe variationnel en électrostatique

Le principe de la méthode variationnelle repose sur le choix d'une "famille de fonctions d'essais" supposée représenter une quantité physique à partir de laquelle on calcule une propriété déterminée. La minimisation de la valeur correspondante par rapport à un ou plusieurs paramètres variationnels fournit les conditions qui s'avèrent les plus voisines de la réalité physique parmi les fonctions d'essais choisies.

En l'absence de charges, la minimisation de l'énergie électrostatique conduit à l'équation de Laplace. En présence de charges, le problème se complique et tout réside dans la fonctionnelle qui va jouer le rôle de l'action. Suivant Feynman, postulons le principe suivant :

Le potentiel électrostatique réel est celui qui minimise la quantité $U^*[\phi(\mathbf{r})]$ *:*

$$U^*[\phi(\mathbf{r})] = \frac{1}{2}\varepsilon_0 \int (\vec{\nabla}\phi)^2 \,\mathrm{d}^3r - \int \rho(\mathbf{r})\phi(\mathbf{r})\,\mathrm{d}^3r. \tag{2.5}$$

En effet, écrire que $\delta U^* = 0$ au premier ordre en $\delta\phi$ entraîne l'équation de Poisson : on pose

$$\phi(\mathbf{r}) = \phi^0(\mathbf{r}) + \delta\phi(\mathbf{r})$$

de sorte que $\delta\phi(\mathbf{r})$ s'annule aux limites A et B. Ecrivons

$$U^*[\phi(x,y,z)] = \int u(\phi, \vec{\nabla}\phi)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

$$u(\phi, \vec{\nabla}\phi) = -\rho(\mathbf{r})\phi(\mathbf{r}) + \frac{1}{2}\varepsilon_0(\vec{\nabla}\phi)^2$$

et écrivons la variation fonctionnelle [3] :

$$\delta U^* = \int_A^B \left[\frac{\partial u}{\partial\phi}\delta\phi + \frac{\partial u}{\partial(\vec{\nabla}\phi)}\delta(\vec{\nabla}\phi)\right] \mathrm{d}^3r$$

avec la notation abrégée

$$\frac{\partial u}{\partial(\vec{\nabla}\phi)}\delta(\vec{\nabla}\phi) = \frac{\partial u}{\partial\left(\frac{\partial\phi}{\partial x}\right)}\,\delta\left(\frac{\partial\phi}{\partial x}\right) + \frac{\partial u}{\partial\left(\frac{\partial\phi}{\partial y}\right)}\,\delta\left(\frac{\partial\phi}{\partial y}\right) + \frac{\partial u}{\partial\left(\frac{\partial\phi}{\partial z}\right)}\,\delta\left(\frac{\partial\phi}{\partial z}\right).$$

Les bornes A et B sont par exemple les deux armatures d'un condensateur entre lesquelles on calcule U^*. Comme $\vec{\nabla}\phi$ se déduit de ϕ, on élimine $\delta(\vec{\nabla}\phi) = \vec{\nabla}(\delta\phi)$ par intégration par parties (ou par le théorème de la divergence)

[3] En termes de dérivée fonctionnelle comme on le verra, on a directement

$$\frac{\delta U^*}{\delta\phi(\mathbf{r})} = \frac{\partial u}{\partial\phi} - \vec{\nabla}\frac{\partial u}{\partial(\vec{\nabla}\phi)} = 0.$$

$$\int_A^B \frac{\partial u}{\partial(\vec{\nabla}\phi)}\delta(\vec{\nabla}\phi)\,\mathrm{d}^3r = \left[\frac{\partial u}{\partial(\vec{\nabla}\phi)}\delta\phi\right]_A^B - \int_A^B \vec{\nabla}\left(\frac{\partial u}{\partial(\vec{\nabla}\phi)}\right)\delta\phi\,\mathrm{d}^3r$$

(trois termes différents sont rassemblés ici). Avec les conditions aux limites sur $\delta\phi$ (les limites peuvent d'ailleurs être à l'infini et dans ce cas $\phi \longrightarrow 0$ si la distribution de charges est d'extension non infinie), le terme intégré disparaît, il reste

$$\delta U^* = 0 = \int_A^B \left[\frac{\partial u}{\partial\phi} - \vec{\nabla}\left(\frac{\partial u}{\partial(\vec{\nabla}\phi)}\right)\right]\delta\phi\,\mathrm{d}^3r$$

d'où l'équation

$$\frac{\partial u}{\partial\phi} = \vec{\nabla}\left[\frac{\partial u}{\partial(\vec{\nabla}\phi)}\right]. \tag{2.6}$$

Cela donne ici

$$\frac{\partial u}{\partial\phi} = -\rho(\mathbf{r}) \quad \text{et} \quad \frac{\partial u}{\partial(\vec{\nabla}\phi)} = \varepsilon_0\vec{\nabla}\phi$$

soit $-\rho(\mathbf{r}) = \varepsilon_0\vec{\nabla}^2\phi$, c'est-à-dire l'équation de Poisson

$$\vec{\nabla}^2\phi = -\frac{\rho(\mathbf{r})}{\varepsilon_0}. \tag{2.7}$$

Dans cet exemple simple, on démontre l'équation de Poisson moyennant l'hypothèse qu'une certaine quantité intégrale a une variation fonctionnelle nulle. Cette dernière proposition est donc érigée en principe, puisque sa déduction logique, l'équation de Poisson, est satisfaite dans la nature. L'aspect non trivial de cette formulation variationnelle consiste à déterminer la quantité intégrale à minimiser.

Formalisme lagrangien

3

Formulation lagrangienne de la dynamique des milieux continus

La mécanique analytique propose des formulations alternatives des lois de la dynamique comme la dynamique lagrangienne ou hamiltonienne, les crochets de Poisson, ... Dans le cas des systèmes continus, la nouveauté réside avant tout dans la forme du lagrangien. En effet, le lagrangien dépend alors naturellement d'un champ et de sa dérivée temporelle, mais aussi, pour rendre compte des interactions, des dérivées spatiales du champ. Cela modifie un peu le formalisme que nous allons présenter ici.

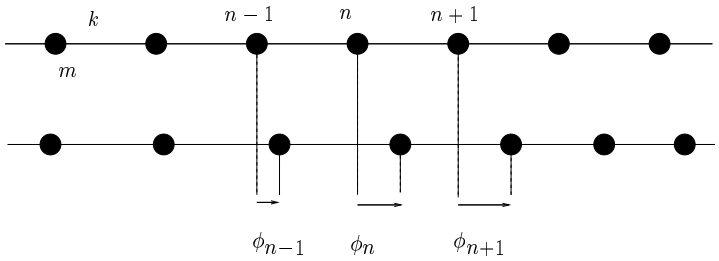
Passage du discret au continu

▷ *Lagrangien d'une chaîne de masses liées par des ressorts.* On considère un système discret de N corps en interaction, par exemple un système de masses identiques couplées par des ressorts de même raideur. On note ϕ_n l'écart à la position d'équilibre de la masse située au site n. A l'équilibre, les masses sont distantes de a. En négligeant les effets de bords (il faudrait sinon préciser les conditions aux limites) la masse en n étant couplée à ses deux voisines, l'énergie potentielle élastique vaut $\frac{1}{2}k(\phi_{n+1}-\phi_n)^2+\frac{1}{2}k(\phi_n-\phi_{n-1})^2$. Pour un système de N sites, on somme toutes les contributions

$$V=\frac{1}{2}\sum_n k(\phi_{n+1}-\phi_n)^2.$$

Pour l'énergie cinétique on a de même $\frac{1}{2}m\dot{\phi}_n^2$ pour la contribution de la masse située au site n, avec $\dot{\phi}\equiv\frac{\mathrm{d}\phi}{\mathrm{d}t}$, soit au total

$$T=\frac{1}{2}\sum_n m\dot{\phi}_n^2.$$



Le lagrangien du système est donné par

$$\begin{aligned} L(\{\phi_n(t)\},\{\dot{\phi}_n(t)\}) &= T - V \\ &= \frac{1}{2}\sum_n a\left[\frac{m}{a}\dot{\phi}_n^2 - ka\left(\frac{\phi_{n+1}-\phi_n}{a}\right)^2\right]. \end{aligned}$$

La notation $L(\{\phi_n(t)\},\{\dot{\phi}_n(t)\})$ précise que le lagrangien dépend de l'ensemble des variables $\phi_n(t)$ et de leurs dérivées temporelles. Le terme entre crochets au second membre joue le rôle d'une densité lagrangienne (ici par unité de longueur). On introduit la masse volumique (ici masse par unité de longueur) $\mu = m/a$ et le module d'Young Y. Pour un corps élastique, la loi de Hooke stipule que l'allongement par unité de longueur ξ (ici $\frac{\phi_{n+1}-\phi_n}{a}$) est proportionnel à la force appliquée F, soit $F = Y\xi$[4] . Ce qui joue le rôle de force est le terme $k(\phi_{n+1}-\phi_n)$, soit $Y = ka$. On peut re-exprimer le lagrangien comme

$$L(\{\phi_n(t)\},\{\dot{\phi}_n(t)\}) = \frac{1}{2}\sum_n a\left[\mu\dot{\phi}_n^2 - Y\left(\frac{\phi_{n+1}-\phi_n}{a}\right)^2\right]. \tag{3.1}$$

▷ *Limite continue.* Introduisons maintenant la variable $x = na$. Dans la limite où $a \to 0$ (c'est une variable microscopique), x devient une variable continue et les valeurs discrètes $\phi_n(t)$ sont décrites par un champ continu $\phi(x,t)$:

$$\begin{aligned} \phi_n(t) &\to \phi(na,t) = \phi(x,t) \\ \phi_{n\pm1}(t) &\to \phi((n\pm1)a,t) = \phi(x\pm a,t) \end{aligned}$$

On peut développer $\phi(x\pm a,t) = \phi(x,t) \pm a\phi'(x,t) + O(a^2)$ de sorte que le terme d'interaction dans le lagrangien fait apparaître une dérivée spatiale dans la limite continue :

$$\lim_{a\to0}\frac{\phi_{n\pm1}-\phi_n}{a} = \pm\phi'(x,t) \equiv \pm\frac{\mathrm{d}\phi}{\mathrm{d}x}.$$

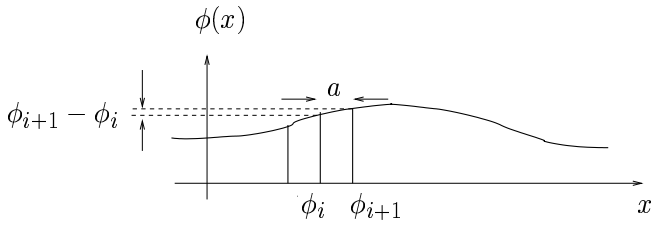
[4] C'est l'approximation linéaire d'une relation générale entre F et ξ.

Dans le même temps, la somme devient une intégrale

$$\sum_n a \to \int \mathrm{d}x$$

et le lagrangien prend la forme d'une fonctionnelle *sur l'espace*[5]

$$L[\phi(x,t)] = \frac{1}{2}\int \mathrm{d}x \left[\mu\left(\frac{\partial\phi}{\partial t}\right)^2 - Y\left(\frac{\partial\phi}{\partial x}\right)^2\right]. \tag{3.2}$$



On introduit souvent une notation abrégée pour les dérivées partielles,

$$\begin{aligned} L[\phi(x,t)] &= \frac{1}{2}\int \mathrm{d}x \left[\mu(\partial_t\phi)^2 - Y(\partial_x\phi)^2\right] \\ &= \int \mathrm{d}x\, \mathcal{L}(\partial_t\phi, \partial_x\phi) \end{aligned}$$

où il est sous-entendu que ϕ est un champ dépendant à la fois de la position et du temps. La quantité $\mathcal{L}$ est une densité lagrangienne. C'est une fonction ordinaire. Dans un cas plus général elle pourrait également dépendre de la valeur du champ lui-même,

$$\texttt{Densité lagrangienne} : \mathcal{L}(\phi(x,t), \partial_t\phi, \partial_x\phi).$$

La forme des équations du mouvement

Dans le problème discret original, les équations du mouvement sont données par application des équations d'Euler-Lagrange,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\phi}_n} - \frac{\partial L}{\partial \phi_n} = 0,$$

soit

$$m\ddot{\phi}_n = k(\phi_{n+1} - 2\phi_n + \phi_{n-1}).$$

[5] On entend par là que l'intégration porte sur l'espace et pas le temps, come par exemple dans le cas de l'action.

Lorsque l'on passe à la limite continue, on a

$$\phi_{n+1} - 2\phi_n + \phi_{n-1} \to a^2\phi''(x,t) + O(a^3),$$

d'où l'équation du mouvement $m\ddot{\phi}(x,t) = ka^2\phi''(x,t)$, soit en divisant par a,

$$\mu\partial_t^2\phi - Y\partial_x^2\phi = 0. \tag{3.3}$$

C'est une équation d'ondes se propageant à la vitesse $v = \sqrt{Y/\mu}$.

Voyons comment obtenir cette équation à partir du lagrangien. Les dérivées de la densité lagrangienne donnent $\frac{\partial\mathcal{L}}{\partial(\partial_t\phi)} = \mu\partial_t\phi$ et $\frac{\partial\mathcal{L}}{\partial(\partial_x\phi)} = -Y\partial_x\phi$, de sorte que la dynamique du problème est donnée apparemment non pas par

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial(\partial_t\phi)} - \frac{\partial\mathcal{L}}{\partial\phi} = 0 \qquad \texttt{FAUX},$$

mais par une généralisation de la forme

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial(\partial_t\phi)} + \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial\mathcal{L}}{\partial(\partial_x\phi)} - \frac{\partial\mathcal{L}}{\partial\phi} = 0 \qquad \texttt{CORRECT}.$$

Formulation lagrangienne pour les systèmes continus

Considérons toujours un problème unidimensionnel, mais défini par un lagrangien plus général dépendant a priori d'un champ continu $\phi(x,t)$, mais aussi de ses dérivées temporelle et spatiale, $\partial_t\phi(x,t)$ et $\partial_x\phi(x,t)$. L'action vaut [6]

$$S[\phi(x,t)] = \int \mathrm{d}t\, L = \int_{t_1}^{t_2}\mathrm{d}t\int_{x_1}^{x_2}\mathrm{d}x\,\mathcal{L}(\phi,\partial_t\phi,\partial_x\phi) \tag{3.4}$$

On admet maintenant le principe d'Hamilton qui stipule que la dynamique du champ est gouvernée par les équations d'Euler-Lagrange obtenues en minimisant l'action par rapport à des variations arbitraires. On paramètre le champ par un réel α, qui mesure la variation,

$$\phi(x,t;\alpha) = \underbrace{\phi(x,t;0)}_{\text{minimise } S} + \underbrace{\alpha\zeta(x,t)}_{\text{variation}}$$

Par hypothèse, la variation $\zeta(x,t)$ s'annule aux bords en t et en x, par exemple $\zeta(x,t_1) = \zeta(x,t_2) = 0$. Le calcul de la dérivée ordinaire de l'action par rapport au paramètre α conduit à

$$\frac{\mathrm{d}S}{\mathrm{d}\alpha} = \int_{t_1}^{t_2}\mathrm{d}t\int_{x_1}^{x_2}\mathrm{d}x\left[\frac{\partial\mathcal{L}}{\partial\phi}\frac{\partial\phi}{\partial\alpha} + \frac{\partial\mathcal{L}}{\partial(\partial_t\phi)}\frac{\partial}{\partial\alpha}\left(\frac{\partial\phi}{\partial t}\right) + \frac{\partial\mathcal{L}}{\partial(\partial_x\phi)}\frac{\partial}{\partial\alpha}\left(\frac{\partial\phi}{\partial x}\right)\right].$$

[6] Il s'agit cette fois d'une fonctionnelle *sur l'espace et le temps* car l'intégration porte sur les deux types de coordonnées.

L'objectif est d'obtenir l'intégrale d'une quantité en facteur de la variation $\frac{\partial\phi}{\partial\alpha} = \zeta$. De cette manière, comme cette variation est arbitraire à l'intérieur du domaine d'intégration, pour assurer une variation nulle de l'action il faudra imposer que la quantité en facteur de $\frac{\partial\phi}{\partial\alpha}$ soit nulle. Le premier terme sous l'intégrale fait déjà apparaître le facteur $\frac{\partial\phi}{\partial\alpha}$. Pour le second on procède à une intégration par parties, soit

$$u = \frac{\partial\mathcal{L}}{\partial(\partial_t\phi)} \qquad \mathrm{d}u = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial(\partial_t\phi)}\right)\mathrm{d}t$$

$$\mathrm{d}v = \underbrace{\frac{\partial}{\partial\alpha}\left(\frac{\partial\phi}{\partial t}\right)\mathrm{d}t}_{\frac{\partial}{\partial t}\left(\frac{\partial\phi}{\partial\alpha}\right)\mathrm{d}t} \qquad v = \frac{\partial\phi}{\partial\alpha}$$

qui permet d'écrire

$$\int_{t_1}^{t_2}\mathrm{d}t\int_{x_1}^{x_2}\mathrm{d}x\,\frac{\partial\mathcal{L}}{\partial(\partial_t\phi)}\frac{\partial}{\partial\alpha}\left(\frac{\partial\phi}{\partial t}\right) = \int_{x_1}^{x_2}\mathrm{d}x\left[\frac{\partial\mathcal{L}}{\partial(\partial_t\phi)}\frac{\partial\phi}{\partial\alpha}\right]_{t_1}^{t_2} - \int_{t_1}^{t_2}\mathrm{d}t\int_{x_1}^{x_2}\mathrm{d}x\,\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial(\partial_t\phi)}\right)\frac{\partial\phi}{\partial\alpha}.$$

Le terme intégré sur le temps est nul en raison de la condition sur la variation du champ aux bornes et seul subsiste la seconde intégrale. On procède de même pour la dernière intégrale de $\frac{\mathrm{d}S}{\mathrm{d}\alpha}$ où subsiste uniquement le terme

$$-\int_{t_1}^{t_2}\mathrm{d}t\int_{x_1}^{x_2}\mathrm{d}x\,\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial\mathcal{L}}{\partial(\partial_x\phi)}\right)\frac{\partial\phi}{\partial\alpha}$$

et finalement il vient

$$\frac{\mathrm{d}S}{\mathrm{d}\alpha} = \int_{t_1}^{t_2}\mathrm{d}t\int_{x_1}^{x_2}\mathrm{d}x\left[\frac{\partial\mathcal{L}}{\partial\phi} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial(\partial_t\phi)}\right) - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial\mathcal{L}}{\partial(\partial_x\phi)}\right)\right]\frac{\partial\phi}{\partial\alpha}.$$

La condition d'annulation $\frac{\mathrm{d}S}{\mathrm{d}\alpha} = 0$ impose les équations d'Euler-Lagrange,

$$\texttt{Equations d'Euler-Lagrange :}\ \frac{\partial\mathcal{L}}{\partial\phi} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial(\partial_t\phi)}\right) - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial\mathcal{L}}{\partial(\partial_x\phi)}\right) = 0, \quad (3.5)$$

où l'on voit qu'il apparaît effectivement en plus des dérivées de la densité lagrangienne par rapport au champ et à sa dérivée temporelle, un terme de dérivée par rapport aux variations spatiales du champ.

Fonctionnelle et dérivée fonctionnelle

Une fonction projette un scalaire sur un scalaire, c'est par exemple une application de $\mathbb{R}$ dans $\mathbb{R}$, notée

$$\begin{aligned} f\ :\ \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) \end{aligned}$$

Une fonctionnelle généralise cette notion [7] , mais l'espace de départ est un espace de fonctions satisfaisant certaines conditions qui assurent la convergence de la fonctionnelle. Par exemple

$$\begin{aligned} F \ : \ \mathcal{E} &\longrightarrow \mathbb{R} \\ g(\ ) &\longmapsto F[g(x)] \end{aligned}$$

Comme dans le cas des fonctions, où les physiciens utilisent par abus de langage la notation $f(x)$ pour représenter une fonction, on appelera $F[g(x)]$ ou parfois simplement $F[g]$ une fonctionnelle.

