Séminaire de Théorie spectrale et géométrie

Luc Rozoy

De $E = mc^2$ à l'équation de Dirac : une introduction heuristique aux spineurs

Séminaire de Théorie spectrale et géométrie, tome 14 (1995-1996), p. 121-148
<http://www.numdam.org/item?id=TSG_1995-1996__14__121_0>

© Séminaire de Théorie spectrale et géométrie (Grenoble), 1995-1996, tous droits réservés.

L'accès aux archives de la revue « Séminaire de Théorie spectrale et géométrie » implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

Numdam
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire de théorie spectrale et géométrie
GRENOBLE
1995–1996 (121–148)

DE $E = mc^2$ À L'ÉQUATION DE DIRAC : UNE INTRODUCTION HEURISTIQUE AUX SPINEURS

Luc ROZOY

But : une leçon d'histoire (essai)

I. En mécanique classique

La variété utilisée est $V = \mathbb{R}^3 \times \mathbb{R}$. On pose en axiome qu'il existe des repères privilégiés, dit Galiléens, dans lesquels, pour une particule $\overrightarrow{F} = m\overrightarrow{\gamma}$ est vraie (on postule la donnée de $\begin{array}{l}\mathbb{R} \mapsto \mathbb{R}^3 \\ t \mapsto \overrightarrow{F}(t)\end{array}$ et on en déduit la trajectoire d'une particule de masse m, constante, située en M grâce à $m\frac{d^2\overrightarrow{OM}}{dt^2} = \overrightarrow{F}(t)$). Les repères galiléens se déduisent d'un seul repère galiléen par le groupe des translations rectilignes uniformes, ou groupe de Galilée. (Un élément générique du groupe de Galilée à (t, x, y, z) de $\mathbb{R}^4$ fait correspondre (t', x', y', z') de $\mathbb{R}^4$ avec $t' = t + a$ et

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = Q\left[\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + t\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}\right]$$

où Q est une matrice orthogonale $\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$ et $\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$ deux vecteurs de $\mathbb{R}^3$). Une autre manière de présenter cette démarche consiste à imposer que les seuls changements de cartes possibles de la variété soient ceux définis par les translations rectilignes uniformes. Cette structure de variété affine très spéciale est sous jacente au contenu physique de la théorie : nous excluons volontairement tous les autres changements de cartes possibles. Alors

---

Classification math. : 01A60, 81-03, 83-03.

l'équation $\overrightarrow{F} = m\overrightarrow{\gamma}$ est invariante, reste la même, quand on change de repère, quand on change de cartes.

Résumé. — $\overrightarrow{F} = m\overrightarrow{\gamma}$ est invariante par le groupe de Galilée.

II. Les lois de l'électricité et du magnétisme classique

Elles reviennent à introduire dans le cadre de la mécanique classique, deux vecteurs de l'espace $\mathbb{R}^3$, dépendant du temps, $\overrightarrow{E}$ le vecteur champ électrique et $\overrightarrow{B}$ le vecteur champ magnétique, astreints à vérifier

$$\left\{\begin{array}{r} \overrightarrow{\mathrm{rot}\,\overrightarrow{B}} - \dfrac{1}{c}\dfrac{\partial \overrightarrow{E}}{\partial t} = 4\pi\mu\dfrac{\overrightarrow{V}}{c} \\ \mathrm{div}\,\overrightarrow{E} = 4\pi\mu \end{array}\right\} \quad 1^{\mathrm{er}} \text{ groupe des équations de Maxwell}$$

$$\left\{\begin{array}{r} \overrightarrow{\mathrm{rot}\,\overrightarrow{E}} + \dfrac{1}{c}\dfrac{\partial \overrightarrow{H}}{\partial t} = 0 \\ \mathrm{div}\,\overrightarrow{B} = 0 \end{array}\right\} \quad 2^{\mathrm{e}} \text{ groupe des équations de Maxwell}$$

(μ est la densité (volumique) de charge et $\overrightarrow{V}$ le vecteur vitesse d'espace de la charge, la quantité d'électricité occupant le volume élémentaire de V_0 étant μdV_0, pour le vide faire $\mu = 0$).

En introduisant sur la variété $\mathbb{R}^3 \times \mathbb{R}$ de la mécanique classique, dans un repère galiléen orthonormé, la forme quadratique $ds^2 = c^2dt^2 - (dx^2 + dy^2 + dz^2)$ il est remarquable que ces équations deviennent (pour le vide par exemple)

$$\left\{\begin{array}{ll} \delta F = 0 & 1^{\mathrm{er}} \text{ groupe des équations de Maxwell} \\ dF = 0 & 2^{\mathrm{e}} \text{ groupe des équations de Maxwell} \end{array}\right\}$$

si F est la 2-forme définie par la matrice suivante dans un repère galiléen orthonormé direct :

$$\begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & +B_y & -B_x & 0 \end{pmatrix}.$$

d ne dépend que de la structure différentiable de la variété mais δ fait intervenir la structure lorentzienne plate $ds^2 = c^2dt^2 - (dx^2 + dy^2 + dz^2)$ introduite.

(La variété n'est pas compacte et la métrique non définie positive. Ces deux raisons contraignent à l'abandon de la définition de δ par dualité avec d à travers le produit scalaire et à poser

$$\delta = (-1)^p *^{-1} d * \quad \text{si } \delta \text{ agit sur les } p \text{ formes}$$

où $*^{-1} = (-1)^{(\text{signe de la pseudo-métrique})}(-1)^{p(n-p)}*$ si $*^{-1}$ agit sur une p-forme et où $*$ est l'adjonction de Hodge définie grâce à la forme volume.)

Conséquence. — Les équations de Maxwell sont invariantes par le groupe de Lorentz L ensemble de toutes les transformations linéaires de $\mathbb{R}^4$ qui laissent invariante la forme bilinéaire

$$xy = x^0y^0 - x^1y^1 - x^2y^2 - x^3y^3 = (x^0x^1x^2x^3)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}.$$

Ce groupe *n'est pas* dans le groupe de Galilée. Cela conduit à un conflit entre la mécanique classique et les équations de Maxwell. Si les deux théories sont vraies simultanément les ondes électromagnétiques devraient particulariser des repères galiléens spéciaux. Il devrait exister "un vent d'éther". Les expériences de Michelson et Morley montrent que cela n'est pas le cas. Il fallait ou bien modifier les équations de Maxwell (et cela conduit à un échec) ou bien modifier la cinématique de la mécanique classique. C'est l'origine de la relativité restreinte.

III. La relativité restreinte

Premier axiome.

La variété utilisée est réelle de dimension 4. *Les changements de cartes sont restreints au groupe de Poincaré* (cf. ci-dessous). La variété est munie d'une structure pseudo-riemannienne plate. (Dans chaque carte la restriction sur les changements de cartes conduit à ce que la structure pseudo-riemannienne puisse s'écrire :

$$ds^2 = c^2dt^2 - (dx^2 + dy^2 + dz^2) = (dx^0)^2 - \big((dx^1)^2 + (dx^2)^2 + (dx^3)^2\big).$$

Définition du groupe de Lorentz L (ou groupe de Lorentz étendu). — C'est le sous-groupe de $GL(\mathbb{R}^4)$ qui laisse invariant la forme $xy = x^0y^0 - x^1y^1 - x^2y^2 - x^3y^3$.

$L_+^\uparrow$ est le groupe de Lorentz restreint, (c'est la composante connexe de l'identité de $SO(3,1)$).

$$\Lambda \in L \text{ (groupe de Lorentz)} \iff {}^t\Lambda G\Lambda = G \text{ avec } G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

${}^t\Lambda G\Lambda = G$ conduit à $(\Lambda_0^0)^2 = 1 + (\Lambda_0^1)^2 + (\Lambda_0^2)^2 + (\Lambda_0^3)^2$ si Λ_b^a désigne l'élément de Λ qui se trouve sur la a^{ième} ligne et la b^{ième} colonne. Voici quelques notations classiques :

$$\begin{aligned}
L_+^\uparrow &= \{\Lambda \in L, \det\Lambda = 1 \quad \Lambda_0^0 \geq 1\}.\\
L_+^\downarrow &= \{\Lambda \in L, \det\Lambda = 1 \quad \Lambda_0^0 \leq -1\}.\\
L_-^\uparrow &= \{\Lambda \in L, \det\Lambda = -1 \quad \Lambda_0^0 \geq 1\}.\\
L_-^\downarrow &= \{\Lambda \in L, \det\Lambda = -1 \quad \Lambda_0^0 \leq -1\}.
\end{aligned}$$

Définition du groupe de Poincaré (étendu). — Ensemble des transformations de $\mathbb{R}^4$ dans $\mathbb{R}^4$ de la forme $x' = \Lambda x + a$ où $\Lambda \in L$ et $a \in \mathbb{R}^4$.

$SL(2,\mathbb{C})$ est le revêtement universel de $L_+^\uparrow$.

Une construction de ce revêtement universel peut se faire de la manière suivante qui sera utilisée systématiquement par la suite. (Aucune construction n'est plus canonique qu'une autre. Il en fallait bien une).

1re étape : une bijection entre $\mathbb{R}^4$ et les matrices hermitiennes $(2,2)$. Toute matrice hermitienne $(2,2)$ se décompose sur la base

$$\tau_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tau_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$$

(matrices de Pauli). À tout $x \in \mathbb{R}^4 \quad x = (x^0, x^1, x^2, x^3)$ on associe la matrice hermitienne $(2,2)$ $\underset{\sim}{x} = \sum_{\mu=0}^{3} x^\mu \tau_\mu$. À toute matrice hermitienne $(2,2)$ $\underset{\sim}{x}$ on associe x de $\mathbb{R}^4$ par $x^\lambda = \frac{1}{2}\,\mathrm{Trace}(\tau_\lambda \underset{\sim}{x})$. Se trouvent ainsi réalisées une bijection et sa réciproque entre $\mathbb{R}^4$ et les matrices hermitiennes $(2,2)$.

2e étape : Soit $A \in SL(2,\mathbb{C})$, considérons l'application linéaire de $\mathbb{R}^4$ dans $\mathbb{R}^4$ définie par :

$$\mathbb{R}^4 \longrightarrow \{\text{ matrices hermitiennes } (2,2)\} \longrightarrow \{\text{ matrices hermitiennes } (2,2)\} \longrightarrow \mathbb{R}^4$$

$$x \longrightarrow \underset{\sim}{x} = \sum_{\mu=0}^{3} x^\mu \tau_\mu \longrightarrow \underset{\sim}{y} = A\,\underset{\sim}{x}\,A^* \longrightarrow y.$$

Le résultat intéressant est que cette application linéaire appartient à $L_+^\uparrow$ et que cette construction constitue un revêtement à 2 feuillets (universel) de $L_+^\uparrow$. (Preuve faite dans les exposés de Hubert Pesce).

Théorème. — *Toute matrice A de $SL(2,\mathbb{C})$ s'écrit $A = VH$ ou V est unitaire et H*

hermitienne avec :

$$H = (\pm)e^{-\frac{\varphi}{2}\sum_{i=1}^{3} n^i \tau_i} \qquad \vec{n} = \sum_{i=1}^{3} n^i \vec{e}_i \in \mathbb{R}^3 \text{ avec } \|\vec{n}\| = 1,$$

$$V = (\pm)e^{i\frac{\theta}{2}\sum_{j=1}^{3} m^j \tau_j} \qquad \vec{m} = \sum_{j=1}^{3} m^j \vec{e}_i \in \mathbb{R}^3 \text{ avec } \|\vec{m}\| = 1.$$

Cette décomposition polaire de toute matrice de $SL(2,\mathbb{C})$ et le revêtement présenté de $L_+^\uparrow$ par $SL(2,\mathbb{C})$ nous donne une décomposition des éléments de $L_+^\uparrow$.

