Chapitre 3

La radioactivité β. Le neutrino.

La théorie de Fermi

1 La radioactivité β

La fin du siècle dernier a été marquée par une série de découvertes importantes: les rayons X, la radioactivité (α, β, γ). La radioactivité β nous a fait connaître une nouvelle forme d'interaction: les interactions faibles.

La radioactivité β a été observée pour la première fois en physique nucléaire. Par exemple, un noyau de ^{234}Th se transforme en un noyau de ^{234}Pa avec émission d'un électron:

$$^{234}Th \rightarrow ^{234} Pa \; e^- \tag{1}$$

Si cela était le mode de désintégration on devrait observer un électron monoénergétique, son énergie étant pratiquement la différence de masse ΔM entre le noyau initial et le noyau final (l'énergie de recul du noyau final étant négligeable).

Or, l'expérience montre que le spectre d'énergie des électrons est un spectre continu. Il y a apparemment violation de la conservation de l'énergie. Cette violation est-elle réelle, ou a-t-on mal compris le phénomène? Pour certains, comme Niels Bohr, on pourrait accepter une violation de la conservation de l'énergie, puisque nous sommes en face d'un phénomène nouveau: il peut être dangereux de penser qu'un phénomène nouveau obéit aux lois de la physique connue jusqu'alors. Peut être que la conservation de l'énergie est valable seulement en moyenne, pour des ensembles de particules, mais pas pour chaque processus individuel. Si c'est le cas, des électrons pourraient avoir des énergies supérieures à ΔM, et d'autres auraient des énergies inférieures à ΔM. Mais l'énergie devrait être distribuée autour de la valeur moyenne ΔM. Il existe une expérience qui permet de vérifier ce point. Une source radioactive β est placée dans un calorimètre qui absorbe totalement l'énergie des électrons. La quantité de chaleur dégagée divisée par le nombre de désintégrations fournit la valeur moyenne recherchée. Elle est largement inférieure à ΔM, ce qui prouve qu'il n'y a pas conservation de l'énergie, même pour un ensemble de particules.

2 Le neutrino

Il y a bien un moyen de rétablir la conservation de l'énergie: on admet qu'une partie de l'énergie s'échappe du calorimètre. C'est ce que proclame Pauli en inventant le "neutrino".

En fait le mot “neutrino” a été inventé plus tard et pendant quelque temps on a utilisé le mot “neutron”.

Si l’on admet l’existence du neutrino, la radioactivité β devient:

$$^{234}Th \rightarrow ^{234} Pa\ e^- \nu \tag{2}$$

En examinant l’énergie maximale du spectre β on remarque que la masse du neutrino doit être très petite, beaucoup plus petite que la masse de l’électron. Elle pourrait même être nulle, et cette idée a fait son chemin (une particule de masse nulle ne devrait pas nous étonner, le photon a bien une masse nulle!). Aujourd’hui on ne sait toujours pas quelle est la masse du neutrino.

Le neutron libre subit la désintégration β:

$$n \rightarrow pe^- \nu \tag{3}$$

Nous verrons plus loin que le neutrino émis dans la désintégration β^+ est différent du neutrino émis dans la désintégration β^-. Les hélicités sont différentes: $+1$ pour le neutrino émis dans la désintégration β^+ et -1 pour le neutrino émis dans la désintégration β^-. On peut aussi postuler une loi de conservation des leptons. Dans ce cas le neutrino de la désintégration β^- serait un antineutrino et le neutrino de la désintégration β^- serait un neutrino.

Si la charge électrique du neutron est rigoureusement nulle, et si les charges électriques du proton et du neutron sont rigoureusement égales en valeur absolue, la charge électrique du neutrino doit être nulle. Ceci est vérifié expérimentalement de façon très précise. On considérera donc le neutrino comme électriquement neutre. Il pourrait cependant posséder un moment magnétique. On a établit une limite supérieure de $1.9 \times 10^{-10}\mu$B.

Si le neutrino n’a aucun moment électromagnétique on peut se poser la question de la distinction entre neutrino et antineutrino. Si cette distinction est réelle on aura quatre états possibles, avec deux degrés de liberté pour l’hélicité et deux degrés de liberté à cause de la dualité particule-antiparticule:

$$\nu_L \qquad \nu_R \qquad \overline{\nu}_L \qquad \overline{\nu}_R \tag{4}$$

Dans ce cas il faudra quatre composantes pour décrire le neutrino, et on a affaire à un neutrino “de Dirac”.

Majorana a proposé une théorie dans laquelle le neutrino est identique à son antiparticule. Dans ce cas on n’aura que deux états avec les deux hélicités opposées.

$$\nu_L \qquad \nu_R \tag{5}$$

Dans ce cas deux composantes suffisent pour décrire le neutrino, et on a affaire à un neutrino “de Majorana”.

Nous reviendrons sur cette question. Aujoud’hui on est toujours incapable de dire si le neutrino est “de Dirac” ou de ”Majorana”.

3 Interaction de contact de quatre fermions

Que pouvait-on faire, en 1934, pour développer une théorie quantique des interactions faibles, dont la seule manifestation connue était la radioactivité β? Dans le cas des interactions électromagnétiques on connaissait le point de départ, l'électromagnétisme classique. Et on savait comment s'y prendre pour quantifier le champ électromagnétique. Rien de semblable pour les interactions faibles, il fallait donc partir de zéro.

C'est alors que Fermi eut une idée de génie, consistant à affirmer qu'il y avait une profonde analogie entre les interactions faibles et les interactions électromagnétiques. Peut-on dire qu'il fut un des premiers artisans de l'unification de ces interactions?

En électromagnétisme, on décrit l'interaction d'une particule électriquement chargée avec le champ électromagnétique comme le produit scalaire de deux quadrivecteurs. La densité hamiltonienne, qui est un scalaire de Lorentz, s'ecrit:

$$H_{E.M.} = eJ^\mu A_\mu \tag{6}$$

où e est la charge électrique élementaire, qui joue le rôle de constante de couplage, eJ^μ le courant électromagnétique et A_μ le potentiel électromagnétique.

Le courant électromagnétique:

$$eJ^\mu = e\overline{\psi}\gamma^\mu\psi \tag{7}$$

se décompose en une partie temporelle:

$$eJ^0 = \psi^\dagger\psi \tag{8}$$

et une partie spatiale:

$$e\vec{J} = \psi^\dagger\vec{\alpha}\psi \tag{9}$$

On a utilisé ici une représentation particulière pour les matrices γ^μ (voir chapitre 1).

La composante temporelle est la densité de charge $e\rho$. En se rappelant la signification de l'opérateur $\vec{\alpha}$, on voit que la partie spatiale est la densité de courant, dont l'expression classique est $e\rho\vec{v}$, où $\vec{v}$ est la vitesse de la particule.

Le quadrivecteur A^μ se décompose en une partie temporelle qui est improprement désignée par potentiel "scalaire":

$$A_0 = -V \tag{10}$$

et une partie spatiale qui est improprement désignée par potentiel "vecteur" $\vec{A}$. Le quadrivecteur A_μ a pour composantes V et $-\vec{A}$. Et l'hamitonien d'interaction a pour expression:

$$H_{E.M.} = e(-\rho V + \vec{J}\cdot\vec{A}) \tag{11}$$

Fermi suppose alors que les interactions faibles aussi ont un caractère vectoriel. S'appuyant sur l'hypothèse du neutrino de Pauli, il propose de coupler un courant avec un potentiel.

