

Vous avez dit boson de Higgs ?

Christophe Roland

1 Introduction

Ce document s'adresse aux personnes qui veulent comprendre ce qu'est le boson de Higgs sans entrer dans les détails les plus complexes mais sans non plus se contenter d'une simple vulgarisation. Tout ce qu'il faut au préalable pour comprendre ce qui va suivre, c'est un niveau élémentaire en mathématique (dérivées, intégrales, nombres complexes, ...) et une connaissance basique de la relativité restreinte est préférable.

2 Préliminaire

On va commencer par quelques rappels et/ou compléments avant d'entrer dans le vif du sujet, ce sera aussi l'occasion de fixer quelques conventions.

2.1 Relativité restreinte

Nous vivons dans un espace-temps à quatre dimensions (trois dimensions d'espace et une de temps). On peut bien sûr y définir des vecteurs, et dans ce cas pour distinguer les vecteurs (à quatre dimensions) des vecteurs « ordinaires » de l'espace à trois dimensions, on les nomme *quadrivecteurs*. On peut ainsi (par exemple) définir des *quadrivecteurs points-événements* : un tel quadrivecteur repère, par rapport à un point d'origine, un point dans l'espace de coordonnées (x, y, z) à un moment précis t . Par convention, la première composante du quadrivecteur est la composante temporelle, les trois autres composantes étant les composantes spatiales. Mais pour avoir quelque chose de cohérent, on multiplie la composante temporelle par la vitesse de la lumière, pour avoir toutes les composantes en unité de longueur.

Les composantes (ct, x, y, z) vont être notées (x^0, x^1, x^2, x^3) , et le quadrivecteur lui même va être noté \bar{x} , ou par abus de notation x^μ (qui est en fait la μ -ième composante de \bar{x}). Ainsi, on note :

$$x^\mu = (ct, x, y, z) = (x^0, x^1, x^2, x^3)$$

Bien sûr, les composantes du quadrivecteur sont obtenues dans un certain système de coordonnées. Dans un autre système de coordonnées, on aura des composantes $(x^\mu)'$ différentes, mais il s'agit bien sûr du *même* quadrivecteur. Ainsi, si on note $\bar{e}_\mu = (\bar{e}_0, \bar{e}_1, \bar{e}_2, \bar{e}_3)$ les vecteurs de bases, on a :

$$\bar{x} = \sum_{\mu=0}^3 x^\mu \bar{e}_\mu = \sum_{\mu=0}^3 (x^\mu)' (\bar{e}_\mu)'$$

On va introduire ici la *convention d'Einstein* : pour simplifier les notations, on n'écrit pas le signe somme (Σ), et on effectue la somme sur les indices répétés en haut et en bas. Explicitement :

$$\bar{x} = x^\mu \bar{e}_\mu = x^0 \bar{e}_0 + x^1 \bar{e}_1 + x^2 \bar{e}_2 + x^3 \bar{e}_3$$

Cette convention sera toujours utilisée par la suite.

Remarquons que l'on peut toujours changer l'indice de sommation, ainsi par exemple on a : $x^\mu \bar{e}_\mu = x^\nu \bar{e}_\nu$.

Quel est le lien entre x^μ et $(x^\mu)'$? On va se limiter ici aux transformations linéaires des coordonnées, qui peuvent donc s'écrire sous la forme :

$$(x^\mu)' = \Lambda^\mu{}_\nu x^\nu \quad (1)$$

Où Λ est une matrice 4×4 . L'indice supérieur (μ) est l'indice de ligne et l'indice inférieur (ν) est l'indice de colonne. Pour que ce soit plus clair, prenons un exemple. Supposons que la matrice Λ s'écrive :

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

avec $\beta = \frac{v}{c}$ et $\gamma = \frac{1}{\sqrt{1-\beta^2}}$. La matrice n'est bien sûr pas choisie au hasard mais j'y reviendrai. Utilisons l'équation (1) pour calculer $(x^0)'$:

$$\begin{aligned} (x^0)' &= \Lambda^0{}_\nu x^\nu \\ &= \Lambda^0{}_0 x^0 + \Lambda^0{}_1 x^1 + \Lambda^0{}_2 x^2 + \Lambda^0{}_3 x^3 \\ &= \gamma \cdot x^0 - (\beta\gamma) \cdot x^1 + 0 \cdot x^2 + 0 \cdot x^3 \\ &= \gamma (x^0 - \beta x^1) \end{aligned}$$

C'est-à-dire :

$$ct' = \gamma (ct - \beta x)$$

de même on a : $x' = \gamma (x - \beta ct)$ et les autres composantes sont inchangées : $y = y'$ et $z = z'$. L'ensemble peut se mettre sous forme matricielle :

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

On ne peut pas choisir n'importe quelle matrice Λ . En effet, il faut respecter le postulat d'invariance de la vitesse de la lumière. Pour cela on demande que si on a deux points-événements, la quantité :

$$\Delta s^2 = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$$

soit identique, quelque soit le référentiel dans lequel on la calcule. Supposons que les points-événements sont le point de départ et le point d'arrivée d'une particule, alors sa vitesse est :

$$v = \frac{\Delta r}{\Delta t} = \frac{\sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}}{\Delta t} = \sqrt{\frac{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}{(\Delta t)^2}} = \sqrt{c^2 - \frac{\Delta s^2}{(\Delta t)^2}}$$

Si $\Delta s^2 = 0$, alors $v = c$ dans n'importe quel référentiel, puisque l'on suppose Δs^2 invariant. On a donc bien l'invariance de la vitesse de la lumière.

On va maintenant définir un nouvel objet (en fait un tenseur) que l'on appelle la *métrique*, et que l'on va noter $g_{\mu\nu}$. Les composantes de la métrique sont par définition les produits scalaires des vecteurs de bases :

$$g_{\mu\nu} = \bar{e}_\mu \cdot \bar{e}_\nu \quad (2)$$

Notez que comme le produit scalaire est commutatif (i.e. $\bar{e}_\mu \cdot \bar{e}_\nu = \bar{e}_\nu \cdot \bar{e}_\mu$), on dit que la métrique est un tenseur *symétrique* c'est à dire que $g_{\mu\nu} = g_{\nu\mu}$.

On peut donner un exemple simple de métrique : dans l'espace traditionnel à trois dimensions, dans une base orthonormée, on obtient :

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

En relativité restreinte, on va choisir le produit scalaire de manière à ce que la « norme » (en fait pseudo-norme) d'un quadrivecteur correspond à l'invariant Δs^2 , on a donc une base dans lequel la métrique (dite métrique de Minkowski, l'espace dans lequel on travaille étant appelé espace de Minkowski) s'écrit :

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Par convention, le premier indice est l'indice de ligne et le deuxième l'indice de colonne (bien que dans le cas présent cela n'a aucune importance puisque la matrice est diagonale).

Attention que cette expression est correcte pour les systèmes de coordonnées cartésiens, mais ce n'est pas une expression générale, si par exemple, on travaille en coordonnées sphériques on aura une expression un peu plus compliquée.

Si on a deux quadrivecteurs \bar{x} et \bar{y} , on peut calculer leur produit scalaire (en utilisant (2)) :

$$\bar{x} \cdot \bar{y} = (x^\mu \bar{e}_\mu) \cdot (y^\nu \bar{e}_\nu) = x^\mu y^\nu (\bar{e}_\mu \cdot \bar{e}_\nu) = g_{\mu\nu} x^\mu y^\nu$$

En particulier, la pseudo-norme d'un quadrivecteur est :

$$\|\bar{x}\|^2 = \bar{x} \cdot \bar{x} = g_{\mu\nu} x^\mu x^\nu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$$

On retrouve bien quelque chose qui correspond à l'invariant Δs^2 . Tout les ingrédients sont enfin réunis pour savoir quelles matrices Λ laissent Δs^2 invariant.