Si A de $SL(2,\mathbb{C})$ s'écrit $A = V$, la transformation Λ de $L_+^\uparrow$ est liée à une rotation des axes de coordonnées $\vec{e}_1, \vec{e}_2, \vec{e}_3$ autour de $\vec{m}$ d'un angle θ. On retrouve que $SU_2(\mathbb{C})$ est le revêtement universel de $SO_3(\mathbb{R})$.

Si A de $SL(2,\mathbb{C})$ s'écrit $A = H$, il nous faut interpréter l'élément Λ de $L_+^\uparrow$ obtenu. Prenons $\vec{n} = \vec{e}_1$. On obtient $\Lambda = \begin{pmatrix} ch\varphi & -sh\varphi & 0 & 0 \\ -sh\varphi & ch\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

En posant $th\varphi = \frac{v}{c}$ et en utilisant $ch\varphi = \frac{1}{\sqrt{1-th^2\varphi}}$, $sh\varphi = \frac{th\varphi}{\sqrt{1-th^2\varphi}}$ l'élément de Λ de $L_+^\uparrow$ fait passer d'une première carte où les coordonnées sont (t, x, y, z) à une deuxième carte où les coordonnées sont (t', x', y', z') avec les relations :

$$x' = \frac{x - vt}{\sqrt{1-\frac{v^2}{c^2}}}; \quad y' = y; \quad z' = z; \quad t' = \frac{t - x\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Physiquement les deux cartes représentent les référentiels des espaces et des temps de deux observateurs différents. Le premier utilise (t, x, y, z) dans le référentiel R et le deuxième (t', x', y', z') dans le référentiel R'. (Quand $c = +\infty$ on retrouve une transformation de Galilée associée à une translation rectiligne uniforme de vitesse v le long de l'axe x).

Dans R' soit M une particule de masse m' se déplaçant à la vitesse u' le long de l'axe y', donc $x' = 0$

$$y'_M(t') = u't' = y = u'\frac{t - x\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} = u'\frac{t - vt\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} = u'\sqrt{1-\frac{v^2}{c^2}}\,t.$$

La vitesse u' dans R' devient la vitesse $u = u'\sqrt{1-\frac{v^2}{c^2}}$ dans R. Comme dans chacun des repères R et R' globalement les axes y et y' sont les mêmes, on a envie que la quantité de mouvement soit conservée le long de y (nous sommes en train d'essayer de reconstruire une cinématique par "induction"). Alors il faut que $m'u' = mu = mu'\sqrt{1-\frac{v^2}{c^2}}$

donc que $m' = m\sqrt{1-\frac{v^2}{c^2}}$ si m' est la masse dans le repère R' et m la masse dans le repère R. Cette formule n'est pas symétrique entre les repères R et R'. Nous sommes alors conduit à postuler que si la particule M est au repos dans un repère R_0 qui appartient à l'atlas de l'axiome 1 et que sa masse y est m_0 alors dans un repère R qui se déduit de R_0 par une transformation de Lorentz (correspondant à une $\pm A$ de $SL(2,\mathbb{C})$ avec $A = VH$ où H hermitienne utilise $th\varphi = \frac{v}{c}$) sa masse sera donnée par $m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$. (Il faut bien postuler cette loi, *elle n'est pas incluse dans l'axiome 1*). Quitte à faire des rotations dans R et R_0 nous sommes ramenés à

$$x = \frac{x_0 - vt}{\sqrt{1-\frac{v^2}{c^2}}}\,; \quad y = y_0\,; \quad z = z_0\,; \quad t = \frac{t_0 - x_0\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Alors $m_0^2 = m^2(1-\frac{v^2}{c^2})$ donc $m^2c^2 - m^2v^2 = m_0^2c^2$ donc $c^2m\frac{dm}{dt} = mv\frac{d}{dt}(mv)$. Comparons cette équation au calcul suivant en mécanique classique $\frac{d}{dt}(\frac{1}{2}mv^2) = \frac{dE}{dt} = m\vec{v}\frac{d}{dt}(\vec{v})$ où $m = C^{\text{te}}$. D'où $m\vec{v}\frac{d}{dt}(m\vec{v}) = m\frac{dE}{dt}$.

$$\left\{\begin{array}{rl} \text{Mécanique classique} & m\vec{v}\dfrac{d}{dt}(m\vec{v}) = m\dfrac{dE}{dt} \\ \text{Ici} & m\vec{v}\dfrac{d}{dt}(m\vec{v}) = c^2m\dfrac{dm}{dt} \end{array}\right\}.$$

Il est donc naturel de poser $\frac{dE}{dt} = c^2\frac{dm}{dt}$ d'où $E = mc^2 + C^{\text{te}}$. Le choix de la constante n'est pas indifférent.

En mécanique classique $\left(\frac{d}{dt}(mv^1), \frac{d}{dt}(mv^2), \frac{d}{dt}(mv^3)\right) = (f^1, f^2, f^3)$ et $\frac{dE}{dt} = \vec{F}\cdot\vec{V}$ sont des équations qui n'ont pas un lien direct avec la géométrie de la variété utilisée. Ici si nous choisissons la constante d'intégration de manière à ce que $E = mc^2$ alors en posant $p_1 = \frac{d}{dt}(mv^1)$, $p_2 = \frac{d}{dt}(mv^2)$, $p_3 = \frac{d}{dt}(mv^3)$, ces équations se rassemblent en $\frac{d}{dt}(\frac{E}{c}, p_1, p_2, p_3) = F$ où F est un champ de (quadri)vecteur de la variété lorentzienne de l'axiome 1 et le (quadri)vecteur $(\frac{E}{c}, p_1, p_2, p_3)$ a alors la propriété remarquable que sa pseudo-norme pour la "métrique" lorentzienne est constante puisque alors :

$$\frac{d}{dt}\Big(E^2 - c^2(p_1^2+p_2^2+p_3^2)\Big) = 2\Big(E\frac{dE}{dt} - c^2\sum_{i=1}^{3} p_i\frac{dp_i}{dt}\Big)$$

$$= \frac{d}{dt}(E^2 - m^2\sum_{i=1}^{3} v_i^2) = \frac{d}{dt}(E^2 - m^2v^2) = \frac{d}{dt}(m^2c^4 - m^2v^2)$$

$$= \frac{d}{dt}\Big(m^2c^4(1-\frac{v^2}{c^2})\Big) = \frac{d}{dt}(m_0^2c^4) = 0$$

et donc $E^2 - c^2p^2 = m_0^2c^4 = c^2\Big((\frac{E}{c})^2 - p^2\Big)$.

Axiomes suivants et commentaires sur le contenu physique de la théorie.

On reconstruit la cinématique de manière à ce que pour la "métrique" lorentzienne et pour une particule matérielle $(\frac{E}{c}, p_1, p_2, p_3)$ soit un vecteur de pseudo-norme $m_0 c$ constante. La variété utilisée est lorentzienne *plate*, on peut ou bien absorber dans la formulation hamiltonienne les forces (et alors dans l'espace des phases on utilise une notion de variété avec courbure) ou bien généraliser la notion de force classique pour décrire les mouvements classiques (non quantiques) avec F. Ainsi la force de Lorentz classique $q\vec{E} + \vec{V} \wedge \vec{B}$ donne naissance à un (quadri)vecteur de la variété de l'axiome 1. Dans les deux cas la variété physique sous-jacente ("géométrique" et "physique") *est plate*. Les seuls changements de cartes qui ont un sens physique sont ceux correspondants à une transformation appartenant au groupe de Poincaré. Si l'on reste dans le cadre de la relativité restreinte, une transformation canonique de la formulation hamiltonnienne n'a pas de sens physique.

Dans une carte donnée, que représente le t?

Disposez en chaque point de coordonnées spatiales de la carte une horloge et imposez aux horloges la condition suivante : l'horloge placée en A et l'horloge placée en B diffusent leur temps grâce aux ondes électromagnétiques, si vous vous placez au milieu de AB, vous devez recevoir le même temps en "écoutant" l'horloge A qu'en "écoutant" l'horloge B. Cette condition est bien réalisable pour toutes les horloges de la partie spatiale de la carte à cause des propriétés d'homogénéité spatiale de la diffusion des ondes électromagnétiques, conséquences des équations de Maxwell. Alors une carte de l'atlas théorique de la variété lorentzienne utilisée représente physiquement un espace spatial $\mathbb{R}^3$ possédant en chaque point une horloge et toutes les horloges doivent être synchronisées de la manière indiquée. L'invariance des équations de Maxwell par le groupe de Poincaré nous permet d'affirmer que cette synchronisation étant réalisée dans une carte le calcul abstrait en un point d'une autre carte des valeurs de $x'y'z'$ et de t' dans cette nouvelle carte à partir des valeurs de xyz et de t dans l'ancienne carte, donnera un résultat cohérent et, en particulier, que les horloges de la nouvelle carte seront aussi synchronisées dans le sens expliqué plus haut. La physique de la relativité restreinte affirme que $x'y'z't'$ sont les temps et espaces de l'autre référentiel. Un changement de carte correspondant à un mouvement classique hélicoïdal, par exemple, n'est pas dans le groupe de Poincaré. *Dans une telle carte aucune notion relativiste (restreinte) de temps n'a de sens.* En particulier au sens de la relativité restreinte, *il n'existe pas de notion de temps* associée aux points d'un manège en mouvement de rotation par rapport aux directions fixes des étoiles. Cette restriction *cruciale* au groupe de Poincaré pour les changements de cartes est assez méconnue sous cette forme. (La plupart des "réfutations théoriques" de la relativité utilise des changements de cartes qui ne sont pas dans le groupe de Poincaré). Seule une telle réduction pour les changements de cartes permet de conserver une cohérence physique à la théorie. En relativité (restreinte) la simultanéité de deux événements n'est pas conservée quand on change de

cartes, d'observateurs.

Si vous introduisez d'autres cartes, la notion de temps devient incohérente. Si vous voulez quand même utiliser une variété se permettant d'autres changements de cartes que ceux permis par le groupe de Poincaré, sachez que son interprétation physique ne fut pas simple à découvrir (relativité générale) mais qu'elle est basée sur la structure lorentzienne plate de l'espace tangent en un point identifié à l'espace de la relativité restreinte où de nouveau vous n'aurez droit qu'aux transformations de Poincaré. Le même problème se retrouve en relativité générale : elle est décrite par une variété de classe C^3 par morceaux, C^1 globalement. Alors un théorème de Whitney vous assure que vous pouvez la réaliser comme une sous variété analytique d'un certain $\mathbb{R}^n$. Toutes les cartes nouvelles que vous introduisez ainsi n'ont pas non plus de sens physique. C'est en analysant précisément quels sont les changements de cartes qui ont un sens physique et les autres qui n'ont de sens que mathématique que Lichnerovicz a réussi à comprendre pourquoi en relativité générale une particule matérielle non chargée suit une géodésique. Le but de l'exposé est d'introduire les spineurs qui sont dans un premier temps des sections sur un fibré au-dessus de la relativité restreinte. Leur sens physique ne peut être compris qu'en se restreignant au groupe de Poincaré, dans une première approche.