Le courant est défini par analogie avec le courant électromagnétique. Pour un proton sans structure forte (on suppose que les interaction fortes ont été débranchées):

$$J^{\mu}_{E.M.} = \overline{\psi}_P \gamma^{\mu} \psi_P \tag{12}$$

On annihile un proton et on le crée dans un nouvel état, d'énergie-impulsion différent. On dit que l'on a affaire a un courant "neutre", avec

$$\Delta Q = 1 - 1 = 0 \tag{13}$$

On a défini ΔQ comme:

$$\Delta Q = Q_i - Q_f \tag{14}$$

où les indices i et f se rapportent à l'état initial et l'état final, respectivement.

Lors de la désintégration β^- on annihile un neutron et on crée un proton:

$$J^{\mu}_{-} = \overline{\psi}_P \gamma^{\mu} \psi_N \tag{15}$$

et ce courant est un courant chargé, car

$$\Delta Q = 0 - 1 = -1 \tag{16}$$

Lors de la désintégration β^+ on annihile un proton et on crée un neutron:

$$J^{\mu}_{+} = \overline{\psi}_N \gamma^{\mu} \psi_P \tag{17}$$

et ce courant est un courant chargé, car

$$\Delta Q = 1 - 0 = +1 \tag{18}$$

Mais que mettre à la place du potentiel A_{μ}? Il doit s'agir d'un champ que l'on peut construire à partir de l'électron (positron) et le neutrino (antineutrino). Si l'on suppose que le neutrino est différent de l'antineutrino, et si l'on postule une loi de conservation pour les leptons, on écrira pour l'hamiltonien de la désintégration β^-:

$$H_{-} = G(\overline{\psi}_P \gamma^{\mu} \psi_N)(\overline{\psi}_{e^-} \gamma_{\mu} \psi_{\nu}) \tag{19}$$

L'annihilation d'un neutrino équivaut à la création d'un antineutrino.

$$N \to P e^- \overline{\nu} \tag{20}$$

équivaut à

$$N \nu \to P e^- \tag{21}$$

Ce processus existe, c'est une réaction induite par un neutrino sur un neutron appartenant à un noyau atomique.

Pour la désintégration β^+ on a:

$$H_+ = G(\overline{\psi}_N \gamma^\mu \psi_P)(\overline{\psi}_\nu \gamma_\mu \psi_{e^-}) \tag{22}$$

L'annihilation d'un électron équivaut à la création d'un positron. Ainsi,

$$P \to N e^+ \nu \tag{23}$$

équivaut à

$$P e^- \to N \nu \tag{24}$$

Ce processus est connu depuis longtemps, il s'agit de la capture d'un électron orbital par un noyau. Un exemple est:

$$^7Be\ e^- \to\ ^7Li\ \nu \tag{25}$$

La capture d'électron orbital entre en compétition avec la désintégration β^+.

La constante de couplage G caractérise la "force" des interactions faibles. Elle a pour dimension EL^3 et peut donc s'exprimer en J-m^3. Il est souvent plus pratique d'utiliser d'autres unités. Par exemple (Particle Data Booklet[1]):

$$\frac{G}{(\hbar c)^3} = 1.16639(2) \times 10^{-5}\ \mathrm{GeV}^{-2} \tag{26}$$

Il faut noter que l'on a introduit seulement les spineurs des leptons, mais pas de termes contenant des dérivées, par souci de simplicité. L'introduction de termes dérivatifs a été proposée plus tard par Konopinski et Uhlenbeck, mais n'a pas été utilisée par la suite.

Il faut aussi noter que les quatre spineurs sont évalués ont même point de l'espace-temps. On a affaire à une interaction "de contact".

4 Transitions β permises

Toutes les grandeurs observables de la désintégration β s'obtiennent à partir de la densité d'hamiltonien, en utilisant la théorie des perturbations:

$$P_{i\to f} = \frac{2\pi}{\hbar} |\langle i|H|f\rangle|^2 \rho(E_f) \tag{27}$$

L'utilisation de la théorie des perturbations est légitime, puisqu'il s'agit d'une interaction faible.

Nous allons d'abord calculer le terme:

$$\rho(E_f) = \frac{dn}{dE_f} \tag{28}$$

qui représente la densité d'états par intervalle d'énergie finale. Le bilan cinématique s'écrit:

$$M_i = M_f + E_e + E_\nu + T_R \tag{29}$$

$$0 = \vec{p}_e + \vec{p}_\nu + \vec{p}_R \tag{30}$$

$$E_e = \sqrt{p_e^2 + m_e^2} \tag{31}$$

et si l'on donne une masse au neutrino:

$$E_\nu = \sqrt{p_\nu^2 + m_\nu^2} \tag{32}$$

avec les notations de la table 1.

Particule	Masse	Impulsion	Énergie totale	Énergie cinétique
Noyau initial	M_i	0	M_i	0
Noyau final	M_f	$\vec{p}_R$	$T_R + M_f$	T_R
Électron	m_e	$\vec{p}_e$	E_e	$E_e - m_e$
Neutrino	m_ν	$\vec{p}_\nu$	E_ν	$E_\nu - m_\nu$

Table 1: Variables cinématiques pour la désintégration β.

****
Exercice: Évaluez, de façon approximative, l'énergie de recul T_R du noyau final. Placez vous dans la configuration cinématique qui rend cette énergie maximale.
****

Le calcul montre que l'énergie cinétique de recul du noyau est négligeable (mais pas son impulsion). On aura donc pour l'énergie dans l'état final:

$$E_f = M_i - M_f = W_0 \tag{33}$$

La densité des états est donnée par:

$$dn = p_e^2 dp_e p_\nu^2 dp_\nu = p_e E_e dE_e p_\nu E_\nu dE_\nu \tag{34}$$

Faisons un changement de variables, de (E_e, E_ν) à (E_e, E_f). On trouve:

$$\rho(E_f) = \frac{dn}{dE_f} = p_e E_e (W_0 - E_e)\left[(W_0 - E_e)^2 - m_\nu^2\right]^{1/2} dE_e \tag{35}$$

Si la masse du neutrino est nulle, l'expression se simplifie:

$$\rho(E_f) = \frac{dn}{dE_f} = p_e E_e (W_0 - E_e)^2 dE_e \tag{36}$$

L'élément de matrice contient une partie leptonique avec les spineurs de l'électron (émis) et de l'antineutrino (absorbé). Négligeant l'interaction faible de l'antineutrino avec le noyau final (et l'électron) nous utiliserons, pour la dépendance spatiale, une onde plane:

$$\psi_\nu \sim u_\nu \exp(-i\vec{p}_\nu \cdot \vec{r}) \tag{37}$$

Le signe $-$ traduit le fait que l'émission d'un antineutrino avec une impulsion $\vec{p}_\nu$ équivaut à l'absorption d'un neutrino d'impulsion $-\vec{p}_\nu$. En ce qui concerne l'électron, une onde plane sera une mauvaise approximation puisque l'électron émis se trouve dans le champ coulombien du noyau final. Faisons tout de même cette approximation afin de ne pas occulter la physique derrière plusieurs pages de calcul. Nous verrons plus tard les corrections à apporter. La partie spatiale sera pour l'électron:

$$\psi_e \sim u_e \exp(i\vec{p}_e \cdot \vec{r}) \tag{38}$$

Il faut maintenant penser à la partie hadronique. Les nucléons étant non-relativistes, les termes les plus grands dans l'élément de matrice seront ceux qui proviennent des opérateurs non-relativistes. Dans le cas de l'interaction proposée par Fermi c'est l'opérateur γ^0, dont la limite non-relativiste est l'opérateur I (identité).