Si on a un intervalle Δx^μ , dans un autre référentiel on aura :

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

On veut Δs^2 invariant, donc (rappelez-vous que l'on peut toujours changer les indices de sommation) :

$$\begin{aligned} \Delta s^2 &= g_{\mu\nu} \Delta x^\mu \Delta x^\nu = g_{\rho\sigma} (\Delta x^\rho)' (\Delta x^\sigma)' \\ &= g_{\rho\sigma} (\Lambda^\rho{}_\mu \Delta x^\mu) (\Lambda^\sigma{}_\nu \Delta x^\nu) \\ &= g_{\rho\sigma} \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu \Delta x^\mu \Delta x^\nu \end{aligned}$$

On en déduit que :

$$g_{\mu\nu} = g_{\rho\sigma} \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu \quad (3)$$

Pour ceux qui n'aiment pas les indices, on peut écrire ceci sous forme matricielle. Si on note Λ^T la matrice transposée de Λ , on peut montrer que (3) se ramène à :

$$g = \Lambda^T g \Lambda$$

L'ensemble des matrices qui vérifient la relation (3) forment un groupe : le *groupe de Lorentz*, il est noté $O(3, 1)$. Les transformations correspondantes : $(x^\mu)' = \Lambda^\mu{}_\nu x^\nu$ laissent Δs^2 invariant.

On s'est limité aux transformations linéaires (i.e. de la forme $(x^\mu)' = \Lambda^\mu{}_\nu x^\nu$), ce sont des transformations qui envoient l'origine du référentiel \mathcal{R} sur l'origine du référentiel \mathcal{R}' . On a donc dans ce groupe des transformations qui correspondent à :

- des rotations dans l'espace,
- des boosts, c'est-à-dire des transformations qui relient deux référentiels en mouvement rectilignes uniformes (l'exemple donné plus haut de matrice Λ correspond à un boost dans la direction x),
- toutes les combinaisons de rotations de boosts.

Si on considère toutes les isométries possibles de l'espace de Minkowski, et non pas seulement les transformations linéaires dans les coordonnées, on a un groupe plus large que le groupe de Lorentz, que l'on appelle *groupe de Poincaré*.

Il est utile de définir la métrique «avec les indices en hauts» par¹ :

$$g_{\mu\nu} g^{\nu\sigma} = g^{\sigma\nu} g_{\nu\mu} = \delta_\mu^\sigma \quad (4)$$

où δ_μ^σ est le *symbole de Kronecker* défini par (en utilisant ici les indices μ et ν pour suivre la convention) :

$$\delta_\mu^\nu = \begin{cases} 0 & \text{si } \nu \neq \mu \\ 1 & \text{si } \nu = \mu \end{cases}$$

Si on veut écrire (4) sous forme matricielle, on voit que la matrice g avec les indices en haut est l'inverse de la matrice g avec les indices en bas. En effet, le symbole de Kronecker correspond à la matrice unité $\mathbb{1}$. La définition (4) est une définition générale, mais dans les systèmes de coordonnées que l'on utilise ici, on a égalité : $g^{\mu\nu} = g_{\mu\nu}$, ce que l'on supposera donc dans la suite.

Nous avons noté les composantes des quadrivecteurs avec les indices en haut. Ces composantes sont appelées *composantes contravariantes*. Elles correspondent aux composantes vectorielles « habituelles ». On peut définir aussi ce que l'on appelle les *composantes covariantes*, que l'on note avec un indice en bas, et qui se définissent comme ceci :

$$x_\mu = g_{\mu\nu} x^\nu$$

dans le système de coordonnées (cartésien) que l'on utilise on a simplement :

$$x_\mu = (ct, -x, -y, -z)$$

La métrique permet de passer des coordonnées contravariantes aux coordonnées covariantes, la métrique inverse fait donc le travail de passer des coordonnées covariantes aux coordonnées contravariantes :

$$x^\mu = g^{\mu\nu} x_\nu$$

On dit que la métrique fait « monter ou descendre les indices ». On peut maintenant écrire le produit de deux quadrivecteurs de plusieurs façons différentes en faisant monter ou descendre les indices :

$$\bar{x} \cdot \bar{y} = g_{\mu\nu} x^\mu y^\nu = x^\mu y_\mu = x_\mu y^\mu$$

En effet :

$$x^\mu y_\mu = x^\mu (g_{\mu\nu} y^\nu) = g_{\mu\nu} x^\mu y^\nu$$

1. Cette définition peut sembler arbitraire, mais justifier rigoureusement ce qu'on va faire maintenant serait inutilement long et compliqué.

et (rappelons que l'on peut renommer l'indice de sommation et que $g_{\mu\nu} = g_{\nu\mu}$) :

$$x_\mu y^\mu = x_\nu y^\nu = (g_{\nu\mu} x^\mu) y^\nu = g_{\mu\nu} x^\mu y^\nu$$

Supposons qu'on ait une fonction qui dépend des coordonnées : $f(x^0, x^1, x^2, x^3)$. On peut prendre la dérivée partielle de cette fonction par rapport à l'une des composantes x^μ , et on note :

$$\partial_\mu f = \frac{\partial f}{\partial x^\mu}$$

Ainsi, on a $\partial_0 f = \frac{\partial f}{\partial(ct)}$, $\partial_1 f = \frac{\partial f}{\partial(x)}$, et ainsi de suite.

On définit de même :

$$\partial^\mu f = \frac{\partial f}{\partial x_\mu}$$

Avec, comme précédemment, $\partial^\mu = g^{\mu\nu} \partial_\nu$.

Enfin, on définit le D'Alembertien \square :

$$\square = \partial^\mu \partial_\mu = g^{\mu\nu} \partial_\nu \partial_\mu = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right)$$

2.2 Mécanique quantique

On va maintenant introduire quelques notions élémentaires de mécanique quantique. Une notion dont vous avez certainement déjà entendu parler est la *dualité onde-corpuscule*. C'est le fait que les objets physique présentent à la fois des caractéristiques d'ondes et de particules.

Des quantités qui apparaissent naturellement lorsque l'on parle de particules sont l'énergie E et la quantité de mouvement $p = mv$, tandis que pour les ondes apparaissent naturellement la fréquence ν et la longueur d'onde λ . On peut en fait établir un lien entre les quantités « particules » et les quantités « ondes ». C'est la relation de Planck-Einstein qui donne l'énergie d'une particule :

$$E = h\nu \tag{5}$$

et la relation de Louis de Broglie :

$$p = \frac{h}{\lambda} \tag{6}$$

h est une constante de la nature appelée *constante de Planck*.

Faisons maintenant un petit rappel sur les ondes. Quand on considère une onde plane, comme par exemple en électromagnétisme, on a quelque chose comme :

$$f(\vec{r}, t) = A \cos(\vec{k} \cdot \vec{r} - \omega t)$$

avec $\omega = 2\pi\nu$ (la *pulsation*), A l'*amplitude* de l'onde, et où \vec{k} (le *vecteur d'onde*) est un vecteur qui pointe dans la direction de propagation et dont la norme est $k = \frac{2\pi}{\lambda}$.