Des exemples élémentaires sont par exemple l'intégrale définie

$$I[g] = \int_{-\infty}^{+\infty} g(x)\,\mathrm{d}x,$$

la valeur d'une fonction en un point

$$F_0[g] = g(0),$$

le maximum d'une fonction sur un intervalle

$$M[g] = \max_{0\leq x\leq 1}[g(x)],$$

l'entropie de g

$$E[g] = -\int g(x)\ln|g(x)|\,\mathrm{d}x$$

(si $g(x)$ est positive et normée à l'unité, $\int \mathrm{d}x\, g(x) = 1$), ou encore l'action

$$S[q] = \int L(q(t),\dot{q}(t))\,\mathrm{d}t.$$

La fonctionnelle $I[g]$ est correctement définie si g est bornée et décroît assez vite à l'infini, $F_0[g]$ et $M[g]$ si g est continue, etc. On peu définir une fonctionnelle par une intégrale à plusieurs dimensions sur des fonctions test, comme la fonctionnelle de Ginzburg-Landau

$$F[\varphi] = \int \mathrm{d}^3r \left(\frac{1}{2}a|\varphi(\mathbf{r})|^2 + \frac{1}{4}b|\varphi(\mathbf{r})|^4 + \frac{1}{2}c|\vec{\nabla}\varphi(\mathbf{r})|^2\right).$$

Comme la dérivée est définie par la variation de la valeur de la fonction pour une variation infinitésimale dans l'espace de départ,

$$\frac{\mathrm{d}f}{\mathrm{d}x} \equiv f'(x) = \lim_{\epsilon\to 0}\frac{f(x+\epsilon)-f(x)}{\epsilon},$$

[7] Les fonctionnelles ont été introduites par le mathématicien italien Vito Volterra.

la dérivée fonctionnelle est définie par par une variation de la fonction test en un point y,

$$\frac{\delta F}{\delta g(y)} = \lim_{\epsilon \to 0} \frac{F[g(x) + \epsilon\delta(x-y)] - F[g(x)]}{\epsilon}$$

Cette définition vaut par exemple dans le cas de la fonctionnelle

$$F[g(x)] = \int_a^b f(g(x))\,\mathrm{d}x$$

où f est une fonction ordinaire (qu'on peut éventuellement appeler densité associée à F). Notons que la dérivée fonctionnelle est une fonction de y et une fonctionnelle de g. On peut aussi généraliser la développement de Taylor d'une fonction

$$f(x+\epsilon) = f(x) + \epsilon\frac{\mathrm{d}f}{\mathrm{d}x} + O(\epsilon^2)$$

au cas des fonctionnelles,

$$F[g+\epsilon h] = F[g] + \epsilon \int \frac{\delta F}{\delta h(x)} h(x)\,\mathrm{d}x + O(\epsilon^2).$$

Traitons le cas de l'action pour rester dans le contexte de la mécanique analytique. La méthode est maintenant bien connue. $S[q]$ est une fonctionnelle de $q(t)$, le lagrangien jouant un rôle de densité. Pour t' tel que $t_a < t' < t_b$, lorsque l'on change $q(t)$ en $q(t) + \epsilon\delta(t-t')$, la dérivée change comme $\dot{q}(t) \to \dot{q}(t) + \epsilon\frac{\mathrm{d}}{\mathrm{d}t}\delta(t-t')$. Le lagrangien varie alors de

$$\begin{aligned} L(q(t),\dot{q}(t)) &\to L\left(q(t) + \epsilon\delta(t-t'), \dot{q}(t) + \epsilon\frac{\mathrm{d}}{\mathrm{d}t}\delta(t-t')\right) \\ &= L(q(t),\dot{q}(t)) + \frac{\partial L}{\partial q}\epsilon\delta(t-t') + \frac{\partial L}{\partial \dot{q}}\epsilon\frac{\mathrm{d}}{\mathrm{d}t}\delta(t-t') + O(\epsilon^2). \end{aligned}$$

La variation de la fonctionnelle est donc

$$\begin{aligned} S[q(t) + \epsilon\delta(t-t')] = &\int_{t_a}^{t_b} L(q(t),\dot{q}(t))\,\mathrm{d}t \\ &+ \int_{t_a}^{t_b} \frac{\partial L}{\partial q}\epsilon\delta(t-t')\,\mathrm{d}t + \int_{t_a}^{t_b} \frac{\partial L}{\partial \dot{q}}\epsilon\frac{\mathrm{d}}{\mathrm{d}t}\delta(t-t')\,\mathrm{d}t + O(\epsilon^2) \end{aligned}$$

Le dernier terme s'intègre par parties et donne

$$\left[\frac{\partial L}{\partial \dot{q}}\epsilon\delta(t-t')\right]_{t=t_a}^{t=t_b} - \int_{t_a}^{t_b} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)\epsilon\delta(t-t')\,\mathrm{d}t.$$

Le terme de bords disparait grâce à la condition $t_a < t' < t_b$, et il vient finalement

$$\begin{aligned} S[q(t)+\epsilon\delta(t-t')] &= \int_{t_a}^{t_b} L(q(t),\dot{q}(t))\,\mathrm{d}t + \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)\right)\epsilon\delta(t-t')\,\mathrm{d}t + O(\epsilon^2) \\ &= S[q(t)] + \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)\right)\epsilon + O(\epsilon^2) \\ &= S[q(t)] + \epsilon\frac{\delta S}{\delta q(t')} + O(\epsilon^2) \end{aligned}$$

d'où le résultat finalement

$$\frac{\delta S}{\delta q(t')} = \frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right).$$

Dans le cas où l'action est définie à partir d'une densité lagrangienne, par exemple à une dimension spatiale,

$$S[\phi(x)] = \int_{t_a}^{t_b} \mathrm{d}t \int \mathcal{L}(\phi,\dot{\phi},\partial_x\phi)\,\mathrm{d}x$$

on calcule la dérivée fonctionnelle à partir de la variation lorsque le champ $\phi(x,t)$ varie de $\epsilon\delta(x-x',t-t')$, c'est-à-dire en faisant intervenir la distribution δ à deux dimensions. On trouve alors que

$$\frac{\delta S}{\delta\phi(x',t')} = \frac{\partial\mathcal{L}}{\partial\phi} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\right) - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial\mathcal{L}}{\partial(\partial_x\phi)}\right).$$

On notera que de la définition

$$L[\phi(x)] = \int \mathcal{L}(\phi,\dot{\phi},\partial_x\phi)\,\mathrm{d}x$$

on déduit également

$$\frac{\delta L}{\delta\phi(x',t')} = \frac{\partial\mathcal{L}}{\partial\phi} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial\mathcal{L}}{\partial(\partial_x\phi)}\right).$$

De manière très générale, pour une fonctionnelle de la forme

$$F[\rho] = \int f(\mathbf{r},\rho,\vec{\nabla}\rho,\vec{\nabla}^2\rho,\ldots)\,\mathrm{d}^3r$$

la dérivée fonctionnelle peut s'écrire

$$\frac{\delta F}{\delta\rho} = \frac{\partial f}{\partial\rho} - \vec{\nabla}\frac{\partial f}{\partial(\vec{\nabla}\rho)} + \vec{\nabla}^2\frac{\partial f}{\partial(\vec{\nabla}^2\rho)} - \ldots$$

Tableau 3.1 Quelques dérivées fonctionnelles.

Fonctionnelle	dérivée fonctionnelle
$I[g] = \int_{-\infty}^{+\infty} g(x)\,\mathrm{d}x$	$\frac{\delta I}{\delta g(y)} = 1$
$F_0[g] = g(0)$	$\frac{\delta F_0}{\delta g(y)} = \delta(y)$
$M[g] = \max_{0\leq x\leq 1}[g(x)]$	$\frac{\delta M}{\delta g(y)} = \delta(y - y_M)$
$F[\varphi] = \int \mathrm{d}^3r \left(\frac{1}{2}a\lvert\varphi(\mathbf{r})\rvert^2 + \frac{1}{4}b\lvert\varphi(\mathbf{r})\rvert^4 + \frac{1}{2}c\lvert\vec{\nabla}\varphi(\mathbf{r})\rvert^2\right)$	$\frac{\delta F}{\delta\varphi(\mathbf{r})} = a\varphi(\mathbf{r}) + b\lvert\varphi(\mathbf{r})\rvert^2\varphi(\mathbf{r}) - c\vec{\nabla}^2\varphi(\mathbf{r})$
$J[\rho] = \int \left(\frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{\lvert\mathbf{r}-\mathbf{r}'\rvert}\,\mathrm{d}^3r'\right)\mathrm{d}^3r$	$\frac{\delta J}{\delta\rho(\mathbf{r}')} = \int \frac{\rho(\mathbf{r})}{\lvert\mathbf{r}-\mathbf{r}'\rvert}\,\mathrm{d}^3r$
$T[\rho] = \frac{1}{8}\int \frac{(\vec{\nabla}\rho(\mathbf{r}))^2}{\rho(\mathbf{r})}\,\mathrm{d}^3r$	$\frac{\delta T}{\delta\rho(\mathbf{r}')} = -\frac{1}{8}\frac{(\vec{\nabla}\rho(\mathbf{r}'))^2}{\rho(\mathbf{r}')^2} - \frac{1}{4}\vec{\nabla}\left(\frac{\vec{\nabla}\rho(\mathbf{r}')}{\rho(\mathbf{r}')}\right)$

Tableau 3.2 Dérivées fonctionnelles utiles en mécanique analytique.

$\frac{\delta}{\delta\phi}\int \mathcal{L}(\phi,\dot{\phi},\vec{\nabla}\phi)\,\mathrm{d}^3r = \frac{\partial\mathcal{L}}{\partial\phi} - \vec{\nabla}\left(\frac{\partial\mathcal{L}}{\partial(\vec{\nabla}\phi)}\right)$
$\frac{\delta}{\delta\phi}\int \mathrm{d}t\int \mathcal{L}(\phi,\dot{\phi},\vec{\nabla}\phi)\,\mathrm{d}^3r = \frac{\partial\mathcal{L}}{\partial\phi} - \vec{\nabla}\left(\frac{\partial\mathcal{L}}{\partial(\vec{\nabla}\phi)}\right) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\right)$
$\frac{\delta}{\delta\phi}\int \mathcal{H}(\phi,\vec{\nabla}\phi,\pi)\,\mathrm{d}^3r = \frac{\partial\mathcal{H}}{\partial\phi} - \vec{\nabla}\left(\frac{\partial\mathcal{H}}{\partial(\vec{\nabla}\phi)}\right)$
$\frac{\delta}{\delta\pi}\int \mathcal{H}(\phi,\vec{\nabla}\phi,\pi)\,\mathrm{d}^3r = \frac{\partial\mathcal{H}}{\partial\pi}$

4

Invariance de jauge

Théorème de Noether

En théorie des champs, on définit donc le lagrangien L à partir d'une densité lagrangienne

$$L = \int_{\mathbb{R}^3} \mathcal{L}(\psi, \dot{\psi}, \vec{\nabla}\psi)\, \mathrm{d}^3 r$$

où $\psi(\mathbf{r}, t)$ est un champ qui varie dans l'espace et le temps. L'intégration porte sur l'espace $\mathbb{R}^3$ ou sur une partie de cet espace dont la frontière est notée $\partial\mathbb{R}^3$. La dépendance en $\vec{\nabla}\psi$ de la densité lagrangienne reflète l'existence d'interactions à courte portée. La minimisation de l'action

$$S = \int_{t_1}^{t_2} \mathrm{d}t \int_{\mathbb{R}^3} \mathcal{L}(\psi, \dot{\psi}, \vec{\nabla}\psi)\, \mathrm{d}^3 r \equiv \int_{\Omega} \mathcal{L}(\psi, \dot{\psi}, \vec{\nabla}\psi)\, \mathrm{d}^3 r\, \mathrm{d}t$$

conduit aux équations d'Euler-Lagrange si l'on a imposé des conditions assez douces aux variations de ψ aux limites $\partial\mathbb{R}^3$ de $\mathbb{R}^3$. On effectue en effet la variation

$$\begin{aligned}\delta S = 0 = \int_{t_1}^{t_2} \mathrm{d}t \int_{\mathbb{R}^3} \mathrm{d}^3 r \left[\frac{\partial \mathcal{L}}{\partial \psi} - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\psi}}\right) - \vec{\nabla}\left(\frac{\partial \mathcal{L}}{\partial(\vec{\nabla}\psi)}\right)\right] \delta\psi \\ + \left[\int_{\mathbb{R}^3} \mathrm{d}^3 r \frac{\partial \mathcal{L}}{\partial \dot{\psi}} \delta\psi\right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \mathrm{d}t \oint_{\partial\mathbb{R}^3} \mathrm{d}\mathbf{S} \frac{\partial \mathcal{L}}{\partial(\vec{\nabla}\psi)} \delta\psi\end{aligned}$$

et en supposant que les termes de bords sont tous nuls, il vient

$$\frac{\partial \mathcal{L}}{\partial \psi} - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\psi}}\right) - \vec{\nabla}\left(\frac{\partial \mathcal{L}}{\partial(\vec{\nabla}\psi)}\right) = 0.$$

Supposons maintenant une transformation, gouvernée par un paramètre ϵ, au cours de laquelle les champs varient au premier ordre de $\psi \to \psi + \delta\psi = \psi + \frac{\partial \psi}{\partial \epsilon}\,\mathrm{d}\epsilon$, $\dot{\psi} \to \dot{\psi} + \delta\dot{\psi} = \dot{\psi} + \frac{\partial \dot{\psi}}{\partial \epsilon}\,\mathrm{d}\epsilon = \dot{\psi} + \frac{\partial}{\partial t}\frac{\partial \psi}{\partial \epsilon}\,\mathrm{d}\epsilon$, $\vec{\nabla}\psi \to \vec{\nabla}\psi + \delta\vec{\nabla}\psi = \vec{\nabla}\psi + \frac{\partial(\vec{\nabla}\psi)}{\partial \epsilon}\,\mathrm{d}\epsilon = \vec{\nabla}\psi + \vec{\nabla}\frac{\partial \psi}{\partial \epsilon}\,\mathrm{d}\epsilon$. Le lagrangien $\mathcal{L}(\psi, \dot{\psi}, \vec{\nabla}\psi, \epsilon)$ varie de $\mathrm{d}\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \epsilon}\,\mathrm{d}\epsilon$ sous l'effet de cette transformation, et sous l'effet d'une variation des champs, on a également :

$$\begin{aligned}
\mathrm{d}\mathcal{L} &= \frac{\partial \mathcal{L}}{\partial \epsilon}\,\mathrm{d}\epsilon \\
&= \frac{\partial \mathcal{L}}{\partial \psi}\delta\psi + \frac{\partial \mathcal{L}}{\partial \dot{\psi}}\delta\dot{\psi} + \frac{\partial \mathcal{L}}{\partial(\vec{\nabla}\psi)}\delta\vec{\nabla}\psi \\
&= \frac{\partial \mathcal{L}}{\partial \psi}\frac{\partial \psi}{\partial \epsilon}\,\mathrm{d}\epsilon + \frac{\partial \mathcal{L}}{\partial \dot{\psi}}\frac{\partial}{\partial t}\left(\frac{\partial \psi}{\partial \epsilon}\right)\mathrm{d}\epsilon + \frac{\partial \mathcal{L}}{\partial(\vec{\nabla}\psi)}\vec{\nabla}\left(\frac{\partial \psi}{\partial \epsilon}\right)\mathrm{d}\epsilon \\
&= \left[\frac{\partial \mathcal{L}}{\partial \psi}\frac{\partial \psi}{\partial \epsilon} + \frac{\partial \mathcal{L}}{\partial \dot{\psi}}\frac{\partial}{\partial t}\left(\frac{\partial \psi}{\partial \epsilon}\right) + \frac{\partial \mathcal{L}}{\partial(\vec{\nabla}\psi)}\vec{\nabla}\left(\frac{\partial \psi}{\partial \epsilon}\right)\right]\mathrm{d}\epsilon.
\end{aligned}$$

En remplaçant le premier terme par l'équation d'Euler-Lagrange, on complète les dérivées partielles,

$$\mathrm{d}\mathcal{L} = \left[\frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\psi}}\frac{\partial \psi}{\partial \epsilon}\right) + \vec{\nabla}\left(\frac{\partial \mathcal{L}}{\partial(\vec{\nabla}\psi)}\frac{\partial \psi}{\partial \epsilon}\right)\right]\mathrm{d}\epsilon.$$

Finalement, en faisant la différence entre première et dernière ligne, on obtient une loi de conservation qui s'écrit sous la forme

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}\mathbf{j} = \frac{\partial \mathcal{L}}{\partial \epsilon}$$

avec

$$\begin{aligned}
\rho &= \frac{\partial \mathcal{L}}{\partial \dot{\psi}}\frac{\partial \psi}{\partial \epsilon} \\
\mathbf{j} &= \frac{\partial \mathcal{L}}{\partial(\vec{\nabla}\psi)}\frac{\partial \psi}{\partial \epsilon}.
\end{aligned}$$

C'est le théorème de Noether.

Dans le cas où le lagrangien n'a pas de dépendance explicite en fonction de la variable ϵ (on parle de symétrie interne ou de symétrie d'isospin), on en déduit l'équation de continuité

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

qui correspond à la loi de conservation [8]

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = 0, \quad \text{avec} \quad Q = \int_{\mathbb{R}^3} \rho \, \mathrm{d}^3 r.$$

En effet, toujours avec l'hypothèse de champs qui s'annulent aux bords,

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = -\int_{\mathbb{R}^3} \vec{\nabla}\mathbf{j} \, \mathrm{d}^3 r = -\oint_{\partial\mathbb{R}^3} \mathbf{j} \cdot \mathrm{d}\mathbf{S} = 0.$$

A titre d'exemple considérons le cas où le paramètre ϵ est le temps. On peut alors regrouper les dérivées par rapport au temps pour écrire

$$\frac{\partial}{\partial t}\left[\frac{\partial \mathcal{L}}{\partial \dot{\psi}}\dot{\psi} - \mathcal{L}\right] + \vec{\nabla}\left[\frac{\partial \mathcal{L}}{\partial(\vec{\nabla}\psi)}\dot{\psi}\right] = 0.$$

Si la densité lagrangienne ne dépend pas explicitement du temps, on en déduit que la quantité conservée est l'énergie totale,

$$\frac{\mathrm{d}H}{\mathrm{d}t} = 0$$

avec

$$H = \int_{\mathbb{R}^3} \mathcal{H} \, \mathrm{d}^3 r,$$
$$\mathcal{H} = \pi\dot{\psi} - \mathcal{L},$$
$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}}.$$

En effet, dans ce cas on a $\rho = \pi\dot{\psi}$ et la densité hamiltonienne apparaît par le terme $\rho - \mathcal{L}$. La conservation de l'énergie apparaît comme une conséquence de l'invariance dans la transformation $t \to t + \mathrm{d}t$, c'est-à-dire la translation dans le temps.