Nous allons définir comme "transitions β permises" celles qui se produisent avec la plus grande probabilité. Nous retiendrons donc seulement l'opérateur γ^0, puis sa limite non-relativiste, I. Nous obtenons donc pour l'élément de matrice:

$$\langle f|H|i\rangle \;=\; \langle f|(\overline{\psi_p}\gamma^0\psi_n)(\overline{\psi_e}\gamma^0\psi_\nu)|i\rangle \;=\; \langle f|(\psi_p^\dagger\psi_n)(\psi_e^\dagger\psi_\nu)|i\rangle \tag{39}$$

La partie leptonique contient le terme:

$$\psi_e^\dagger\psi_\nu \sim \exp(-i\vec{p}_l \cdot \vec{r}) = 1 - i\vec{p}_l \cdot \vec{r} + ... \tag{40}$$

où $\vec{p}_l$ est l'impulsion emportée par les leptons:

$$\vec{p}_l = \vec{p}_e + \vec{p}_\nu \tag{41}$$

Il est facile de montrer que dans le volume d'intégration (le volume du noyau puisque c'est là que les quatre fermions subissent l'interaction de contact) les termes dans le développement de l'exponentielle décroissent rapidement. En effet, la longueur d'onde de de Broglie correspondant à p_l:

$$\lambda = \frac{\hbar c}{c p_l} \tag{42}$$

est beaucoup plus grande que le rayon d'un noyau.

En effet, dans une transition β typique, l'impulsion emportée par les leptons est de l'ordre que quelques MeV/c. Le rayon R d'un noyau est de l'ordre de quelques femtomètres (fm). D'autre part[1]:

$$\hbar c \sim 197.327 \text{ MeV-fm} \tag{43}$$

La quantité $p_l R$ est alors:

$$p_l R \sim \frac{R}{\lambda} \sim \frac{1}{200} \tag{44}$$

Pour les transitions permises on retiendra donc seulement le terme 1 dans le développement de l'exponentielle. Finalement l'élément de matrice se simplifie:

$$M_{if} = \langle f|\psi_p^\dagger \psi_n|i\rangle \tag{45}$$

On annihile un neutron dans l'état initial et on crée un proton dans l'état final. Dans le langage de l'isospin cette action équivaut à celle de l'opérateur τ_+ (qui transforme un neutron en un proton). D'où:

$$M_{if} = \langle f|\sum_j \tau_+^{(j)}|i\rangle \tag{46}$$

La somme sur j est étendue à tous les nucléons du noyau (les Z protons et les N neutrons).

On constate que cet élément de matrice est simplement le recouvrement des fonctions d'ondes spatiales des deux noyaux (initial et final) dès lors que la règle de sélection sur l'isospin est satisfaite:

$$Z_f = Z_i + 1 \qquad N_f = N_i - 1 \tag{47}$$

On caractérise cet élément de matrice nucléaire par "élément de matrice de Fermi".

Examinons maintenant les règles de sélection sur le moment angulaire et la parité, puisque ce sont de bons nombres quantiques pour des états nucléaires (en tout cas, cela est une bonne approximation). La partie spatiale d'un état nucléaire sera représentée par:

$$|\alpha JMP\rangle \tag{48}$$

où J est le moment angulaire du noyau, M sa composante suivant Oz, et P la parité. Les autres nombres quantiques, sans intérêt pour la présente discussion, sont désignés par α.

En fait, si l'on ne polarise pas les noyaux et si l'on ne détecte pas de polarisation, on devra sommer sur toutes les composantes M et on aura affaire à l'élément de matrice "réduit" tel que défini par le théorème de Wigner-Eckart:

$$\langle \alpha_f J_f P_f||T(k,P)||\alpha_i J_i P_i\rangle \tag{49}$$

où $T(k,P)$ est un opérateur d'ordre tensoriel k et de parité P. Les règles de sélection sur le moment angulaire et la parité sont alors:

$$|J_f - J_i| \leq k \leq J_f + J_i \quad \text{et} \quad P = P_f P_i \tag{50}$$

Mais revenons à l'élément de matrice de Fermi. Dans ce cas l'opérateur $T(k,P)$ est l'opérateur identité. C'est un opérateur scalaire ($k=0$) et de parité $P=+$. Les règles de sélection sont donc:

$$J_f = J_i \quad \text{et} \quad P_f = P_i \tag{51}$$

On définit:

$$\Delta J = |J_i - J_f| \quad \text{et} \quad \Delta P = P_i P_f \tag{52}$$

Les règles de sélection sont alors:

$$\Delta J = 0 \quad \text{et} \quad \Delta P = + \tag{53}$$

Ces règles de sélection sont appelées "règles de sélection de Fermi" et les transitions correspondantes sont appelées "transitions de Fermi".

La probabilité de transition $\lambda(E_e)dE_e$ par unité de temps qui découle de la théorie a été calculée par Fermi. C'est une probabilité différentielle. Son intégrale sur l'énergie

$$\lambda = \int_{m_e}^{W_0} \lambda(E_e)dE_e \tag{54}$$

est, dans le cas simple d'une transition entre un état initial i et un état final f, la constante de décroissance radioactive λ, définie par:

$$N(t) = N_0 \exp(-\lambda t) \tag{55}$$

où N_0 est le nombre de noyaux du type i présents au temps $t = 0$ et $N(t)$ le nombre de noyaux présents au temps t. On définit souvent la période radioactive T, le temps au bout duquel le nombre de noyaux a décru d'un facteur 2. On a:

$$T = \frac{\ln 2}{\lambda} \tag{56}$$

En physique des particules on définit plutôt la vie moyenne τ:

$$\tau = \frac{1}{\lambda} \tag{57}$$

En supposant que le neutrino a une masse nulle, la constante λ est donnée par:

$$\lambda(E_e)dE_e = \frac{G^2}{2\pi^3}|M_{if}|^2 F(Z, E_e) p_e E_e (W_0 - E_e)^2 dE_e \tag{58}$$

La constante des interactions faibles apparaît évidemment au carré. Plusieurs facteurs (2π) proviennent, entre autres, du terme $\rho(E_f)$. On retrouve l'élément de matrice nucléaire au carré. La fonction $F(Z, E_e)$ tient compte de l'interaction électromagnétique de l'électron avec le noyau final. Son calcul est possible analytiquement pour un potentiel coulombien pur (en $1/r$) et pour un noyau ponctuel. En réalité, le noyau a une certaine étendue, et la distribution de charge n'y est pas uniforme. À l'extérieur du noyau il faut tenir compte des électrons atomiques qui produisent un "effet d'écran". Sans compter plusieurs effets plus subtils sur lesquels nous reviendront. La fonction de Fermi devra donc être calculée numériquement sur ordinateur. Il en existe des tables.

La formule précédente nous donne le nombre d'électrons émis avec une énergie appartenant à l'intervalle $(E_e, E_e + dE_e)$ (spectre β). C'est une quantité mesurable expérimentalement avec un spectromètre β.

Il existe une méthode bien pratique pour analyser les résultats expérimentaux. Elle consiste à linéariser la formule du spectre β. On pose:

$$K(E_e) = \left[\frac{\lambda(E_e)}{F(Z, E_e)p_e E_e}\right]^{1/2} = C(W_0 - E_e) \tag{59}$$

où C est une constante. Le graphe de K en fonction de E_e est une droite qui coupe l'axe des énergies au point W_0. Cette méthode est due à Kurie (ne pas confondre avec Curie) et la droite correspondante est appelée "droite de Kurie". Elle est utilisée de deux façons. D'une part, on doit vérifier que les points expérimentaux sont bien alignés sur une droite. Ceci constitue un test de la théorie. D'autre part, le point d'intersection de cette droite avec l'axe des énergies donne la valeur de W_0, ou encore de la différence de masse $M_i - M_f$.