Pour simplifier les calculs, il est courant de passer en nombres complexes et de remplacer $f(\vec{r}, t)$ par :

$$\tilde{f}(\vec{r}, t) = A e^{i(\vec{k} \cdot \vec{r} - \omega t)}$$

où i est l'«unité imaginaire» tel que $i^2 = -1$. En effet, on sait en mathématique que :

$$e^{ix} = \cos(x) + i \sin(x)$$

dans notre cas, on a alors :

$$\tilde{f}(\vec{r}, t) = A \left[\cos(\vec{k} \cdot \vec{r} - \omega t) + i \sin(\vec{k} \cdot \vec{r} - \omega t) \right]$$

On peut donc travailler avec \tilde{f} tout le long du calcul, puis ne prendre à la fin que la partie réelle (ou imaginaire). Un des avantages est que l'on peut factoriser la partie temporelle :

$$\tilde{f}(\vec{r}, t) = Ae^{-i\omega t} e^{i\vec{k}\cdot\vec{r}}$$

On peut faire une observation intéressante. Si on regarde uniquement ce qui se passe en un point précis \vec{r}_0 , on obtient ceci (j'enlève maintenant le tilde sur le f pour la lisibilité) :

$$f(t) = A'e^{-i\omega t}$$

où $A' = Ae^{i\vec{k}\cdot\vec{r}_0}$. Or, il est facile de montrer que c'est une solution de l'équation différentielle :

$$\frac{d^2 f(t)}{dt^2} + \omega^2 f(t) = 0 \quad (7)$$

En effet :

$$\frac{d^2 f(t)}{dt^2} = \frac{d}{dt} \frac{df(t)}{dt} = \frac{d}{dt} (-i\omega A'e^{-i\omega t}) = -\omega^2 A'e^{-i\omega t} = -\omega^2 f(t)$$

Un système qui obéit à l'équation différentielle (7) est appelé un *oscillateur harmonique*. Un exemple physique simple où l'on rencontre cette équation est le mouvement d'une masse sur un ressort. La force produite par un ressort est (approximativement) proportionnelle à son élongation (ou contraction) par rapport à sa taille au repos. On applique donc la deuxième loi de Newton :

$$m \frac{d^2 x(t)}{dt^2} = -kx(t)$$

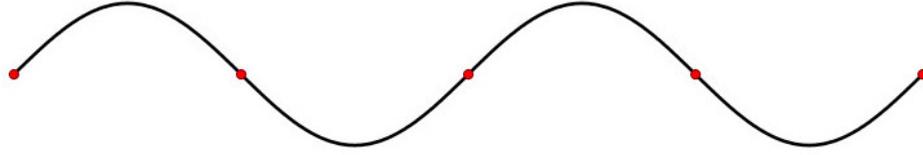
où k est une constante qui dépend du ressort (qu'on appelle la raideur du ressort) et où le signe moins dans le membre de droite vient du fait que la force du ressort s'oppose au sens dans lequel on l'a étiré ou contracté au départ. Si on compare avec (7), on voit que l'on a une équation équivalente si l'on pose $\omega = \frac{k}{m}$. L'oscillateur harmonique se retrouve dans d'autres systèmes physiques comme les oscillations libres d'un circuit LC².

Ce qui est intéressant, et on y reviendra plus tard, c'est qu'alors que l'amplitude de l'oscillation est arbitraire, la pulsation $\omega = 2\pi\nu$ et donc la fréquence est fixée dans l'équation (7) et dépend en principe d'autres paramètres physiques. Ainsi, dans le cas d'une masse qui oscille au bout d'un ressort, on doit avoir $\omega = \frac{k}{m}$, et il se fait aussi par exemple que dans un circuit LC, on doit avoir $\omega = \frac{1}{\sqrt{LC}}$. Cette fréquence particulière est appelée *fréquence de résonance*. Si on soumet le système à une force extérieure qui varie en intensité et en direction suivant cette fréquence alors l'amplitude ne va cesser d'augmenter, mais ce ne sera pas le cas si la fréquence de cette force extérieure est différente. Tout ceux qui ont déjà poussé quelqu'un sur une balançoire connaît ce phénomène : il faut pousser à intervalles réguliers à la bonne fréquence, être donc synchronisé avec le mouvement sinon l'amplitude ne va pas augmenter.

Le traitement de l'oscillateur harmonique que nous avons fait est classique, c'est à dire que l'on n'a pas tenu compte des effets quantiques. Qu'est ce qui se passe quand on en tient compte ?

La chose importante est que l'énergie n'est pas arbitraire mais qu'il existe des niveaux d'énergie. Ceci peut se comprendre par la dualité onde-corpuscule et par l'analogie suivante : une onde stationnaire, sur une corde, doit obligatoirement avoir une longueur totale entre les deux extrémités de la corde qui soit un multiple entier de la demi-longueur d'onde. On peut le voir facilement sur le dessin suivant :

2. Si vous ne savez pas que ce c'est, ne vous faites pas de soucis, cet exemple n'est pas essentiel pour comprendre la suite.



Comme seules certaines longueurs sont autorisées, ainsi en est-il de l'énergie. Cet argument fonctionne aussi dans le cas de l'oscillateur harmonique quantique, bien que ce soit alors bien plus compliqué, mais l'idée principale est la même. En fait, on peut montrer que dans ce cas, les niveaux d'énergies sont donnés par :

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right)$$

où n est un entier positif ou nul. Remarquez que la différence d'énergie entre le niveau n et le niveau $n + 1$ est :

$$\Delta E = E_{n+1} - E_n = \hbar\omega \left(n + 1 + \frac{1}{2} \right) - \hbar\omega \left(n + \frac{1}{2} \right) = \hbar\omega = \frac{h}{2\pi} 2\pi\nu = h\nu$$

à comparer avec la relation de Planck-Einstein $E = h\nu$. Chaque niveau est donc séparé du suivant par un «quantum d'énergie» $h\nu$ ce qui n'est pas trop surprenant. Ce qui l'est bien plus est que l'énergie minimal n'est pas zéro mais :

$$E_0 = \frac{1}{2} \hbar\omega$$

Cette énergie est appelé *énergie du point zéro*. En mécanique classique, on pouvait avoir $E = 0$ quand la masse au bout de ressort ne bougeait pas et restait au point d'équilibre. En mécanique quantique, il faut tenir compte du fameux principe d'incertitude d'Heisenberg :

$$\Delta x \cdot \Delta p \geq \frac{\hbar}{2}$$

où Δx et Δp sont les incertitudes sur la position et sur la quantité de mouvement de la particule. Plus on localise une particule avec précision, plus sa quantité de mouvement est incertaine et vice-versa. Ainsi, il est impossible d'avoir une particule localisée qui ne bouge pas du tout, il reste donc toujours une énergie résiduelle quoi que l'on fasse. Une image (un peu simplifiée et naïve) possible pour se représenter cela est d'imaginer que la particule bouge légèrement et de façon aléatoire autour de la position d'équilibre.

On a esquissé un lien entre les ondes électromagnétiques et l'oscillateur harmonique. En utilisant l'oscillateur harmonique *quantique*, on peut alors entrevoir intuitivement l'existence des photons! On a trouvé en effet des niveaux d'énergies qui coïncident avec l'énergie des photons. Mais le lien est encore un peu flou, aussi on va le préciser mais dans un cas plus simple : au lieu de s'occuper du champs électromagnétique on va se concentrer sur le cas d'un champ *scalaire*, c'est à dire qu'à tout point de l'espace on associe un nombre.

3 Théorie quantique des champs

On entre maintenant dans la théorie quantique des champs, où l'on met l'accent sur les champs et non les particules, et on interprète les particules comme des «vibrations» de ces champs.

Pour décrire un champs, il faut se donner une équation auquel il obéit, par exemple pour le champs électromagnétique ce sont les équations de Maxwell. On va essayer de deviner une équation pour un champs scalaire. On va noter la valeur du champs au point-événement \bar{x} de la façon suivante :

$$\varphi(\bar{x})$$

Il faut trouver une équation qui soit à la fois invariante par transformation de Lorentz, et qui admette des ondes planes comme solutions. La solution la plus simple est (rappelons que $\square = \partial^\mu \partial_\mu$) :

$$\square \varphi(\bar{x}) = 0$$

Cette équation suffit à remplir les conditions que l'on s'est fixé, mais on peut aussi généraliser en ajoutant un terme proportionnel à $\varphi(\bar{x})$:

$$(\square + \mu^2)\varphi(\bar{x}) = 0 \quad (8)$$

où μ est un certain paramètre réel ou imaginaire, de sorte que son carré soit réel (le fait de mettre au carré sera justifié plus loin). Cette équation est appelée *équation de Klein-Gordon*.

On va maintenant montrer que l'on a bien des solutions en ondes planes. Pour cela, on va introduire un nouveau quadrivecteur : le *quadrivecteur énergie-impulsion* :

$$p^\mu = \left(\frac{E}{c}, p^1, p^2, p^3 \right)$$

La première composante est l'énergie (en fait l'énergie divisée par c pour que les unités concordent), les composantes p^1, p^2, p^3 étant les composantes du vecteur \vec{p} à trois dimensions.