Pour les mathématiciens géomètres cette première approche est sans grand intérêt et c'est l'équation de Dirac courbe qui les stimulent. Pour en obtenir une interprétation physique, il faut alors étudier la théorie de la relativité générale, et ses généralisations Kazula-Klein puis Yang et Mill. L'interprétation physique des cartes utilisées n'est pratiquement jamais donnée dans ces théories et est vraiment délicate. Dans cet exposé élémentaire, il n'est pas question d'exposer cette interprétation. (Retenez que les cartes mathématiques introduites sont trop nombreuses, en général, et détruisent la cohérence des notions d'espace et de temps que la démarche d'Einstein fait *survivre* de concert avec l'invariance dans l'espace tangent en un point des équations de Maxwell sous l'action du groupe de Poincaré).

Pour les mathématiciens de théorie spectrale par contre l'équation de Dirac sur un espace plat pseudo-lorentzien s'associe avec un Laplacien plus potentiel. Il est donc instructif de suivre l'évolution historique de ce sujet qui est un grand classique en physique (1920–1935).

IV. Mécanique quantique classique

Vous trouverez dans l'appendice une construction mathématique d'un cadre logique axiomatisé pour définir la mécanique quantique classique. Ce système est cohérent

au sens mathématique et logique. La mécanique quantique devient alors une série de méthodes permettant d'évaluer les probabilités d'obtenir certains résultats quand on fait des expériences. La géométrie de l'espace temps a complètement disparue. Reste un espace de Hilbert abstrait sensé représenter les états du système physique en question. Comment choisir cet espace de Hilbert pour une situation donnée?

Nous allons étudier un cas particulier, simple, où le "remplissage" du formalisme abstrait de la mécanique quantique permet de refaire un pont avec la géométrie et suivre l'élaboration historique de la notion de spin en physique. Plaçons nous en mécanique ondulatoire avec une seule particule. Dans cette version primitive, l'espace des états de la particule est $L^2_{\mathbb{C}}(\mathbb{R}^3)$ où le $\mathbb{R}^3$ qui intervient est celui de la mécanique classique. Cela signifie que nous nous plaçons dans un repère galiléen classique (où au sens classique le mouvement de la particule est réglé par $\overrightarrow{F} = m\overrightarrow{\gamma}$ où $\overrightarrow{F}$ nous est donnée). Dans ce repère une fonction $\psi \in L^2_{\mathbb{C}}(\mathbb{R}^3)$ est sensée représenter la particule de la manière suivante. Soit Δ un borélien de $\mathbb{R}^3$, alors si nous effectuons une mesure de la position de la particule, la probabilité de la trouver dans Δ vaut $\frac{\int_\Delta \psi(x)\overline{\psi}(x)dx}{\int_{\mathbb{R}^3} \psi(x)\overline{\psi}(x)dx}$.

Soit maintenant $\mathcal{F}\psi$ la transformée de Fourier de ψ définie par :

$$\mathcal{F}\psi(\xi) = \int_{\mathbb{R}^3} e^{-i\frac{h}{2\pi}x\xi}\psi(x)dx = \widehat{\psi}(\xi).$$

Si nous effectuons une mesure de l'impulsion de la particule, la probabilité de trouver cette impulsion dans Γ borélien de $\mathbb{R}^3$ (espace des impulsions donc des vecteurs (mv^1, mv^2, mv^3) de $\mathbb{R}^3$ associés à la particule libre ou espace des $\left(\frac{\partial \mathcal{L}}{\partial x^1}, \frac{\partial \mathcal{L}}{\partial x^2}, \frac{\partial \mathcal{L}}{\partial x^3}\right)$ si la particule a un mouvement classique réglé par le lagrangien $\mathcal{L}$) vaut

$$\frac{\int_\Gamma \widehat{\psi}(\xi)\overline{\widehat{\psi}}(\xi)d\xi}{\int_{\mathbb{R}^3} \widehat{\psi}(\xi)\overline{\widehat{\psi}}(\xi)d\xi}.$$

Les propriétés de la transformation de Fourier permettent de montrer que les inégalités classiques de Heisenberg $(\Delta p)(\Delta x) \geq \hbar$ prennent alors une interprétation probabiliste (non exposée ici).

Appliquons ce modèle à l'atome d'hydrogène dans lequel nous ne ferons intervenir que l'attraction coulombienne entre le (noyau-proton) et l'électron. Pour la mécanique classique la représentation du mouvement fait intervenir un repère galiléen placé au centre de gravité du système et le mouvement des 2 corps se ramène à un mouvement d'une particule fictive placée dans un champ de force centrale attractif. Prenons donc ce repère galiléen comme repère de référence et traitons la particule fictive de la mécanique classique de manière quantique. Calculons l'hamiltonien $\mathcal{H}$ classique du système et imposons si nous nous intéressons aux états stables de l'atome d'hydrogène à la fonction d'onde de vérifier l'équation de Schrödinger $Op_{\frac{1}{2}}(\mathcal{H})(\psi) = E \cdot \psi$ où $Op_{\frac{1}{2}}$ est une opération qui permet de définir à partir d'une quantité classique un opérateur aux dérivées partielles. Ici dans notre

cas simple l'hamiltonien classique est

$$\mathcal{H} = \frac{px^2 + py^2 + pz^2}{2\mu} - \frac{q^2}{4\pi\varepsilon_0}\frac{1}{r} = \frac{px^2 + py^2 + pz^2}{2\mu} - \frac{e^2}{r}$$

où μ est la masse réduite $\mu = \frac{m_e m_p}{m_e + m_p} \simeq m_e(1 - \frac{1}{1800})$ où m_p est la masse du proton $m_p = 1,710^{-27}$ kg et $q = 1,610^{-19}$ coulomb la charge du proton (=moins celle de l'électron) et r la distance de la particule fictive au centre de gravité de l'ensemble.

La fonction d'onde $\psi(x, y, z) = \widetilde{\psi}(r, \theta, \varphi)$ est à proprement parler la fonction d'onde de la particule fictive (appelée aussi relative) et non la fonction d'onde de l'électron. Par abus de language on parle de fonction d'onde de l'électron mais cela n'est pas trop gênant.

Dans ce cas simple, la quantification de la fonction classique $\mathcal{H}$ est sans ambiguïté définie par le passage de $\mathcal{H}$ à l'opérateur

$$-\frac{\hbar^2}{2\mu}\Delta - U_0(r) = H\Big(\text{ où } U_0(r) = \frac{e^2}{r}\Big).$$

La décomposition spectrale associée au spectre discret de H présente des dégénérescences cruciales dans l'élaboration de la notion de spin. Il nous faut l'analyser. Pour cela on cherche d'autres opérateurs qui commutent avec H et qui décomposent les sous espaces propres de H associés aux valeurs propres de H en sous espaces de dimension 1. En mécanique classique le moment cinétique d'une particule matérielle est

$$m\overrightarrow{OM} \wedge \overrightarrow{V(M)} = \begin{pmatrix} myz - mzy \\ mzx - mxz \\ mxy - myz \end{pmatrix} = \mathcal{L}_x \vec{i} + \mathcal{L}_y \vec{j} + \mathcal{L}_z \vec{k} = \vec{\mathcal{L}}.$$

En faisant un calcul techniquement long qui consiste à faire le changement de variables $\left\{\begin{array}{l} x = r\sin\theta\cos\varphi \\ y = r\sin\theta\sin\varphi \\ z = r\cos\theta \end{array}\right\}$ dans les opérateurs donc *après* avoir quantifié $\mathcal{L}_x$, $\mathcal{L}_y$ et $\mathcal{L}_z$ on obtient :

$$L_x = i\hbar\Big(\sin\varphi\frac{\partial}{\partial\theta} + \frac{\cos\varphi}{tg\theta}\frac{\partial}{\partial\varphi}\Big)$$

$$L_y = i\hbar\Big(-\cos\varphi\frac{\partial}{\partial\theta} + \frac{\sin\varphi}{tg\theta}\frac{\partial}{\partial\varphi}\Big)$$

$$L_z = \frac{\hbar}{i}\frac{\partial}{\partial\varphi}$$

$$L^2 = L_x^2 + L_y^2 + L_z^2 = -\hbar^2\Big(\frac{\partial^2}{\partial\theta^2} + \frac{1}{tg\theta}\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\Big)$$

$$L_+ = L_x + iL_y = \hbar e^{i\varphi}\Big(\frac{\partial}{\partial\theta} + i\operatorname{cotg}\theta\frac{\partial}{\partial\varphi}\Big)$$

$$L_- = L_x - iL_y = \hbar e^{-i\varphi}\Big(\frac{-\partial}{\partial\theta} + i\operatorname{cotg}\theta\frac{\partial}{\partial\varphi}\Big).$$

Attention. — Les trois opérateurs L_+, L_- et L_z ne sont pas affectés par l'inversion quantification-changement de variables. Ce n'est pas le cas pour L^2. En effet $\mathcal{L}^2 = \frac{1}{\sin^2\theta}p_\varphi^2 + p_\theta^2$ dont le quantifié est $-\hbar^2\left(\frac{\partial^2}{\partial\theta^2} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right)$ et il manque le terme $\frac{1}{tg\theta}\frac{\partial}{\partial\theta}$.

Nous sommes alors en position pour trouver la décomposition spectrale associée à H.

1re étape.

En utilisant L_+, L_- et L^2 avec une technique classique on obtient :

THÉORÈME. — *L'harmonique sphérique Y_ℓ^m est la seule fonction de θ et φ, normalisée à l'unité, qui vérifie*

$$L^2 Y_\ell^m = \ell(\ell+1)\hbar^2 Y_\ell^m$$
$$L_z Y_\ell^m = m\hbar Y_\ell^m$$

et la décomposition de $L^2(\theta,\varphi)$ en intégrale hilbertienne associée aux opérateurs commutant L^2 et L_z est simplement la décomposition de toute fonction de $L^2(\theta,\varphi)$ sur la base dénombrable des harmoniques sphériques $\{Y_\ell^m(\theta,\varphi)\}_{\substack{\ell=0,1,2\ldots\\ -\ell\le m\le\ell}}$ définies par

$$Y_\ell^m(\theta,\varphi) = \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell+m)!}{(\ell-m)!}}P_\ell^{-m}(\cos\theta)e^{im\varphi} \text{ si } m<0$$

et

$$Y_\ell^m(\theta,\varphi) = (-1)^m\sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}P_\ell^m(\cos\theta)e^{im\varphi} \text{ si } m>0$$

où le polynôme de Legendre $P_\ell^m(u)$ est défini par

$$P_\ell^m(u) = \sqrt{(1-u^2)^m}\frac{d^m}{du^m}(P_\ell(u)) \quad (-1\le u\le 1).$$

2e étape.