Évidemment, si la masse du neutrino n'est pas nulle, le graphe de K en fonction de E_e ne sera plus une droite. On s'attend à ce que la masse du neutrino soit très petite, ce qui a deux conséquences:

- le graphe de Kurie coupera l'axe des énergies au point

$$E_e = W_0 - m_\nu \tag{60}$$

- ce graphe sera encore pratiquement linéaire, sauf au voisinage de l'énergie maximale, où il s'incurvera vers le bas pour présenter, au point d'intersection avec l'axe des énergies, une tangente verticale.

****
Exercice: Étudier le comportement de la tangente au graphe de Kurie au voisinage du point d'énergie maximale. Examiner ce qui se passe pour $m_\nu = 0$.
****

Nous reviendrons plus tard sur la mesure de la forme des spectres β car elle peut, en principe, conduire à une mesure de la masse du neutrino.

5 La théorie est insuffisante

Si l'on en croit la théorie de Fermi, les transitions permises ne devraient se produire qu'entre des noyaux de même moment angulaire et de même parité ($\Delta J = 0$, $\Delta P = +$). Si des transitions peuvent se produire en violant les règles de sélection de Fermi, c'est parce que l'on va chercher des termes d'ordres supérieurs dans le développement de l'exponentielle, ou des opérateurs relativistes, ou les deux. Mais alors on obtient des constantes radioactives λ beaucoup plus petites. On a affaire à des transitions "interdites". En accumulant des facteurs d'interdiction on arrive ainsi à des transitions "une fois interdites", "deux fois interdites", etc... Ceci n'est pas sans rappeler ce qui se passe pour les transitions électromagnétiques, ou l'on observe aussi des degrés d'interdiction.

Il existe un moyen simple de reconnaître si une transition est permise. L'intégration de la formule du spectre donne:

$$\lambda = \frac{\ln 2}{T} = \frac{G^2}{2\pi^3}|M_{if}|^2 f(Z, W_0) \tag{61}$$

avec:

$$f(Z, W_0) = \int_{m_e}^{W_0} F(Z, E_e) p_e E_e (W_0 - E_e)^2 dE_e \tag{62}$$

****

<u>Exercice</u>: Calculez $f(Z, W_0)$ en supposant que la fonction de Fermi $F(Z, E_e)$ est constante et égale a 1. Ce résultat est une bonne approximation pour des transitions β entre des noyaux très légers.

****

On obtient le résultat:

$$f(Z, W_0)T = \frac{2\pi^3 \ln 2}{G^2|M_{if}|^2} \tag{63}$$

On voit que ce produit fT fournit une valeur de l'élément de matrice nucléaire. Pour des transitions permises on obtiendra la valeur maximale de fT. Effectivement, les valeurs expérimentales de fT se regroupent dans plusieurs régions, permettant de reconnaître les transitions permises, une fois interdites, etc... Il faut noter que la constante λ ne permet pas de comparaison valable car elle dépend de façon critique de l'énergie maximale W_0 (en gros comme W_0^5).

****

<u>Exercice</u>: Montrez que λ varie grossièrement comme W_0^5 (on pourra utiliser le résultat de l'exercice précédent).

****

Mais revenons à la théorie de Fermi. On ne tarda pas à découvrir une transition β permise (avec un fT caractéristique) violant les règles de sélection de Fermi. C'est la transition:

$$^6He \rightarrow \ ^6Li \ e^- \ \overline{\nu} \tag{64}$$

Le noyau ^6He a un spin $J = 0^+$ alors que le noyau ^6Li a un spin $J = 1^+$. Il s'agit donc d'une transition avec $\Delta J = 1$ et $\Delta P = +$.

Il faut en conclure que la théorie, sous sa forme primitive, est incomplète. Pour induire une transition permise avec $\Delta J = 1$ il faut un opérateur d'ordre tensoriel $k = 1$ et dont la parité est positive. On voit que l'opérateur $\vec{\sigma}$ fera l'affaire. Les règles de sélection pour un opérateur vectoriel sont:

$$|J_i - J_f| \leq k \leq J_i + J_f \quad \text{et} \quad P = P_i P_f \tag{65}$$

Ces règles de sélection sont appelées "règles de Gamow-Teller", et les transitions correspondantes sont appelées "transitions de Gamow-Teller". Les règles de sélection de Gamow-Teller autorisent les transitions suivantes:

$$\Delta J = 1 \qquad \Delta P = + \tag{66}$$

et:

$$\Delta J = 0 \qquad \Delta P = + \qquad \text{en excluant } (J=0) \to (J=0) \tag{67}$$

Les règles de sélection sont résumées dans la table 2.

ΔJ	ΔP	Type
1	+	Gamow-Teller pure
0 sauf $0 \to 0$	+	Mixte Fermi Gamow-Teller
$0 \to 0$	+	Fermi pure

Table 2: Règles de sélection pour les transitions β permises.

D'une manière générale, l'élément de matrice nucléaire d'une transition permise peut comporter deux termes, M_F et M_{GT}, avec:

$$M_F = \langle i | \sum_j \tau_+^{(j)} | f \rangle \tag{68}$$

$$M_{GT} = \langle i | \sum_j \vec{\sigma}^{(j)} \tau_+^{(j)} | \rangle \tag{69}$$

Ceci pour une transition β^-. Pour une transition β^+ il faut remplacer τ_+ par τ_-.

Comment introduire l'opérateur $\vec{\sigma}$? Cet opérateur est la limite non-relativiste de l'opérateur $\beta\vec{\sigma}$ qui se trouve dans l'interaction T. On le trouve aussi dans l'interaction A. Comme on ne sait pas laquelle de ces interactions est la bonne, ou si elles sont présentes toutes les deux, on va les retenir toutes les deux. Mais on note alors que l'opérateur identité peut s'obtenir comme limite non-relativiste de l'opérateur β, qui se trouve dans l'interaction scalaire. On doit donc retenir les interactions S, V, T et A pour engendrer les transitions permises. Quant à l'interaction P, rien ne prouve qu'on n'en aura pas besoin pour engendrer les transitions interdites. On va donc introduire les cinq types d'interaction:

$$H_- = G \sum_i g_i (\overline{\psi}_p O^i \psi_n)(\overline{\psi}_e O_i \psi_\nu) \tag{70}$$

avec $i = S, V, T, A, P$ et les opérateurs donnés dans la table 3.

Interaction	Constante de couplage	Opérateur O^i
Scalaire	g_S	I
Vectorielle	g_V	γ^μ
Tensorielle	g_T	$\sigma^{\mu\nu}$
Axiale	g_A	$\gamma^\mu\gamma^5$
Pseudoscalaire	g_P	γ^5

Table 3: Les différents couplages pour les interactions faibles.

Interaction	Constante de couplage	Opérateur A^i Non rel.	Opérateur A^i Rel.
Scalaire	g_S	β	0
Vectorielle	g_V	I	$i\vec{\alpha}$
Tensorielle	g_T	$\beta\vec{\sigma}$	$i\beta\vec{\alpha}$
Axiale	g_A	$\vec{\sigma}$	$i\gamma^5$
Pseudoscalaire	g_P	0	$\beta\gamma^5$

Table 4: Les différents couplages pour les interactions faibles.

Si l'on utilise une représentation particulière bien adaptée à la transition vers l'approximation non-relativiste, on écrira:

$$H_- = G\sum_i g_i(\psi_p^\dagger A^i\psi_n)(\psi_e^\dagger A_i\psi_\nu) \tag{71}$$

où les opérateurs A_i sont donnés dans la table 4, avec leur partie "non-relativiste" et leur partie "relativiste".

À ce stade on semble avoir perdu l'analogie avec les transitions électromagnétiques, qui ne font intervenir qu'une interaction V. Nous verrons plus loin comment on retrouve cette analogie.