On peut effectuer le produit scalaire de \vec{p} avec un quadrivecteur point-événement \bar{x} :

$$\vec{p} \cdot \bar{x} = g_{\mu\nu} p^\mu x^\nu = Et - p^1 x^1 - p^2 x^2 - p^3 x^3 = Et - \vec{p} \cdot \vec{r}$$

On utilise alors les relations (5) et (6), on remplace donc E par $h\nu$ et \vec{p} par $\vec{p} = \frac{h}{\lambda} \vec{n}$ avec \vec{n} un vecteur normé orienté dans le sens du mouvement. On peut donc remplacer :

$$\vec{p} \cdot \bar{x} = h\nu t - \frac{h}{\lambda} \vec{n} \cdot \vec{r} = \hbar (\omega t - \vec{k} \cdot \vec{r})$$

en utilisant les définitions vues précédemment de ω et \vec{k} . Une onde plane $Ae^{i(\vec{k} \cdot \vec{r} - \omega t)}$ peut donc s'écrire en fonction des quadrivecteurs \vec{p} et \bar{x} :

$$Ae^{i(\vec{k} \cdot \vec{r} - \omega t)} = Ae^{-\frac{i}{\hbar} \vec{p} \cdot \bar{x}}$$

On cherche donc une solution de (8) de la forme : $\varphi(\bar{x}) = Ae^{-\frac{i}{\hbar} p_\mu x^\mu}$. On a :

$$\partial_\mu \varphi(\bar{x}) = \frac{\partial}{\partial x^\mu} \left(Ae^{-\frac{i}{\hbar} p_\mu x^\mu} \right) = -i \frac{p_\mu}{\hbar} Ae^{-\frac{i}{\hbar} p_\mu x^\mu} = -i \frac{p_\mu}{\hbar} \varphi(\bar{x})$$

et par un calcul tout à fait similaire (en sachant que $p_\mu x^\mu = p^\mu x_\mu$) :

$$\partial^\mu \varphi(\bar{x}) = -i \frac{p^\mu}{\hbar} \varphi(\bar{x})$$

On peut donc calculer le d'Alembertien de $\varphi(\bar{x})$:

$$\square \varphi(\bar{x}) = \partial^\mu (\partial_\mu \varphi(\bar{x})) = -i \frac{p_\mu}{\hbar} \partial^\mu \varphi(\bar{x}) = -\frac{p_\mu p^\mu}{\hbar^2} \varphi(\bar{x})$$

L'équation de Klein-Gordon nous dit que $\square\varphi(\bar{x}) = -\mu^2\varphi(\bar{x})$, donc :

$$-\frac{p^\mu p_\mu}{\hbar^2}\varphi(\bar{x}) = -\mu^2\varphi(\bar{x})$$

On doit donc avoir :

$$p^\mu p_\mu = \mu^2\hbar^2$$

Or, si on calcule le produit scalaire $p^\mu p_\mu$, on a :

$$\left(\frac{E}{c}\right)^2 - (p^1)^2 - (p^2)^2 - (p^3)^2 = \frac{E^2}{c^2} - p^2$$

Comme ceci doit être égale à $\mu^2\hbar^2$, on a :

$$E^2 = p^2 c^2 + \mu^2 \hbar^2 c^2$$

Ce qui montre que les ondes planes sont bien solutions de l'équation de Klein-Gordon, à condition que la condition que l'on vient d'écrire soit respectée. On peut la comparer avec une relation très importante venue de la relativité restreinte³ :

$$E^2 = p^2 c^2 + m^2 c^4$$

En comparant les deux expressions, on trouve un lien entre le paramètre μ et la masse m , qui est simplement $\mu = \frac{mc}{\hbar}$ ce qui justifie le choix de l'avoir mis au carré dans l'équation de Klein-Gordon.

Supposons que nous ayons une particule de masse nulle (on a alors $E = pc$), qui se dirige en ligne droite. Pour la simplicité, on choisit un référentiel où la particule passe par l'origine à $t = 0$. Dans ce cas, le vecteur position \vec{r} est colinéaire à \vec{p} , et le produit scalaire $\vec{p} \cdot \vec{r}$ est simplement le produit des normes pr . Si on reprend l'expression de $\bar{p} \cdot \bar{x}$, on a :

$$\bar{p} \cdot \bar{x} = Et - \vec{p} \cdot \vec{r} = Et - pr = E \left(t - \frac{r}{c} \right)$$

À l'origine ($t = 0, r = 0$), on a $\bar{p}_0 \cdot \bar{x}_0 = 0$. On a donc $\varphi(\bar{x}_0) = Ae^{-\frac{i}{\hbar}\bar{p}_0 \cdot \bar{x}_0} = Ae^0 = A$. L'onde va ensuite continuer sa propagation de telle sorte que la zone d'amplitude A à l'origine à $t = 0$ va se retrouver à une certaine distance r de l'origine au bout d'un certain temps t . En ce point-événement \bar{x}_1 , on sait que $\varphi(\bar{x}_1) = Ae^{-\frac{i}{\hbar}\bar{p}_1 \cdot \bar{x}_1} = A$ (l'onde progresse sans se déformer, elle garde la même amplitude), ce qui impose que $\bar{p}_1 \cdot \bar{x}_1 = 0$. Mais comme on sait que $\bar{p}_1 \cdot \bar{x}_1 = E(t_1 - \frac{r_1}{c})$, et qu'une particule a forcément une énergie non nulle, on doit avoir $t_1 = \frac{r_1}{c}$. L'onde progresse donc à la vitesse $v = \frac{r_1}{t_1} = c$, ce qui tout simplement la vitesse de la lumière.

On a donc là un argument pour dire qu'une particule de masse nulle, *doit* se déplacer à la vitesse de la lumière.

L'équation de Klein-Gordon (8) est linéaire c'est-à-dire que si $\varphi_1(\bar{x})$ et $\varphi_2(\bar{x})$ sont des solutions de cette équation, alors :

$$\lambda_1\varphi_1(\bar{x}) + \lambda_2\varphi_2(\bar{x})$$

est aussi une solution de l'équation, où λ_1 et λ_2 sont deux nombres réels. On peut ainsi superposer un nombre arbitrairement grand d'ondes de fréquences différentes (ici donc de \bar{p} différents). On aura une somme de la forme :

$$\varphi(\bar{x}) = \sum_k \lambda_k \varphi_k(\bar{x}) = \sum_k \lambda_k A_k e^{\frac{i}{\hbar}\bar{p}_k \cdot \bar{x}}$$

3. Le cas particulier où $p = 0$ donne $E = mc^2$ qui est probablement l'équation la plus célèbre du monde.

On peut même sommer sur une somme infinie continue d'ondes de différents \vec{p} , en remplaçant la somme par une intégrale, en introduisant la fonction $\varphi(\vec{p}, t)$, et en rappelant que $\vec{p} \cdot \vec{x} = Et - \vec{p} \cdot \vec{r}$:

$$\varphi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \varphi(\vec{p}, t) e^{\frac{i}{\hbar} \vec{p} \cdot \vec{r}}$$

Le facteur $1/(2\pi)^3$ est ici pour des raisons de convention. Notez que le facteur $e^{-\frac{i}{\hbar} Et}$ a été implicitement incorporé dans la fonction $\varphi(\vec{p}, t)$, qui est appelée transformée de Fourier de la fonction $\varphi(\vec{x}, t) = \varphi(\vec{x})$. On peut montrer que l'on a :

$$\varphi(\vec{p}, t) = \int d^3x \varphi(\vec{x}, t) e^{-\frac{i}{\hbar} \vec{p} \cdot \vec{x}}$$