En utilisant le résultat précédent et en utilisant le laplacien en coordonnées sphériques on obtient l'équation radiale

$$\left(-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2} + \frac{\ell(\ell+1)}{2\mu r^2}\hbar^2 - U_0(r)\right)u_{\lambda,k,\ell}(r) = E_{k\ell}u_{\lambda,k,\ell}(r)$$

à résoudre, et on est donc ramené à une équation de Schrödinger en dimension 1 avec un potentiel effectif

$$V_{\text{eff}}(r) = -\underset{\circ}{U}(r) + \frac{\ell(\ell+1)}{2\mu r^2}\hbar^2 \quad \text{(répulsif pour } \ell\neq 0 \text{ au voisinage de 0).}$$

Comme on sait résoudre cette équation différentielle et étudier les comportements limites obtenus on obtient en définitive (l'équation radiale dépend de ℓ et non de m) :

THÉORÈME. — *Soient*

$$\psi^m_{\ell,k}(r,\theta,\varphi)=c'_0\left[e^{-a_0^{-1}\frac{r}{k+\ell}}r^\ell\sum_{q=0}^{k-1}\left(\frac{-1}{a_0}\right)^q\left(\frac{2}{k+\ell}\right)^q\frac{(k-1)!}{(k-q-1)!}\frac{(2\ell+1)!}{(q+2\ell+1)!}r^q\right]Y^m_\ell(\theta,\varphi).$$

Ces fonctions (au facteur multiplicatif c'_0 près) sont les seules fonctions vérifiant :

$$\begin{cases} H\psi^m_{\ell,k} = -\dfrac{E_I}{(k+\ell)^2}\psi^m_{\ell,k} \\ L^2\psi^m_{\ell,k} = \ell(\ell+1)\hbar^2\psi^m_{\ell,k}, \\ L_z\psi^m_{\ell,k} = m\hbar\psi^m_{\ell,k}. \end{cases}$$

$(E_I = \frac{\mu e^4}{2\hbar^2} \simeq 13,6eV, a_0 = \frac{\hbar^2}{\mu e^2} \simeq 0,52\ \mathring{A})$.

Si nous ne nous intéressons qu'aux états où l'électron est lié au proton et forme l'atome d'hydrogène, l'espace de Hilbert est

$$\bigoplus_{k=0}^{\infty}\bigoplus_{\ell=0}^{\infty}\bigoplus_{m=-\ell}^{m=+\ell} H(k,\ell,m) = \text{Hilbert}$$

où $H(k,\ell,m)$ est le sous espace de dimension 1 de $L^2_{\mathbb{C}}(r,\theta,\varphi)$ engendré par $\psi^m_{\ell,k}$. Cet espace de Hilbert joue un grand rôle historique en mécanique quantique.

En posant $n = k+\ell$ on retrouve les niveaux d'énergie du modèle de Bohr *n est le nombre quantique principal* caractérisant une couche électronique : le sous espace propre de $L^2(\theta,\varphi)$ associé à la valeur propre $E_n = -\frac{E_I}{n^2}$ de l'opérateur H.

ℓ est le nombre quantique azimutal. Comme $k \geq 1, \ell$ ne prend que les valeurs $\ell = 0, \ell = 1, \ldots, \ell = n-1$. *La couche caractérisée par n comprend n sous couches indexées par ℓ.*

m est le *nombre quantique magnétique*. Dans tous les cas $-\ell \leq m \leq +\ell$. Il s'agit de la *dégénérescence essentielle* liée au fait que l'équation radiale ne dépend que de ℓ et non de m. On appelle dégénérescence accidentelle la dégénérescence associée au fait que $k'+\ell' = k+\ell$ est possible avec deux valeurs différentes de ℓ. En notation spectroscopique ($\ell = 0 \leftrightarrow s, \ell = 1 \leftrightarrow p, \ell = 2 \leftrightarrow d, \ell = 3 \leftrightarrow f, \ell = 4 \leftrightarrow g, \cdots$ ordre alphabétique).

Confrontation avec l'expérience.

L'étude expérimentale des raies spectrales des atomes fait apparaître une structure fine : chaque raie comporte en fait plusieurs composantes de fréquences très voisines.

Pour l'atome d'hydrogène la formule $E = -\frac{E_I}{n^2}$ ne donne que les énergies moyennes des différents groupes.

Si un atome est plongé dans un champ magnétique uniforme chacune de ses raies de structure fine se divise en un certain nombre de raies plus ou moins équidistantes, séparées par un intervalle proportionnel au champ magnétique : c'est *l'effet Zeeman*.

Sans entrer plus en détail dans le calcul qui permettrait d'approcher théoriquement l'effet Zeeman, retenons que cet effet permet de connaître expérimentalement la dimension du sous espace propre associé à une valeur propre de H. Pour les atomes de numéro atomique pair la théorie précédente prédit bien l'effet Zeeman (effet Zeeman dit "normal").

Pour les atomes de numéro atomique impair (hydrogène par exemple) *l'expérience montre que les niveaux se divisent en un nombre pair de sous niveaux Zeeman* alors qu'avec la théorie précédente le champ magnétique $\overrightarrow{B}$ uniforme traité comme perturbateur ne fait que distinguer les $(2\ell + 1)$ dégénérescences essentielles dues au fait que l'équation radiale ne dépend que de ℓ et pas de m. Il faudrait des valeurs demi-entières de ℓ. Les valeurs demi-entières de ℓ ont été exclues parce qu'en recherchant une fonction $\psi_{\lambda,\ell,m} = |\lambda, \ell, m\rangle$ vérifiant $L_z|\lambda, \ell, m\rangle = m\hbar|\lambda, \ell, m\rangle$, en coordonnées sphériques on obtient $-i\frac{\partial}{\partial\varphi}|\lambda, \ell, m\rangle = m|\lambda, \ell, m\rangle$ qui s'intègre en $|\lambda, \ell, m\rangle(\theta, \varphi) = F_{\lambda,\ell,m}(\theta)e^{im\varphi}$ et on veut que $e^{2im\pi} = 1$ d'où m entier.

En mécanique classique pour un solide $\mathcal{S}$ de centre de gravité G de vecteur rotation instantané $\overrightarrow{\Omega}$ le moment cinétique est :

$$\overrightarrow{\mathcal{L}}_0(\mathcal{S}) = \int_{\mathcal{S}} \overrightarrow{OM} \wedge \overrightarrow{V(M)}dm = m\overrightarrow{OG} \wedge \frac{d\overrightarrow{OG}}{dt} + \overline{\square}_G \cdot \overrightarrow{\Omega}$$

où $\overline{\square}_G$ est l'opérateur d'inertie en G du solide $\mathcal{S}$. Il faut trois coordonnées cartésiennes et trois angles d'Euler pour représenter la position de $\mathcal{S}$. Si on veut quantifier cette situation en conservant aux trois angles d'Euler leur interprétation géométrique on retombe sur le quantifié de $\overrightarrow{\mathcal{L}}_G$ et pour les mêmes raisons que précédemment $|\overrightarrow{\mathcal{L}}_G|^2$ et $\overrightarrow{\mathcal{L}}_z = \overrightarrow{\mathcal{L}}_G \cdot \overrightarrow{z}$ ne prendront que les valeurs $\ell(\ell + 1)\hbar^2$ et $m\hbar$ et on aura rien gagné. Alors :

• *ou bien* on introduit des opérateurs moments cinétiques de spin S abstraits tels que dans une base orthonormée particulière $[S_x, S_y] = i\hbar S_z$ (et les deux autres relations déduites par permutations circulaires sur xyz) et on impose que $S^2 = S_x^2 + S_y^2 + S_z^2$ ne prenne qu'une seule valeur appelée spin s de la particule avec $s = 0, \frac{1}{2}, 1, \frac{3}{2}, 2 \cdots$ et on décide que S agit dans l'espace des états de spin où S^2 et S_z constitue un ensemble complet d'observables qui commutent (pour l'électron $s = \frac{1}{2}$, on connaît des spins jusqu'à $\frac{11}{2}$) ;

• *ou bien*, en regardant ce que donne la démarche précédente, $\langle S^2|s, m\rangle = s(s + 1)\hbar^2|s, m\rangle$ et $S_z|s, m\rangle = m\hbar|s, m\rangle$ avec $s = 0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$) au lieu d'associer à l'électron (pour $s = \frac{1}{2}$) *une* fonction d'onde on lui associe *deux* fonctions d'onde en utilisant l'interprétation géométrique du revêtement de $SO_3(\mathbb{R})$ par $SU_2(\mathbb{C})$ de la manière suivante. Souvenons-nous que notre manière de voir la mécanique classique particularise des repères orthonormés spéciaux posés en axiomes dans la théorie, les repères galiléens. Dans un repère galiléen particulier, utilisons la construction que nous avons décrite du revêtement de $L_+^\uparrow$ par $SL(2, \mathbb{C})$. Elle nous donne une construction d'un revêtement de $SO_3(\mathbb{R})$ par $SU_2(\mathbb{C})$. Alors au lieu de prendre la fonction d'onde dans $L^2_{\mathbb{C}}(\mathbb{R}^3)$ prenons la dans

$L^2_{\mathbb{C}^2}(\mathbb{R}^3)$ où le $\mathbb{C}^2$ qui intervient est le $\mathbb{C}^2$ sur lequel agit $SU_2(\mathbb{C})$. Cela veut dire que si nous faisons subir à notre repère galiléen une famille à un paramètre λ de rotation R_λ agissant sur $\mathbb{R}^3$, nous pourrons en suivant par continuité dans la construction donnée du revêtement de $SO_3(\mathbb{R})$ par $SU_2(\mathbb{C})$ suivre la transformation qui agit sur l'espace vectoriel complexe de dimension 2 sur $\mathbb{C}$ sur lequel agit $SU_2(\mathbb{C})$. Notons la $\underset{\sim}{R}\ \lambda$.

Alors la fonction d'onde $\binom{\psi_1}{\psi_2}$ à deux composantes devra se transformer dans cette rotation suivant $\underset{\sim}{R}\ \lambda$. Cette construction ne fonctionne pas parce que le revêtement de $SO_3(\mathbb{R})$ par $SU_2(\mathbb{C})$ choisit n'a rien d'intrinsèque et est à deux feuillets. Mais si l'on sait que *nous n'avons droit qu'aux repères galiléens*, en fixant dans un repère particulier la correspondance, nous la connaissons dans tous les repères galiléens au signe près. Quelque soit le revêtement de $SO_3(\mathbb{R})$ par $SU_2(\mathbb{C})$ choisi, à la fin on obtient un facteur multiplicatif dans l'évaluation des fonctions d'onde et ce facteur s'élimine quand on évalue les quotients du genre $\frac{\int_\Delta \psi_1\overline{\psi}_1 dx}{\int_{\mathbb{R}^3} \psi_1\overline{\psi}_1 dx}$. La démarche revient à prendre la fonction d'onde sur un espace vectoriel de dimension finie sur lequel agit de manière irréductible $SU_2(\mathbb{C})$ et de prendre comme changements de cartes des éléments de $SU_2(\mathbb{C})$ (et non de $SO_3(\mathbb{R})$) agissant entre les espaces vectoriels de dimension finie sur lesquels agit de manière irréductible $SU_2(\mathbb{C})$, la variété sous jacente restant de dimension 3 réelle via un choix particulier d'un revêtement de $SO_3(\mathbb{R})$ par $SU_2(\mathbb{C})$ dans un repère galiléen particulier. Comme $SU_2(\mathbb{C})$ laisse invariante la forme $\psi_1\overline{\psi}_1 + \psi_2\overline{\psi}_2$ la masse totale proportionnelle à $\int(\psi_1\overline{\psi}_1 + \psi_2\overline{\psi}_2)d^3x$ reste bien constante comme il se doit en mécanique classique. Sans aller plus avant dans la modification de la variété utilisée et en revenant à la variété $\mathbb{R}^3 \times \mathbb{R}$ de la mécanique classique, nous obtenons l'axiomatique de Pauli dont voici un petit résumé. Historiquement, c'est comme cela que le spin est né, avec la présentation bancale du deuxième "ou bien".

Axiomatique de Pauli de la mécanique ondulatoire classique avec spin.