Les constantes de couplages g_i devront être déterminées à partir de l'expérience. Malheureusement, c'est un travail ardu. Certaines observables ne disent rien sur les valeurs relatives des constantes de couplages. Par exemple, les interactions V et A se manifesteront dans les produits fT par la combinaison:

$$g_V^2 M_F^2 + g_A^2 M_{GT}^2 \tag{72}$$

Les interactions S et T donnent un résultat semblable:

$$g_S^2 M_F^2 + g_T^2 M_{GT}^2 \tag{73}$$

La forme des spectres β présente le même genre d'ambiguïté. Il y a cependant une observable qui permet, en principe, de trancher. C'est la corrélation angulaire électron-neutrino. On peut montrer que cette dernière est caractéristique du type d'interaction. Il s'agit de la probabilité de détecter un électron et un antineutrino emis à l'angle θ, en fonction de θ. Pour une transition permise cette corrélation s'écrit:

$$W(\theta) = N\left(1 + A\,\frac{v_e}{c}\cos\theta\right) \tag{74}$$

où v_e est la vitesse de l'électron et A un facteur de normalisation sans intérêt. Les valeurs de A sont données dans la table 5.

Interaction	$A =$
S	-1
V	$+1$
T	$-1/3$
A	$+1/3$

Table 5: Le coefficient de corrélation angulaire électron-neutrino, pour les différents couplages.

La forme générale du coefficient A est:

$$A = \left(\frac{1}{\xi}\right)\left\{\left[|g_V|^2 - |g_S|^2\right]|M_F|^2 + \frac{1}{3}\left[|g_T|^2 - |g_A|^2\right]|M_{GT}|^2\right\} \tag{75}$$

avec:

$$\xi = \left[|g_V|^2 + |g_S|^2\right]|M_F|^2 + \left[|g_T|^2 + |g_A|^2\right]|M_{GT}|^2 \tag{76}$$

Mais il est impossible de détecter des neutrinos d'aussi basse énergie. Il y a tout de même une façon de s'en tirer. Dans l'état final il y a trois corps: le noyau final, l'électron et l'antineutrino. Les trois impulsions ont une résultante nulle si le noyau initial était au repos dans le laboratoire:

$$\vec{p}_R = -(\vec{p}_\nu + \vec{p}_e) \tag{77}$$

Considérons l'interaction S. Elle favorise l'émission des leptons avec des impulsions antiparallèles, donc l'impulsion p_R de recul du noyau sera petite. Au contraire, l'interaction V favorise l'émission des leptons avec des impulsions parallèles. L'impulsion de recul sera grande. Les interactions T et A donneront des effets plus petits à cause du facteur $1/3$.

Nous avons déjà montré que les énergies de recul des noyaux sont très petites, donc le noyau final aura un parcours très petit dans la matière. L'expérience ne pourra donc se faire qu'en milieu gazeux. Il s'agit d'expériences très difficiles, mais elles avaient été faites dans les années cinquante.

En exploitant les résulats expérimentaux on était arrivé à la conclusion que l'interaction faible était une combinaison S et T, et que les constantes g_S et g_T étaient à peu près égales. On sait maintenant que cette conclusion est fausse.

6 La révolution de la parité

Il s'est passé, en 1956, des événements qui ont profondément modifié nos connaissances sur les interactions faibles, et qui ont permis des progrès considérables.

6.1 Le mystère des mésons θ et τ ($\theta - \tau$ puzzle)

On connaissait l'existence de deux mésons θ et τ de spin 0, qui avaient la même masse et la même vie moyenne. Pourquoi ces deux mésons n'étaient-ils pas une seule et même particule? Parce que leurs modes de désintégration sont différents?

$$\theta^+ \rightarrow \pi^+\pi^0 \tag{78}$$

$$\tau^+ \rightarrow \pi^+\pi^-\pi^+ \tag{79}$$

Mais on connaît de nombreuses particules qui ont plusieurs modes de désintégration. Pourquoi le θ et le τ ne seraient pas une seule et même particule avec deux modes de désintégration en deux pions et trois pions.

La difficulté vient du fait que, si l'on accepte la conservation de la parité dans les interactions faibles, le θ et le τ ont des parités intrinsèques différentes, et doivent donc être deux particules distinctes.

Etudions d'abord le système de deux pions. Ces particules sont des bosons de spin 0. Le θ ayant spin 0, le moment orbital relatif $\vec{l}$ des deux pions ne peut être que 0. En effet:

$$\vec{S}_\theta = \vec{S}_{\pi^+} + \vec{S}_{\pi_-} + \vec{l} \tag{80}$$

La fonction d'onde spatiale des deux pions a donc une parité $(-)^l$ positive. Les pions ont une parité intrinsèque négative. Le système des deux pions produit par la désintégration du θ a donc une parité positive:

$$P = (P_\pi)^2(-)^l = + \tag{81}$$

Si la parité est conservée par l'interaction faible, le θ a une parité positive.

Pour le système de trois pions c'est un peu plus compliqué. Il faut introduire deux moments orbitaux. Traditionnellement, on introduit un premier moment orbital $\vec{l}$ entre deux corps (par exemple les deux pions positifs), et un autre moment orbital $\vec{L}$ entre le centre de gravité des deux corps précédents et le troisième. On a alors:

$$\vec{S}_\theta = \vec{S}_{\pi^+} + \vec{S}_{\pi_-} + \vec{S}_{\pi^+} + \vec{l} + \vec{L} \tag{82}$$

Ce qui donne:

$$\vec{l} + \vec{L} = 0 \tag{83}$$

On a donc $l = L$ et la parité de la fonction d'onde spatiale $(-)^{l+L}$ est positive. La parité du système des trois pions produit par la désintégration du τ est donc négative:

$$P = (P_\pi)^3(-)^{l+L} = - \tag{84}$$

Si la parité est conservée par l'interaction faible, le τ a une parité négative.

Il est évidemment déplaisant d'avoir deux particules aussi semblables, mais de parités différentes, sans être capable d'expliquer pourquoi. L'idée d'une coïncidence n'est pas une explication satisfaisante. Les théoriciens ont retourné cette question dans tous les sens, sans arriver à une explication simple et logique. Deux théoriciens chinois, T.D. Lee et C.N. Yang, posèrent la question iconoclaste: tout ce raisonnement repose sur l'hypothèse de la conservation de la parité dans les interactions faibles. Et que savons nous exactement sur la conservation de la parité dans les interactions faibles? La réponse: on ne sait strictement rien; on a seulement des préjugés.

Lee et Yang proposèrent alors plusieurs expériences en vue de tester la conservation de la parité. C'est une chinoise, C.S. Wu, qui réalisa la première expérience qui démontra non seulement que la parité est violée dans les interaction faibles, mais aussi que cette violation est maximale. L'idée est simple: si la parité est conservée, les seules observables non nulles sont des quantités scalaires: une vie moyenne, un spectre d'énergie, une corrélation angulaire du type $\vec{p}_1 \cdot \vec{p}_2$, une corrélation spin-spin du type $\vec{S}_1 \cdot \vec{S}_2$, etc... Toute quantité pseudoscalaire du type $\vec{S} \cdot \vec{p}$ est rigoureusement nulle: la physique-image vue dans un miroir est identique à la physique-objet, toute quantité qui change de signe dans l'opération parité est donc formellement interdite. Par contre, si on abandonne la conservation de la parité, on admet que la physique-image vue dans un miroir n'existe tout simplement pas. On peut alors accepter l'existence d'observables pseudoscalaires.