Cette décomposition de $\varphi(\vec{x})$ en une «somme continue» de fonctions à différentes fréquences (rappelons encore que dans ce cas la fréquence est représentée par la quantité de mouvement \vec{p}) est intéressante, mais le fait que l'équation de Klein-Gordon soit linéaire ne garantit pas que toute fonction $\varphi(\vec{x})$ qui puisse s'écrire sous une telle forme soit automatiquement une solution de cette équation. Une intégrale et une somme finie sont deux choses bien différentes. Il faut donc réintroduire cette expression dans l'équation de Klein-Gordon. Rappelons d'abord comment on peut calculer le D'Alembertien :

$$\square \varphi(\vec{x}) = \frac{1}{c^2} \frac{\partial^2 \varphi(\vec{x})}{\partial t^2} - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \varphi(\vec{x})$$

on a donc (en dérivant sous le signe intégral) :

$$(\square + \mu^2) \varphi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} e^{\frac{i}{\hbar} \vec{p} \cdot \vec{r}} \left(\frac{1}{c^2} \frac{\partial^2 \varphi(\vec{p}, t)}{\partial t^2} + \frac{p^2}{\hbar^2} \varphi(\vec{p}, t) + \mu^2 \varphi(\vec{p}, t) \right) = 0$$

On a donc comme équation pour $\varphi(\vec{p}, t)$:

$$\frac{\partial^2 \varphi(\vec{p}, t)}{\partial t^2} + c^2 \left(\frac{p^2}{\hbar^2} + \mu^2 \right) \varphi(\vec{p}, t) = 0$$

On reconnaît là l'équation d'un oscillateur harmonique (voir 7) de fréquence $\omega_{\vec{p}}^2 = c^2 \left(\frac{p^2}{\hbar^2} + \mu^2 \right)$, c'est à dire :

$$\omega_{\vec{p}} = \sqrt{c^2 \frac{p^2}{\hbar^2} + c^2 \frac{m^2 c^2}{\hbar^2}} = \frac{1}{\hbar} \sqrt{p^2 c^2 + m^2 c^4}$$

Notez qu'avec $E^2 = p^2 c^2 + m^2 c^4$, on retrouve $E = \hbar \omega = h \nu$. On a vu les niveaux d'énergies de l'oscillateur harmonique quantique, ici cela prend tout son sens : il existe des niveaux d'excitations élémentaires du champs, ces excitations élémentaires sont tout simplement appelés des particules !!

En plus des particules, il y a des fluctuations du champs comme nous l'avons mentionnés quand nous avons parlé de l'oscillateur harmonique quantique, relié au principe d'incertitude. On a alors une énergie minimale, même en l'absence de particules, l'énergie du point zéro $E_0 = \frac{1}{2} \hbar \omega$. Problème, il y a ici autant d'oscillateurs qu'il y a de fréquences possibles, c'est à dire une infinité. On devrait alors aboutir à la conclusion que l'énergie du vide est infinie ! Bien sûr c'est un peu plus compliqué que ça, mais même en faisant les choses plus rigoureusement, on trouve encore une valeur 10^{120} fois plus grande que celle effectivement observée...

Pour la suite et par commodité nous allons poser $c = \hbar = 1$ (cela revient à faire un choix d'unité). On aura alors en particuliers $\mu = m$.

3.1 Formalisme lagrangien

On peut formuler toute la mécanique classique en se basant sur les lois de Newton. Mais il est possible de trouver d'autres formulations intéressantes, comme celle du formalisme lagrangien.

Supposons une particule qui se déplace d'un point \vec{r}_i au temps initial t_i jusqu'au point \vec{r}_f au temps final t_f . Sur toutes les trajectoires possibles entre ces deux points, une seule correspond à la trajectoire physique, la *trajectoire naturelle*.

On va, en plus de la trajectoire naturelle, considérer des trajectoires non-physiques, qui respectent les deux conditions suivantes :

1. elles passent en \vec{r}_i au temps t_i et en \vec{r}_f au temps t_f ,
2. à chaque instant, elles diffèrent de la trajectoire physique par un *déplacement virtuel*. C'est un déplacement infinitésimal, instantané, compatible avec les contraintes physiques. Par exemple, pour un pendule avec une tige rigide, on doit toujours avoir la même distance entre le point de pivot et le poids au bout du pendule (c'est bien la trajectoire de ce dernier que l'on étudie).

On définit alors le *lagrangien* comme la différence entre l'énergie cinétique T et l'énergie potentielle V :

$$L = T - V$$

Et une fonction (ou fonctionnelle) qui à chaque trajectoire associe un nombre, l'*action* :

$$S = \int_{t_i}^{t_f} L dt$$

On peut alors montrer que la trajectoire naturelle est celle qui rend l'action stationnaire, ce que l'on note $\delta S = 0$. Pour comprendre ce que l'on veut dire par là, on peut utiliser une analogie. Supposons une fonction f de \mathbb{R} dans \mathbb{R} , on dit qu'elle est stationnaire au point x_0 si sa dérivée s'annule en ce point : $f'(x_0) = 0$. Ici, S est une fonction de l'espace des trajectoire vers l'ensemble \mathbb{R} , donc cette notion est mathématiquement différente bien qu'analogue, et l'on dit que l'action est stationnaire suivant la trajectoire naturelle.

Lorsqu'un système est soumis à des contraintes, il est utile de travailler en utilisant les *coordonnées de Lagrange*. Prenons l'exemple d'un pendule simple, le poids au bout du pendule est à distance fixe du point de pivot et son mouvement s'effectue dans un plan vertical. Il suffit donc d'un seul paramètre, l'angle entre le pendule et la verticale, pour fixer la position du pendule, au lieu des trois coordonnées spatiales que l'on doit utiliser en l'absence de contraintes. On va noter ces coordonnées q^α , et leurs dérivées totales par rapport au temps \dot{q}^α . Travailler dans ces coordonnées permet de ne pas se préoccuper de la deuxième condition sur les trajectoires puisque les contraintes sont automatiquement satisfaites.

Nous sommes maintenant prêt à formuler la mécanique lagrangienne. L'idée est d'introduire une certaine fonction, appelé lagrangien, qui dépend des q^α , \dot{q}^α et de t . Le lagrangien est dans ce contexte une donnée de la théorie : il est en fait *postulé*. On oublie donc la définition précédente pour quelque chose de plus général.

Comme précédemment, on postule que les trajectoires naturelles rendent l'action stationnaire. On a donc :

$$\begin{aligned} \delta S &= \int_{t_i}^{t_f} \delta L dt \\ &= \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q^\alpha} \delta q^\alpha + \frac{\partial L}{\partial \dot{q}^\alpha} \delta \dot{q}^\alpha \right) dt \end{aligned} \quad (9)$$

Comme $\delta\dot{q}^\alpha = \delta \frac{d}{dt} q^\alpha = \frac{d}{dt} \delta q^\alpha$, on a :

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^\alpha} \delta q^\alpha \right) = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^\alpha} \right) \delta q^\alpha + \frac{\partial L}{\partial \dot{q}^\alpha} \delta \dot{q}^\alpha$$

en utilisant la règle de Leibniz. On peut donc isoler le dernier terme et remplacer dans (9) :

$$\begin{aligned} \delta S &= \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q^\alpha} \delta q^\alpha + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^\alpha} \delta q^\alpha \right) - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^\alpha} \right) \delta q^\alpha \right) dt \\ &= \left[\frac{\partial L}{\partial \dot{q}^\alpha} \delta q^\alpha \right]_{t_i}^{t_f} + \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q^\alpha} \delta q^\alpha - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^\alpha} \right) \delta q^\alpha \right) dt \\ &= \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q^\alpha} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^\alpha} \right) \delta q^\alpha dt \end{aligned}$$

On a utilisé la première condition sur les trajectoire : $\delta q^\alpha(t_i) = \delta q^\alpha(t_f) = 0$. Comme on doit avoir $\delta S = 0$ quelque soit δq^α , on a :

$$\frac{\partial L}{\partial q^\alpha} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^\alpha} = 0 \quad (10)$$

Ce sont les *équations d'Euler-Lagrange*.

Prenons un exemple : supposons que le lagrangien s'écrive :

$$L = \frac{1}{2} m v^2 - q\varphi(\vec{r}, t)$$

On suppose l'absence de contraintes, aussi on travaille avec les coordonnées cartésiennes usuelles. Dans ce lagrangien, φ est une fonction quelconque qui ne dépend que de la position et du temps.