Étant donné un repère galiléen particulier, l'espace des états de la particule est $L^2_{\mathbb{C}}(\mathbb{R}^3) \otimes \mathcal{E}_s$ où $\mathcal{E}_s$ est l'espace des variables de spin. *$\mathcal{E}_s$ par définition est un espace de Hilbert de dimension finie* $2s+1$ *sur* $\mathbb{C}$.

La donnée d'un élément de cet espace équivaut à la donnée de $(2s+1)$ fonctions de $L^2_{\mathbb{C}}(\mathbb{R}^3)$. La matrice colonne $\psi = \begin{pmatrix} \psi_s \\ \vdots \\ \psi_{-s} \end{pmatrix}$ s'appelle un spineur à $(2s+1)$ composante. Il existe un opérateur moment cinétique de spin noté S, opérateur vectoriel qui dans le repère galiléen particulier vérifie $[S_x, S_y] = i\hbar S_z$, $[S_y, S_z] = i\hbar S_x$, $[S_z, S_x] = i\hbar S_y$ et $S^2 = S_x^2 + S_y^2 + S_z^2$ ne prend qu'une seule valeur la valeur s ($s = 0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$ fixée une fois pour toute) (cette constance de s correspond à la longueur constante en mécanique classique du moment cinétique).

Un spineur dépend de (x, y, z) ou de (p_x, p_y, p_z) suivant que l'on travaille avec (x, y, z, S^2, S_z) ou $(P_x, P_y, P_z, S^2, S_z)$. On appelle *opérateur de spin* un opérateur qui n'agit que sur la variable ε valeur propre de S_z telle que $-s \leq \varepsilon \leq s$, ε variant de 1 en 1. Par exemple pour $s = \frac{1}{2}$, $S_+ = S_x + iS_y$ et $S_- = S_x - iS_y$ sont tels que $S_+\binom{\psi_{\frac{1}{2}}}{\psi_{\frac{-1}{2}}} = \hbar\binom{\psi_{\frac{-1}{2}}}{0}$ et $S_-\binom{\psi_{\frac{1}{2}}}{\psi_{\frac{-1}{2}}} = \hbar\binom{0}{\psi_{\frac{1}{2}}}$ ce que l'on note $S_+ = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $S_- = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

On appelle *opérateur orbital* un opérateur qui n'agit que sur les variables (x, y, z) (ou $(p_x p_y p_z)$ ou (k, ℓ, m) dans une autre représentation). Par exemple $X = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}$, $P_X = \frac{\hbar}{i}\begin{pmatrix} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial x} \end{pmatrix}$.

Le passage de $(X_1, X_2, X_3, S^2 S_z)$ à $(P_{X_1}, P_{X_2}, P_{X_3}, S^2, S_z)$ est $(2s+1)$ fois la transformation de Fourier sur ψ_ε où $\varepsilon \in \{-s, -s+1, \ldots, s\}$. $\frac{\int_\Delta |\psi_\varepsilon|^2 dx}{\int_{\mathbb{R}^3} |\psi_\varepsilon|^2 dx}$ est la probabilité de trouver la particule dans Δ avec ε comme valeur de son spin.

Même interprétation avec la mesure de l'impulsion. Ces définitions étant données dans un repère galiléen particulier il faut décrire comment elles se transforment sous l'action d'un élément du groupe de Galilée et tout revient à décrire comment elles se transforment sous l'action d'un élément de $SO_3(\mathbb{R})$ (on décide que la formulation est la même sous la transformation $t' = t + \alpha, x' = x_0 + x, y' = y_0 + y, z' = z_0 + z$).

Mathématiquement il n'est pas possible de faire agir $SO_3(\mathbb{R})$ sur l'espace des spineurs d'une dimension donnée finie sur $\mathbb{C}$. Comme le revêtement $SU_2(\mathbb{C}) \to SO_3(\mathbb{R})$ est à deux feuillets cette action est forcément multivaluée (un signe $+$ ou $-$ intervient forcément). Seule l'action de $SU_2(\mathbb{C})$ aura un sens. En faisant vivre la fonction d'onde à plusieurs composantes complexes sur la variété de la mécanique classique un spineur ne peut être défini qu'au signe près et dire que $SO_3(\mathbb{R})$ agit sur un spineur revient à projeter l'action d'un élément de $SU_2(\mathbb{C})$ sur ce spineur. Cette action de $SU_2(\mathbb{C})$ sur le spineur est définie comme étant l'action de la représentation irréductible dite spinorielle de $SU_2(\mathbb{C})$. Pour $s = \frac{1}{2}$ donnons une forme possible de cette action en utilisant les matrices de Pauli $\tau_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$; $\tau_x = \tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; $\tau_2 = \tau_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ et $\tau_3 = \tau_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Soit $\vec{u} = u_1\vec{e}_1 + u_2\vec{e}_2 + u_3\vec{e}_3$ un vecteur de norme 1 dans $\mathbb{R}^3$ rapporté à notre repère de Galilée initial. La rotation d'angle α autour de $\vec{u}$ fera intervenir l'élément suivant de $SU_2(\mathbb{C})$ (défini au signe près)

$$R_{\vec{u},\alpha} = \begin{pmatrix} \cos\frac{\alpha}{2} - iu_z \sin\frac{\alpha}{2} & (-iu_x - u_y)\sin\frac{\alpha}{2} \\ (-iu_x + u_y)\sin\frac{\alpha}{2} & \cos\frac{\alpha}{2} + iu_z \sin\frac{\alpha}{2} \end{pmatrix} = e^{-\frac{i}{\hbar}\alpha S\cdot u} = \cos\frac{\alpha}{2} I - i(\tau\cdot u)\sin\frac{\alpha}{2}$$

(formule déduite du choix du revêtement de $L_+^\uparrow$ par $SL_2(\mathbb{C})$ que nous avons fait).

Il faut pour terminer dire ce que devient l'équation de Schrödinger. Pour $s = \frac{1}{2}$

Pauli impose pour φ spineur à deux composantes l'équation

$$i\hbar\frac{\partial\varphi}{\partial t} = \Big(\frac{1}{2m}(\vec{p} - \frac{e}{c}\widetilde{A})^2 - \frac{e\hbar}{2mc}\tau\cdot B + e\Phi\Big)\varphi$$

où

$$\frac{1}{i\hbar}\vec{p} = \sum_{i=1}^{3}\frac{\partial}{\partial x}\cdot\tau_i = \begin{pmatrix} \frac{\partial}{\partial z} & \frac{\partial}{\partial x} - i\frac{\partial}{\partial y} \\ \frac{\partial}{\partial x} + i\frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{pmatrix}.$$

Le terme $e\Phi$ est le potentiel coulombien habituel. Le terme $-e\frac{\hbar}{2mc}\tau\cdot B = -e\frac{\hbar}{2mc}\sum_{i=1}^{3}B_i\tau_i$ où $\vec{B}$ est le champ magnétique est rajouté à la main par Pauli pour pouvoir expliquer les spectres des atomes connus et A est le potentiel du champ électromagnétique provenant de l'intégration locale de $dF = 0$, deuxième groupe des équations de Maxwell, en $F = dA$ et $\widetilde{A} = i\hbar\sum_{i=1}^{3}A_i\cdot\tau_i$. Pour les autres valeurs de s nous verrons comment obtenir les équations correspondantes. Surtout nous déduirons cette équation de Pauli à partir de l'équation de Dirac. Noter bien le coefficient 2 au dénominateur. Dans la littérature, c'est ce qui s'appelle "le rapport gyromagnétique". Pauli devait mettre ce coefficient $\frac{1}{2}$ sans autres justifications que de retrouver les résultats expérimentaux à l'aide de son équation et de l'arsenal de la mécanique ondulatoire classique à spin. Le succès fut grand, et le pouvoir prédictif énorme.

V. L'équation de Dirac

Continuons à suivre l'histoire de la physique. D'un côté la relativité (restreinte) et de l'autre la mécanique quantique classique avec spin de Pauli qui n'est pas invariante par le groupe de Poincaré mais par le groupe de Galilée. Schrödinger explique qu'il a d'abord tenté d'écrire une équation invariante relativiste, mais devant l'échec à préféré écrire une équation classique, invariante par le groupe de Galilée, et en tirer une formulation cohérente de la mécanique quantique (classique donc). Dirac motive son équation par l'envie d'une équation invariante par le groupe de Poincaré. Voilà la présentation ancienne de l'équation de Dirac. Elle est là pour la cohérence historique de l'exposé. Elle permet de retrouver l'équation de Pauli et donne ainsi le contenu physique des spineurs au sens de l'histoire.

Équation de Dirac : ancienne version.

Pour quantifier classiquement on considère le temps comme un paramètre et on associe à $a_t(x, y, z, p_x, p_y, p_z)$ un opérateur par $Op_{\frac{1}{2}}$.

En relativité restreinte l'espace et le temps ne sont pas deux entités séparées mais simplement une vision relative de l'espace temps.

L'énergie d'une particule libre est $E = \sqrt{p^2c^2 + m_0^2c^4}$. Il est donc naturel d'adopter la quantification $\begin{cases} E \to i\hbar\frac{\partial}{\partial t} \\ p_{x_\mu} \to i\hbar\frac{\partial}{\partial x_\mu} \end{cases}$ puisque $(\frac{E}{c}, p_1, p_2, p_3)$ est un (quadri) vecteur (de pseudo-norme m_0c constante).

Simplement si l'on prend $Op_{\frac{1}{2}}$ de fonction $a(ct, x, y, z; E, p_x, p_y, p_z)$, on obtient alors pour une particule libre $i\hbar\frac{\partial\psi}{\partial t} = \sqrt{-\hbar^2c^2\Delta + m_0^2c^4}\psi$ qui est impraticable à cause de $\sqrt{}$ (à moins que de considérer le développement formel $i\hbar\frac{\partial\psi}{\partial t} = m_0c^2 + \frac{\Delta}{2m_0} - \frac{\Delta^2}{8m_0c^2} + \dots$ ce qui montre que l'équation de Schrödinger $i\hbar\frac{\partial\psi}{\partial t} = \frac{\Delta}{2m_0}\psi$ peut être perturbée pour tenir compte des effets relativistes. Ceci ne conduit qu'à des théories classiques avec termes correctifs pour tenir compte de la variation de la masse avec la vitesse et non à des théories invariantes relativistes).

En ne prenant pas la racine carrée $E^2 = p^2c^2 + m_0^2c^4$ et en quantifiant on obtient $-\hbar^2\frac{\partial^2\psi}{\partial t^2} = (-\hbar^2c^2\Delta^2 + m_0^2c^4)\psi$ ce qui donne *l'équation de Klein Gordon* $\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} + (\frac{m_0c^2}{\hbar})^2\right)\psi = 0$. On note le dalembertien $\Box = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}$ d'où $\left(\Box + (\frac{m_0c}{\hbar})^2\right)\psi = 0$. Cette équation est du second ordre. Cherchons *a priori* une équation du premier ordre pour une particule libre. Si on pose $g^{00} = -g^{11} = -g^{22} = -g^{33} = +1$, le dalembertien s'écrit $-\Box = \sum_{n=0}^{3} g^{nn}P^nP^n$ en posant $P^n = ig^{nn}\frac{\partial}{\partial x^n}$. Cherchons P combinaison linéaire des P^n, avec des coefficients γ^k constants $P = \sum_k g^{kk}P^k\gamma^k$, tels que par analogie avec la formule de décomposition $a^2 - b^2 = (a-b)(a+b)$ on ait

$$\left(P - \frac{m_0c}{\hbar}\right)\left(P + \frac{m_0c}{\hbar}\right) = \sum_{n=0}^{3} g^{nn}P^nP^n - \left(\frac{m_0c}{\hbar}\right)^2$$

ce qui est (au signe près) l'opérateur de Klein Gordon $\Box + \left(\frac{mc}{\hbar}\right)^2$.