Il est facile d'imaginer plusieurs observables pseudoscalaires du type $\vec{S} \cdot \vec{p}$:

- Dans l'état initial on a un ensemble de noyaux polarisés. Supposons que les spins pointent préférentiellement vers le haut. Cette direction préférentielle définit un pseudovecteur $\vec{S}_i$. On mesure les impulsions des électrons de désintégration. Supposons que celles-ci ont une direction préférentielle qui définit un vecteur $\vec{P}$. La quantité $\vec{S}_i \cdot \vec{P}$ est un pseudoscalaire. Si la parité est conservée il doit y avoir autant d'électrons allant vers le haut que vers le bas ($\vec{P} = 0$).

- On détecte les électrons émis dans une direction (par exemple Oz). Ceci définit un vecteur $\vec{P}$. On mesure alors la polarisation des noyaux dans l'état final. Celle-ci définit un pseudovecteur $\vec{S}_f$. La quantité $\vec{S}_f \cdot \vec{P}$ est un pseudoscalaire. Si la parité est conservée l'ensemble des noyaux qui ont émis des électrons suivant Oz n'est pas polarisé ($\vec{S}_f = 0$).

- L'hélicité des électrons (ou des positrons) est aussi un pseudoscalaire de la forme $\vec{S} \cdot \vec{p}$ Elle doit être nulle si la parité est conservée.

Ces considérations donnent lieu a plusieurs types d'expériences:

7 L'asymétrie des électrons de désintégration β

L'expérience de C.S. Wu consiste à polariser des noyaux de ^{60}Co et à observer les électrons de désintégration. En fait on ne déplace pas le détecteur, qui se trouve à l'intérieur du cryostat, mais on renverse la polarisation en changeant le sens du champ magnétique appliqué aux noyaux de cobalt. On enregistre les nombres d'électrons N^+ et N^- détectés alors que la polarisation des noyaux est dirigée suivant $+Oz$ ou $-Oz$. On forme le rapport:

$$\eta = \frac{N^+ - N^-}{N^+ + N^-} \tag{85}$$

La prédiction théorique pour la distribution angulaire des électrons est:

$$W(\theta) = 1 + \frac{v_e}{c} P_N A \cos\theta \tag{86}$$

où v_e est la vitesse de l'électron, P_N le taux de polarisation des noyaux de Co, et A un terme qui dépend de la structure des interactions faibles[2].

Le terme A doit être nul si la parité est conservée. Or l'expérience montre qu'il n'est pas nul, et même qu'il a la valeur maximale compatible avec la théorie. On dit que la violation de la parité est "maximale".

Comme conséquence de ce résultat expérimental il a fallu généraliser l'expression de la densite hamiltonienne. En prenant la désintégration β^- comme exemple:

$$H_- = G \sum_i g_i(\overline{\psi}_p O^i \psi_n)(\overline{\psi}_e O_i \psi_\nu) \tag{87}$$

avec $i = S, V, T, A, P$.

Cet hamiltonien contient des termes du type $S \cdot S$, $V \cdot V$, etc... C'est donc un scalaire. Toutes les observables calculées à partir de cet hamiltonien ne pourront être que des scalaires. Or on observe aussi des observables pseudoscalaires.

L'hamiltonien:

$$H'_- = G \sum_i g'_i(\overline{\psi}_p O^i \psi_n)(\overline{\psi}_e O_i \gamma^5 \psi_\nu) \tag{88}$$

contient des termes du type $S \cdot P$, $V \cdot A$, etc... C'est donc un pseudoscalaire. Cependant, toutes les observables calculées à partir de cet hamiltonien seront encore des scalaires, et seront identiques à celles calculées à partir de l'hamiltonien H_-. C'est parce que l'on calcule le carré de l'élément de matrice.

Par contre, si l'on mélange les hamiltoniens H_- et H'_- on pourra prédire l'existence d'observables scalaires et pseudoscalaires. On utilisera donc l'hamiltonien:

$$\mathcal{H}_- = H_- + H'_- = G \sum_i (\overline{\psi}_p O^i \psi_n) \left[(\overline{\psi}_e O_i (g_i + \gamma^5 g'_i) \psi_\nu) \right] \tag{89}$$

On obtient, pour l'asymétrie des électrons de désintégration β^- du ^{60}Co (transition de Gamow-Teller pure):

$$A = 2\frac{\mathcal{R}e(g_T g_T'^* - g_A g_A'^*)}{|g_T|^2 + |g_T'|^2 + |g_A|^2 + |g_A'|^2} \tag{90}$$

Même si l'expérience n'est pas très précise, elle suggère que A est compatible avec le valeur -1. Or, cette valeur ne peut être atteinte que si:

$$g_T = g_T' = 0 \quad \text{et} \quad g_A = g_A' \tag{91}$$

Ceci démontre que la violation de la parité est maximale, mais aussi que la violation de l'invariance C est maximale.

Cette expérience nécessite l'utilisation des techniques de polarisation, ce qui n'est pas à la portée de tout laboratoire.

8 La corrélation électron-polarisation circulaire γ

Une autre solution consiste à utiliser la polarisation dans l'état final, qui se manifeste par la polarisation des photons de désexcitation, si l'état final est un état excité, qui retourne à l'état fondamental par émission γ. Il n'est pas nécessaire de polariser les noyaux, mais par contre il faut pouvoir mesurer la polarisation des photon de désexcitation en coïncidence avec les électrons. Cette technique de corrélation β-polarisation circulaire γ a été utilisée dans de nombreux laboratoires. Le principe est le suivant: partant d'un ensemble de noyaux non polarisés dans l'état initial, la violation de la parité introduit une corrélation entre la direction des électrons et la polarisation des noyaux dans l'état final. En observant les électrons dans une certaine direction, on sélectionne un sous-ensemble de l'état final qui est polarisé. Si ces noyaux sont dans un état excité, et si la désexcitation par emission γ conduit à un état de spin zéro (donc non polarisé), on comprend que les rayons γ doivent emporter une certaine polarisation. Ainsi, les rayons γ qui sont émis dans la direction opposée aux électrons sont polarisés circulairement. Il reste donc à mesurer la polarisation circulaire des rayons γ en coïncidence avec les électrons.

Pour la mesure de la polarisation circulaire des photons, on utilise la diffusion Compton des photons sur des électrons orientés par un champ magnétique. La section efficace de diffusion Compton comporte un terme qui dépend de l'orientation relative des vecteurs "polarisation" du photon et de l'électron. Par exemple, en considérant seulement la polarisation circulaire du photon:

$$\sigma(E_\gamma, \theta) = \sigma_0(E_\gamma, \theta) + (\vec{P}_c^{(\gamma)} \cdot \vec{P}^e)\sigma_1(E_\gamma, \theta) \tag{92}$$

Les sections efficaces σ_0 (insensible à la polarisation) et σ_1 (sensible à la polarisation) dépendent de l'énergie du rayon γ et de l'angle de diffusion Compton.

Pour obtenir des électrons orientés (ici polarisés) on soumet un bloc de fer (ou autre matériau ferromagnétique) à un champ magnétique. On peut ainsi obtenir un taux de polarisation de l'ordre de 7%. On peut renverser le vecteur polarisation de l'électron en inversant le sens du courant de magnétisation. On mesure alors le rapport:

$$A = \frac{N_1 - N_2}{N_1 + N_2} \tag{93}$$

où N_1 et N_2 sont les nombres de photons observés pour les deux sens 1 et 2 du champ magnétique. On a:

$$A = \frac{\sigma_1}{\sigma_0} f P_c^{(\gamma)} \tag{94}$$

où f est la fraction des électrons orientés ($\sim$ 7%). La mesure de A détermine la polarisation circulaire du photon, les valeurs des sections efficaces σ_0 et σ_1 étant données par la théorie de l'effet Compton (électrodynamique quantique).

