

Une théorie des électrons et des protons. P. A. M. Dirac

1. Nature de la difficulté de l'énergie négative.

La théorie de la relativité quantique d'un électron bougeant dans un champ électromagnétique, bien qu'ayant remporté du succès à prédire les propriétés de spin de l'électron, rencontre cependant encore une sérieuse difficulté qui montre que des modifications fondamentales doivent lui être apportées avant que nous puissions la regarder comme décrivant précisément la nature. Cette difficulté est liée au fait que la fonction d'onde, qui est de la forme

$$(1) \quad \left[\frac{W}{c} + \frac{e}{c}A_0 + \rho_1 \left(\sigma, \mathbf{p} + \frac{e}{c}\mathbf{A} \right) + \sigma_3 mc \right] \psi = 0,$$

a, en plus des solutions souhaitées pour lesquelles l'énergie cinétique de l'électron est positive, un nombre égal de solutions non souhaitées avec une énergie cinétique négative pour l'électron, qui semblent n'avoir aucun sens physique.

Par conséquent, si nous nous autorisons par facilité un champ électromagnétique stable, l'équation (1) admettra des solutions périodiques de la forme

$$(2) \quad \psi = ue^{-iEt/h},$$

où u est indépendant de t , représentant les états stationnaires, E étant l'énergie totale de l'état, incluant le terme de la relativité mc^2 . Il existera alors des solutions (2) à valeurs négatives pour H aussi bien que des solutions à valeurs positives; en fait, si nous prenons une représentation matricielle des opérateurs $\rho_1\sigma_1, \rho_1\sigma_2, \rho_1\sigma_3$ avec des éléments matriciels tous réels, alors le complexe conjugué de toute solution de (1) sera solution de l'équation d'onde obtenue à partir de (1) en renversant le signe des potentiels A , et soit la fonction d'onde originale, soit sa complexe conjuguée doit faire référence à un E négatif.

La difficulté n'est pas une difficulté spécifique reliée à la théorie quantique de l'électron, mais une difficulté générale qui apparaît dans toutes les théories de la relativité, ainsi que dans la théorie classique. Elle provient du fait fondamental qu'en relativité, dans l'équation de l'Hamiltonien de la théorie classique, notamment,

$$(3) \quad \left(\frac{W}{c} + \frac{e}{c}A_0 \right)^2 - \left(\mathbf{p} + \frac{e}{c}\mathbf{A} \right)^2 - m^2c^2 = 0,$$

il y a une ambiguïté dans le signe de W , ou plutôt de $W + eA_0$. Bien que l'opérateur de la fonction d'onde dans (1) soit linéaire en W , il est également, pour le dire grossièrement, équivalent au côté gauche de (3) et l'ambiguïté de signe persiste. La difficulté n'est pas importante dans la théorie classique, puisque là les variables dynamiques doivent toujours varier continument, de telle façon qu'il y aura une légère distinction nette entre ces solutions des équations de mouvement pour lesquelles $W + eA_0 \geq mc^2$ et celles pour lesquelles $W + eA_0 \leq -mc^2$, et nous pouvons simplement ignorer ces dernières.

P. A. M. Dirac, Lycée St. John, Cambridge. (Communiqué par R. H. Fowler, F.R.S., Reçu le 6 décembre 1929.)
Référence : Proc. R. Soc. Lond. A 1930 126 , 360-365.
traduction Denise Vella-Chemla, janvier 2021.