Remarque. — Les notations à propos du dalembertien ne sont pas unifiées et $\Box$ est parfois $\frac{\partial^2}{\partial t^2} - \Delta$ et parfois $\Delta - \frac{\partial^2}{\partial t^2}$ cela dépend des auteurs.

Pour réaliser ce programme il est impossible de choisir les γ^k comme de simples nombres complexes. *Recherchons l'ordre de la représentation irréductible de ces grandeurs* c'est-à-dire étudions l'algèbre A sur le corps des nombres complexes, formée, par tous les produits possibles et toutes les combinaisons linéaires à coefficients complexes possibles, à partir des quatre γ^k. (Les produits doivent *a priori* être cherchés non commutatifs). Tout d'abord il faut que

$$\left(P - \frac{mc}{\hbar}\right)\left(P + \frac{mc}{\hbar}\right) = \sum_{0}^{3} g^{nn}P^nP^n - \left(\frac{m_0c^2}{\hbar}\right)^2$$

donc que

$$\sum_{n=0}^{3} g^{nn}P^nP^n = P\cdot P = \Big(\sum_k g^{kk}\gamma^k i g^{kk}\frac{\partial}{\partial x^k}\Big)\Big(\sum_{k'} g^{k'k'}\gamma^{k'} i g^{k'k'}\frac{\partial}{\partial x^{k'}}\Big).$$

Comme $g^{kk}g^{kk} = +1$ alors

$$\begin{aligned} P\cdot P &= \Big(\sum_k i\gamma^k\frac{\partial}{\partial x^k}\Big)\Big(\sum_{k'} i\gamma^{k'}\frac{\partial}{\partial x^{k'}}\Big) \\ &= -\Big(\gamma^0\frac{\partial}{\partial t}+\gamma^1\frac{\partial}{\partial x^1}+\gamma^2\frac{\partial}{\partial x^2}+\gamma^3\frac{\partial}{\partial x^3}\Big)\Big(\gamma^0\frac{\partial}{\partial t}+\gamma'\frac{\partial}{\partial x'}+\gamma^2\frac{\partial}{\partial x^2}+\gamma^3\frac{\partial}{\partial x^3}\Big) \\ &= \sum_0^3 g^{nn}P^nP^n = -\sum_0^3\Big(\frac{\partial^2}{\partial t^2}-\frac{\partial^2}{\partial x'^2}-\frac{\partial^2}{\partial x^{2^2}}-\frac{\partial^2}{\partial x^{3^2}}\Big) \end{aligned}$$

d'où $\gamma^n\gamma^m + \gamma^m\gamma^n = 2g^{mn}$.

À partir des 4 matrices γ^k on peut construire 16 matrices linéairement indépendantes

$*I = \sum_k g^{kk}\gamma^k\gamma^k$ une matrice unité ;

$*$ quatre matrices γ^k ;

$*$ six matrices $\Gamma^{k\ell} = \gamma^k\gamma^\ell \quad (k<\ell)$;

$*$ quatre matrices $D^{k\ell m} = \gamma^k\gamma^\ell\gamma^n \quad (k<\ell<n)$;

$*$ une matrice $\gamma^5 = \gamma^0\gamma^1\gamma^2\gamma^3$ (on réserve γ^4 pour $\gamma^4 = i\gamma^0$).

En tripatouillant ces 16 matrices obtenues avec les propriétés des traces on se rend compte que ces 16 matrices sont linéairement indépendantes et engendrent toute l'algèbre recherchée. Donc on choisira de représenter les γ^k par des matrices (4, 4).

La relation de départ $\gamma^n\gamma^m + \gamma^m\gamma^n = 2g^{mn}$ et les tripatouillages sont invariants par rapport à une transformation unitaire (une matrice $\theta(4,4)$ inversible unitaire) $\gamma'^k = \theta\gamma^k\theta^{-1}$.

La convention habituelle est de se placer dans une représentation où on retrouve les matrices de Pauli :

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} ; \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} ;$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} ; \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

En notant :

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_1 ; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma_2 ; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_3$$

on note :

$$\alpha_i = \left(\begin{array}{cc|c} 0 & 0 & \sigma_i \\ 0 & 0 & \\ \hline \sigma_i & & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \end{array}\right)_{i=1,2,3} \quad \text{ou bien} \quad \gamma^i = \left(\begin{array}{cc|c} 0 & 0 & \sigma_i \\ 0 & 0 & \\ \hline -\sigma_i & & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \end{array}\right)$$

suivant les auteurs. Les opérateurs $P - \frac{m_0 c}{\hbar}$ et $P + \frac{m_0 c}{\hbar}$ sont les opérateurs associés respectivement à l'électron et au position (couple particule antiparticule, l'antiparticule d'une particule sans charge est la particule elle-même).

À l'aide des α_ℓ l'équation de Dirac de l'électron libre est :

$$i\hbar \frac{\partial \psi_j}{\partial t} = \frac{\hbar c}{i} \left(\sum_{\substack{\ell=1 \\ k=1}}^{\substack{k=4 \\ \ell=3}} \alpha_\ell^{kj} \frac{\partial \psi_k}{\partial x^\ell} \right) + \gamma^0 m_0 c^2 \psi_j$$

avec $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}$.

Définissons $\psi^+ = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)$ et multiplions l'équation ci-dessus à gauche par ψ^+ puis à droite par ψ et faisons la différence on obtient :

$$i\hbar \frac{\partial}{\partial t}(\psi^+ \psi) = \sum_{k=1}^{3} \frac{\hbar c}{i} \frac{\partial}{\partial x^k}(\psi^+ \alpha^k \psi).$$

En posant $\rho = \psi^+ \psi$ $\quad (= \sum_1^4 \psi_\sigma^* \psi_\sigma)$ et $j^k = c\psi^+ \alpha^k \psi$ $\quad k = 1, 2, 3$ on obtient $\frac{\partial \rho}{\partial t} + \operatorname{div} \vec{j} = 0$ à rapprocher des équations classiques de conservation. Il faudrait vérifier que le vecteur $\vec{j}$ est bien défini intrinsèquement et est "invariant sous les rotations de l'espace $\mathbb{R}^3$" pour avoir un équivalent classique. Il est préférable de se poser le problème plus général : $\begin{pmatrix} \rho \\ j_1 \\ j_2 \\ j_3 \end{pmatrix}$ forme-t-il un quadrivecteur invariant par transformation de Lorentz ?

La réponse est OUI de même que l'équation de Dirac est, elle aussi, Lorentz invariante et plus exactement $SU_2(\mathbb{C})$-invariante (voir la discussion dans le cadre du spin classique). (La partie non $SU_2(\mathbb{C})$ du groupe de Lorentz transforme bizarrement l'équation ancienne de Dirac. C'est l'origine de beaucoup de débats... et d'un retour géométrique du spin...)

Équation de Dirac en présence d'un champ électromagnétique.

Grâce à l'introduction du champ électromagnétique dans le formalisme hamiltonien, pour obtenir l'équation de Dirac pour un électron (dont l'analogue classique est une

particule de charge e et de masse m) faisons la substitution $p^i \to p^i - \frac{e}{c}A^i$ dans l'équation de Dirac pour une particule libre

$$i\hbar\frac{\partial\psi}{\partial t} = \Big(c\sum_{\ell=1}^{3}\alpha_\ell\big(P^\ell - \frac{e}{c}A^\ell\big) + \gamma^0 m_0 c^2 + e\Phi\Big)\psi.$$

(Φ, A^1, A^2, A^3) forme un quadrivecteur invariant de Lorentz défini avec le tenseur de Faraday. Φ est le potentiel scalaire électrique et apparaît comme opérateur de multiplication dans la représentation de position utilisée ici et A^1, A^2, A^3 sont les potentiels vecteurs magnétiques et apparaissent comme opérateur de dérivation. Bien entendu cette distinction électrique-magnétique est une distinction classique. Par une transformation de Lorentz $(\Phi, A^1, A^2, A^3) \to (A'^0, A'^1, A'^2, A'^3)$ et pour le nouvel observateur dans le nouveau repère de Lorentz A'^0 son champ électrique résulte d'un mélange d'électrique et de magnétisme de l'observateur dans l'ancien repère de Lorentz.

(On trouverai bien sûr exactement la même équation de Dirac pour un électron plongé dans un champ électromagnétique (le champ électromagnétique dans cette description n'est pas quantifié) en ajoutant à l'hamiltonien d'une particule libre $\frac{p^2}{2m_0}$ l'hamiltonien $-\frac{e}{c}\vec{V}\cdot\vec{A} + e\Phi$ d'interaction électromagnétique et en adoptant la correspondance que l'opérateur associé à V_i est $c\alpha_i$).

Limite non relativiste de l'équation de Dirac.

Choisissons la représentation des γ^k comme étant $\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ et $\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}$ et écrivons $\psi = \binom{\tilde{\varphi}}{\tilde{\chi}}$ où $\tilde{\varphi}$ et $\tilde{\chi}$ sont des matrices colonnes à 2 composantes. Alors

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix} = c\sum_{i=1}^{3}\sigma_i(p^i - \frac{e}{c}A^i)\begin{pmatrix}\tilde{\chi}\\ \tilde{\varphi}\end{pmatrix} + e\Phi\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix} + m_0c^2\begin{pmatrix}\tilde{\varphi}\\ -\tilde{\chi}\end{pmatrix}.$$

Soit le changement de variables $\binom{\tilde{\varphi}}{\tilde{\chi}} = e^{-\frac{im_0c^2}{\hbar}t}\binom{\varphi}{\chi}$. L'équation devient

$$i\hbar\frac{\partial}{\partial t}\frac{\varphi}{\chi} = c\sum_{i=1}^{3}\sigma_i\big(p^i - \frac{e}{c}A^i\big)\begin{pmatrix}\chi\\ \varphi\end{pmatrix} + e\Phi\begin{pmatrix}\varphi\\ \chi\end{pmatrix} - 2m_0c^2\begin{pmatrix}0\\ \chi\end{pmatrix}.$$

Considérons la deuxième de ces 2 équations. m_0c^2 l'énergie du repos dans l'approximation relativiste est grande par rapport aux autres termes :

$$i\hbar\frac{\partial\chi}{\partial t} = c\sum_{i=1}^{3}\sigma_i\big(p^i - \frac{e}{c}A^i\big)\varphi + e\Phi\chi - 2m_0c^2\chi$$

admet donc à l'approximation non relativiste la solution $\chi = \frac{\sum_{i=1}^{3} \sigma_i(p^i - \frac{e}{c}A^i)\varphi}{2m_0 c}$, à l'approximation non relativiste $\frac{v}{c} \ll 1$ donc χ est "petite" devant φ. La première équation devient en remplaçant χ par $\frac{\sum_{i=1}^{3} \sigma_i(p^i - \frac{e}{c}A^i)\varphi}{2m_0 c}$

$$i\hbar\frac{\partial\varphi}{\partial t} = \Big(\frac{\sum_{i=1}^{3} \sigma_i\big(p^i - \frac{e}{c}A^i\big) \circ \sum_{\ell=1}^{3} \sigma_\ell\big(p^\ell - \frac{e}{c}A^\ell\big)}{2m_0} + e\Phi\Big)\varphi.$$

Formulaire sur les matrices de Pauli pour terminer le calcul.