Il existe deux méthodes[2]:

- La méthode par transmission. Dans ce cas on mesure le nombre de photons transmis sans changement de direction, donc ceux qui n'ont pas subi de diffusion Compton.
- La méthode par diffusion. Dans ce cas on mesure le nombre de photons qui ont subi la diffusion Compton (vers l'avant ou vers l'arrière).

Le choix de la méthode expérimentale dépend de l'énergie des photons.

9 Mesure de l'hélicité des électrons (positrons).

On dispose de plusieurs méthodes:

- La diffusion Mott.
- La diffusion Møller (électrons) ou la diffusion Bhabha (positrons).
- La mesure de la polarisation circulaire du rayonnement de freinage.
- L'annihilation des positrons avec des électrons.

Références

[1] Particle Data Group, *"Review of Particle Properties"*, Phys. Rev. D54, 1996, page 65.

[2] H. Schopper, *"Weak interactions and Nuclear Beta Decay"*, North Holland, 1966.

10 Appendice 1: La polarisation (électrons, photons)

10.1 Introduction

On parle de plus en plus de la *physique du spin.* Il s'agit de la physique qui est faite avec des faisceaux polarisés et des cibles polarisées. Pourquoi?

Les interactions fortes et électrofaibles sont sensibles aux variables de spin des particules. L'utilisation de faisceaux polarisés et de cibles polarisées augmente considérablement le nombre d'expériences possibles et permet de mieux préciser les caractéristiques de ces interactions. La physique du spin est particulièrement importante en ce qui concerne la vérification des propriétés de symétrie des interactions. Par exemple, la violation de la parité dans les interactions faibles a été découverte lors d'expériences avec des particules polarisées. En physique des hautes énergies il a été surprenant de constater que les variables de spin continuent à jouer un rôle important.

Du côté technique, beaucoup d'efforts ont été déployés pour obtenir des faisceaux polarisés et des cibles polarisées.

Nous allons considérer ici la polarisation des électrons et des photons car c'est le même formalisme mathématique qui peut être utilisé dans les deux cas. Il s'agit en effet de particules possédant deux états de polarisation. On note que l'électron est une particule de spin 1/2 alors que le photon est plutôt une particule de spin 1 (il faut être prudent avec la notion de spin pour le photon).

10.2 Position du problème

Un cible polarisée contient environ 10^{23} particules (de l'ordre du nombre d'Avogadro). Un faisceau de particules peut transporter par seconde des microampères (10^{13} particules par seconde) ou même davantage. Une description microscopique d'un tel système exigerait l'utilisation d'un nombre énorme de variables, ce qui est impossible. Ce n'est pas non plus nécessaire car lors d'une expérience on mesure toujours des valeurs moyennes. La question est donc la suivante: combien de paramètres sont nécessaires pour décrire l'état de polarisation d'un ensemble d'électrons ou de photons? La réponse: on a seulement besoin de trois paramètres qui peuvent être identifiés avec les trois composantes d'un vecteur réel, le *vecteur de Stokes* (ce vecteur avait été introduit au siècle dernier par Stokes pour décrire la polarisation de la lumière). Un formalisme équivalent est celui de la *matrice densité.*

10.3 Matrice densité et vecteur de Stokes

Commençons par considérer le système propre de la particule (le système du laboratoire si la particule y est au repos).

10.3.1 Définition

Pour une particule ayant deux états de polarisation, l'état de polarisation le plus général est représenté comme une combinaison linéaire de deux états de base:

$$|\pi\rangle = \sum_{j=1}^{2} c_j |j\rangle \qquad\qquad \sum_{j=1}^{2} |c_j|^2 = 1 \tag{95}$$

Considérons une observable A. Sa valeur moyenne sur un état de polarisation (pour une seule particule) sera:

$$\langle\pi|A|\pi\rangle = \sum_{j=1}^{2}\sum_{k=1}^{2} c_j c_k^* \langle k|A|j\rangle = \sum_{j=1}^{2}\sum_{k=1}^{2} c_j c_k^* A_{kj} \tag{96}$$

Ce qui peut être mesuré expérimentalement, c'est la valeur moyenne de $\langle\pi|A|\pi\rangle$ sur l'ensemble, soit:

$$\bar{A} = \frac{1}{N}\sum_{n=1}^{N}\langle\pi^{(n)}|A|\pi^{(n)}\rangle = \frac{1}{N}\sum_{n=1}^{N}\sum_{j=1}^{2}\sum_{k=1}^{2} c_j^{(n)} c_k^{(n)*} A_{kj} = \sum_{j=1}^{2}\sum_{k=1}^{2}\frac{1}{N}\sum_{n=1}^{N} c_j^{(n)} c_k^{(n)*} A_{kj} \tag{97}$$

où n sert à numéroter les N particules.

Définissons la matrice densité ρ:

$$\rho_{jk} = \overline{c_j c_k^*} \tag{98}$$

avec:

$$\overline{c_j c_k^*} = \frac{1}{N}\sum_{n=1}^{N} c_j^{(n)} c_k^{(n)*} \tag{99}$$

Il en résulte que:

$$\bar{A} = \sum_{j=1}^{2}\sum_{k=1}^{2} \rho_{jk} A_{kj} = \mathrm{Trace}\,[\rho A] \tag{100}$$

La matrice ρ décrit complétement l'ensemble puisqu'elle permet de calculer la valeur moyenne de n'importe quelle observable.

10.3.2 Propriétés de la matrice densité

On peut écrire la matrice ρ sous la forme:

$$\rho = \begin{pmatrix} \overline{c_1 c_1^*} & \overline{c_1 c_2^*} \\ \overline{c_2 c_1^*} & \overline{c_2 c_2^*} \end{pmatrix} \tag{101}$$

Calculons la trace:

$$\text{Trace}\,[\rho] = \overline{|c_1|^2 + |c_2|^2} = 1 \tag{102}$$

Comme toute matrice 2×2 la matrice ρ peut s'écrire:

$$\rho = aI + \sum_{\nu=1}^{3} b_\nu \sigma_\nu \tag{103}$$

où les σ_ν sont les matrices de Pauli et I la matrice unité. On peut calculer a et les b_ν:

$$\text{Trace}\,[\rho] = 2a = 1 \tag{104}$$

$$\text{Trace}\,[\rho\sigma_\lambda] = a\text{Trace}\,[\sigma_\lambda] + \sum_{\nu=1}^{3} b_\nu \text{Trace}\,[\sigma_\nu\sigma_\lambda] = \sum_{\nu=1}^{3} b_\nu \text{Trace}\,[\delta_{\nu\lambda} + i\epsilon_{\nu\lambda\tau}\sigma_\tau] = 2b_\lambda \tag{105}$$

On peut finalement mettre la matrice ρ sous la forme:

$$\rho = \frac{1}{2}(1 + \vec{\xi} \cdot \vec{\sigma}) \tag{106}$$

Le vecteur $\vec{\xi}$ (vecteur de Stokes) a pour composantes:

$$\xi_x = \overline{c_1 c_2^*} + \overline{c_2 c_1^*} = 2\mathcal{R}e\left[\overline{c_1 c_2^*}\right] \tag{107}$$

$$\xi_y = i(\overline{c_1 c_2^*} - \overline{c_2 c_1^*}) = 2\mathcal{I}m\left[\overline{c_2 c_1^*}\right] \tag{108}$$

$$\xi_z = \overline{c_1 c_1^*} - \overline{c_2 c_2^*} = \overline{|c_1|^2} - \overline{|c_2|^2} \tag{109}$$