On a donc (en notant x^i et v^i les composantes de \vec{r} et \vec{v} respectivement) :

$$\frac{\partial L}{\partial v^i} = m v^i = p^i$$

et :

$$\frac{\partial L}{\partial x^i} = -q \frac{\partial \varphi}{\partial x^i} = -q E^i$$

où l'on a posé⁴ $E = -\nabla\varphi$. Les équations d'Euler-Lagrange donnent donc :

$$-q E^i - \frac{d}{dt} p_i = 0$$

soit, en notation vectorielle :

$$\frac{d\vec{p}}{dt} = q\vec{E}$$

qui est l'équation du mouvement d'une particule plongée dans un champ électrique \vec{E} .

4. On définit ∇ de la manière suivante : $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$ c'est-à-dire que l'on a :

$$E = \left(-\frac{\partial \varphi}{\partial x}, -\frac{\partial \varphi}{\partial y}, -\frac{\partial \varphi}{\partial z} \right)$$

Dans le cadre de la théorie des champs, on introduit la *densité lagrangienne* \mathcal{L} qui est telle que :

$$L = \int \mathcal{L} d^3x$$

et l'action s'écrit alors :

$$S = \int \mathcal{L} d^4x$$

Mais par abus de langage, on appelle souvent \mathcal{L} le lagrangien. Comme on travaille sur des champs, le lagrangien va dépendre des champs $\varphi_A(x^\mu)$ et de leurs dérivées $\partial_\mu \varphi_A(x^\mu)$ (l'indice A sert à distinguer les différents champs). Les équations d'Euler-Lagrange (10) deviennent :

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_A)} - \frac{\partial \mathcal{L}}{\partial \varphi_A} = 0$$

Supposons le lagrangien suivant :

$$\mathcal{L} = \frac{1}{2}(\partial^\mu \varphi)(\partial_\mu \varphi) - \frac{1}{2}m^2 \varphi^2$$

Le premier terme est appelé terme cinétique et le deuxième est appelé terme de masse.

On a donc :

$$\frac{\partial \mathcal{L}}{\partial \varphi} = -m^2 \varphi$$

et :

$$\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} = \partial^\mu \varphi$$

L'équation d'Euler-Lagrange donne alors :

$$\partial_\mu (\partial^\mu \varphi) = -m^2 \varphi$$

ce qui peut se réécrire, en reconnaissant l'expression du d'Alembertien :

$$\left(\square + m^2\right) \varphi = 0$$

on retrouve l'équation de Klein-Gordon (8).

4 Théories de Jauge

Supposons le lagrangien suivant :

$$\mathcal{L} = (\partial^\mu \varphi^*)(\partial_\mu \varphi) - m^2 \varphi^* \varphi$$

Cela ressemble au lagrangien donné précédemment, à la différence que l'on n'a plus un champ réel mais un champ complexe (on a noté φ^* le complexe conjugué⁵ de φ). On peut montrer par un calcul analogue au précédent, que φ (et φ^*) obéit à l'équation de Klein-Gordon.

Supposons que nous l'on remplace le champ φ par un champ $\tilde{\varphi}$:

$$\varphi \longrightarrow \tilde{\varphi} = e^{i\alpha} \varphi \tag{11}$$

5. Rappelons que le complexe conjugué d'un nombre complexe $a + bi$ est $a - bi$. Notez que le conjugué de e^{ix} où x est un réel est $(e^{ix})^* = (\cos(x) + i \sin(x))^* = \cos(x) - i \sin(x) = \cos(x) + i \sin(-x) = e^{-ix}$.

Si on suppose que α est une constante, alors en remplaçant φ par $\tilde{\varphi}$ dans le lagrangien, on a :

$$\begin{aligned}\mathcal{L} &= (\partial^\mu(e^{i\alpha}\varphi)^*)(\partial_\mu(e^{i\alpha}\varphi)) - m^2(e^{i\alpha}\varphi)^*(e^{i\alpha}\varphi) \\ &= (\partial^\mu(e^{-i\alpha}\varphi^*))(\partial_\mu(e^{i\alpha}\varphi)) - m^2e^{-i\alpha}\varphi^*e^{i\alpha}\varphi \\ &= (\partial^\mu\varphi^*)(\partial_\mu\varphi) - m^2\varphi^*\varphi\end{aligned}$$

On retombe sur le lagrangien de départ. On ne peut donc pas « distinguer » φ et $\tilde{\varphi}$, il sont physiquement complètement équivalents. On dit que l'on a une *symétrie de jauge* et la transformation (11) est appelé *transformation de jauge*.

Pour montrer l'invariance de jauge, on a supposé α constante, de telle sorte que l'on puisse sortir l'exponentielle de la dérivée. La même transformation se fait partout dans l'espace-temps, c'est une transformation *globale*. Peut-on être plus exigeant que cela et demander une symétrie locale? À priori c'est impossible, mais peut-être peut-on le faire en modifiant le lagrangien. Remarquons que :

$$\partial_\mu\tilde{\varphi} = \partial_\mu(e^{i\alpha}\varphi) = e^{i\alpha}\partial_\mu\varphi + i(\partial_\mu\alpha)e^{i\alpha}\varphi = \widetilde{\partial_\mu\varphi} + i(\partial_\mu\alpha)\tilde{\varphi}$$

Le terme qui nous dérange est le terme $i(\partial_\mu\alpha)\tilde{\varphi}$. Il est possible de ne plus avoir de terme de ce type en remplaçant la dérivée « ordinaire » par ce qu'on appelle une *dérivée covariante*, qui s'écrit :

$$D_\mu = \partial_\mu - ieA_\mu$$

e est un simple nombre (une constante) et l'on a introduit un champs vectoriel A_μ . Pour restaurer l'invariance de jauge, il faut que $\widetilde{\partial_\mu\varphi} = \widetilde{D_\mu\varphi}$ ou encore :

$$\begin{aligned}\partial_\mu(e^{i\alpha}\varphi) - ie\widetilde{A}_\mu e^{i\alpha}\varphi &= e^{i\alpha}(\partial_\mu\varphi - ieA_\mu\varphi) \\ e^{i\alpha}\partial_\mu\varphi + i(\partial_\mu\alpha)e^{i\alpha}\varphi - ie\widetilde{A}_\mu e^{i\alpha}\varphi &= e^{i\alpha}(\partial_\mu\varphi - ieA_\mu\varphi) \\ \partial_\mu\varphi + i(\partial_\mu\alpha)\varphi - ie\widetilde{A}_\mu\varphi &= \partial_\mu\varphi - ieA_\mu\varphi \\ i(\partial_\mu\alpha)\varphi - ie\widetilde{A}_\mu\varphi &= -ieA_\mu\varphi\end{aligned}$$

En isolant \widetilde{A}_μ , on a :

$$\widetilde{A}_\mu = A_\mu - \frac{1}{e}\partial_\mu\alpha$$

On a donc introduit un nouveau champs A_μ qui est un champs vectoriel. Mais pour qu'il y ait le moindre intérêt physique, il faut ajouter un terme cinétique pour A_μ , il doit y avoir une dynamique. Mais il faut respecter l'invariance de Jauge et l'invariance de Lorentz, donc il faut bien choisir ce terme. Je vais donner ici (sans démonstration) le choix le plus simple. Définissons le tenseur de Faraday par :

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

Alors le terme cinétique est donné par $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$.