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ; \quad \sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} ; \quad \sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$\det \sigma_j = -1$; $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$; $\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z$ et de même sur les permutations, circulaires sur xyz

$$\sigma_j\sigma_k = i\sum_{\ell} \varepsilon_{ijk}\sigma_\ell + \delta_{jk}I \quad \varepsilon_{\ell jk} = \begin{cases} 0 & \text{si tous les indices ne sont pas} \neq s \\ 1 & jk\ell \text{ permutation paire de } 123 \\ -1 & jk\ell \text{ permutation impaire de } 123 \end{cases}$$

$[\sigma_x, \sigma_y] = 2i\sigma_z, \sigma_x\sigma_y + \sigma_y\sigma_x = 0$

$$(\sigma \cdot A)(\sigma \cdot B) = \sum_{j,k} \sigma_j A_j \sigma_k B_k = \sum_{j,k} A_j B_k \Big(\sum_{\ell} \varepsilon_{jk\ell}\sigma_\ell + \delta_{jk}I\Big)$$

$$\Longrightarrow \quad \boxed{\sum_{\ell} \sigma_\ell \Big(\sum_{j,k} \varepsilon_{jk\ell} A_j B_k\Big) + \sum_j A_j B_j I = (\sigma \cdot A)(\sigma \cdot B)}$$

que l'on écrit parfois $(\sigma \cdot A)(\sigma \cdot B) = AB \cdot I + i\sigma(A \wedge B)$ dont il faut se méfier parce que $\vec{A} \wedge \vec{A} \neq \vec{0}$ (*a priori* les composantes de A ne commutent pas) en prenant $A = B = \sigma(p - \frac{e}{c}A)$ la moitié de l'équation de Dirac dans le domaine non relativiste donne l'équation de Pauli et justifie le rapport gyromagnétique imposé *a priori* dans la théorie de Pauli (φ spineur à 2 composantes).

$$\boxed{i\hbar\frac{\partial\varphi}{\partial t} = \Big(\frac{(p - \frac{e}{c}A)^2}{2m_0} - \frac{e\hbar}{2mc}\sigma \cdot B + e\Phi\Big)\varphi}$$

Équation de Dirac, les spineurs, essai plus moderne.

Devant l'extraordinaire pouvoir de prédiction (même dans le cadre classique où le champ électromagnétique est traité comme champ classique, exerçant une "quadriforce" (au sens de la relativité restreinte) sur la particule, beaucoup de travaux mathématiques ont été effectués pour extraire un formalisme plus satisfaisant de cette équation.

Voici une formulation donnée par H. Pesce et G. Besson dans le groupe de travail sur l'équation de Dirac. Soit $V = \mathbb{R}^n$ et Q une forme quadratique positive ou négative, en complexifiant, on obtient $V_{\mathbb{C}}$ et $\widetilde{Q}$ une forme quadratique à partir de laquelle on construit l'algèbre de Clifford complexifiée de V, $\widetilde{\mathcal{C}\ell}(V) = \mathcal{C}\ell(\widetilde{V})$. La construction de l'algèbre de Clifford procède de la volonté d'injecter V dans une algèbre A de sorte que les conjugaisons par des éléments $B \subset A$ soient les isométries de V pour sa forme quadratique Q. A sera une algèbre qui dépend de la forme quadratique. Commençons par considérer $T(V) = (\hbar 1) \oplus (V) \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \ldots$ ($h = \mathbb{R}$ ou $\mathbb{C}$) et on s'impose $v^2 = q(v) \cdot 1$ (ce qui conduit à $vw + wv = 2\langle v|w\rangle$). On considère l'idéal bilatère engendré par les éléments de la forme $v \otimes v - q(v)1$, $qdv \in V$

$$I(V,q) = \Big\{ \sum_{\text{somme finie}} a_i \otimes (v_i \otimes v_i - q(v_i)1) \otimes b_i \Big\}.$$

Par définition $\mathcal{C}\ell(V,q) = T(V)/I(V,q)$ est si $n = \dim V$, $\dim \mathcal{C}\ell(V) = 2^n$.

Dans $\mathcal{C}\ell(V,q)$, $uv = 2\langle u|v\rangle - vu$. Si un tenseur sur V avait un degré, dans l'algèbre de Clifford il ne subsiste que la parité $\mathcal{C}\ell(V,q) = \mathcal{C}\ell^0(V,q) \oplus \mathcal{C}\ell^1(V,q)$ où $\mathcal{C}\ell^0$ est engendré par des produits en nombre pair d'éléments d'une base de V et $\mathcal{C}\ell^1$ par des produits de nombres impairs.

On définit $\alpha(x) = x$ si $x \in \mathcal{C}\ell^0$ et $\alpha(x) = -x$ si $x \in \mathcal{C}\ell^1$. α est un automorphisme et avec ${}^t(x_1 \otimes \cdots \otimes x_2) = x_2 \otimes \cdots \otimes x_1$ la transposition et α commutent.

Alors $\Gamma = \Big\{ x \in \mathcal{C}\ell(V,q) \mid \alpha(x) V x^{-1} = V \Big\}$ est un groupe stable par α et par transposition et en posant $\rho : \Gamma \longrightarrow GL(V)$

$$\rho(x)v = \alpha(x) v x^{-1}$$

on démontre que $\ker \rho = \mathbb{R}^* \cdot 1$ puisque $\operatorname{Im} \rho \subseteq \mathcal{O}(V)$ d'où $\Gamma = \{v_1 \cdots v_k \mid q(v_i) \neq 0\}$.

En posant $\operatorname{Pin}(V,q) = \{v_1 \cdots v_k,\ q(v_i) = \pm 1\}$ alors $\ker \rho$ (restreint à) $\operatorname{Pin}(V,q) = \{\pm 1\}$ et $\rho : \operatorname{Pin}(V,q) \longrightarrow \mathcal{O}(q)$ est surjective et en quotientant on obtient

$$\rho : \operatorname{Spin}(V,q) \longrightarrow SO(q).$$

Dans le cas où n la dimension de V est paire $(e_1 \cdots e_n)$ une base orthonormée de V alors définissons

$$V_{\mathbb{C}} = M \oplus P \quad \text{où } m_k = \frac{1}{2}(e_{2k-1} - ie_{2k}) \text{ engendre } M,$$
$$p_k = \frac{1}{2}(e_{2k-1} + ie_{2k}) \text{ engendre } P.$$

et par définition l'espace des spineurs est $S = \Lambda M$ l'algèbre extérieure de M et

THÉORÈME.

$$\widetilde{\mathcal{C}\ell}(V) \simeq \operatorname{End}_{\mathbb{C}}(S).$$

Pour $m \in M$ on pose $\rho(m)s = m \wedge s$ et $p \in P$, $\rho(p)s = 2p \lrcorner s$, ρ s'étend à une action de $\widetilde{\mathcal{C}\ell}(V)$, ρ modifie la parité et si $S^+ = \Lambda^+ M$ sont les spineurs pairs et $S^- = \Lambda^- M$ les spineurs impairs alors ρ restreinte à $V_{\mathbb{C}}$ envoie S^+ sur S^-. Si $\widetilde{\mathcal{C}\ell}_0(V)$ sont les éléments pairs de $\widetilde{\mathcal{C}\ell}(V)$ ρ restreinte à $\widetilde{\mathcal{C}\ell}_0$ préserve S^+ et S^- et $\widetilde{\mathcal{C}\ell}_0(V) \simeq \mathrm{End}_{\mathbb{C}}\, S^+ \oplus \mathrm{End}_{\mathbb{C}}\, S^-$.

Dans le cas où n la dimension de V est impaire $n = 2m - 1$ alors $\widetilde{\mathcal{C}\ell}(\mathbb{R}^n) \simeq \widetilde{\mathcal{C}\ell}_0(\mathbb{R}^{n+1})$ en utilisant

$$\rho : \mathbb{R}^n \longrightarrow \mathcal{C}\ell_0(\mathbb{R}^{n+1})$$
$$e_i \longrightarrow e_i e_{n+1}$$

(la multiplication par e_{n+1} enlace S^+ et S^- et on obtient qu'une représentation irréductible).

Après toute cette construction algébrique, il faut réaliser un fibré sur la variété considérée dont la fibre en chaque point est S l'espace des spineurs $= \Lambda M$ construit à partir de l'espace tangent $V_{\mathbb{C}}$ de la variété, complexifiée, muni de la forme quadratique issue d'une structure existante sur la variété, que l'on complexifie elle aussi.

Une telle construction n'existe pas toujours. Si elle est possible (la 2e classe de S.W. étant nulle) alors une section de ce fibré sera un spineur.

Dans le cas où la variété sous-jacente est plate, l'opérateur $\Sigma e_i \cdot \partial_{e_i}$ où le produit $\cdot$ est le produit de Clifford donne l'opérateur de Dirac recherché (Ouf!).

(À partir de la dérivation (plate ici) dans la variété, il faut définir une dérivation sur les spineurs. Cela peut se faire en utilisant une base de l'espace tangent en chaque point (un repère mobile local) et en étendnant la définition "naturellement", puis vérifier que la démarche est cohérente... et s'étend aussi aux représentations ; l'idée est simple, sa formulation est lourde.)

Il faut avec ce formalisme retrouver les spineurs de Pauli et puis les spineurs de Dirac. La variété choisie est la variété *plate* de la relativité restreinte et on suppose que la structure globale permet la construction ci-dessus (donc que sa 2e classe de S.W. soit nulle). En chaque point, on dispose de la forme quadratique de Lorentz (comme ensuite on complexifie puisqu'en mécanique quantique les fonctions d'onde sont à valeurs complexes, on obtient la même chose qu'avec une forme quadratique définie > 0 réelle) ce qui permet de construire l'espace $S = \Lambda M$ dans l'espace tangent en chaque point de la variété.

Comme la construction est faite de sorte que les conjugaisons par les éléments de Γ soient les isométries de V pour sa forme quadratique QW et que $\mathrm{Pin}(V, q)$ et $\mathrm{Spin}(V, s)$ s'obtiennent en normalisant un élément générique de Γ de la forme $v_1, \ldots, v_k$ avec $q(v_i) = \pm 1$, il s'agit bien d'une formalisation de la volonté d'écrire une expression invariante par les isométries de V donc dans le cadre de la relativité restreinte une théorie construite sur les transformations laissant la forme de Lorentz invariante. En dimension 4 réelle avec la

forme quadratique de Lorentz puis en complexifiant on obtient M de dimension 2 complexe et donc l'espace des spineurs est ΛM et est de dimension $2^2 = 4$ complexe. C'est bien la dimension minimale pour les matrices γ recherchées par Dirac.

L'opérateur de Dirac s'exprime à l'aide d'une base par $(\Sigma e_i \cdot \nabla_{e_i})$ et on vérifie que cet opérateur est défini de manière intrinsèque indépendant de la base. Le fibré de spineurs sur la variété plate de la relativité restreinte est somme directe des spineurs positifs S^+ et spineurs négatifs S^-. L'opérateur de Dirac $\mathcal{D}$ échange S^+ et S^- et nous retrouvons bien dans ce cas particulier la forme historique de l'équation de Dirac. Mais ainsi exprimé seul le Spin $\frac{1}{2}$ est possible. Alors en utilisant le modèle utilisé par Pauli (la dimension 3 avec la forme euclidienne, puis complexification) par analogie nous sommes conduits à envisager les représentations *de dimension finie* de $SL(2, \mathbb{C})$. Rappelons que si G est un groupe, T est une représentation de G si T associe à tout élément g de G un élément $T_g \in \mathrm{Aut}\,(F$ espace vectoriel de dimension n sur le corps $K)$ si T vérifie $T_{g_1 g_2} = T_{g_1} \circ T_{g_2}$. $X \subset F$ est dit invariant par T si $\forall g \in G$, $T_g(X) \subset X$. T est dite irréductible si les seules parties invariantes de F sont $\{0\}$ de F. Ces rappels pour dire que les représentations ne doivent pas être unitaires puisque la masse n'est pas constante.