Invariance de jauge en électromagnétisme classique

L'équation de Maxwell pour les charges magnétiques

$$\vec{\nabla}\mathbf{B} = 0$$

permet d'écrire la champ magnétique sous la forme

$$\mathbf{B} = \vec{\nabla} \wedge \mathbf{A}$$

ce qui assure que $\mathbf{B}$ est à divergence nulle ($\vec{\nabla}(\vec{\nabla}\wedge\mathbf{A}) \equiv 0.$). Si l'on ajoute au potentiel vecteur un gradient arbitraire (changement de jauge),

$$\mathbf{A} \to \mathbf{A} + \vec{\nabla}\Lambda$$

[8] Après intégration de $\rho(\mathbf{r}, t)$ sur l'espace, on obtient une grandeur, Q, qui ne dépend plus que du temps, d'où la dérivée "droite".

le champ magnétique est inchangé car le rotationnel d'un gradient est identiquement nul, $\vec{\nabla} \wedge (\mathbf{A} + \vec{\nabla}\Lambda) = \vec{\nabla} \wedge \mathbf{A} = \mathbf{B}$. De manière similaire, la loi de Faraday-Lenz

$$\vec{\nabla} \wedge \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

écrite sous la forme

$$\vec{\nabla} \wedge (\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}) = 0$$

invite à écrire

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\vec{\nabla}\phi.$$

Le changement de jauge introduit ci-dessus pour le potentiel vecteur doit donc s'accompagner d'un changement de jauge pour le potentiel scalaire afin de préserver le champ électrique inchangé, comme le champ magnétique,

$$\phi \to \phi - \frac{\partial \Lambda}{\partial t}.$$

Invariance de jauge globale

▷ *Lagrangien de Schrödinger.* La densité lagrangienne de Schrödinger (due à Jordan et Wigner) est donnée par

$$\mathcal{L}(\psi, \dot{\psi}, \psi^\star, \dot{\psi}^\star, \vec{\nabla}\psi, \vec{\nabla}\psi^\star) = i\hbar\psi^\star\dot{\psi} - \frac{\hbar^2}{2m}(\vec{\nabla}\psi^\star)\cdot(\vec{\nabla}\psi) - V(\mathbf{r},t)\psi^\star\psi$$

Ce n'est pas une quantité hermitienne. Les champs ψ et $\psi^\star$ sont fonction de l'espace et du temps, $\psi = \psi(\mathbf{r},t)$. Par application des équations d'Euler-Lagrange à ψ, on obtient

$$-i\hbar\dot{\psi}^\star = -\frac{\hbar^2}{2m}\vec{\nabla}^2\psi^\star + V(\mathbf{r},t)\psi^\star$$

et par application à $\psi^\star$,

$$i\hbar\dot{\psi} = -\frac{\hbar^2}{2m}\vec{\nabla}^2\psi + V(\mathbf{r},t)\psi,$$

c'est-à-dire l'équation de Schrödinger et sa conjuguée. Le moment canonique est donné par

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\hbar\psi^\star$$

et la densité hamiltonienne

$$\mathcal{H} = \pi\dot{\psi} - \mathcal{L} = \frac{-i\hbar}{2m}(\vec{\nabla}\pi)\cdot(\vec{\nabla}\psi) - \frac{i}{\hbar}V(\mathbf{r},t)\pi\psi.$$

L'hamiltonien (classique) total vaut

$$H = \int \left[\frac{\hbar^2}{2m} (\vec{\nabla}\psi^\star) \cdot (\vec{\nabla}\psi) + V(\mathbf{r},t)\psi^\star\psi \right] \mathrm{d}^3 r.$$

Par intégration par parties du premier terme on obtient la forme habituelle

$$H = \int \psi^\star \left[-\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\mathbf{r},t) \right] \psi \, \mathrm{d}^3 r = \langle \psi | \hat{H} | \psi \rangle.$$

On retrouve ainsi l'élément de matrice de l'Hamiltonien de Schrödinger dans l'état $|\psi\rangle$, résultat finalement très naturel.

▷ *Conservation de la charge électrique.* Considérons la transformation qui modifie la fonction d'onde par un facteur de phase global

$$\psi(\mathbf{r},t) \to \psi'(\mathbf{r},t) = \mathrm{e}^{-i\frac{q}{\hbar}\alpha}\psi(\mathbf{r},t).$$

Cette transformation n'affecte pas les propriétés physiques qui dépendent du module de la fonction d'onde et pas de sa phase, $|\psi'(\mathbf{r},t)| = |\psi(\mathbf{r},t)|$. De l'expression $\mathrm{d}\psi(\mathbf{r},t) = -i\frac{q}{\hbar}\psi(\mathbf{r},t)\,\mathrm{d}\alpha$, on déduit à l'aide du théorème de Noether que la quantité suivante est conservée,

$$\int_{\mathbb{R}^3} \rho(\mathbf{r},t)\,\mathrm{d}^3 r = \int_{\mathbb{R}^3} q|\psi(\mathbf{r},t)|^2\,\mathrm{d}^3 r.$$

A ce niveau, q n'est encore qu'un scalaire que nous identifierons à la charge électrique de la particule représentée par la fonction d'onde (la densité de probabilité de présence de cette particule est $|\psi(\mathbf{r},t)|^2$).

Invariance de jauge locale

▷ *Couplage minimal.* La question qui se pose ensuite est la conséquence d'un changement de phase local, c'est-à-dire fonction de l'espace et du temps, $\alpha(\mathbf{r},t)$. Soit la transformation

$$\psi(\mathbf{r},t) \to \psi'(\mathbf{r},t) = \mathrm{e}^{-i\frac{q}{\hbar}\alpha(\mathbf{r},t)}\psi(\mathbf{r},t).$$

Calculons

$$\frac{1}{2m}\hat{\mathbf{p}}^2\psi'(\mathbf{r},t) - i\hbar\frac{\partial}{\partial t}\psi'(\mathbf{r},t).$$

On développe tout d'abord le premier terme,

$$\begin{aligned}
\hat{\mathbf{p}}^2\psi'(\mathbf{r},t) &= (-i\hbar\vec{\nabla})^2\psi'(\mathbf{r},t) \\
&= -\hbar^2\vec{\nabla}[-i\frac{q}{\hbar}(\vec{\nabla}\alpha)\psi\,\mathrm{e}^{-i\frac{q}{\hbar}\alpha} + (\vec{\nabla}\psi)\,\mathrm{e}^{-i\frac{q}{\hbar}\alpha}] \\
&= -\hbar^2[-i\frac{q}{\hbar}(\vec{\nabla}^2\alpha)\psi - 2i\frac{q}{\hbar}(\vec{\nabla}\alpha)(\vec{\nabla}\psi) - \frac{q^2}{\hbar^2}(\vec{\nabla}\alpha)^2\psi + \vec{\nabla}^2\psi]\,\mathrm{e}^{-i\frac{q}{\hbar}\alpha} \\
&= [-\hbar^2\vec{\nabla}^2\psi + i\hbar q\vec{\nabla}(\psi\vec{\nabla}\alpha) + i\hbar q(\vec{\nabla}\alpha)\vec{\nabla}\psi + q^2(\vec{\nabla}\alpha)^2\psi]\,\mathrm{e}^{-i\frac{q}{\hbar}\alpha}.
\end{aligned}$$

Cette expression rappelle immédiatement l'expression de l'énergie cinétique d'une particule chargée sous champ

$$\frac{1}{2m}\hat{\mathbf{p}}^2 \to \frac{1}{2m}(-\hbar^2\vec{\nabla}^2 + iq\hbar\vec{\nabla}(\mathbf{A}\cdot) + iq\hbar\mathbf{A}\vec{\nabla} + q^2\mathbf{A}^2)$$

et suggère donc de poser pour identifier les deux expressions

$$\vec{\nabla}\alpha(\mathbf{r},t) = \mathbf{A}(\mathbf{r},t).$$

Le potentiel vecteur apparaît ainsi généré par l'action de $\hat{\mathbf{p}}$ sur la jauge locale. Pour la dérivée temporelle on obtient

$$i\hbar\frac{\partial}{\partial t}\psi'(\mathbf{r},t) = \left[i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) + q\frac{\partial\alpha}{\partial t}\psi(\mathbf{r},t)\right] e^{-i\frac{q}{\hbar}\alpha}$$

que l'on identifie à l'hamiltonien d'une charge dans un potentiel scalaire,

$$i\hbar\frac{\partial}{\partial t} - q\phi$$

moyennant la correspondance

$$\frac{\partial\alpha(\mathbf{r},t)}{\partial t} = -\phi(\mathbf{r},t)$$

qui montre que c'est la dérivée temporelle qui génère le potentiel scalaire.

▷ *L'invariance de jauge comme principe dynamique.* Finalement, en supposant satisfaite l'équation de Schrödinger pour une particule libre avec la fonction d'onde $\psi'(\mathbf{r},t)$,

$$i\hbar\frac{\partial}{\partial t}\psi'(\mathbf{r},t) = \frac{1}{2m}\hat{\mathbf{p}}^2\psi'(\mathbf{r},t),$$

où ψ se déduit de la transformation de jauge locale

$$\psi(\mathbf{r},t) \to \psi'(\mathbf{r},t) = \mathrm{e}^{-i\frac{q}{\hbar}\alpha(\mathbf{r},t)}\psi(\mathbf{r},t)$$

on obtient que la fonction d'onde originale obéit à l'équation de Schrödinger sous champ

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) = \frac{1}{2m}(\hat{\mathbf{p}} - q\mathbf{A})^2\psi(\mathbf{r},t) + q\phi\psi(\mathbf{r},t),$$

moyennant la correspondance

$$\vec{\nabla}\alpha(\mathbf{r},t) = \mathbf{A}(\mathbf{r},t)$$
$$\frac{\partial\alpha(\mathbf{r},t)}{\partial t} = -\phi(\mathbf{r},t).$$

Il faut noter qu'aux transformations des potentiels qui assurent l'invariance de jauge de l'électromagnétisme classique, il faut ajouter la transformation de la fonction d'onde pour préserver l'invariance de jauge en théorie quantique [9] ,

$$\begin{aligned}
&\mathbf{A} \to \mathbf{A} + \vec{\nabla}\Lambda \\
&\phi \to \phi - \frac{\partial \Lambda}{\partial t} \\
&\psi(\mathbf{r}, t) \to \mathrm{e}^{-i\frac{q}{\hbar}\Lambda(\mathbf{r},t)}\psi(\mathbf{r}, t).
\end{aligned}$$

On constate que le couplage avec le champ électromagnétique apparaît comme le résultat de l'invariance de jauge locale c'est-à-dire que la symétrie implique la dynamique. On peut citer Salam et Ward [10]

Our basic postulate is that it should be possible to generate strong, weak and electromagnetic interaction terms (...) by making local gauge transformations on the kinetic-energy terms of the free Lagrangian.

La transformation précédente est appelée couplage minimal. On définit les opérateurs de dérivée covariante [11]

$$\begin{aligned}
\vec{\mathcal{D}} &= \vec{\nabla} - i\frac{q}{\hbar}\mathbf{A} \\
\mathcal{D}_t &= \partial_t + i\frac{q}{\hbar}\phi
\end{aligned}$$

qui permettent d'écrire l'équation sous champ de manière analogue à l'équation de la particule libre,

$$i\hbar\mathcal{D}_t\psi(\mathbf{r}, t) = -\frac{\hbar^2}{2m}(\vec{\mathcal{D}})^2\psi(\mathbf{r}, t).$$

---

[9] On reprend la notation $\Lambda(\mathbf{r}, t)$ pour la fonction qui produit le changement de jauge.

[10] Voir à ce propos hep-lat/0001283.

[11] On écrit aussi

$$-i\hbar\vec{\mathcal{D}} = -i\hbar\vec{\nabla} - q\mathbf{A} \quad \text{et} \quad i\hbar\mathcal{D}_t = i\hbar\partial_t - q\phi$$

Formalisme hamiltonien

5

Formulation hamiltonienne de la dynamique des milieux continus

Le formalisme hamiltonien nécessite également quelques modifications pour être applicable aux milieux continus. Le prix à payer est de nouveau une forme un peu différente pour les équations d'Hamilton, mais la notion de dérivée fonctionnelle restitue la forme habituelle aux équations dynamiques.

Passage au continu

Revenons au cas de la chaîne linéaire de masses couplées par des ressorts. Le lagrangien est donné par

$$L = \sum_n aL_n = \frac{1}{2}\sum_n a\left(\mu\dot{\phi}_n^2 - Y\left(\frac{\phi_{n+1}-\phi_n}{a}\right)^2\right).$$

Dans le cas d'un système discret, on définit les moments généralisés p_n par

$$p_n = \frac{\partial L}{\partial \dot{\phi}_n} = a\frac{\partial L_n}{\partial \dot{\phi}_n}. \tag{5.1}$$

L'intérêt de cette définition est de permettre de passer simplement par une transformation de Legendre de l'expression du lagrangien comme fonction des ϕ_n et des $\dot{\phi}_n$,

$$L(\phi_n, \dot{\phi}_n),$$

à une fonction X dépendant des ϕ_n et d'une nouvelle série de variables, les p_n. La dépendance en $\dot{\phi}_n$ en revanche a disparu. On procède donc comme suit,

$$X \equiv L(\phi_n, \dot{\phi}_n) - p_n \dot{\phi}_n,$$

et la condition $\frac{\partial X}{\partial \dot{\phi}_n} = 0$ est assurée à condition de définir $p_n = \frac{\partial L}{\partial \dot{\phi}_n}$. On a bien

$$X \equiv X(p_n, \phi_n).$$

On définit l'hamiltonien par l'opposé de cette quantité X,

$$\texttt{Hamiltonien} : H(p_n, \phi_n) = p_n \dot{\phi}_n - L(\phi_n, \dot{\phi}_n). \tag{5.2}$$

Les équations du mouvement qui en découlent sont données par

$$\texttt{Equations} \qquad \frac{\partial H}{\partial p_n} = \dot{\phi}_n$$

$$\texttt{d'Hamilton} : \quad \frac{\partial H}{\partial \phi_n} = -\frac{\partial L}{\partial \phi_n} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\phi}_n}\right) = -\dot{\phi}_n.$$

Ce sont deux séries d'équations différentielles du premier ordre couplées entre elles.

Nous allons maintenant suivre une démarche analogue dans le cas des milieux continus. De l'expression $L = \sum_n a L_n$, on déduit d'abord le moment conjugué $p_n = \frac{\partial L}{\partial \dot{\phi}_n} = a \frac{\partial L_n}{\partial \dot{\phi}_n}$ puis l'hamiltonien

$$H = \sum_n (p_n \dot{\phi}_n - a L_n),$$

soit encore

$$H = \sum_n a \left(\frac{\partial L_n}{\partial \dot{\phi}_n}\dot{\phi}_n - L_n\right) \longrightarrow_{a\to 0} \int \mathrm{d}x \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\dot{\phi} - \mathcal{L}\right).$$

Dans la limite $a \to 0$ les p_n tendent vers zéro, mais on peut définir une densité associée qui reste finie dans cette limite,

$$\texttt{Moment conjugué} : \pi \equiv \lim_{a\to 0} \frac{p_n}{a} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_n}, \tag{5.3}$$

de sorte que l'hamiltonien s'exprime à partir d'une densité

$$\texttt{Hamiltonien} : H[\phi(x), \pi(x)] = \int \mathrm{d}x \, (\pi\dot{\phi} - \mathcal{L}) = \int \mathrm{d}x \, \mathcal{H}. \tag{5.4}$$

C'est une fonctionnelle des champs $\phi(x)$ et $\pi(x)$. On constate ici que le traitement de la coordonnée temporelle est particulier. Il apparaît notamment une dissymétrie

avec le traitement des coordonnées spatiales, contrairement à ce que l'on a vu avec le formalisme lagrangien. On peut donc s'attendre à ce qu'il soit difficile d'incorporer le formalisme hamiltonien dans une théorie relativiste. Cela justifie que les théories des champs aient davantage recours au formalisme lagrangien.

Equations d'Hamilton

A partir de l'expression

$$\texttt{Densité d'Hamiltonien : } \mathcal{H} = \pi\dot{\phi} - \mathcal{L}(\phi, \dot{\phi}, \partial_x\phi)$$

on montre que la densité d'hamiltonien ne dépend pas de $\dot{\phi}$:

$$\frac{\partial \mathcal{H}}{\partial \dot{\phi}} = \pi - \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = 0 \quad \text{si} \quad \pi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}}.$$

On a bien

$$\mathcal{H} \equiv \mathcal{H}(\pi, \phi, \partial_x\phi).$$

Connaissant les équations de Lagrange,

$$\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}} - \frac{\partial}{\partial x}\frac{\partial \mathcal{L}}{\partial(\partial_x\phi)} = 0,$$

on obtient les équations d'Hamilton de la façon habituelle. On a tout d'abord $\frac{\partial \mathcal{H}}{\partial \pi} = \dot{\phi}$. On peut ensuite évaluer la dérivée partielle par rapport aux variations spatiales du champ ϕ, $\frac{\partial \mathcal{H}}{\partial(\partial_x\phi)} = -\frac{\partial \mathcal{L}}{\partial(\partial_x\phi)}$, puis finalement $\frac{\partial \mathcal{H}}{\partial \phi} = -\frac{\partial \mathcal{L}}{\partial \phi}$ où l'on utilise les équations de Lagrange pour transformer cette expression en

$$-\frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}} - \frac{\partial}{\partial x}\frac{\partial \mathcal{L}}{\partial(\partial_x\phi)} = -\dot{\pi} + \frac{\partial}{\partial x}\frac{\partial \mathcal{H}}{\partial(\partial_x\phi)}.$$

Finalement les équations d'Hamilton prennent la forme

$$\texttt{Equations d'Hamilton :}$$

$$\begin{aligned}
\frac{\partial \mathcal{H}}{\partial \pi} &= \dot{\phi}, \\
\frac{\partial \mathcal{H}}{\partial \phi} &= -\dot{\pi} + \frac{\partial}{\partial x}\frac{\partial \mathcal{H}}{\partial(\partial_x\phi)}, \\
\frac{\partial \mathcal{H}}{\partial(\partial_x\phi)} &= -\frac{\partial \mathcal{L}}{\partial(\partial_x\phi)}.
\end{aligned}$$

Elles ont une forme plus compliquée que dans le cas des systèmes discrets et encore une fois cela provient de la dépendance de la densité d'hamiltonien en fonction des variations spatiales du champ.

Equations d'Hamilton et dérivées fonctionnelles

A l'aide des dérivées fonctionnelles les équations d'Hamilton prennent leur forme habituelle. On écrit

$$\frac{\delta H}{\delta \phi(x)} = \frac{\partial \mathcal{H}}{\partial \phi} - \frac{\partial}{\partial x}\frac{\partial \mathcal{H}}{\partial(\partial_x \phi)},$$
$$\frac{\delta H}{\delta \pi(x)} = \frac{\partial \mathcal{H}}{\partial \pi},$$

et les équations d'Hamilton deviennent

$$\frac{\delta H}{\delta \phi(x)} = -\dot{\pi}(x),$$
$$\frac{\delta H}{\delta \pi(x)} = \dot{\phi}(x).$$

Calcul tensoriel

6

Eléments de calcul tensoriel

Les tenseurs ont été introduits par les mathématiciens italiens Ricci et Levi-Civita. La nouvelle branche des mathématiques ainsi créée s'appelait alors le calcul différentiel absolu. Dans ce chapitre on introduit les tenseurs et on présente les rudiments de base du calcul tensoriel[12]

Quadrivecteurs contravariants et covariants

En relativité, la structure de l'espace-temps suggère l'introduction de quadrivecteurs ou 4-vecteurs, A^μ ayant une composante temporelle A^0 et trois composantes spatiales A^i. On adopte à partir de maintenant la convention selon laquelle les indices grecs peuvent prendre 4 valeurs, $\mu = 0, 1, 2, 3$, la valeur 0 étant toujours réservée aux composantes temporelles, et les indices latins peuvent prendre 3 valeurs, $i = 1, 2, 3$ et sont réservés aux composantes spatiales exclusivement[13] . On peut écrire indifféremment un quadrivecteur comme

$$A^\mu \equiv (A^0, A^i) \equiv (A^0, A^1, A^2, A^3). \tag{6.1}$$

La notation en exposant est réservée au label des composantes. Si l'on veut former le carré ou une autre puissance d'une quantité, on utilise des parenthèses, par exemple $(A^i)^2$, sauf s'il n'y a aucun risque de confusion, comme dans $x^2 + y^2 + z^2$. Lorsque

[12] On pourra consulter L.D. Landau et E. Lifshitz, *Théorie des champs*, 3ème édition, Editions MIR, Moscou 1970, §6 ou les notes de cours de David Sénéchal, *Ondes électromagnétiques*, §4.2.