On a donc le lagrangien complet :

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D^\mu\varphi^*)(D_\mu\varphi) - m^2\varphi^*\varphi$$

Comment interpréter le champs A_μ ? Le calcul n'est pas compliqué mais fort long (aussi on ne va le faire ici), mais on peut montrer qu'en écrivant l'équation d'Euler-Lagrange pour ce champs, on retombe sur les équations de Maxwell, si on écrit :

$$A^\mu = \left(\frac{\varphi}{c}, \vec{A}\right)$$

et qu'on identifie φ et \vec{A} comme les potentiels électriques et magnétiques respectivement. C'est à dire que l'on a⁶ :

$$\vec{E} = -\nabla\varphi - \frac{\partial\vec{A}}{\partial t} \quad \text{et} \quad \vec{B} = \nabla \times \vec{A}$$

Nous avons donc montré que le champs électromagnétique doit être présent si on impose notre invariance de Jauge. Question : peut-on ajouter un terme de masse au lagrangien pour ce champs ? La réponse est non, ce n'est pas possible. Un tel terme devrait s'écrire $\frac{1}{2}m_A^2 A_\mu A^\mu$ ce qui briserait l'invariance de Jauge. La particule associée à ce champs (ici le photon) doit donc être obligatoirement de masse nulle. Le photon est dit le «boson vecteur» de l'électromagnétisme, et on peut montrer que c'est une particule de spin un⁷. Le champs électromagnétique est relié au groupe de symétrie $U(1)$, qui est l'ensemble des nombres complexes qui peuvent s'écrire $e^{i\alpha}$ avec $\alpha \in \mathbb{R}$ (c.f. invariance de Jauge (11)).

5 Théorie électrofaible et masses des bosons vecteurs

Compliquons un petit peu les choses. Supposons deux champs de Klein-Gordon complexes φ_1 et φ_2 . Alors, on a le lagrangien :

$$\mathcal{L} = (\partial^\mu \varphi_1^*)(\partial_\mu \varphi_1) + (\partial^\mu \varphi_2^*)(\partial_\mu \varphi_2) - m^2 \varphi_1^* \varphi_1 - m^2 \varphi_2^* \varphi_2$$

Ce lagrangien peut se réécrire comme si l'on avait qu'un seul champs Φ . Il suffit de poser : $\Phi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}$, et de définir Φ^\dagger comme étant la transposée de la conjuguée de Φ , c'est-à-dire que l'on remplace chaque élément de la matrice Φ par son complexe conjugué et on échange les lignes et les colonnes, ici on a simplement $\Phi^\dagger = \begin{pmatrix} \varphi_1^* & \varphi_2^* \end{pmatrix}$. On a alors :

$$\mathcal{L} = (\partial^\mu \Phi^\dagger)(\partial_\mu \Phi) - m^2 \Phi^\dagger \Phi$$

En effet, on a : $\Phi^\dagger \Phi = \begin{pmatrix} \varphi_1^* & \varphi_2^* \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = \varphi_1^* \varphi_1 + \varphi_2^* \varphi_2$

Le lagrangien est invariant si on remplace Φ par :

$$\tilde{\Phi} = U\Phi \quad \text{avec} \quad U^\dagger U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

En effet⁸ :

$$\tilde{\Phi}^\dagger \tilde{\Phi} = \Phi^\dagger U^\dagger U \Phi = \Phi^\dagger \Phi$$

Tout comme on avait pour le cas précédent le groupe $U(1)$, ici on utilise le groupe $SU(2)$, qui est l'ensemble des matrices U deux fois deux tel que $U^\dagger U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et de déterminant un. Si on n'impose pas cette dernière condition, on a l'ensemble $U(2)$.

6. On définit $\nabla \times \vec{A}$ par :

$$\nabla \times \vec{A} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}, \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}, \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right)$$

7. On ne va pas discuter ici du spin car je veux garder ce texte aussi court que possible, mais comprendre ce qu'est vraiment le spin n'est pas essentiel ici.

8. On peut montrer mathématiquement que $(AB)^\dagger = B^\dagger A^\dagger$.

Il reste à trouver comment construire la dérivée covariante, trouver combien de champs supplémentaires il faut postuler, et trouver un terme cinétique. On ne va pas le faire ici car ce serait trop long et pas assez instructif. Contrairement au cas précédent, il se fait qu'il faut postuler trois champs supplémentaires (et non un seul).

On a donc : trois bosons vecteurs pour $SU(2)$ et un pour $U(1)$. Aucun d'entre eux ne peut avoir de masse, comme nous l'avons vu pour l'électromagnétisme, ajouter un terme de masse briserait l'invariance de jauge. Il se fait qu'il existe une interaction entre certaines particules, appelée interaction faible, qui a trois bosons vecteurs : deux bosons de charges plus et moins (notés W^+ et W^-) et un boson neutre (noté Z^0). Si on ajoute à cela l'électromagnétisme, on a quatre bosons en tout, et même mieux que ça : un boson pour une interaction, et trois pour un autre. Cela semble confirmer ce que nous avons fait jusqu'ici. Mais il y a un gros problème : les bosons W^\pm et Z^0 ont une masse, ce dernier ayant une masse légèrement supérieure à celle des bosons W^+ et W^- qui ont tous les deux la même masse. Comment résoudre ce problème ?

6 Brisure spontanée de symétrie

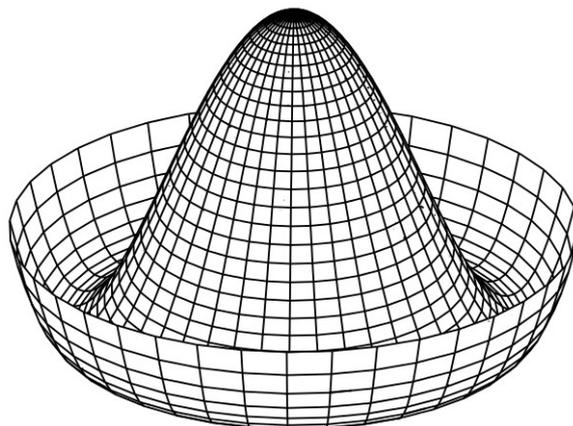
On va ici donner une vision très simplifiée du mécanisme de brisure de symétrie, avec un champs de Klein-Gordon plus un champs de Jauge A_μ . On reprend donc le Lagrangien que l'on a écrit lorsque l'on a discuté de l'électromagnétisme, mais cette fois en y ajoutant un potentiel :

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D^\mu\varphi^*)(D_\mu\varphi) - V(\varphi)$$

On n'a pas mis de terme de masse au champs A_μ car on a vu que l'on a pas le droit de le faire si on veut respecter la symétrie de Jauge. Le potentiel $V(\varphi)$ doit lui-même «respecter» cette invariance. Un choix possible (et judicieux) d'un tel potentiel est ceci :

$$\frac{\lambda}{4}(\varphi^*\varphi - v^2)^2$$

où λ et v sont des paramètres. Il est facile de voir que remplacer φ par $e^{i\alpha}\varphi$ ne change rien, on a donc bien invariance. La «forme» du potentiel est la suivante :



Remarquons qu'en développant le carré, on a un terme (le terme croisé) qui vaut $-\frac{\lambda}{4}2\varphi^*\varphi(-v^2) = \frac{\lambda}{2}v^2\varphi^*\varphi$. Si on compare cela avec un terme de masse éventuel pour φ qui s'écrirait alors $-m_\varphi^2\varphi^*\varphi$, on doit alors en déduire que la particule φ a une masse au carré de :

$$m_\varphi^2 = -\frac{\lambda}{2}v^2$$

tandis que la masse de A_μ est nulle. Notez que le carré de la masse est négatif!

Maintenant, réfléchissons un peu sur la situation. Quelle est la valeur moyenne du champs φ dans un espace vide? On pourrait imaginer que dans un espace vide, cette valeur est simplement nulle. Mais dans ce cas, on se retrouve au sommet de la «bosse» du potentiel $V(\varphi)$. C'est clairement une position d'équilibre instable. Un vide dans cette configuration va évoluer spontanément vers un *autre* vide où la valeur moyenne du champs est de module v . On se retrouvera quelque part sur le cercle de rayon v centré autour de zéro dans le plan complexe, qui correspond au minimum du potentiel.