Soit E un espace vectoriel de dimension 2 sur $\mathbb{C}$. C'est un ensemble muni d'une loi $+$ et d'une loi de multiplication externe par $\lambda \in \mathbb{C}$. Alors E muni de la même loi $+$ et de la multiplication $(\lambda, x) \longrightarrow \overline{\lambda} \cdot x$ est aussi un espace vectoriel de dimension 2 sur $\mathbb{C}$ noté $\overline{E}$.

On définit de manière naturelle une représentation de $SL(2, \mathbb{C})$ dans $\bigotimes_p E \bigotimes_q \overline{E}$ mais elle n'est pas irréductible. Pour qu'elle le soit, il faut la faire agir sur $\overset{s}{\bigotimes_p} E \overset{s}{\bigotimes_q} \overline{E}$ où $\overset{s}{\bigotimes}$ est le produit tensoriel symétrisé. On obtient ainsi les dites "représentations spinorielles" de $SL(2, \mathbb{C})$. Naimark a démontré que toute représentation irréductible de dimension finie sur $\mathbb{C}$ de $SL(2, \mathbb{C})$ était équivalente à une représentation spinorielle (T agissant sur F et T' agissant sur F' sont dites équivalentes s'il existe une application inversible V de F dans F' telle que pour tout $g \in G$ $T'_g = V T_g V^{-1}$). Alors un spineur est un tenseur sur $\overset{s}{\bigotimes_p} E \overset{s}{\bigotimes_q} \overline{E}$ qui laisse invariant $s(x \otimes \bar{y}) = y \otimes \bar{x}$ appelée "inversion complexe" chez Landau-Lifchitz. De cette manière, on retrouve la structure formelle des spineurs ΛM si $\dim_{\mathbb{C}} E = 2$. Mais l'on peut aussi de cette manière avoir des spins avec toutes les valeurs $0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ et une équation de Dirac associée. Alors on peut jouer au jeu formel suivant : on dispose de spineurs positifs et négatifs. Dans l'équation de Dirac classique par exemple, on associe un spineur positif et un spineur négatif $\psi = \binom{\xi}{\eta}$ et $\mathcal{D}_+$ associe à ξ, spineur positif, un spineur négatif $\mathcal{D}_+ \xi$ et à η, spineur négatif $\mathcal{D}_-$ associe un spineur positif. L'équation classique de Dirac devient

$$\begin{cases} \mathcal{D}_+ \xi = m\eta \\ \mathcal{D}_- \eta = m\xi \end{cases}.$$

Si $m = 0$, on se rend compte que l'on a besoin que d'un spineur (positif ou négatif puisque η est associé à ξ par "inversion complexe") pour écrire l'équation correspondante. C'est ainsi que l'existence neutrino a été postulée (avant son observation). Pour terminer reprenons l'équation de Maxwell à l'origine de toute cette démarche qui dans le vide s'écrit $dF = \delta F = 0$ où F est une 2-forme sur la variété plate de la relativité restreinte et δ est associé à la forme de Lorentz. En prenant (ξ, η) un couple de 2 spineurs se correspondant dans "l'inversion" avec $p = q = 2$ alors les équations de Maxwell dans le vide deviennent une équation de Dirac

$$\mathcal{D}_- \eta = 0 \text{ ou } \mathcal{D}_+ \xi = 0 .$$

APPENDICE

Axiomatique par le calcul des propositions

La description d'un système physique nécessite deux types d'éléments : les observables et les états du système. Étant donné une observable A, un état α et un borélien Δ de $\mathbb{R}$, le but de la théorie est de calculer la probabilité pour que la mesure de A sur le système dans l'état α donne un résultat appartenant à Δ.

La méthode la plus employée consiste à utiliser les observables comme concept de base. Résumons l'approche introduite par Birkhoff et Von Neumann suivis récemment par Finkelstein, Mackey, Piron et Guenin connue sous le nom de calcul des propositions. Des observables qui ont, soit 1, soit 0, comme valeurs possibles s'appellent des propositions. Soit $\mathcal{L}$ l'ensemble de toutes les propositions (expériences OUI-NON) d'un système physique. $\mathcal{L}$ est partiellement ordonné pour $a \subseteq b$ si a implique b.

* $\subseteq$ vérifie $a \subseteq a$, $a \subseteq b$ et $b \subseteq a$ implique $a = b$, $a \subseteq b$ et $b \subseteq c$ implique $a \subseteq c$.
* si $(a_i)_{i \in I}$ est une famille d'élément de I, il existe une proposition $\bigcap_I a_i$ telle que $(x \subseteq a_i \text{ pour tout } i) \Longleftrightarrow (x \subseteq \bigcap_I a_i)$.

Comme très souvent en mécanique quantique (mesure a et puis b) ou bien (b et puis a) ne donne pas les mêmes résultats. Cet axiome introduit des éléments idéaux dans la description des systèmes physiques

* Pour tout a, il existe a' tel que $(a')' = a$, $a' \cap a = \emptyset$, $a \subseteq b \Leftrightarrow b'' \subseteq a'$.

On obtient un treillis complet orthocomplémenté. a et b sont compatibles si $\{a, b\}$ génère un sous-treillis de boule.

* On postule que $a \subseteq b \Rightarrow a$ et b sont compatibles.
* et que pour $a \neq \emptyset$, il existe une proposition minimale P tel que $P \subseteq a$ et si Q est une proposition minimale alors $a \subseteq x \subseteq a \cup Q \Rightarrow x = a$ ou $x = a \cup Q$.

Si $\mathcal{L}$ vérifie tous ces axiomes $\mathcal{L}$ s'appelle un système de proposition. Les différents travaux de cette axiomatisation montrent que tous les dispositifs expérimentaux définissent un système de proposition $\mathcal{L}$. Un état du système est alors une application p de $\mathcal{L}$ dans $[0, 1]$ telle que $p(I) = 1$ (I plus grand élément de $\mathcal{L}$), $p(\emptyset) = 0$ et $p(\bigvee_i a_i) = \sum_i p(a_i)$ où $a_i \subseteq a'_j$.

Piron a montré qu'on pouvait construire alors un espace de Hilbert H et un iso-

morphisme entre $\mathcal{L}$ et H qui au plus grand élément de $\mathcal{L}$ fait correspondre H au plus petit élément fait correspondre $\{0\}$ où 0 est l'élément nul de H et à chaque élément de $\mathcal{L}$ correspond un sous-espace clos de H, si à a correspond R_a et à b correspond R_b alors à $a \wedge b$ correspond $R_a \cap R_b$, à $a \wedge b$ correspond $\overline{R_a \cup R_b}$, si $a \subseteq b$ alors $R_a \subseteq R_b$ et à $\neg a$ correspond $(R_a)^{\perp}$.

Les résultats fondamentaux sont que :

- tout système irréductible de propositions (irréductible = il n'y a pas de questions non triviales qui soit compatible avec une autre) est isomorphe à l'ensemble des opérateurs de projection sur tous les sous-espace clos possibles de H.
- tout système de question est une union directe de systèmes irréductibles.

Les observables peuvent être reconstruites à partir des propositions ; avec les résultats ci-dessus, cela correspond à la reconstruction d'opérateurs self-adjoints exprimé par le théorème de décomposition spectrale. L'ensemble des observables d'un système irréductible est identifié avec l'ensemble de tous les opérateurs self-adjoints sur H, des observables compatibles et sont représentés par des opérateurs commutant fortement (toutes leurs projections spectrales $E(\Delta)$, pour tout Δ borélien, commutent). Un état p du système (cf. haut de la page), par un théorème dû à Gleason, est identifié à un opérateur de densité et un état pur à un sous-espace de dimension 1 de H.

Axiomatique hilbertienne. Point de vue de Dirac pour le cas classique

H est un espace de Hilbert séparé complet de dimension quelconque (dénombrable au plus).

Postulat 1. — À un instant fixé t_0, l'état d'un système physique est défini par la donnée de $\psi(t_0)$ appartenant à l'espace des états H.

Plus rigoureusement c'est le sous-espace engendré par ψ qui définit l'état notation de Dirac $|\psi(t_0)\rangle$.

Postulat 2. — Toute grandeur physique mesurable $\mathcal{A}$ est décrite par un opérateur A autoadjoint de H.

Postulat 3. — La mesure d'une grandeur physique $\mathcal{A}$ ne peut donner comme résultat qu'une des valeurs propres de l'observable A correspondante.

Postulat 4. — La probabilité, lorsqu'on mesure la grandeur physique $\mathcal{A}$ sur un système dans l'état ψ, d'obtenir une valeur appartenant à Δ (borélien en $\mathbb{R}$) est $\frac{1}{\|\psi\|_H} \int_\Delta |A\psi(\lambda)|^2 \, d\mu(\lambda)$ où $H \longleftrightarrow \int_\Lambda^\oplus \mathbb{C}^{n(\lambda)} \, d\mu(\lambda)$ est la décomposition en intégrale hilbertienne de H associée à l'opérateur A et $\psi \longleftrightarrow (\lambda \mapsto \psi(\lambda))$ l'image de ψ dans $\int_\Lambda^\oplus \mathbb{C}^{n(\lambda)} \, d\mu(\lambda)$.

Postulat 5. — Si la mesure de la grandeur physique $\mathcal{A}$ sur le système dans l'état ψ donne un résultat appartenant à Δ, l'état du système, immédiatement après la mesure, exprimé dans $\int_\Lambda^\oplus \mathbb{C}^{n(\lambda)} \, d\mu(\lambda)$ est $1_\Delta(\lambda)\psi(\lambda)$.

Postulat 6. — L'évolution dans le temps du vecteur d'état $\psi(t)$ est réglée par l'équation de Schrödinger $i\hbar\frac{d}{dt}(\psi(t)) = H(t)\psi(t)$ où $H(t)$ est l'observable associée à l'énergie totale du système.

En fait, ce postulat nécessiterait un exposé à lui tout seul : on peut considérer soit l'équation de Heisenberg, soit celle de Schrödinger, soit un opérateur d'évolution. Sous certaines hypothèses, ces trois points de vue sont équivalents.

Bibliographie

[1] A. Einstein. — *The meaning of relativity*, Princeton University Press, 1953.

[2] P.A.M. Dirac. — *The principle of quantum mechanics*, Clarendon Oxford, 1958.

[3] J. Von Neumann. — *Mathematical foundation of quantum mechanics*, Princeton U.P., Princeton New Jersey, 1955.

[4] W. Pauli. — , Z. Phys. 31 (1925), 765.

[5] R. Jost. — *Pauli memorial volume*, London: interscience, 1960.

Luc ROZOY
INSTITUT FOURIER
Laboratoire de Mathématiques
UMR5582 CNRS–UJF
BP 74
38402 St MARTIN D'HÈRES Cedex (France)

Luc.Rozoy@ujf-grenoble.fr