[13] Les notations ici sont assez diverses. Certains auteurs (comme Landau et Lifshitz) réservent les lettres latines plutôt que grecques pour les indices spatio-temporels, d'autres font courir les indices de 1 à 4, la quatrième composante étant temporelle, elle est parfois complexe (comme chez Pauli qui note $x^\mu \equiv (x^1, x^2, x^3, x^4) = (x, y, z, ict)$ là où nous écrirons $x^\mu \equiv (x^0, x^1, x^2, x^3) = (ct, x, y, z)$). Cela implique un choix de métrique qui détermine le signe de la norme invariante.

c'est nécessaire on peut également utiliser la notation vectorielle habituelle pour rassembler les composantes spatiales,

$$A^\mu = (A^0, \mathbf{A}).$$

Il se peut qu'il soit nécessaire de distinguer le vecteur ordinaire $\mathbf{A}$ des composantes spatiales du quadrivecteur, qui dans ce cas sont notées $\mathbf{A}$. Ce type de distinction est utile, par exemple pour distinguer les composantes spatiales $\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\tau}$ de la quadrivitesse en fonction de la vitesse ordinaire $\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$, $\mathbf{v} = \gamma_{\mathbf{v}}\mathbf{v}$.

Par définition [14] , un quadrivecteur *contravariant* est l'ensemble de 4 scalaires A^μ (la position de l'indice en haut est très importante) obéissant à la loi de transformation de Lorentz par changement de référentiel inertiel :

$$\begin{aligned} A'^0 &= \gamma(A^0 - \beta A^1) \\ A'^1 &= \gamma(A^1 - \beta A^0) \\ A'^2 &= A^2 \\ A'^3 &= A^3 \end{aligned}$$

avec $\beta = |\mathbf{u}|/c$ et par convention, si rien n'indique le contraire, la vitesse (constante) $\mathbf{u}$ du référentiel primé par rapport au référentiel non primé est portée par l'axe 1, $\mathbf{u} = (u, 0, 0) = u\,\hat{\mathbf{u}}_1$, où $\hat{\mathbf{u}}_1$ est unitaire dans la direction 1. On dit que les A^μ se transforment comme les composantes d'un vecteur à 4 dimensions dans une rotation de Lorentz, ce que l'on écrit de la façon suivante[15] :

$$A'^\mu = \Lambda^\mu{}_\nu A^\nu. \tag{6.2}$$

Cette écriture est très importante ; la convention d'Einstein de sommation sur les indices répétés en position haute et basse est sous-entendue. Après sommation sur cet indice (ν ici), la quantité au second membre ne dépend plus de cet indice, il est dit muet. L'écriture ci-dessus résume donc 4 équations (une pour chacune des 4 valeurs possibles pour l'indice μ) et chacune de ces équations est une somme de 4 termes, par exemple

$$A'^2 = \Lambda^2{}_0 A^0 + \Lambda^2{}_1 A^1 + \Lambda^2{}_2 A^2 + \Lambda^2{}_3 A^3.$$

C'est donc équivalent à une notation matricielle. La matrice de rotation de Lorentz prend la forme (on note entre crochets $[\Lambda^\mu{}_\nu]$ la matrice associée à la transformation) :

$$[\Lambda^\mu{}_\nu] = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{6.3}$$

[14] Les lecteurs un peu rigoureux peuvent consulter W. Appel, *Mathématiques pour la physique et les physiciens !*, H & K éditions, Paris 2002, chap.15. Cet ouvrage est d'ailleurs conseillé aussi pour ses autres chapitres...

[15] Les positions des indices (haut et bas, avant et arrière) sont importantes.

Un exemple de quadrivecteur contravariant est donné par

$$\begin{aligned} x^\mu &= (x^0, x^i) = (x^0, \mathbf{r}) \\ &= (ct, \mathbf{r}) = (ct, x, y, z), \end{aligned}$$

ou, pour des variations infinitésimales de coordonnées,

$$\mathrm{d}x^\mu = (\mathrm{d}x^0, \mathrm{d}x^i) = (c\,\mathrm{d}t, \mathrm{d}\mathbf{r}) = (c\,\mathrm{d}t, \mathrm{d}x, \mathrm{d}y, \mathrm{d}z).$$

La partie spatiale du 4-vecteur, $\mathbf{r}$, coïncide dans ce cas avec le vecteur position ordinaire $\mathbf{r}$.

L'opération de sommation sur un indice muet est appelée *contraction*. Un exemple particulièrement important est obtenu lorsque l'on exprime la variation d'une quantité scalaire Q lors d'un déplacement élémentaire dans l'espace-temps :

$$\delta Q = \frac{\partial Q}{\partial x^\mu} \delta x^\mu.$$

La quantité au premier membre étant scalaire, celle du second membre doit l'être également, il s'agit dont d'une contraction sur l'indice muet μ qui exige que l'on note la dérivée avec un indice en bas,

$$\frac{\partial Q}{\partial x^\mu} \equiv \partial_\mu Q.$$

On écrit donc

$$\delta Q = \partial_\mu Q \delta x^\mu$$

où la convention d'Einstein rend la sommation explicite. La dérivée quadridimensionnelle ou quadri-gradient (l'opérateur gradient auquel on a ajouté la dérivée temporelle)

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} \equiv \left(\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right) \tag{6.4}$$

se transforme différemment d'un quadrivecteur contravariant. On peut en effet faire apparaître explicitement un changement de référentiel inertiel :

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \underbrace{\frac{\partial x'^\nu}{\partial x^\mu}}_{\Lambda^\nu{}_\mu} \frac{\partial}{\partial x'^\nu} = \Lambda^\nu{}_\mu \partial'_\nu$$

où l'on a utilisé la définition

$$x'^\nu = \Lambda^\nu{}_\mu x^\mu$$

pour identifier $\frac{\partial x'^\nu}{\partial x^\mu} = \Lambda^\nu{}_\mu$. On a donc

$$\partial_\mu = \Lambda^\nu{}_\mu \partial'_\nu$$

et par inversion, en faisant $\mathbf{u} \to -\mathbf{u}$, on a

$$\partial'_\mu = (\Lambda^{-1})^\nu{}_\mu \partial_\nu \tag{6.5}$$

avec la matrice de transformation inverse de Lorentz (qui est d'ailleurs également une tranformation de Lorentz)

$$[(\Lambda^{-1})^\nu{}_\mu] = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.6}$$

et toujours avec la définition $\beta = |\mathbf{u}|/c$. Le quadrivecteur ∂_μ se transforme avec la matrice de transformation inverse, $[\Lambda^{-1}]$, contrairement aux quadrivecteurs contravariants qui se transforment avec la matrice directe $[\Lambda]$. On dit que ∂_μ est un quadrivecteur *covariant*. Par extension, tous les quadrivecteurs pour lesquels l'indice de composante est noté en bas sont des quadrivecteurs covariants. La quadri-dérivée d'une quantité scalaire par rapport aux composantes contravariantes des coordonnées, $\partial_\mu Q$, définit donc un quadrivecteur covariant. C'est un résultat très général. On note également

$$\partial_\mu Q = Q_{,\mu}. \tag{6.7}$$

Soit A^μ un 4-vecteur contravariant, $A^\mu = (A^0, A^1, A^2, A^3)$. Par définition, il se transforme avec la matrice directe $[\Lambda]$. Le 4-vecteur covariant A_μ qui lui est associé, $A_\mu = (A_0, A_1, A_2, A_3)$, devant se transformer avec la matrice inverse, on a donc

$$\begin{aligned} A'^0 &= \gamma(A^0 - \beta A^1), & A'_0 &= \gamma(A_0 + \beta A_1), \\ A'^1 &= \gamma(A^1 - \beta A^0), & A'_1 &= \gamma(A_1 + \beta A_0), \\ A'^2 &= A^2, & A'_2 &= A_2, \\ A'^3 &= A^3, & A'_3 &= A_3. \end{aligned}$$

Les objets A^μ et A_μ représentent la même grandeur physique, leurs composantes sont donc liées et pour être conformes aux transformations ci-dessus, les composantes temporelles doivent être identiques alors que les composantes spatiales doivent être changées de signe [16] :

$$A_0 = A^0,\ A_1 = -A^1,\ A_2 = -A^2,\ A_3 = -A^3.$$

Pour éviter toute confusion, on choisit généralement, lorsqu'on explicite les composantes, de travailler soit avec les composantes contravariantes, soit avec les composantes covariantes. On choisit ici d'utiliser le plus souvent possible les composantes contravariantes, on écrira par exemple

$$A^\mu = (A^0, \mathbf{A}), \quad A_\mu = (A^0, -\mathbf{A}). \tag{6.8}$$

[16] On se limite ici à des systèmes de coordonnées cartésiennes. Cette propriété n'est plus exacte dans un système de coordonnées arbitraire, mais alors la relation $A_\mu = g_{\mu\nu}A^\nu$ que l'on verra un peu plus loin reste vraie.

Des exemples de quadrivecteurs covariants sont donnés par

$$x_\mu = (ct, -\mathbf{r}) = (ct, -x, -y, -z),$$
$$\mathrm{d}x_\mu = (c\,\mathrm{d}t, -\mathrm{d}\mathbf{r}) = (c\,\mathrm{d}t, -\mathrm{d}x, -\mathrm{d}y, -\mathrm{d}z).$$

On notera la particularité de la quadri-dérivée pour ce qui concerne le signe de la partie spatiale.

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} \equiv \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right),$$
$$\partial^\mu \equiv \frac{\partial}{\partial x_\mu} \equiv \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right).$$

On définit le carré (on parle parfois de norme) d'un quadrivecteur par la quantité

$$\begin{aligned} A_\mu A^\mu \equiv A^\mu A_\mu &= A_0A^0 + A_1A^1 + A_2A^2 + A_3A^3 \\ &= (A_0)^2 - (A_1)^2 - (A_2)^2 - (A_3)^2 \\ &= (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2 \\ &= (A^0)^2 - |\mathbf{A}|^2. \end{aligned} \tag{6.9}$$

Cette quantité est une contraction sur l'indice muet μ, elle ne dépend donc plus d'aucun indice et constitue un scalaire invariant (ce ne serait pas le cas de combinaisons scalaires du type $(A^0)^2 + (A^1)^2 + (A^2)^2 + (A^3)^2$ par exemple). Par extension, on définit le produit invariant des quadrivecteurs [17] A^μ et B^μ par (on ne précise plus cette fois les différentes formes équivalentes)

$$A_\mu B^\mu = A^0B^0 - \mathbf{AB}. \tag{6.10}$$

On vérifie bien qu'il s'agit d'un *scalaire invariant de Lorentz* (ou scalaire de Lorentz) en écrivant le produit dans le référentiel primé :

$$\begin{aligned} A'_\mu B'^\mu &= A'^0B'^0 - \mathbf{A'B'} \\ &= \gamma^2(A^0 - \beta A^1)(B^0 - \beta B^1) - \gamma^2(A^1 - \beta A^0)(B^1 - \beta B^0) - A^2B^2 - A^3B^3 \\ &= A^0B^0 - \mathbf{AB} \\ &= A_\mu B^\mu. \end{aligned}$$

Le tenseur métrique

A partir du quadrivecteur contravariant A^μ, on définit le tenseur métrique $g_{\mu\nu}$ qui fait passer au quadrivecteur covariant associé

$$A_\mu = g_{\mu\nu}A^\nu \tag{6.11}$$

[17] Encore une fois, l'égalité ci-dessous vaut pour des systèmes de coordonnées orthogonales, mais l'expression $A_\mu B^\mu$ est bien un invariant dans tout système de coordonnées.

par sommation sur l'indice ν. On précisera au paragraphe suivant la définition d'un tenseur, on se contente pour le moment de la dénomination de tenseur métrique sans davantage de précision. On voit que $g_{\mu\nu}$ doit être la matrice diagonale qui assure que $A_0 = A^0$ et $A_i = -A^i$, soit

$$[g_{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{6.12}$$

Dans une expression telle que $A_\mu = g_{\mu\nu}A^\nu$, le second membre dépend a priori de trois indices, mais il s'agit en fait d'une contraction, puisque l'un des indices (ν en l'occurence) est répété en positions haute et basse, il disparaît donc après sommation et la quantité ainsi définie ne dépend que de l'indice restant, μ, ce qui en fait un quadrivecteur. On notera que l'indice non répété dans le membre de droite, μ, occupe la même position basse dans les deux membres de l'expression.

On peut également effectuer la transformation inverse

$$A^\mu = g^{\mu\nu}A_\nu, \tag{6.13}$$

$$[g^{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{6.14}$$

L'égalité $g_{\mu\nu} = g^{\mu\nu}$ est très particulière. Elle n'est possible qu'en coordonnées cartésiennes.

Le carré invariant d'un quadrivecteur peut maintenant s'écrire à l'aide uniquement de la forme contravariante et du tenseur métrique (ou à l'aide de la forme covariante uniquement) :

$$A_\mu A^\mu = g_{\mu\nu}A^\mu A^\nu = g^{\mu\nu}A_\mu A_\nu. \tag{6.15}$$

L'introduction du tenseur métrique vient au départ de ce qu'il définit la métrique par l'intervalle infinitésimal

$$\begin{aligned} \mathrm{d}s^2 &= \mathrm{d}x_\mu \, \mathrm{d}x^\mu \\ &= g_{\mu\nu} \, \mathrm{d}x^\mu \, \mathrm{d}x^\nu \\ &= c^2 \, \mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2, \end{aligned}$$

soit sign $(g_{\mu\nu}) = (1, -1, -1, -1)$ où sign (pour signature) signifie que l'on ne donne que les éléments diagonaux (en toute rigueur, leur signe seulement). Cette écriture est très intéressante, car elle rend manifeste l'invariance de l'intervalle : comme $\mathrm{d}s^2$ représente un carré par l'intermédiaire d'une contraction, c'est automatiquement un scalaire invariant de Lorentz.

On peut noter que la forme du tenseur métrique dépend du système de coordonnées. En coordonnées sphériques, on aura $\mathrm{d}x^\mu = (c\,\mathrm{d}t, \mathrm{d}r, \mathrm{d}\theta, \mathrm{d}\varphi)$, et $\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 - \mathrm{d}r^2 - r^2\,\mathrm{d}\theta^2 - r^2\sin^2\theta\,\mathrm{d}\varphi^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu$, soit

$$[g_{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix}. \tag{6.16}$$

Dans ce cas, $g^{\mu\nu}$ et $g_{\mu\nu}$ ne sont plus représentés par la même matrice. En effet, comme $\mathrm{d}x_\mu = g_{\mu\nu}\,\mathrm{d}x^\nu$ et $\mathrm{d}x^\mu = g^{\mu\nu}\,\mathrm{d}x_\nu$ l'intervalle s'écrit

$$\mathrm{d}s^2 = \mathrm{d}x_\mu\,\mathrm{d}x^\mu = g_{\mu\nu}g^{\mu\nu}\underbrace{\mathrm{d}x_\nu\,\mathrm{d}x^\nu}_{\mathrm{d}s^2},$$

soit $g_{\mu\nu}g^{\mu\nu} = 1$ d'où l'on déduit

$$g^{\mu\nu} = \frac{1}{g}G^{\mu\nu}$$

où g est le déterminant de $g_{\mu\nu}$ et $G^{\mu\nu}$ la matrice de ses cofacteurs ($g_{\mu\nu}$ étant symétrique, il n'est pas nécessaire de transposer la matrice des cofacteurs). On a donc

$$[g^{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -\frac{1}{r^2} & 0 \\ 0 & 0 & 0 & -\frac{1}{r^2\sin^2\theta} \end{pmatrix}.$$

Tenseurs de rang arbitraire

On définit un tenseur de rang deux par l'introduction de deux indices. Un tenseur deux fois contravariant est un objet $A^{\mu\nu}$ qui obéit, par changement de référentiel inertiel, à la transformation de Lorentz généralisée aux deux indices :

$$A'^{\mu\nu} = \Lambda^\mu{}_\sigma\Lambda^\nu{}_\tau A^{\sigma\tau}. \tag{6.17}$$

Pour un tenseur deux fois covariant on a de même

$$A'_{\mu\nu} = (\Lambda^{-1})^\sigma{}_\mu(\Lambda^{-1})^\tau{}_\nu A_{\sigma\tau}. \tag{6.18}$$

Pour un tenseur mixte de rang deux,

$$A'^\mu{}_\nu = \Lambda^\mu{}_\sigma(\Lambda^{-1})^\tau{}_\nu A^\sigma{}_\tau. \tag{6.19}$$

On note toujours les mêmes règles en calcul tensoriel : les indices contractés disparaissent et les indices non sommés conservent leur position des deux côtés de

l'égalité. Un quadrivecteur est un tenseur de rang 1. On peut former des tenseurs de rang arbitraire par produits de tenseurs de rang plus faible :

$$
\begin{aligned}
&A^{\mu\nu} = a^\mu b^\nu, \\
&B^\mu{}_\nu = a^\mu c_\nu, \\
&C^\mu{}_{\nu\rho} = a^\mu b_\nu d_\rho, \\
&D_\mu{}^{\nu\rho} = b_\mu A^{\nu\rho}, \ldots
\end{aligned}
$$

En ce qui concerne le passage des composantes covariantes ou mixtes aux composantes contravariantes par exemple, on utilise à nouveau le tenseur métrique. Par exemple de $g_{\nu\mu}A^{\mu\sigma} = A_\nu{}^\sigma$ on déduit $A_0{}^1 = g_{0\mu}A^{\mu 1} = A^{01}$ et de $g_{\nu\mu}A^{\sigma\nu} = A^\sigma{}_\mu$ on déduit $A^0{}_1 = g_{\nu 1}A^{0\nu} = -A^{01}$, ce qui justifie la nécessité de distinguer en général $A^\mu{}_\nu$ et $A_\nu{}^\mu$. De manière générale, l'élévation ou l'abaissement d'un indice temporel laisse la valeur de la composante inchangée alors que l'élévation ou l'abaissement d'un indice spatial en change le signe. Voici quelques exemples :

$$
\begin{aligned}
&A_{00} = A^{00}, \\
&A_{01} = -A^{01}, \\
&A_{11} = A^{11}, \\
&A_0{}^0 = A^{00}, \\
&A_0{}^1 = A^{01}, \\
&A^0{}_1 = -A^{01}, \\
&A^1{}_1 = -A^{11} \ldots
\end{aligned}
$$

On peut résumer ces transformations en notation matricielle, en introduisant une notation simplifiée,

$$
[A_{\mu\nu}] = \begin{pmatrix} \boxed{A_{00}} & \boxed{A_{0j}} \\ \boxed{A_{i0}} & \boxed{A_{ij}} \end{pmatrix} = \begin{pmatrix} A_{00} & A_{0j} \\ A_{i0} & A_{ij} \end{pmatrix},
$$

$$
[A_{\mu\nu}] = \begin{pmatrix} A_{00} & A_{0j} \\ A_{i0} & A_{ij} \end{pmatrix} = \begin{pmatrix} A^{00} & -A^{0j} \\ -A^{i0} & A^{ij} \end{pmatrix}, \tag{6.20}
$$

$$
[A^\mu{}_\nu] = \begin{pmatrix} A^0{}_0 & A^0{}_j \\ A^i{}_0 & A^i{}_j \end{pmatrix} = \begin{pmatrix} A^{00} & -A^{0j} \\ A^{i0} & -A^{ij} \end{pmatrix}, \tag{6.21}
$$

$$[A_\mu{}^\nu] = \begin{pmatrix} A_0{}^0 & A_0{}^j \\ A_i{}^0 & A_i{}^j \end{pmatrix} = \begin{pmatrix} A^{00} & A^{0j} \\ -A^{i0} & -A^{ij} \end{pmatrix}. \tag{6.22}$$