Il faut bien comprendre que cet état est bien un vide : ce n'est pas parce la valeur moyenne du champs est différente de zéro qu'il y a des particules. Une particule, est comme nous l'avons vu une vibration d'un champ, peu importe la valeur moyenne de celui-ci. Il est donc utile de séparer le champ φ en une partie constante de module v , et une partie «oscillante». Par symétrie de Jauge, on peut toujours supposer (par simplicité) le champ φ réel. En effet, on se retrouve quelque part sur le cercle correspondant au minimum du potentiel, mais où que l'on soit sur le cercle la situation est équivalente, on peut poser que l'on est sur l'axe réel. On écrit alors la décomposition :

$$\varphi(\bar{x}) = v + \eta(\bar{x})$$

On peut alors remplacer dans le lagrangien :

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (\partial^\mu\eta - ieA^\mu(v + \eta))(\partial_\mu\eta + ieA_\mu(v + \eta)) - \frac{\lambda}{4}((v + \eta) - v)^2$$

Considérons d'abord le deuxième terme :

$$(\partial^\mu\eta - ieA^\mu(v + \eta))(\partial_\mu\eta + ieA_\mu(v + \eta)) = (\partial_\mu\eta)(\partial^\mu\eta) + e^2A_\mu A^\mu(v^2 + 2v\eta + \eta^2)$$

et le troisième :

$$-\frac{\lambda}{4}((v + \eta) - v)^2 = -\frac{\lambda}{4}(\eta^2 + 2v\eta)^2 = -\frac{\lambda}{4}(\eta^4 + 4v\eta^3 + 4v^2\eta^2)$$

Au total, le lagrangien s'écrit :

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (\partial_\mu\eta)(\partial^\mu\eta) + e^2v^2A_\mu A^\mu - \lambda v^2\eta^2 + \dots$$

où nous avons omis d'écrire certains termes d'interaction qui ne nous intéressent pas ici. On voit apparaître un terme de masse pour A_μ et un terme pour η . Le signe du terme de masse pour A_μ est bien correct, à cause de notre convention pour la métrique, la partie spatiale de $A_\mu A^\mu$ fait apparaître un signe moins, et maintenant, le signe pour η est correct lui aussi : on n'a pas de masse au carré négative. Notez le tour de force : on n'a à aucun moment violé la symétrie de Jauge, et pourtant il apparaît un terme de masse pour le boson de Jauge! C'est parce que la symétrie de Jauge est maintenant «cachée», elle existe dans un certains sens mais n'est plus manifeste, on dit qu'il y a *brisure spontanée de symétrie*. La symétrie est respectée parce qu'aucune direction (dans le plan complexe) n'est privilégiée au départ (voir la symétrie circulaire du potentiel), mais une fois que le vide a transité vers un état de plus basse énergie (et stable), cette symétrie n'est plus manifeste.

Le terme de masse de A_μ doit s'écrire $\frac{1}{2}m_A^2 A_\mu A^\mu$ tandis que celui de η doit s'écrire simplement $-m_\eta^2 \eta^2$, on a donc :

$$m_A = \sqrt{2}ev \quad \text{et} \quad m_\eta = \sqrt{\lambda}v$$

Passons maintenant à quelque chose de plus réaliste : l'interaction électrofaible. Je vais là encore passer les détails mathématiques (et donc omettre certaines choses intéressantes comme

les bosons de Goldstone, mais ce texte est déjà assez long comme ça). Je vais juste illustrer sur un problème particuliers que les choses sont beaucoup plus compliquées que ce que nous venons de voir. On a à la fois la symétrie $SU(2)$ et la symétrie $U(1)$, et on a trois champs de Jauge pour $SU(2)$ (que l'on note $A_\mu^1, A_\mu^2, A_\mu^3$) et un pour $U(1)$ (que l'on note B_μ). À la symétrie $SU(2)$ correspond ce que l'on appelle une constante de couplage que l'on note g , et de même pour $U(1)$ on a une constante notée g' . Ces constantes correspondent intuitivement à la «force» de l'interaction, elles apparaissent dans l'expression de la dérivée covariante.

Maintenant, vous vous attendez peut-être à ce que les particules A_μ^i acquièrent chacun une masse et correspondent aux bosons W^\pm et Z^0 , et que le B_μ soit en fait le photon (sans masse bien sûr). Eh bien non, pas du tout ! Encore une fois, je ne fait pas le calcul pour ne pas compliquer plus que nécessaire, mais en fait les bosons W^\pm , Z^0 , et le photon sont des «mélanges» des champs A_μ^i et B_μ . On a, en effet :

$$W^\pm = \frac{1}{\sqrt{2}}(A_\mu^1 \mp iA_\mu^2)$$

pour les bosons W^\pm et ils ont une masse de $gv/2$. On a :

$$Z^0 = \frac{1}{\sqrt{g+g'}}(gA_\mu^3 - g'B_\mu)$$

pour le boson Z^0 qui a une masse de $\sqrt{g^2+g'^2}(v/2)$. Remarquez qu'il a une masse un peu supérieure au bosons W^\pm , effectivement expérimentalement ceux-ci ont une masse de ~ 80.4 GeV, alors que le Z^0 a une masse de ~ 91.2 GeV. Enfin le photon s'écrit :

$$A_\mu = \frac{1}{\sqrt{g+g'}}(g'A_\mu^3 + gB_\mu)$$

Remarquez la similitude avec le boson Z^0 .

7 Questions - réponses

Pour terminer, voici des réponses à quelques questions que vous pourriez avoir.

7.1 D'où vient la masse de l'électron ?

L'électron, tout comme ses «cousins» le muon et le tau (qui sont identiques à l'électrons à part qu'ils sont plus lourds), ainsi que les quarks, acquièrent leur masse grâce au champs de Higgs. Cela peut paraître étrange puisque l'on a vu pourquoi il faut donner une masse aux bosons W^\pm et Z^0 , mais pour les électrons, ne peut-on pas juste ajouter un terme de masse, sans introduire une interaction avec le champs de Higgs ? En fait, on ne peut pas. Ce serait long (et trop compliqué) à expliquer ici, mais de toute façon l'explication est similaire : on ne peut pas simplement ajouter de terme de masse sans détruire l'invariance de Jauge (et avoir des tas des problèmes, voire même des contradictions). Le champs de Higgs permet de contourner le problème.

7.2 Notre masse vient-elle du champs de Higgs ?

Non. Seul une infime partie de notre masse provient du mécanisme de Higgs. En effet, dans un atome, presque toute la masse est concentrée dans le noyau. Les noyaux d'atomes sont formés de protons et de neutrons, leurs masses proviennent de l'énergie de l'interaction entre les quarks

qui les composent (les quarks eux-même ont une faible masse, de part le mécanisme de Higgs, mais là encore c'est négligeable). En effet, comme Einstein l'a montré, il y a un lien étroit entre l'énergie et la masse (le fameux $E = mc^2$). De façon plus précise, l'interaction entre les quarks dans un proton (ou un neutron) est tellement forte que l'on ne peut pas se contenter de dire qu'il y a simplement trois quarks en interactions, ou du moins c'est une vision trop naïve. En réalité, il faut imaginer une «mer» de quarks, d'anti-quarks et de gluons, toutes ces particules étant «virtuelles» (sauf pour les trois quarks dits de «valence»). Le qualificatif de «virtuelles» veut dire qu'elles ne respectent pas la relation $E^2 = p^2c^2 + m^2c^4$, en réalité ce ne sont pas (vraiment) des particules, mais des mouvements incessants dans les champs de quarks de gluons.

7.3 Champs de Higgs ou boson de Higgs ?

C'est le champs de Higgs qui donne de la masse à certaines particules. Le boson de Higgs est une excitation (vibration) de ce champs.

7.4 Le champs de Higgs agit-il comme un frein ?

C'est une image classique de la façon dont le champs de Higgs agit, en «freinant» certaine particules. Bien sûr, ce n'est qu'une analogie, aussi il est intéressant de soulever les limites de cette analogie. Les particules libres gardent en effet une vitesse constante et ne freinent pas, mais ne peuvent plus aller à la vitesse de la lumière. Aussi il ne faut pas croire (comme pourrait indiquer cette analogie) que cela veut dire quelque chose que de se déplacer par rapport au champs de Higgs : toute vitesse est relative !