Toutes ces composantes sont des nombres réels. Calculons le module carré de ξ (il est évidemment positif):

$$\xi^2 = \xi_x^2 + \xi_y^2 + \xi_z^2 = 1 - 4\left[\overline{c_1 c_1^*} \cdot \overline{c_2 c_2^*} - \overline{c_1 c_2^*} \cdot \overline{c_2 c_1^*}\right] \tag{110}$$

Les quantités $\overline{c_j c_k^*}$ peuvent être considérées (à un facteur N près) comme des produits scalaires dans un espace à N dimensions et dont les composantes sont les $c_j^{(n)}$ (et complexes conjugués). En invoquant le lemme de Schwartz on peut affirmer que:

$$\overline{c_1 c_1^*} \cdot \overline{c_2 c_2^*} \geq \overline{c_1 c_2^*} \cdot \overline{c_2 c_1^*} \tag{111}$$

Il en résulte que:

$$0 \leq \xi \leq 1 \tag{112}$$

Calculons la Trace de ρ^2:

$$\text{Trace}\left[\rho^2\right] = \frac{1}{4}\text{Trace}\left[1 + 2\vec{\xi} \cdot \vec{\sigma} + \xi^2\right] = \frac{1}{2}(1 + \xi^2) \tag{113}$$

On a donc:

$$\frac{1}{2} \leq \text{Trace}\left[\rho^2\right] \leq 1 \tag{114}$$

10.3.3 Ensemble non-polarisé

Supposons que les modules et phases des c_j soient choisis au hasard (en respectant la condition de normalisation). Si l'on pose:

$$c_j = \gamma_j e^{i\phi_j} \tag{115}$$

on trouve:

$$\xi_x = 2\ \overline{\gamma_1\gamma_2 cos(\phi_2 - \phi_1)} \tag{116}$$

$$\xi_y = 2\ \overline{\gamma_1\gamma_2 sin(\phi_2 - \phi_1)} \tag{117}$$

$$\xi_z = \overline{\gamma_1^2 - \gamma_2^2} \tag{118}$$

Avec des phases et modules choisis au hasard on voit que les composantes du vecteur de Stokes sont toutes nulles. Un ensemble non-polarisé est donc caractérisé par $\vec{\xi} = 0$.

10.3.4 Ensemble complétement polarisé (ou état pur)

Supposons que toutes les particules soient dans le même état de polarisation. Elles sont caractérisées par les mêmes valeurs des c_j. On a alors:

$$\overline{c_j c_k^*} = c_j c_k^* \tag{119}$$

et:

$$\overline{c_1 c_1^*} \cdot \overline{c_2 c_2^*} - \overline{c_1 c_2^*} \cdot \overline{c_2 c_1^*} = c_1 c_1^* c_2 c_2^* - c_1 c_2^* c_2 c_1^* = 0 \tag{120}$$

d'où:

$$\xi = 1 \tag{121}$$

Le module de $\vec{\xi}$ est appelé *degré de polarisation* P. Il varie de 0 (état non polarisé) à 1 (état complétement polarisé). Les ensembles avec $0 \leq \xi \leq 1$ sont *partiellement* polarisés. La direction du vecteur $\vec{\xi}$ définit la nature de l'état de polarisation.

10.4 Représentation des états de polarisation

On peut utiliser une sphère de rayon unité pour représenter les états de polarisation. Évidemment, le centre de la sphère représente un état non polarisé, alors que la surface de la sphère représente les états complétement polarisés. À l'intérieur de la sphère on trouve les états de polarisation partielle.

Supposons maintenant que les états de base sont définis de la manière suivante (pour la lumière):

$|1\rangle$ est un état de polarisation circulaire droite: $c_1 = 1$, $c_2 = 0$, $\xi_x = \xi_y = 0$, $\xi_z = +1$.
$|2\rangle$ est un état de polarisation circulaire gauche: $c_1 = 0$, $c_2 = 1$, $\xi_x = \xi_y = 0$, $\xi_z = -1$.

Les deux pôles correspondent à des états de polarisation circulaire, droite (pôle nord) et gauche (pôle sud).

Considérons les états:

$$c_1 = +1/\sqrt{2} \qquad c_2 = +1/\sqrt{2} \qquad \xi_x = +1 \qquad \xi_y = \xi_z = 0.$$

$$c_1 = +1/\sqrt{2} \qquad c_2 = -1/\sqrt{2} \qquad \xi_x = -1 \qquad \xi_y = \xi_z = 0.$$

Ce sont des états de polarisation *linéaire* horizontale et verticale, respectivement.

Enfin, les états:

$$c_1 = +1/\sqrt{2} \qquad c_2 = +i/\sqrt{2} \qquad \xi_y = +1 \qquad \xi_x = \xi_z = 0.$$
$$c_1 = +1/\sqrt{2} \qquad c_2 = -1/\sqrt{2} \qquad \xi_y = -1 \qquad \xi_x = \xi_z = 0.$$

Ce sont des états de polarisation *linéaire* inclinés de 45^o et 135^o par rapport à l'horizontale.

On voit donc que le plan équatorial correspond à des états de polarisation linéaire partielle ou complète (sur l'équateur). Les états situés dans les deux hémisphères correspondent à des états de polarisation *elliptique* partielle ou complète (à la surface de la sphère). Évidemment un état de polarisation circulaire ou linéaire peut toujours être considéré comme un cas particulier de polarisation elliptique.

10.5 Quelques théorèmes

Mettons le vecteur de Stokes sous la forme:

$$\vec{\xi} = P\vec{n} \tag{122}$$

où $\vec{n}$ est un vecteur unitaire définissant l'état de polarisation et P le degré de polarisation. On peut écrire:

$$\rho = \frac{1}{2}(1 + P\vec{n}\cdot\vec{\sigma}) = (1-P)\frac{1}{2} + P\frac{1}{2}(1 + \vec{n}\cdot\vec{\sigma}) \tag{123}$$

Le premier terme représente un ensemble non-polarisé de population $(1-P)$ et le second un ensemble complétement polarisé de population P caractérisé par le vecteur $\vec{n}$. D'où le théorème: on peut obtenir un ensemble partiellement polarisé de degré P en mélangeant un ensemble non-polarisé et un ensemble complétement polarisé dans les proportions $(1-P)$ et P.

On peut aussi écrire:

$$\rho = \frac{1}{2}(1 + P\vec{n}\cdot\vec{\sigma}) = \frac{1-P}{2}\,\frac{1}{2}(1 - \vec{n}\cdot\vec{\sigma}) + \frac{1+P}{2}\,\frac{1}{2}(1 + \vec{n}\cdot\vec{\sigma}) \tag{124}$$

Les deux termes représentent des états complétement polarisés, mais avec des polarisations opposées. D'où le thèorème: on peut obtenir un ensemble partiellement polarisé de degré P

en mélangeant deux ensembles de polarisations opposées dans les proportions $(1-P)/2$ et $(1+P)/2$.

Comme conséquence, un ensemble non-polarisé peut être obtenu en mélangeant en proportions égales deux ensembles de polarisations opposées:

$$\rho = \frac{1}{2} = \frac{1}{2}\,\frac{1}{2}(1 - \vec{n}\cdot\vec{\sigma}) + \frac{1}{2}\,\frac{1}{2}(1 + \vec{n}\cdot\vec{\sigma}) \tag{125}$$

Par exemple, en mélangeant en proportions égales des photons de polarisations circulaires droite et gauche on obtient un faisceau non-polarisé.

Il faut se rappeler que ces théorèmes ne s'appliquent qu'à des particules dotées de deux états de polarisation seulement.

Remis à jour le 5 janvier 1999.