L'exemple des coordonnées comme quadrivecteur est particulièrement important car il permet d'introduire la matrice de changement de coordonnées de manière simple. Si l'on note x^μ et x'^μ les coordonnées respectives dans les deux référentiels, il existe nécessairement une transformation faisant passer des unes aux autres, $x'^\mu = x'^\mu(x^0, x^1, x^2, x^3)$, soit

$$\begin{aligned} \mathrm{d}x'^\mu &= \frac{\partial x'^\mu}{\partial x^\nu}\,\mathrm{d}x^\nu, \\ &= \frac{\partial x'^\mu}{\partial x^0}\,\mathrm{d}x^0 + \frac{\partial x'^\mu}{\partial x^1}\,\mathrm{d}x^1 + \frac{\partial x'^\mu}{\partial x^2}\,\mathrm{d}x^2 + \frac{\partial x'^\mu}{\partial x^3}\,\mathrm{d}x^3, \end{aligned}$$

ce qui permet d'identifier la transformation

$$\Lambda^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu}, \tag{6.23}$$

puisque $\mathrm{d}x^\mu$ est un quadrivecteur contravariant qui se transforme par définition avec la matrice $[\Lambda^\mu{}_\nu]$. On peut noter que les positions respectives en haut et en bas des indices μ et ν sont conformes aux prescriptions concernant la nature covariante de la quadri-dérivée. Par inversion de la relation de transformation des coordonnées, on peut écrire

$$\mathrm{d}x^\mu = \frac{\partial x^\mu}{\partial x'^\nu}\,\mathrm{d}x'^\nu,$$

ce qui identifie la matrice inverse

$$(\Lambda^{-1})^\mu{}_\nu = \frac{\partial x^\mu}{\partial x'^\nu}. \tag{6.24}$$

On a bien

$$\frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial x^\nu}{\partial x'^\mu} = \Lambda^\mu{}_\nu(\Lambda^{-1})^\nu{}_\mu = 1$$

et c'est bien cette matrice inverse qui assure la transformation des coordonnées covariantes, puisque

$$\begin{aligned} \mathrm{d}s^2 &= \mathrm{d}x_\mu\,\mathrm{d}x^\mu \\ &\qquad \overbrace{\qquad\qquad}^{\mathrm{d}x^\mu = (\Lambda^{-1})^\mu{}_\nu\,\mathrm{d}x'^\nu} \\ &= \mathrm{d}x_\mu(\Lambda^{-1})^\mu{}_\nu\,\mathrm{d}x'^\nu \\ &\quad \underbrace{\qquad\qquad}_{\mathrm{d}x'_\nu = (\Lambda^{-1})^\mu{}_\nu\,\mathrm{d}x_\mu} \\ &= \mathrm{d}x'_\nu\,\mathrm{d}x'^\nu. \end{aligned}$$

Par ailleurs, d'après les règles régissant l'élévation et l'abaissement des indices, on établit la relation

$$[\Lambda_\nu{}^\mu] = \begin{pmatrix} \Lambda_0{}^0 & \Lambda_0{}^j \\ \Lambda_i{}^0 & \Lambda_i{}^j \end{pmatrix} = \begin{pmatrix} \Lambda^0{}_0 & -\Lambda^0{}_j \\ -\Lambda^i{}_0 & \Lambda^i{}_j \end{pmatrix},$$

ce qui prouve, en utilisant les formes des matrices $[\Lambda^\mu{}_\nu]$ et $[(\Lambda^{-1})^\nu{}_\mu]$, que [18]

$$(\Lambda^{-1})^\mu{}_\nu \equiv \Lambda_\nu{}^\mu. \tag{6.25}$$

On note que la distinction entre $T^\mu{}_\nu$ et $T_\nu{}^\mu$ n'est importante que si le tenseur est non diagonal.

Par conséquent, dans les transformations du référentiel initial au référentiel primé, chaque indice contravariant fait apparaître un facteur $\Lambda^*{}_*$ et chaque indice covariant fait intervenir un terme $\Lambda_*{}^*$ et il apparaît autant de facteurs dans la transformation que le rang du tenseur. Les exemples précédents se généralisent aisément, par exemple

$$A'^\mu{}_{\nu\rho} = \Lambda^\mu{}_\sigma \Lambda_\nu{}^\tau \Lambda_\rho{}^\lambda A^\sigma{}_{\tau\lambda}.$$

A propos de $g_{\mu\nu}$, nous avons employé à plusieurs reprises le terme de tenseur métrique. Pour prouver qu'il s'agit bien d'un tenseur, il suffit de vérifier qu'il se transforme convenablement par changement de référentiel. Pour cela, on exprime à nouveau l'invariance de l'intervalle,

$$\begin{aligned} \mathrm{d}s^2 &= g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu \\ &= g_{\mu\nu}\frac{\partial x^\mu}{\partial x'^\sigma}\frac{\partial x^\nu}{\partial x'^\tau}\,\mathrm{d}x'^\sigma\,\mathrm{d}x'^\tau \\ &= g'_{\sigma\tau}\,\mathrm{d}x'^\sigma\,\mathrm{d}x'^\tau, \end{aligned}$$

soit

$$\begin{aligned} g'_{\sigma\tau} &= \frac{\partial x^\mu}{\partial x'^\sigma}\frac{\partial x^\nu}{\partial x'^\tau} g_{\mu\nu} \\ &= \Lambda_\sigma{}^\mu \Lambda_\tau{}^\nu g_{\mu\nu}, \end{aligned}$$

ce qui est bien la transformation d'un tenseur de rang deux, deux fois covariant. Ce point est particulièrement important, car il existe d'autres symboles obéissant aux mêmes règles concernant les indices, les changements de signes relatifs à leurs modifications de position, mais qui ne sont pas des tenseurs car ils n'obéissent pas

[18] Il suffit de noter le changement $\beta \to -\beta$ qui donne

$$[(\Lambda^{-1})^\mu{}_\nu] = \begin{pmatrix} (\Lambda^{-1})^0{}_0 & (\Lambda^{-1})^0{}_j \\ (\Lambda^{-1})^i{}_0 & (\Lambda^{-1})^i{}_j \end{pmatrix} = \begin{pmatrix} \Lambda^0{}_0 & -\Lambda^0{}_j \\ -\Lambda^i{}_0 & \Lambda^i{}_j \end{pmatrix}.$$

à la loi de transformation convenable par changement de référentiel inertiel. Par exemple $\Lambda^{\mu}{}_{\nu}$ est relatif à un changement de référentiel mais pas à une observable dans un référentiel donné, ce n'est pas un tenseur. Un autre tenseur important est le tenseur de Kronecker. Il apparaît par exemple lorsque l'on établit l'invariance du produit scalaire de deux quadrivecteurs,

$$\begin{aligned} A'_{\mu}B'^{\mu} &= \underbrace{\Lambda_{\mu}{}^{\sigma}\Lambda^{\mu}{}_{\tau}}_{\delta^{\sigma}{}_{\tau}} A_{\sigma}B^{\tau} \\ &= A_{\sigma}B^{\sigma}. \end{aligned}$$

$\delta^{\sigma}{}_{\tau}$ est le tenseur unité dont les composantes diagonales valent 1 et les composantes non diagonales sont nulles. Il joue un rôle particulier vis-à-vis du tenseur métrique. En effet, par abaissement du premier indice, on obtient un tenseur deux fois covariant qui n'est autre que $g_{\mu\nu}$ et par élévation du second indice, on obtient un tenseur deux fois contravariant qui coïncide avec $g^{\mu\nu}$:

$$\begin{aligned} \delta_{\mu\nu} &= g_{\mu\nu}, \\ \delta^{\mu\nu} &= g^{\mu\nu}. \end{aligned}$$

On note aussi qu'il n'est pas utile de distinguer les formes $\delta^{\mu}{}_{\nu}$ et $\delta_{\nu}{}^{\mu}$, plus simplement notées δ^{μ}_{ν}, car le tenseur de Kronecker est diagonal.

Exemples

Temps propre : Le temps propre d'un objet en mouvement est le temps tel qu'il s'écoule dans le référentiel propre, solidaire de l'objet. On peut l'exprimer sous diverses formes,

$$\mathrm{d}\tau = \frac{1}{c}\sqrt{\mathrm{d}s^2} = \frac{1}{c}\sqrt{\mathrm{d}x_{\mu}\,\mathrm{d}x^{\mu}} = \mathrm{d}t\sqrt{1-|\mathbf{v}|^2/c^2} = \frac{1}{\gamma_{\mathbf{v}}}\,\mathrm{d}t,$$

où $\mathbf{v}$, la vitesse de l'objet par rapport à un référentiel inertiel dans lequel le temps est noté t, peut éventuellement dépendre du temps (le référentiel propre peut subir une accélération sans pour autant que la relativité restreinte cesse d'être efficace). Le temps propre infinitésimal s'exprime par une contraction, c'est donc un scalaire invariant (tenseur de rang zéro si l'on veut). Par intégration le long de la ligne d'univers on obtient la durée propre entre deux événements,

$$\tau(ab) = \int_a^b \mathrm{d}t\sqrt{1-|\mathbf{v}(t)|^2/c^2}.$$

Cette relation est à comprendre au sens où l'on passe du référentiel d'un observateur (qui mesure t) à toute une succession de référentiels inertiels tangents à la ligne d'univers de la particule test, chacun étant supposé correctement synchronisé à l'instant coïncidant avec le référentiel de l'observateur.

Quadrivitesse : La vitesse ordinaire est définie en référence à un temps absolu,

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}.$$

On définit par analogie la 4-vitesse ou quadrivitesse en référence au seul temps privilégié[19] , le temps propre, et l'on a

$$v^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}. \tag{6.26}$$

Comme $\mathrm{d}\tau$ est une différentielle invariante, v^μ est un quadrivecteur contravariant. En fonction de la vitesse ordinaire, les composantes v^0 et v^i de la quadrivitesse s'écrivent

$$v^\mu = \left(c\frac{\mathrm{d}t}{\mathrm{d}\tau}, \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\tau} \right),$$

soit, avec $\mathrm{d}\tau = \sqrt{1-|\mathbf{v}|^2/c^2}\,\mathrm{d}t$ ou $\mathrm{d}t/\,\mathrm{d}\tau = \gamma_\mathbf{v}$,

$$\begin{aligned} v^\mu &= (v^0, \mathbf{v}) \\ &= (c\gamma_\mathbf{v}, \gamma_\mathbf{v}\mathbf{v}) = \left(\frac{c}{\sqrt{1-|\mathbf{v}|^2/c^2}}, \frac{\mathbf{v}}{\sqrt{1-|\mathbf{v}|^2/c^2}} \right). \end{aligned} \tag{6.27}$$

On constate que l'apparition du facteur $\gamma_\mathbf{v}$ devant la vitesse ordinaire pour les composantes spatiales de la quadrivitesse, v^i, est un effet cinématique dû au passage du temps propre au temps de l'observateur et que ce facteur n'a rien à voir avec un aspect dynamique ou inertiel lié à la masse. Dans ce cours, on utilisera toujours m pour la masse au repos et nous éviterons l'appelation masse en mouvement pour $\gamma_\mathbf{v} m$, comme on le voit parfois.

Disposant de la quadrivitesse, on peut former plusieurs invariants par contraction. Par exemple la norme invariante

$$v_\mu v^\mu = \frac{\mathrm{d}x_\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}s^2}{\mathrm{d}\tau^2} = c^2, \tag{6.28}$$

ou la combinaison

$$v_\mu\,\mathrm{d}x^\mu = \frac{\mathrm{d}x_\mu}{\mathrm{d}\tau}\,\mathrm{d}x^\mu = \frac{\mathrm{d}s^2}{\mathrm{d}\tau} = c^2\,\mathrm{d}\tau$$

qui permet de redéfinir le temps propre

$$\mathrm{d}\tau = \frac{1}{c^2} v_\mu\,\mathrm{d}x^\mu.$$

On peut retrouver la loi de transformation des composantes ordinaires de la vitesse par application de la transformation des composantes d'un quadrivecteur contravariant : $v'^\mu = \Lambda^\mu{}_\nu v^\nu$ donne

$$\begin{pmatrix} v'^0 \\ v'^1 \\ v'^2 \\ v'^3 \end{pmatrix} = \gamma_\mathbf{u} \begin{pmatrix} 1 & -\beta_\mathbf{u} & 0 & 0 \\ -\beta_\mathbf{u} & 1 & 0 & 0 \\ 0 & 0 & 1/\gamma_\mathbf{u} & 0 \\ 0 & 0 & 0 & 1/\gamma_\mathbf{u} \end{pmatrix} \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix},$$

[19] Certains auteurs définissent la quadrivitesse par rapport à l'intervalle, $v^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}s}$, cela simplifie certains facteurs c, mais au prix des dimensions car $v_\mu v^\mu$ par exemple n'est plus égal à c^2, mais vaut 1 (sans dimension). C'est le choix de Landau et Lifshitz par exemple, L.D. Landau et E. Lifshitz, *Théorie des champs*, Editions MIR, Moscou 1970.

$\mathbf{u}$ étant la vitesse du référentiel primé par rapport au référentiel initial [20] , portée comme d'habitude par l'axe 1. En faisant le rapport entre la deuxième et la première équation, on a

$$\frac{v'^1}{v'^0} = \frac{\gamma_{\mathbf{v}'} v'_x}{\gamma_{\mathbf{v}'} c} = \frac{-\beta_{\mathbf{u}} v^0 + v^1}{v^0 - \beta_{\mathbf{u}} v^1}, \quad \text{soit} \quad v'_x = \frac{v_x - u}{1 - u v_x / c^2},$$

ce qui établit la transformation des vitesses longitudinales. Le rapport entre troisième et première donne de même la transformation des vitesses transverses,

$$\frac{v'^2}{v'^0} = \frac{\gamma_{\mathbf{v}'} v'_y}{\gamma_{\mathbf{v}'} c} = \frac{1/\gamma_{\mathbf{u}} v^2}{v^0 - \beta_{\mathbf{u}} v^1}, \quad \text{soit} \quad v'_y = \frac{v_y \sqrt{1 - u^2/c^2}}{1 - u v_x / c^2}.$$

Quadricourant : L'électromagnétisme est compatible avec la relativité restreinte. Ses lois ont donc déjà une forme covariante, même si celle-ci n'est pas forcément manifeste. La loi de conservation de la charge électrique s'écrit

$$\nabla \mathbf{j} + \frac{\partial \rho}{\partial t} = 0 \tag{6.29}$$

et suggère l'introduction d'un quadrivecteur densité de courant

$$j^\mu = (j^0, j^i) = (c\rho, \mathbf{j}), \quad (\mathbf{j} = \mathbf{j}) \tag{6.30}$$

car on a alors une contraction

$$\partial_\mu j^\mu = \frac{1}{c} \frac{\partial}{\partial t} (c\rho) + \nabla \mathbf{j}$$

qui assure que la conservation de la charge électrique

$$\partial_\mu j^\mu = 0 \tag{6.31}$$

est satisfaite dans tous les référentiels inertiels dès qu'elle l'est dans l'un d'entre eux. On dispose ainsi automatiquement de la loi de transformation du quadricourant :

$$j'^\mu = \Lambda^\mu{}_\nu j^\nu .$$

Intérêt du formalisme tensoriel en relativité

L'intérêt essentiel de la notation tensorielle, outre sa forme extrêmement compacte, est qu'elle assure automatiquement qu'une expression obéit aux bonnes lois de transformation par changement de référentiel inertiel : si une loi physique s'exprime par une égalité entre deux tenseurs dans un référentiel donné,

$$G^{\mu\nu} = T^{\mu\nu}$$

[20] Pour éviter toute confusion entre les facteurs relativistes $\gamma_{\mathbf{u}}$ et $\beta_{\mathbf{u}}$ relatifs au changement de référentiel et $\gamma_{\mathbf{v}}$ et $\beta_{\mathbf{v}}$ relatifs au mouvement de la particule vue depuis $\mathcal{R}$, Jackson introduit pour les premiers la notation Γ et B.

par exemple, dans le référentiel primé elle prend la même forme entre les quantités primées correspondantes,

$$G'^{\mu\nu} = T'^{\mu\nu},$$

avec bien sûr $G'^{\mu\nu} = \Lambda^{\mu}{}_{\sigma}\Lambda^{\nu}{}_{\tau}G^{\sigma\tau}$ et $T'^{\mu\nu} = \Lambda^{\mu}{}_{\sigma}\Lambda^{\nu}{}_{\tau}T^{\sigma\tau}$. Une telle expression qui conserve la même forme est dite covariante (sous-entendu par changement de référentiel inertiel) et le formalisme consistant à écrire les lois physiques sous forme tensorielle est dit manifestement covariant (car sous cette forme, la covariance est manifeste). Un autre avantage du formalisme tensoriel est qu'il permet souvent une généralisation relativiste des lois classiques. A l'instar de l'analyse dimensionnelle, il produit en effet des contraintes fortes, puisqu'une égalité tensorielle ne peut intervenir qu'entre tenseurs de même rang et de même nature, ce qui limite en général la forme d'une loi physique, aux constantes sans dimension près. Par exemple le mouvement d'une charge dans un champ électrique est décrit classiquement par la relation fondamentale de la dynamique

$$m\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = q\mathbf{E}.$$

On verra que le champ électrique intervient, sous la forme E_x/c par exemple, dans les composantes d'un tenseur de rang deux, le tenseur champ électromagnétique $F_{\mu\nu}$. On généralise logiquement le premier membre de la relation fondamentale de la dynamique par l'expression $m\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2}$, tenseur contravariant de rang 1. Ecrire alors au second membre $qF^{\mu\nu}$ pose deux difficultés. D'une part du point de vue dimensionnel (c'est E^i/c qui intervient dans $F^{\mu\nu}$), on pourrait alors écrire $qF^{\mu\nu}c$, mais il reste la seconde difficulté, car cette expression est un tenseur de rang 2. Pour en faire un tenseur de rang 1, il faut assurer la contraction d'un des indices contravariants et le plus simple est de remplacer c par v_ν, soit

$$m\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} = qF^{\mu\nu}v_\nu.$$

Cette dernière expression est manifestement covariante, elle est correcte du point de vue dimensionnel, elle a donc toutes les chances d'être exacte, à une constante sans dimension près éventuellement.

7

Formulation covariante de l'électromagnétisme et de l'électrodynamique

L'électromagnétisme et les équations de Maxwell satisfont au principe de relativité. Il est toutefois très utile d'en donner une formulation covariante pour rendre cette propriété manifeste.

Tenseur champ électromagnétique et équations de Maxwell

On cherche à définir un tenseur représentant le champ électromagnétique (qu'on appelle souvent tenseur de Faraday). $A^\mu = (\phi/c, \mathbf{A})$ est un 4-vecteur, comme la jauge de Lorenz le laisse supposer, $\partial_\mu A^\mu = \frac{1}{c^2}\frac{\partial \phi}{\partial t} + \text{div}\,\mathbf{A} = 0$. Par ailleurs la relation $\mathbf{B} = \mathbf{rot}\,\mathbf{A}$ contient des termes de la forme

$$B_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}$$

qui suggèrent de définir un tenseur de rang 2 antisymétrique

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{7.1}$$

On a trivialement des zéros sur la diagonale, $F^{00} = F^{ii} = 0$ et $F^{\mu\nu} = -F^{\nu\mu}$ ou encore $F_{\mu\nu} = -F_{\nu\mu}$. Les composantes non nulles s'expriment explicitement en fonction des champs $\mathbf{E}$ et $\mathbf{B}$.

$$\begin{aligned} F^{0i} &= \partial^0 A^i - \partial^i A^0 \\ &= \partial_0 A^i + \partial_i A^0 \\ &= \frac{1}{c}\frac{\partial}{\partial t}A^i + \frac{\partial}{\partial x^i}\left(\frac{\phi}{c}\right) \\ &\equiv -\frac{1}{c}E^i, \end{aligned}$$

composante ordinaire de $\mathbf{E}$ le long de l'axe i.

$$\begin{aligned} F^{ij} &= \partial^i A^j - \partial^j A^i \\ &= -\partial_i A^j + \partial_j A^i \\ &= -\frac{\partial}{\partial x^i}A^j + \frac{\partial}{\partial x^j}A^i \\ &\equiv \pm B^k \quad \text{et permutations circulaires,} \end{aligned}$$

composante ordinaire de $\mathbf{B}$ (avec permutations circulaires sur les trois indices i, j, k). Finalement on obtient le tenseur champ électromagnétique deux fois contravariant :

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu = \frac{\partial A^\nu}{\partial x_\mu} - \frac{\partial A^\mu}{\partial x_\nu}.$$

$$[F^{\mu\nu}] = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}. \tag{7.2}$$

Le tenseur deux fois covariant

$$F_{\mu\nu} = g_{\mu\sigma} g_{\nu\tau} F^{\sigma\tau}$$

s'en déduit en conservant le signe de $F_{00} = F^{00}$ et des composantes purement spatiales, $F_{ij} = F^{ij}$ et en changeant le signe des composantes mixtes $F_{0j} = -F^{0j}$:

$$[F_{\mu\nu}] = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix}. \tag{7.3}$$

On définit également un tenseur champ électromagnétique dual. Celui-ci est formé au moyen du tenseur de Levi-Civita

$$\epsilon^{\mu\nu\sigma\tau} = \begin{cases} +1 & \text{si } \mu,\nu,\sigma,\tau = 0,1,2,3 \text{ et permutations paires} \\ -1 & \text{si permutations impaires} \\ 0 & \text{si deux indices ou plus sont égaux} \end{cases}$$

En particulier on a $\epsilon_{\mu\nu\sigma\tau} = -\epsilon^{\mu\nu\sigma\tau}$. On définit le tenseur dual $\bar{F}^{\mu\nu}$ par la contraction

$$\bar{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\sigma\tau}F_{\sigma\tau}.$$

Ses composantes s'obtiennent à partir de celles de $F^{\mu\nu}$ en changeant $\mathbf{E}/c$ en $\mathbf{B}$ et $\mathbf{B}$ en $-\mathbf{E}/c$:

$$[\bar{F}^{\mu\nu}] = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z/c & -E_y/c \\ B_y & -E_z/c & 0 & E_x/c \\ B_z & E_y/c & -E_x/c & 0. \end{pmatrix} \tag{7.4}$$

Equations de Maxwell covariantes

Construisons les contractions du tenseur champ électromagnétique, par exemple $\partial_\mu F^{\mu\nu}$. Pour $\nu = 0$, on a

$$\begin{aligned} \partial_\mu F^{\mu 0} &= \partial_0 F^{00} + \partial_1 F^{10} + \partial_2 F^{20} + \partial_3 F^{30} \\ &= 0 + \frac{\partial}{\partial x}(E_x/c) + \frac{\partial}{\partial y}(E_y/c) + \frac{\partial}{\partial z}(E_z/c) \\ &= \frac{1}{c}\operatorname{div}\mathbf{E} = \frac{1}{\varepsilon_0 c^2}\rho c \\ &= \mu_0 j^0. \end{aligned}$$

Pour $\nu = 1$, on a de même

$$\begin{aligned} \partial_\mu F^{\mu 1} &= \partial_0 F^{01} + \partial_1 F^{11} + \partial_2 F^{21} + \partial_3 F^{31} \\ &= \frac{1}{c}\frac{\partial}{\partial t}(-E_x/c) + 0 + \frac{\partial}{\partial y}B_z + \frac{\partial}{\partial z}(-B_y) \\ &= -\frac{1}{c^2}\frac{\partial E_x}{\partial t} + (\mathbf{rot}\,\mathbf{B})\mathbf{u}_x \\ &= \mu_0 j^1, \end{aligned}$$

et les deux dernières composantes spatiales s'en déduisent par permutation. On a donc

$$\partial_\mu F^{\mu\nu} = \mu_0 j^\nu, \tag{7.5}$$

expression qui regroupe les deux équations de Maxwell avec sources,

$$\begin{aligned} \operatorname{div}\mathbf{E} &= \frac{\rho}{\varepsilon_0} \\ \mathbf{rot}\,\mathbf{B} - \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} &= \mu_0 \mathbf{j}. \end{aligned} \tag{7.6}$$

Le second couple d'équations de Maxwell s'obtient plus aisément après l'introduction du tenseur champ électromagnétique dual. Examinons la contraction $\partial_\mu \bar{F}^{\mu\nu}$. Pour $\nu = 0$, on a

$$\begin{aligned}\partial_\mu \bar{F}^{\mu 0} &= \partial_0 \bar{F}^{00} + \partial_1 \bar{F}^{10} + \partial_2 \bar{F}^{20} + \partial_3 \bar{F}^{30} \\ &= 0 + \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} \\ &= \operatorname{div} \mathbf{B}.\end{aligned}$$

Pour $\nu = 1$, on a de même

$$\begin{aligned}\partial_\mu \bar{F}^{\mu 1} &= \partial_0 \bar{F}^{01} + \partial_1 \bar{F}^{11} + \partial_2 \bar{F}^{21} + \partial_3 \bar{F}^{31} \\ &= \frac{1}{c}\frac{\partial}{\partial t}(-B_x) + 0 + \frac{\partial}{\partial y}(-E_z/c) + \frac{\partial}{\partial z}(E_y/c) \\ &= -\frac{1}{c}\frac{\partial B_x}{\partial t} - \frac{1}{c}(\mathbf{rot}\,\mathbf{E})\mathbf{u}_x.\end{aligned}$$

On en déduit l'expression unifiée du second couple d'équations de Maxwell, les équations sans second membre,

$$\partial_\mu \bar{F}^{\mu\nu} = 0. \tag{7.7}$$

Les équations de Maxwell manifestement covariantes s'écrivent donc :

$$\begin{aligned}\partial_\mu F^{\mu\nu} &= \mu_0 j^\nu, \\ \partial_\mu \bar{F}^{\mu\nu} &= \partial_\mu F^{\nu\lambda} + \partial_\nu F^{\lambda\mu} + \partial_\lambda F^{\mu\nu} = 0.\end{aligned} \tag{7.8}$$

Force de Lorentz

Comme on l'a déjà suggéré, la notation covariante est un guide pour l'expression des lois physiques en relativité. Reprenons l'exemple du mouvement d'une charge q dans un champ électromagnétique. On sait maintenant que ce champ s'exprime sous la forme d'un tenseur de rang 2, $F^{\mu\nu}$ et que ses composantes spatio-temporelles F^{0i} font intervenir E_{x_i}/c alors que les composantes purement spatiales sont liées à celles de $\mathbf{B}$. Il est donc logique de généraliser la relation fondamentale de la dynamique

$$m\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = q(\mathbf{E} + \mathbf{v} \wedge \mathbf{B})$$

à une expression covariante,

$$m\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = qF^{\mu\nu}v_\nu.$$

Pour interpréter les composantes de cette équation, on peut l'écrire sous forme matricielle et passer à la limite classique $c \to \infty$.

$$m\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{pmatrix}\gamma_{\mathbf{v}}c \\ \gamma_{\mathbf{v}}v_x \\ \gamma_{\mathbf{v}}v_y \\ \gamma_{\mathbf{v}}v_z\end{pmatrix} = q\begin{pmatrix}0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0\end{pmatrix}\begin{pmatrix}\gamma_{\mathbf{v}}c \\ -\gamma_{\mathbf{v}}v_x \\ -\gamma_{\mathbf{v}}v_y \\ -\gamma_{\mathbf{v}}v_z\end{pmatrix}$$

La première équation s'écrit encore

$$mc\frac{\mathrm{d}\gamma_{\mathbf{v}}}{\mathrm{d}\tau} = \frac{q}{c}\gamma_{\mathbf{v}}\mathbf{E}\cdot\mathbf{v}$$

où il apparaît la puissance classique

$$q\mathbf{E}\cdot\mathbf{v} = \mathcal{P} = -\vec{\nabla}U(\mathbf{r})\cdot\mathbf{v} = -\frac{\mathrm{d}U(\mathbf{r})}{\mathrm{d}t}$$

(la partie magnétique ne travaille pas). En développant la dérivée du premier membre on obtient

$$\frac{m}{1-|\mathbf{v}|^2/c^2}\mathbf{v}\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}\tau} = q\mathbf{E}\cdot\mathbf{v},$$

soit, à la limite $c\to\infty$,

$$m\mathbf{v}\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}m|\mathbf{v}|^2\right) = \mathcal{P} = -\frac{\mathrm{d}U(\mathbf{r})}{\mathrm{d}t}.$$

Le contenu énergétique de la première équation est ainsi lié à la définition de la puissance et au théorème de l'énergie cinétique. Pour ce qui est des équations spatiales, on peut par exemple considérer le cas de la première qui donne

$$m\frac{\mathrm{d}(\gamma_{\mathbf{v}}v_x)}{\mathrm{d}\tau} = q\gamma_{\mathbf{v}}(E_x + B_z v_y - B_y v_z)$$

où l'on voit clairement apparaître la force de Lorentz au second membre.

Champ d'une charge ponctuelle en mouvement arbitraire

▷ *Quadripotentiel A^μ.* On considère une charge q animée d'un mouvement arbitraire par rapport à un observateur situé en $\mathbf{r}_0$. A l'instant t_0 (mesuré dans son propre référentiel), le champ électromagnétique ressenti est celui de la charge lorsqu'elle était à la position retardée $\mathbf{r}_q(t_r)$ telle que

$$|\mathbf{r}_q(t_r) - \mathbf{r}_0| = c(t_0 - t_r)$$

en raison de la vitesse finie de propagation des interactions électromagnétiques. On introduit pour la suite la notation $\mathbf{R}(t_r) = \mathbf{r}_q(t_r) - \mathbf{r}_0$ et la notation quadridimensionnelle associée, $R_\sigma(t_r) = (c(t_0 - t_r), -\mathbf{R}(t_r))$. Ce quadrivecteur est par définition du genre lumière, puisque

$$R_\sigma(t_r)R^\sigma(t_r) = c^2(t_0 - t_r)^2 - |\mathbf{R}(t_r)|^2 = 0.$$

Dans la notation tensorielle, on appelle x_0^μ la quadriposition de l'observateur et $z^\mu(t)$ celle de la charge.

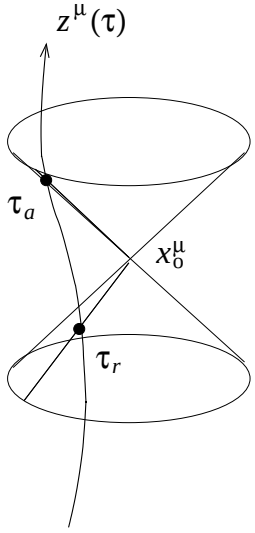


Figure 7.1 Cône de lumière pour un observateur en x_0^μ. La ligne représente la ligne d'univers $z^\mu(\tau)$ d'une charge q en mouvement arbitraire qui intercepte le cône de lumière à l'instant retardé τ_r (mesuré dans son référentiel propre) puis à nouveau à l'instant avancé τ_a.

Pour trouver la forme du quadripotentiel $A^\mu(x_0^\nu)$ au point $\mathbf{r}_0, t_0$ dans le référentiel lié à l'observateur, on utilise la covariance des expressions tensorielles, c'est-à-dire que l'on se place tout d'abord dans le référentiel propre de la charge (le temps propre dans ce référentiel $\mathcal{R}'$ qui accompagne la charge dans son mouvement est noté τ et les quantités exprimées dans ce référentiel sont notées avec des primes) où le quadripotentiel prend une expression simplifiée, puisque seule la partie temporelle est non nulle. On exprime cette partie temporelle sous forme manifestement covariante (c'est-à-dire comme un tenseur contravariant de rang 1) et l'expression correspondante dans le référentiel original en découle immédiatement.

Dans le référentiel propre de la charge, le quadripotentiel se réduit à la partie électrostatique, soit

$$\phi'(\mathbf{r}'_0, \tau_0) = \frac{q}{4\pi\varepsilon_0} \frac{1}{|\mathbf{R}'(\tau_r)|}.$$

Comme on a dans $\mathcal{R}'$, tout comme dans $\mathcal{R}$ d'ailleurs,

$$R'_\sigma(\tau_r) R'^\sigma(\tau_r) = c^2(\tau_0 - \tau_r)^2 - |\mathbf{R}'(\tau_r)|^2 = 0,$$

où $R'^\sigma(\tau_r) = (c(\tau_0 - \tau_r), \mathbf{R}'(\tau_r))$, on peut écrire $|\mathbf{R}'(\tau_r)| = c(\tau_0 - \tau_r)$ et

$$\frac{1}{c}\phi'(\mathbf{r}'_0, \tau_0) = \frac{q}{4\pi\varepsilon_0 c^2(\tau_0 - \tau_r)}. \tag{7.9}$$

Dans ce même référentiel, on note que

$$\mathbf{A}'(\mathbf{r}'_0, \tau_0) = \vec{0}. \tag{7.10}$$

Le point essentiel est que dans le référentiel $\mathcal{R}'$, le quadripotentiel n'a qu'une partie temporelle non nulle, il doit donc être proportionnel à la quadrivitesse qui,

dans ce même référentiel, prend la forme simplifiée $v'^{\mu}(\tau) = (c, \vec{0})$. C'est de plus compatible avec l'exigence d'obtenir pour A'^{μ} un tenseur contravariant de rang 1. Il faut alors faire également apparaître la quadrivitesse au dénominateur pour des raisons dimensionnelles, mais dans une contraction pour obtenir un scalaire au dénominateur. On forme ainsi la quantité

$$\frac{v'^{\mu}(\tau_r)}{v'_{\sigma}(\tau_r)R'^{\sigma}(\tau_r)} = \left(\frac{c}{c^2(\tau_0 - \tau_r)}, \vec{0}\right)$$

qui permet d'écrire

$$A'^{\mu}(x'^{\nu}_0) = \frac{q}{4\pi\varepsilon_0 c}\frac{v'^{\mu}(\tau_r)}{v'_{\sigma}(\tau_r)R'^{\sigma}(\tau_r)} = \frac{q}{4\pi\varepsilon_0 c}\frac{[v'^{\mu}]_{ret.}}{[v'_{\sigma}R'^{\sigma}]_{ret.}},$$

où la notation fréquemment employée $[\ldots]_{ret.}$ signifie que la quantité entre crochets est évaluée au temps retardé.

L'intérêt de cette expression est évident : il s'agit d'une forme tensorielle, donc covariante par changement de référentiel inertiel, elle se généralise immédiatement à tout autre référentiel inertiel, en particulier $\mathcal{R}$.

$$A^{\mu}(x^{\nu}_0) = \frac{q}{4\pi\varepsilon_0 c}\frac{v^{\mu}(t_r)}{v_{\sigma}(t_r)R^{\sigma}(t_r)} = \frac{q}{4\pi\varepsilon_0 c}\frac{[v^{\mu}]_{ret.}}{[v_{\sigma}R^{\sigma}]_{ret.}}. \tag{7.11}$$

En exprimant dans $\mathcal{R}$ les composantes de v^{μ} et le produit invariant $v_{\sigma}R^{\sigma}$, (il n'y a pas d'ambiguïté ici, on notera donc γ pour $\gamma_{\mathbf{v}}$)

$$\begin{aligned} v^{\sigma}(t_r) &= (\gamma c, \gamma\mathbf{v}(t_r)) \\ R^{\sigma}(t_r) &= (c(t_0 - t_r), \mathbf{R}(t_r)) \\ v_{\sigma}(t_r)R^{\sigma}(t_r) &= \gamma c^2(t_0 - t_r) - \gamma\mathbf{v}(t_r)\mathbf{R}(t_r) \\ &= \gamma(|\mathbf{R}(t_r)|c - \mathbf{R}(t_r)\mathbf{v}(t_r)). \end{aligned}$$

on obtient la forme explicite du potentiel électrostatique et du potentiel vecteur dans le référentiel de l'observateur,

$$\begin{aligned} A^{\mu}(\mathbf{r}_0, t_0) &= \frac{q}{4\pi\varepsilon_0 c}\left(\frac{c}{|\mathbf{R}(t_r)|c - \mathbf{R}(t_r)\mathbf{v}(t_r)}, \frac{\mathbf{v}(t_r)}{|\mathbf{R}(t_r)|c - \mathbf{R}(t_r)\mathbf{v}(t_r)}\right) \\ \phi(\mathbf{r}_0, t_0) &= \frac{q}{4\pi\varepsilon_0|\mathbf{R}(t_r)|}\left(1 - \mathbf{R}(t_r)\mathbf{v}(t_r)/|\mathbf{R}(t_r)|c\right)^{-1} \\ \mathbf{A}(\mathbf{r}_0, t_0) &= \frac{q\mathbf{v}(t_r)}{4\pi\varepsilon_0 c^2|\mathbf{R}(t_r)|}\left(1 - \mathbf{R}(t_r)\mathbf{v}(t_r)/|\mathbf{R}(t_r)|c\right)^{-1}. \end{aligned} \tag{7.12}$$

On introduit souvent la notation abrégée $\boldsymbol{\beta}(t) = \mathbf{v}(t)/c$ et le vecteur unitaire dans la direction charge-observateur, $\hat{\mathbf{u}}_R(t) = \mathbf{R}(t)/|\mathbf{R}(t)|$, ce qui conduit à

$$\begin{aligned} \phi(\mathbf{r}_0, t_0) &= \frac{1}{4\pi\varepsilon_0}\frac{q}{|\mathbf{R}(t_r)|}\left(1 - \boldsymbol{\beta}(t_r)\hat{\mathbf{u}}_R(t_r)\right)^{-1} \\ \mathbf{A}(\mathbf{r}_0, t_0) &= \frac{\mu_0}{4\pi}\frac{q\mathbf{v}(t_r)}{|\mathbf{R}(t_r)|}\left(1 - \boldsymbol{\beta}(t_r)\hat{\mathbf{u}}_R(t_r)\right)^{-1}. \end{aligned} \tag{7.13}$$

Ces potentiels sont connus sous le nom de potentiels de Liénard-Wiechert[21] .

▷ *Composantes ordinaires des champs* $\mathbf{E}$ *et* $\mathbf{B}$ *de la charge en mouvement arbitraire.* Après des calculs fastidieux, on exprime le champ électrique

$$\mathbf{E}(\mathbf{r}_0,t_0) = \frac{q}{4\pi\varepsilon_0[|\mathbf{R}|^2]_{ret.}}(1-[\boldsymbol{\beta}\hat{\mathbf{u}}_R]_{ret.})^{-3} \times \left[(1-|\boldsymbol{\beta}|^2)(\hat{\mathbf{u}}_R-\boldsymbol{\beta}) + \frac{|\mathbf{R}|}{c^2}(\mathbf{a}\hat{\mathbf{u}}_R(\hat{\mathbf{u}}_R-\boldsymbol{\beta}) - \mathbf{a}(1-\boldsymbol{\beta}\hat{\mathbf{u}}_R))\right]_{ret.},$$

et en utilisant $\mathbf{a}\hat{\mathbf{u}}_R = a_\parallel$ et $\mathbf{a}-(\mathbf{a}\hat{\mathbf{u}}_R)\hat{\mathbf{u}}_R = \mathbf{a}_\perp$ et de même pour les composantes de $\boldsymbol{\beta}$ on obtient

$$\mathbf{E}(\mathbf{r}_0,t_0) = \frac{q}{4\pi\varepsilon_0[|\mathbf{R}|^2]_{ret.}(1-[\boldsymbol{\beta}\hat{\mathbf{u}}_R]_{ret.})^3} \times \left[(1-|\boldsymbol{\beta}|^2)(\hat{\mathbf{u}}_R-\boldsymbol{\beta}) - \frac{|\mathbf{R}|}{c^2}(\mathbf{a}_\perp - a_\parallel\boldsymbol{\beta} - \beta_\parallel\mathbf{a})\right]_{ret.}. \tag{7.14}$$

Le champ total est constitué d'une partie statique et d'une partie proportionnelle à la vitesse retardée, toutes deux décroissant rapidement en $1/R^2$ et d'une partie faisant intervenir l'accélération, décroissant bien plus lentement en $1/R$ et qui constitue le champ de rayonnement.

Le champ magnétique se déduit du champ électrique par

$$\mathbf{B}(\mathbf{r}_0,t_0) = \frac{[\hat{\mathbf{u}}_R]_{ret.}}{c}\wedge\mathbf{E}(\mathbf{r}_0,t_0). \tag{7.15}$$

Les expressions précédentes des champs $\mathbf{E}$ et $\mathbf{B}$ sont dues à Liénard et Wiechert. Il existe également une autre forme, due à Heaviside et Feynman[22] ,

$$\mathbf{E}(\mathbf{r}_0,t_0) = \frac{q}{4\pi\varepsilon_0}\left(\frac{\hat{\mathbf{u}}_R}{|\mathbf{R}|^2} + \frac{|\mathbf{R}|}{c}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\hat{\mathbf{u}}_R}{|\mathbf{R}|^2} + \frac{1}{c^2}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\hat{\mathbf{u}}_R\right).$$

▷ *Cas limites.* Partant des expressions générales, on peut considérer divers cas limites intéressants. Tout d'abord pour une particule immobile, on retrouve bien le champ coulombien habituel (heureusement !),

$$\mathbf{E}(\mathbf{r}_0,t_0) = \frac{q}{4\pi\varepsilon_0|\mathbf{R}|^2}\hat{\mathbf{u}}_R. \tag{7.16}$$

Pour une particule en mouvement uniforme, on obtient un champ décroissant en $1/|\mathbf{R}|^2$, de type coulombien,

$$\mathbf{E}(\mathbf{r}_0,t_0) = \frac{q}{4\pi\varepsilon_0}\left[\frac{(1-|\boldsymbol{\beta}|^2)(\hat{\mathbf{u}}_R-\boldsymbol{\beta})}{|\mathbf{R}|^2(1-\boldsymbol{\beta}\hat{\mathbf{u}}_R)^3}\right]_{ret.}. \tag{7.17}$$

[21] Ces potentiels ont été obtenus indépendamment avant l'élaboration de la relativité restreinte par Liénard en 1898 et par Wiechert en 1900.

[22] L. Eyges, The classical electromagnetic field, Dover, New York 1972, p. 281 et R.P. Feynman, R.B. Leighton and M. Sands, The Feynman lectures on physics, vol. I chap. 28 et vol. II chap. 21, Addison-Wesley, Reading 1977.

Il est remarquable de noter qu'en l'absence d'accélération (vitesse $\mathbf{v}(t_r) = \mathbf{v}$ uniforme), le champ électrique est centré sur la position actuelle que la particule occupe à l'instant t_0. On a en effet

$$\hat{\mathbf{u}}_R(t_r) - \boldsymbol{\beta}(t_r) = \frac{\mathbf{R}(t_r)}{c(t_0 - t_r)} - \frac{\mathbf{v}}{c},$$

or $\mathbf{R}(t_0) = \mathbf{R}(t_r) - (t_0 - t_r)\mathbf{v}$, d'où

$$\hat{\mathbf{u}}_R(t_r) - \boldsymbol{\beta}(t_r) = \frac{\mathbf{R}(t_0)}{c(t_0 - t_r)}.$$

Il n'y a donc finalement aucun effet de retard et le champ est essentiellement coulombien et centré sur la position instantanée de la charge.

Si l'on considère maintenant une charge accélérée ($\mathbf{a}$ non nul), mais ayant une faible vitesse de sorte que $\beta \ll 1$ soit négligeable, on obtient deux contributions dominantes,

$$\mathbf{E}(\mathbf{r}_0, t_0) = \frac{q}{4\pi\varepsilon_0}\left[\frac{\hat{\mathbf{u}}_R}{|\mathbf{R}|^2}\right]_{ret.} - \frac{q}{4\pi\varepsilon_0 c^2}\left[\frac{\mathbf{a}_\perp}{|\mathbf{R}|}\right]_{ret.}. \tag{7.18}$$

▷ *Tenseur champ électromagnétique $F^{\mu\nu}$.* Le formalisme tridimensionnel, on l'a vu au paragraphe précédent, est lourd à manipuler et les calculs sont finalement plus simples s'ils sont menés à l'aide du formalisme covariant[23] . On calcule les dérivées à l'instant retardé

$$[\partial^\mu A^\nu]_{ret.} = \frac{\mu_0 q}{4\pi}\left[\frac{\partial}{\partial x_\mu}\left(\frac{v^\nu}{\rho}\right)\right]_{ret.},$$

où l'on a posé

$$\rho = v_\sigma R^\sigma / c.$$

On développe la dérivée

$$\begin{aligned}[\partial^\mu A^\nu]_{ret.} &= \frac{\mu_0 q}{4\pi}\left[\frac{1}{\rho}\frac{\partial v^\nu}{\partial x_\mu} + v^\nu \frac{\partial}{\partial x_\mu}\left(\frac{1}{\rho}\right)\right]_{ret.} \\ &= \frac{\mu_0 q}{4\pi}\left[\frac{1}{\rho}\partial^\mu v^\nu + v^\nu \partial^\mu\left(\frac{1}{\rho}\right)\right]_{ret.},\end{aligned}$$

puis on utilise

$$\partial^\mu v^\nu = \frac{\partial v^\nu}{\partial \tau}\frac{\partial \tau}{\partial x_\mu} = a^\nu \partial^\mu \tau$$

$$\partial^\mu\left(\frac{1}{\rho}\right) = \frac{\partial}{\partial \rho}\left(\frac{1}{\rho}\right)\frac{\partial \rho}{\partial x_\mu} = -\frac{1}{\rho^2}\partial^\mu \rho,$$

[23] On en trouve l'exposé détaillé dans F. Rohrlich, Classical charged particles, Addison-Wesley, Redwood 1965, section 4-8, p. 83.

il vient ensuite

$$[\partial^\mu A^\nu]_{ret.} = \frac{\mu_0 q}{4\pi}\left([a^\nu \partial^\mu \tau]_{ret.} - \frac{1}{[\rho^2]_{ret.}}[\partial^\mu \rho]_{ret.}\right),$$

d'où l'on déduit le tenseur champ électromagnétique au point d'observation en fonction de la position retardée de la charge.

$$\begin{aligned} F^{\mu\nu}(x_0^\sigma) = \frac{\mu_0 q}{4\pi}([a^\nu \partial^\mu \tau]_{ret.} - [a^\mu \partial^\nu \tau]_{ret.} \\ - [\partial^\mu \rho]_{ret.}/[\rho^2]_{ret.} + [\partial^\nu \rho]_{ret.}/[\rho^2]_{ret.}). \end{aligned} \tag{7.19}$$

8

Formulation lagrangienne et relativité restreinte

Le lagrangien adapté à l'électromagnétisme (et donc équivalent aux équations de Maxwell) a été obtenu par Schwarzschild en 1903, avant la théorie de la relativité restreinte, mais sa forme n'est pas évidente. En relativité, l'action étant un scalaire, elle doit se mettre sous la forme d'un invariant de Lorentz, c'est-à-dire qu'on doit pouvoir l'écrire à partir d'une contraction, ou plus précisément comme l'intégrale d'une contraction. Ces prescriptions rendent plus facile l'introduction du lagrangien de Schwarzschild.

Le principe de moindre action pour la particule libre

Partons de la particule libre en physique classique, $L(\mathbf{r},\mathbf{v}) = T(\mathbf{v}) = \frac{1}{2}m|\mathbf{v}|^2$ où $\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$, soit

$$S_{class.} = \frac{1}{2}m\int \mathbf{v}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\,\mathrm{d}t = \frac{1}{2}m\int \mathbf{v}\,\mathrm{d}\mathbf{r}$$

et $\delta S_{class.} = 0$ s'écrit encore [24]

$$\delta\left(m\int \mathbf{v}\,\mathrm{d}\mathbf{r}\right) = 0. \tag{8.1}$$

[24] On omet le facteur 1/2 que l'on retrouvera à partir de l'expression relativiste par le développement limité d'une racine carrée.

Or $\mathbf{v}\,\mathrm{d}\mathbf{r}$ est relié à la partie spatiale du produit invariant [25] $v_\mu\,\mathrm{d}x^\mu$, puisque

$$\begin{aligned} \mathrm{d}x^\mu &= (c\,\mathrm{d}t,\,\mathrm{d}\mathbf{r}),\\ v_\mu &= \frac{\mathrm{d}x_\mu}{\mathrm{d}\tau} = \left(c\frac{\mathrm{d}t}{\mathrm{d}\tau}, -\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\tau}\right) = (c\gamma, -\gamma\mathbf{v}),\\ v_\mu\,\mathrm{d}x^\mu &= \gamma c^2\,\mathrm{d}t - \gamma\mathbf{v}\,\mathrm{d}\mathbf{r}. \end{aligned}$$

Comme $\gamma \to 1$ dans la limite non relativiste, on passe du cas classique au cas relativiste en faisant simplement

$$\delta\left(m\int \mathbf{v}\,\mathrm{d}\mathbf{r}\right) = 0 \;\longrightarrow\; \delta\left(-m\int v_\mu\,\mathrm{d}x^\mu\right) = 0. \tag{8.2}$$

On peut donner des formes équivalentes en notant que

$$\begin{aligned} -mv_\mu\,\mathrm{d}x^\mu &= -m\frac{\mathrm{d}x_\mu}{\mathrm{d}\tau}\,\mathrm{d}x^\mu\\ &= -m\frac{\mathrm{d}s^2}{\mathrm{d}\tau}\\ &= -mc\,\mathrm{d}\tau\frac{\mathrm{d}s}{\mathrm{d}\tau}\\ &= -mc\,\mathrm{d}s \end{aligned}$$

ce qui conduit à l'expression traditionnelle de l'action de la particule libre en relativité restreinte

$$S_{rel.} = -mc\int \mathrm{d}s. \tag{8.3}$$

Elle est simplement donnée (au facteur numérique $-mc$ près) par la longueur de l'intervalle le long de la ligne d'univers de la particule libre. C'est une interprétation géométrique élégante où l'on voit que la partie cinétique de l'énergie est totalement intégrée dans la géométrie de l'espace-temps. Cette quantité par ailleurs est un scalaire de Lorentz, ce qui assure en particulier comme c'est un invariant que si $\delta S_{rel.} = 0$ dans un référentiel inertiel, ce sera également vrai dans tous les autres référentiels inertiels.

Tableau 8.1 Principe de moindre action pour la particule libre.

Dynamique newtonienne		Dynamique relativiste
$\delta\left(m\int \mathbf{v}\,\mathrm{d}\mathbf{r}\right) = 0$	$\to$	$\delta\left(-m\int v_\mu\,\mathrm{d}x^\mu\right) = 0$
		$\delta\left(-mc\int \mathrm{d}s\right) = 0$

[25] Le fait qu'il s'agisse d'une forme invariante par changement de référentiel inertiel est bien entendu capital pour assurer la validité de la forme de l'action dans tous ces référentiels équivalents.

En notation tridimensionnelle explicite on développe

$$\begin{aligned} S_{rel.} &= -mc \int \mathrm{d}s \\ &= -mc \int (c^2\,\mathrm{d}t^2 - \mathrm{d}\mathbf{r}^2)^{1/2} \\ &= -mc^2 \int \mathrm{d}t \sqrt{1 - |\mathbf{v}|^2/c^2}, \end{aligned}$$

ce qui conduit au lagrangien

$$\begin{aligned} L &= -mc^2 \sqrt{1 - |\mathbf{v}|^2/c^2} \\ &\simeq -mc^2 \left(1 - \frac{1}{2}\frac{|\mathbf{v}|^2}{c^2} + \dots\right) \\ &\simeq -mc^2 + \frac{1}{2} m|\mathbf{v}|^2 + \dots \end{aligned}$$

A la constante $-mc^2$ près on retrouve bien le lagrangien classique de la particule libre. On peut alors calculer l'impulsion :

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - |\mathbf{v}|^2/c^2}}, \tag{8.4}$$

et l'énergie totale

$$E = \frac{mc^2}{\sqrt{1 - |\mathbf{v}|^2/c^2}}. \tag{8.5}$$

Cette dernière quantité s'exprime également en fonction de l'impulsion pour donner l'hamiltonien,

$$H = c\sqrt{|\mathbf{p}|^2 + m^2c^2}. \tag{8.6}$$

L'équation d'Euler-Lagrange ici donne simplement

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \mathbf{v}}\right) = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \mathbf{0},$$

comme attendu pour la particule libre.

Le principe de moindre action dans le formalisme de Minkowski

On revient à l'expression de l'action exprimée sous forme manifestement covariante,

$$\begin{aligned} S_{rel.} &= -m \int v_\mu \,\mathrm{d}x^\mu \\ &= -mc \int \mathrm{d}s \\ &= -mc \int (\,\mathrm{d}x_\mu \,\mathrm{d}x^\mu)^{1/2}, \end{aligned}$$

soit

$$\delta S_{rel.} = -mc \int \delta(\mathrm{d}x_\mu \mathrm{d}x^\mu)^{1/2},$$

avec

$$\begin{aligned}\delta(\mathrm{d}x_\mu \mathrm{d}x^\mu)^{1/2} &= \frac{1}{2}(\mathrm{d}x_\mu \mathrm{d}x^\mu)^{-1/2} \underbrace{[\delta(\mathrm{d}x_\mu)\,\mathrm{d}x^\mu + \mathrm{d}x_\mu \delta(\mathrm{d}x^\mu)]}_{2\,\mathrm{d}x^\mu \delta(\mathrm{d}x_\mu)}, \\ &= \frac{\mathrm{d}x^\mu}{\mathrm{d}s}\delta(\mathrm{d}x_\mu) \\ &= \frac{1}{c}v^\mu \delta(\mathrm{d}x_\mu).\end{aligned}$$

On a de plus $\delta(\mathrm{d}x_\mu) = \mathrm{d}(\delta x_\mu)$, soit

$$\begin{aligned}\delta S_{rel.} &= -m\int_a^b v^\mu \mathrm{d}(\delta x_\mu) \\ &= -m[v^\mu \delta x_\mu]_a^b + m\int_a^b \mathrm{d}v^\mu \delta x_\mu.\end{aligned}$$

Ici, a et b repèrent les coordonnées d'espace-temps de deux événements entre lesquels on minimise l'action et le terme de bords s'annule. Il est naturel de paramétrer la ligne d'univers par le temps propre, soit

$$\delta S_{rel.} = m\int_a^b \frac{\mathrm{d}v^\mu}{\mathrm{d}\tau}\mathrm{d}\tau \delta x_\mu = 0$$

et cette égalité étant valable pour une variation arbitraire de la trajectoire d'univers, on a

$$a^\mu = \frac{\mathrm{d}v^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} = 0, \tag{8.7}$$

c'est-à-dire une généralisation évidente de l'équation classique de la trajectoire de la particule libre [26] .

[26] On peut noter que s'il s'agit de photons, $\mathrm{d}\tau$ n'est pas défini (puisque $\mathrm{d}s = 0$), et dans ce cas on introduit un paramètre arbitraire λ qui évolue le long de la ligne d'univers et

$$\begin{aligned}\delta S_{rel.} &= m\int_a^b \frac{\mathrm{d}v^\mu}{\mathrm{d}\lambda}\mathrm{d}\lambda \delta x_\mu = 0, \\ \frac{\mathrm{d}v^\mu}{\mathrm{d}\lambda} &= \frac{\mathrm{d}^2x^\mu}{\mathrm{d}\lambda^2} = 0.\end{aligned} \tag{8.8}$$

Particule chargée sous champ

On doit tout d'abord former une action invariante de Lorentz. Pour cela, la seule caractéristique électromagnétique de la particule étant sa charge et la seule caractéristique du champ étant le quadripotentiel, on forme la quantité $qA_\mu \, \mathrm{d}x^\mu$, soit

$$S_q = \int_a^b (-mc \, \mathrm{d}s - qA_\mu \, \mathrm{d}x^\mu)$$

où le signe est justifié par la limite non relativiste et où

$$A^\mu = (\phi/c, \mathbf{A}).$$

On écrit encore

$$S_q = \int_a^b (-mc(\, \mathrm{d}x_\mu \, \mathrm{d}x^\mu)^{1/2} - qA_\mu \, \mathrm{d}x^\mu). \tag{8.9}$$

Sous forme explicite, on a

$$\begin{aligned} S_q &= \int_a^b \left(-mc^2 \, \mathrm{d}t \sqrt{1 - |\mathbf{v}|^2/c^2} - q\frac{\phi}{c} c \, \mathrm{d}t + q\mathbf{A} \, \mathrm{d}\mathbf{r} \right) \\ &= \int_a^b \left(-mc^2 \sqrt{1 - |\mathbf{v}|^2/c^2} - q\phi + q\mathbf{A}\mathbf{v} \right) \mathrm{d}t \end{aligned}$$

ce qui conduit au lagrangien relativiste

$$L = -mc^2 \sqrt{1 - |\mathbf{v}|^2/c^2} - q\phi + q\mathbf{A}\mathbf{v}. \tag{8.10}$$

Dans la limite non relativiste, il devient

$$L \simeq -mc^2 + \frac{1}{2} m|\mathbf{v}|^2 - q\phi + q\mathbf{A}\mathbf{v}.$$

L'impulsion

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1 - |\mathbf{v}|^2/c^2}} + q\mathbf{A} \tag{8.11}$$

s'exprime en fonction de la quantité de mouvement

$$\boldsymbol{\pi} = \frac{m\mathbf{v}}{\sqrt{1 - |\mathbf{v}|^2/c^2}} \tag{8.12}$$

par

$$\mathbf{p} = \boldsymbol{\pi} + q\mathbf{A}. \tag{8.13}$$

L'hamiltonien prend la forme

$$\begin{aligned} H &= \mathbf{v}\frac{\partial L}{\partial \mathbf{v}} - L \\ &= \mathbf{v}(\boldsymbol{\pi} + q\mathbf{A}) + mc^2\sqrt{1-|\mathbf{v}|^2/c^2} + q\phi - q\mathbf{A}\mathbf{v} \\ &= \frac{mc^2}{\sqrt{1-|\mathbf{v}|^2/c^2}} + q\phi. \end{aligned} \tag{8.14}$$

De $\boldsymbol{\pi} = \mathbf{p} - q\mathbf{A}$ on déduit $\pi^\mu = p^\mu - qA^\mu$ dont le carré invariant donne $\pi_\mu\pi^\mu = mv_\mu mv^\mu = m^2c^2$ et de $p^\mu - qA^\mu = ((H-q\phi)/c, \mathbf{p}-q\mathbf{A})$ on déduit de même $(p_\mu - qA_\mu)(p^\mu - qA^\mu) = (H-q\phi)^2/c^2 - |\mathbf{p}-q\mathbf{A}|^2$, soit finalement l'expression invariante

$$(H-q\phi)^2 = m^2c^4 + (\mathbf{p}-q\mathbf{A})^2c^2.$$

L'hamiltonien devient

$$H = \sqrt{m^2c^4 + c^2(\mathbf{p}-q\mathbf{A})^2} + q\phi. \tag{8.15}$$

Aux faibles vitesses on peut simplifier

$$\mathbf{p} - q\mathbf{A} \simeq m\mathbf{v},$$

soit

$$\begin{aligned} H &\simeq mc^2\left(1 + \frac{\mathbf{v}^2}{2c^2}\right) + q\phi \\ &= mc^2 + \frac{1}{2}m\mathbf{v}^2 + q\phi \\ &= mc^2 + \frac{1}{2m}(\mathbf{p}-q\mathbf{A})^2 + q\phi \end{aligned}$$

ce qui constitue une expression non relativiste fréquemment utilisée.

En conservant la notation covariante, on peut exprimer la variation de l'action sous la forme

$$\begin{aligned} \delta S_q &= \int_a^b (-mc\delta(\mathrm{d}x_\mu\,\mathrm{d}x^\mu)^{1/2} - q\delta(A_\mu\,\mathrm{d}x^\mu)) \\ &= -m[v_\mu\delta x^\mu]_a^b + \int_a^b m\,\mathrm{d}v_\mu\delta x^\mu - q\int_a^b (A_\mu\,\mathrm{d}\delta x^\mu + \delta A_\mu\,\mathrm{d}x^\mu) \end{aligned}$$

Or on a aussi

$$\int_a^b A_\mu\,\mathrm{d}\delta x^\mu = [A_\mu\delta x^\mu]_a^b - \int_a^b \mathrm{d}A_\mu\delta x^\mu$$

de sorte qu'il reste finalement

$$\delta S_q = -m[v_\mu \delta x^\mu]_a^b - q[A_\mu \delta x^\mu]_a^b + \int_a^b (m\,\mathrm{d}v_\mu \delta x^\mu + q\,\mathrm{d}A_\mu \delta x^\mu - q\delta A_\mu\,\mathrm{d}x^\mu).$$

Les termes intégrés disparaissent car les variations δx^μ sont nulles aux extrémités de la ligne d'univers. En utilisant de plus

$$\delta A_\mu = \frac{\partial A_\mu}{\partial x^\nu}\delta x^\nu,$$
$$\mathrm{d}A_\mu = \frac{\partial A_\mu}{\partial x^\nu}\,\mathrm{d}x^\nu,$$

il vient

$$\delta S_q = \int_a^b (m\,\mathrm{d}v_\mu \delta x^\mu + q\frac{\partial A_\mu}{\partial x^\nu}\,\mathrm{d}x^\nu \delta x^\mu - q\frac{\partial A_\mu}{\partial x^\nu}\delta x^\nu\,\mathrm{d}x^\mu)$$
$$= \int_a^b \left(m\frac{\mathrm{d}v_\mu}{\mathrm{d}\tau} - q\left(\frac{\partial A_\nu}{\partial x^\mu}v^\nu - \frac{\partial A_\mu}{\partial x^\nu}v^\nu\right)\right)\mathrm{d}\tau\delta x^\mu,$$

d'où les équations du mouvement sous forme tensorielle

$$m\frac{\mathrm{d}v_\mu}{\mathrm{d}\tau} = q(\partial_\mu A_\nu - \partial_\nu A_\mu)v^\nu \tag{8.16}$$

ou encore en faisant apparaître le tenseur champ électromagnétique,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu},$$

$$ma_\mu = qF_{\mu\nu}v^\nu. \tag{8.17}$$

Equations d'Euler-Lagrange covariantes

Le calcul précédent conduit automatiquement à des équations analogues aux équations classiques d'Euler-Lagrange. Il est utile d'écrire directement ces équations sous forme covariante. Partant de l'action

$$S = \int \mathrm{L}_\mu\,\mathrm{d}x^\mu = \int L\,\mathrm{d}\lambda,$$

où $L = \mathrm{L}_\mu\,\mathrm{d}x^\mu/\mathrm{d}\lambda$, le paramètre scalaire λ pouvant être le temps propre τ, celui mesuré par un observateur t, ou encore l'intervalle s. On pose ici $\dot{x}^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$, même

si le paramètre λ n'est pas nécessairement le temps [27] et les équations d'Euler-Lagrange prennent la forme

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial L}{\partial \dot{x}^\mu}\right) - \frac{\partial L}{\partial x^\mu} = 0.$$

Dans le cas d'une charge q, de masse m, en présence d'un champ électromagnétique, l'action prend la forme

$$S = \int \left(-mv_\mu \,\mathrm{d}x^\mu - qA_\mu \,\mathrm{d}x^\mu\right),$$

de sorte que le lagrangien est donné par

$$L = (-mv_\mu - qA_\mu)\dot{x}^\mu.$$

On en déduit l'impulsion covariante [28] ,

$$p_\mu = -\frac{\partial L}{\partial \dot{x}^\mu} = mv_\mu + qA_\mu.$$

On peut noter que cette expression est conforme à ce que nous avons déjà obtenu préalablement, puisque

$$p_\mu = (E/c, -\mathbf{p}) = m(\gamma c, -\gamma \mathbf{v}) + q(\phi/c, -\mathbf{A}).$$

L'équation d'Euler-Lagrange ci-dessus peut s'écrire [29]

$$\dot{p}_\mu = -\frac{\partial L}{\partial x^\mu},$$

où $\dot{p}_\mu = m\dot{v}_\mu + q\dot{A}_\mu$ s'exprime encore plus explicitement en notant que $\mathrm{d}A_\mu = \frac{\partial A_\mu}{\partial x^\nu}\mathrm{d}x^\nu$, soit $\dot{A}_\mu = \frac{\partial A_\mu}{\partial x^\nu}\dot{x}^\nu$ et finalement

$$\dot{p}_\mu = m\dot{v}_\mu + q\partial_\nu A_\mu \dot{x}^\nu.$$

Au second membre il vient

$$\frac{\partial L}{\partial x^\mu} = -q\partial_\mu A_\nu \dot{x}^\nu,$$

---

[27] Nous introduisons ici une notation non conventionnelle pour l'action $\int \mathtt{L}_\mu \,\mathrm{d}x^\mu$ qui permet d'identifier facilement $L = \mathtt{L}_\mu \dot{x}^\mu$ mais n'a pas davantage d'intérêt.

[28] Le signe $-$ dans la définition de p_μ est nécessaire pour retrouver les expressions classiques dans la limite non relativiste. On peut en comprendre l'origine en notant qu'il apparaît forcément un signe négatif dans les composantes spatiales, soit de p_μ, soit, si l'on calcule plutôt $p^\mu = -\frac{\partial L}{\partial \dot{x}_\mu}$, dans celles de x_μ au dénominateur.

[29] Même remarque à propos du signe qui découle ici de la définition choisie pour p_μ.

en mettant à profit un changement d'indice $\mu \to \nu$ dans la contraction, effectuée sur un indice muet, du dernier terme de L. L'équation du mouvement devient immédiatement

$$m\dot{v}_\mu = -q\partial_\nu A_\mu \dot{x}^\nu + \partial_\mu A_\nu \dot{x}^\nu = -qF_{\nu\mu}\dot{x}^\nu = qF_{\mu\nu}\dot{x}^\nu.$$

On a mis à profit l'antisymétrie du tenseur champ électromagnétique dans le dernier terme.

Densité lagrangienne

De manière générale, l'action S, scalaire invariant, peut s'écrire comme une intégrale sur un 4-élément de volume,

$$S = \int \mathrm{d}^4x\, \mathcal{L}(x_\mu, \dot{x}_\mu)$$

où $\mathrm{d}^4x = \mathrm{d}^3r\,\mathrm{d}t$ est l'élément de volume de l'espace-temps et $\mathcal{L}(x_\mu, \dot{x}_\mu)$ est la densité lagrangienne. Le lagrangien est donné par l'intégrale spatiale de la densité lagrangienne,

$$L = \int \mathrm{d}^3r \mathcal{L}(x_\mu, \dot{x}_\mu).$$

L'intérêt de la densité $\mathcal{L}(x_\mu, \dot{x}_\mu)$ est qu'il s'agit d'un scalaire invariant. En effet, d^4x est un invariant car dans $\mathrm{d}^4x = \mathrm{d}^3r\,\mathrm{d}t = \mathrm{d}^2S_\perp\,\mathrm{d}l_\parallel\,\mathrm{d}t$, les sections transverses sont invariables, alors que le temps et les longueurs longitudinales ont des lois de transformation qui se compensent parfaitement.

Considérons maintenant une charge élémentaire $\mathrm{d}q = \rho\,\mathrm{d}^3r$ dans un champ électromagétique. L'action comprend un terme d'interaction charge-champ que nous connaissons,

$$\begin{aligned} S_{\text{int.}} &= -\int \mathrm{d}q A_\mu \,\mathrm{d}x^\mu \\ &= -\int \mathrm{d}q(\phi - \mathbf{A}\mathbf{v})\,\mathrm{d}t \\ &= -\int \rho(\phi - \mathbf{A}\mathbf{v})\,\mathrm{d}^3r\,\mathrm{d}t \\ &= -\int j_\mu A^\mu\,\mathrm{d}^4x \\ &= \int \mathcal{L}_{\text{int.}}\,\mathrm{d}^4x. \end{aligned}$$

Il y a également un terme décrivant le champ libre. Par analogie avec la dépendance en x_μ et $\dot{x}_\mu$ de la densité lagrangienne de la particule libre, ce terme doit faire intervenir a priori une densité lagrangienne de la forme $\mathcal{L}_{\text{chp}}(A_\mu, \partial_\nu A_\mu)$. Comme les équations de Maxwell sont linéaires, il faut au plus une dépendance quadratique

du lagrangien en fonction des champs. On forme la quantité $F_{\mu\nu}F^{\mu\nu}$ et pour être compatible avec les équations de Maxwell, on anticipe le préfacteur, soit

$$\mathcal{L}_{\text{chp}} = -\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu}.$$

Au total on écrit l'action comme

$$S = -\int \left(\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu} + j^\mu A_\mu\right) \mathrm{d}^4x. \tag{8.18}$$

Cela permet bien de retrouver des équations de Maxwell car

$$\delta(F_{\mu\nu}F^{\mu\nu}) = 2F^{\mu\nu}\delta F_{\mu\nu} = 4F^{\mu\nu}\partial_\mu \delta A_\nu,$$

soit, après intégration par parties et élimination des termes de bords,

$$\delta S = -\int \mathrm{d}^4x \left[j^\mu \delta A_\mu - \frac{1}{4\mu_0}4\partial_\mu F^{\mu\nu}\delta A_\nu\right]$$

ou, en changeant les indices du premier terme,

$$\partial_\mu F^{\mu\nu} = \mu_0 j^\nu.$$

L'autre partie des équations de Maxwell (sans sources) provient de la définition même du tenseur de Faraday $F^{\mu\nu}$.

Il est utile d'écrire la densité lagrangienne sous forme tridimensionelle. En l'absence de charges, la densité lagrangienne du champ électromagnétique s'écrit

$$\frac{1}{2}\varepsilon_0\mathbf{E}^2 - \frac{1}{2}\frac{\mathbf{B}^2}{\mu_0}.$$

Le terme densité se justifie, puisque pour obtenir une grandeur extensive on intègre sur l'espace [30] . En présence de charges (immobiles ou en mouvement), la densité lagrangienne s'écrit (Schwarzschild)

$$\mathcal{L} = \frac{1}{2}\varepsilon_0\mathbf{E}^2(\mathbf{r}) - \frac{1}{2\mu_0}\mathbf{B}^2(\mathbf{r}) + \rho(\mathbf{r})(\mathbf{A}(\mathbf{r})\cdot\mathbf{v} - \phi(\mathbf{r})) \tag{8.19}$$

où $\mathbf{v}$ est la vitesse des porteurs de charge, $\rho(\mathbf{r})$ la densité de charges et $\mathbf{j}(\mathbf{r}) = \rho(\mathbf{r})\mathbf{v}$ la densité de courant. En fonction des potentiels, on a

$$\mathcal{L} = \frac{1}{2}\varepsilon_0\left(-\vec{\nabla}\phi - \frac{\partial\mathbf{A}}{\partial t}\right)^2 - \frac{1}{2\mu_0}(\mathbf{rot}\,\mathbf{A})^2 + \rho(\mathbf{A}\cdot\mathbf{v} - \phi). \tag{8.20}$$

[30] $\frac{1}{2}\varepsilon_0\mathbf{E}^2$ par exemple est la densité d'énergie électrique associée à l'existence d'un champ électrique.

Si on note de manière générique ψ_j pour ϕ, A_x, A_y et A_z, les quatre champs scalaires, $\mathcal{L}$ est une fonction de ces champs et de leurs dérivées spatiales et temporelles

$$\mathcal{L}(\psi_j, \dot{\psi}_j, \partial_x\psi_j, \partial_y\psi_j, \partial_z\psi_j).$$

Les équations du mouvement s'écrivent

$$\frac{\partial\mathcal{L}}{\partial\psi_j} - \sum_{\alpha=x,y,z} \frac{\partial}{\partial\alpha}\left[\frac{\partial\mathcal{L}}{\partial(\frac{\partial\psi_j}{\partial\alpha})}\right] - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\psi}_j}\right) = 0. \tag{8.21}$$

Appliquées aux composantes du quadripotentiel, les équations d'Euler – Lagrange permettent de retrouver les équations de propagation pour le potentiel scalaire et les composantes du potentiel vecteur. En coordonnées cartésiennes on a

$$\begin{aligned}\mathcal{L} =& \frac{1}{2}\varepsilon_0\left[\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2 + \left(\frac{\partial\phi}{\partial z}\right)^2 + 2\frac{\partial\phi}{\partial x}\frac{\partial A_x}{\partial t} + 2\frac{\partial\phi}{\partial y}\frac{\partial A_y}{\partial t} + 2\frac{\partial\phi}{\partial z}\frac{\partial A_z}{\partial t}\right.\\ &\left. + \left(\frac{\partial A_x}{\partial t}\right)^2 + \left(\frac{\partial A_y}{\partial t}\right)^2 + \left(\frac{\partial A_z}{\partial t}\right)^2\right] - \frac{1}{2\mu_0}\left[\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)^2 + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)^2\right.\\ &\left. + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)^2\right] + A_x j_x + A_y j_y + A_z j_z - \rho\phi.\end{aligned}$$

- $\psi_j = \phi$, $\frac{\partial\mathcal{L}}{\partial\phi} = -\rho$, $\quad\frac{\partial\mathcal{L}}{\partial\dot{\phi}} = 0$, $\quad\frac{\partial\mathcal{L}}{\partial(\partial_x\phi)} = \varepsilon_0\partial_x\phi + \varepsilon_0\frac{\partial A_x}{\partial t}$ donc

$$\begin{cases}\dfrac{\partial}{\partial x}\left(\dfrac{\partial\mathcal{L}}{\partial(\partial_x\phi)}\right) = \varepsilon_0\left(\dfrac{\partial^2\phi}{\partial x^2} + \dfrac{\partial}{\partial x}\left(\dfrac{\partial A_x}{\partial t}\right)\right)\\ \dfrac{\partial}{\partial y}\left(\dfrac{\partial\mathcal{L}}{\partial(\partial_y\phi)}\right) = \varepsilon_0\left(\dfrac{\partial^2\phi}{\partial y^2} + \dfrac{\partial}{\partial y}\left(\dfrac{\partial A_y}{\partial t}\right)\right)\\ \dfrac{\partial}{\partial z}\left(\dfrac{\partial\mathcal{L}}{\partial(\partial_z\phi)}\right) = \varepsilon_0\left(\dfrac{\partial^2\phi}{\partial z^2} + \dfrac{\partial}{\partial z}\left(\dfrac{\partial A_z}{\partial t}\right)\right)\end{cases}$$

d'où

$$-\rho - \varepsilon_0\left(\frac{\partial^2\phi}{\partial x^2} + \frac{\partial}{\partial t}\left(\frac{\partial A_x}{\partial x}\right) + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial}{\partial t}\left(\frac{\partial A_y}{\partial y}\right) + \frac{\partial^2\phi}{\partial z^2} + \frac{\partial}{\partial t}\left(\frac{\partial A_z}{\partial z}\right)\right) = 0.$$

Dans la parenthèse, on reconnaît $\Delta\phi$ et $\frac{\partial}{\partial t}(\operatorname{div}\mathbf{A})$ qui vaut en jauge de Lorenz $-\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}$, d'où

$$\Delta\phi - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = -\frac{\rho}{\varepsilon_0}.$$

- $\psi_j = A_x$, $\frac{\partial \mathcal{L}}{\partial A_x} = j_x$, $\quad -\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{A}_x} = -\varepsilon_0\left(\frac{\partial}{\partial t}\frac{\partial \phi}{\partial x} + \frac{\partial^2 A_x}{\partial t^2}\right)$

$$\begin{cases} -\dfrac{\partial}{\partial x}\left(\dfrac{\partial \mathcal{L}}{\partial(\partial_x A_x)}\right) = 0 \\ -\dfrac{\partial}{\partial y}\left(\dfrac{\partial \mathcal{L}}{\partial(\partial_y A_y)}\right) = \dfrac{1}{\mu_0}\left(\dfrac{\partial^2 A_x}{\partial y^2} - \dfrac{\partial}{\partial y}\left(\dfrac{\partial A_x}{\partial x}\right)\right) \\ -\dfrac{\partial}{\partial z}\left(\dfrac{\partial \mathcal{L}}{\partial(\partial_z A_z)}\right) = \dfrac{1}{\mu_0}\left(\dfrac{\partial^2 A_x}{\partial z^2} - \dfrac{\partial}{\partial z}\left(\dfrac{\partial A_x}{\partial x}\right)\right) \end{cases}$$

d'où

$$j_x - \varepsilon_0\frac{\partial}{\partial x}\left(\frac{\partial \phi}{\partial t}\right) - \varepsilon_0\frac{\partial^2 A_x}{\partial t^2} + \frac{1}{\mu_0}\left(\frac{\partial^2 A_x}{\partial y^2} + \frac{\partial^2 A_x}{\partial z^2} - \frac{\partial^2 A_x}{\partial x \partial y} - \frac{\partial^2 A_x}{\partial z \partial x}\right) = 0.$$

Avec la jauge de Lorenz, $\frac{\partial \phi}{\partial t}$ donne $-c^2\left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\right)$ qui compense exactement les dérivées croisées et permet de compléter le laplacien. On obtient finalement

$$\Delta A_x - \frac{1}{c^2}\frac{\partial^2 A_x}{\partial t^2} = -\mu_0 j_x.$$

Postuler la minimisation de l'action définie par la densité lagrangienne de Schwarzschild conduit donc aux mêmes équations de propagation que les équations de Maxwell. Cette formulation de l'électromagnétisme est rarement employée en raison de son caractère assez axiomatique. Elle présente toutefois un grand intérêt physique, puisque de nombreux domaines de la physique peuvent être décrits par des principes variationnels qui apparaissent ainsi comme un contexte unificateur.

Généralisation

On a vu que la forme de l'action avait un contenu facile à interpréter,

$$S = \int \mathrm{L}_\mu \, \mathrm{d}x^\mu,$$

où $\mathrm{L}_\mu = -mv_\mu - qA_\mu$, le premier terme donnant la contribution de l'énergie cinétique et le second le couplage (q) à un potentiel externe (A_μ). On omet ici la contribution du champ seul, $-\frac{1}{4\mu_0}\int F_{\mu\nu}F^{\mu\nu}\,\mathrm{d}^4x$. On peut a priori généraliser à des interactions arbitraires, décrites par exemple par des potentiels de rang 2 couplés par une constante g, soit $\mathrm{L}_\mu = -mv_\mu - qA_\mu - gB_{\mu\nu}v^\nu, \ldots$ Les équations d'Euler - Lagrange donnent ensuite une dynamique que l'on peut comparer à l'expérience si on cherche à lui donner un sens physique. On peut ainsi former à volonté des lagrangiens très généraux satisfaisant au principe de covariance relativiste et étudier ensuite la dynamique qu'ils génèrent. Si celle-ci peut être identifiée à des phénomènes se produisant dans la nature, on peut dire que l'on a produit une théorie physique.