Nous ne pouvons pas, cependant, outrepasser la difficulté si facilement dans la théorie quantique. Il est vrai que dans le cas d'un champ électromagnétique stable, nous pouvons faire une distinction entre les solutions de (1) de la forme (2) avec E positif et celles avec E négatif et nous pouvons affirmer que seules les premières ont un sens physique (comme cela a été effectivement fait quand la théorie était appliquée pour déterminer les niveaux d'énergie de l'atome d'hydrogène), mais si une perturbation est appliquée au système, elle peut causer des transitions d'un état à un autre. Dans le cas général d'un champ électromagnétique variable, nous ne pouvons pas séparer franchement et rapidement les solutions de l'équation d'onde entre celles correspondant à une énergie positive et celles correspondant à une énergie négative. De plus, dans la théorie quantique précise dans laquelle le champ électromagnétique est aussi sujet à des lois quantiques, des transitions peuvent avoir lieu dans lesquelles l'énergie de l'électron change d'une valeur positive à une valeur négative même en l'absence de tout champ externe, l'énergie en surplus, au moins égale à $2mc^2$ étant spontanément émise sous la forme d'une radiation. (Les lois de conservation de l'énergie et du moment requièrent au moins que deux quanta de lumière se forment simultanément dans un tel processus). Ainsi nous ne pouvons ignorer les états d'énergie négative sans donner naissance à une ambiguïté dans l'interprétation de la théorie.

Examinons les fonctions d'onde représentant les états d'énergie négative d'un peu plus près. Si nous superposons un certain nombre de ces fonctions d'onde de telle manière que l'on obtienne un paquet d'ondes, le mouvement de ce paquet s'effectuera selon une trajectoire classique donnée par l'Hamiltonien (3) avec $W + eA_0$ négatif. Une telle trajectoire, comme on le voit aisément, est une trajectoire possible pour un électron ordinaire (avec énergie positive) se déplaçant dans le champ électromagnétique avec un signe inverse, ou pour un électron de charge $+e$ (et d'énergie positive) se déplaçant dans le champ électromagnétique original. Ainsi *un électron avec énergie négative se déplace dans un champ externe comme s'il portait une charge positive*.

Ce résultat a amené les gens à suspecter une connexion entre l'électron à énergie négative et le proton ou le noyau de l'hydrogène¹. On ne peut, pourtant, simplement affirmer qu'un électron d'énergie négative est un proton, car cela amènerait aux paradoxes suivants :

- (i) Une transition d'un électron d'un état d'énergie positive à un état d'énergie négative serait interprétée comme une transition d'un électron en un proton, ce qui violerait la loi de conservation de la charge électrique.
- (ii) Bien qu'un électron d'énergie négative se déplace dans un champ externe comme s'il avait une charge positive, à nouveau, comme on peut facilement le voir en considérant la conservation du moment, le champ qu'il produit doit correspondre au fait qu'il a une charge négative, e.g., l'électron de charge négative repousse un électron ordinaire de charge positive bien qu'il soit lui-même attiré par l'électron d'énergie positive.
- (iii) Un électron d'énergie négative aura moins d'énergie plus il se déplace rapidement et il devra absorber l'énergie de façon à être amené au repos. Aucune particule de cette nature n'a jamais été observée.

1. Voir par exemple, Weyl, Z. f. Phys., vol. 56, p. 332 (1929).

Des considérations plus précises sur les conditions que l'on devrait s'attendre à voir vérifiées dans le monde réel suggèrent que la connexion entre les protons et les électrons de charge négative devrait être basée sur quelque chose de différent et ces nouvelles idées vont s'avérer éliminer toutes les difficultés sus-mentionnées.

2. Solution de la difficulté de l'énergie négative.

Les états les plus stables pour un électron (i.e., les états d'énergie la plus basse) sont ceux d'énergie négative et de très grande vitesse. Tous les électrons dans le monde tendent à tomber dans ces états en émettant une radiation. Le principe d'exclusion de Pauli, pourtant, entrera en jeu et empêchera que plus d'un électron ne soit dans un état quelconque. Supposons qu'il y ait tant d'électrons dans le monde que tous les états les plus stables soient occupés, ou, plus précisément que *tous les états d'énergie négative soient occupés exceptés peut-être quelques-uns de petite vitesse*. Quelques électrons avec énergie positive auront maintenant très peu de chances de sauter vers des états d'énergie négative et se comporteront comme les électrons que l'on observe en laboratoire. Nous aurons un nombre infini d'électrons dans des états d'énergie négative, et en effet un nombre infini par unité de volume partout dans le monde, mais si leur distribution est exactement uniforme, nous devrions nous attendre à ce qu'ils soient complètement inobservables. *Nous pourrions juste espérer observer les seuls petits départs de l'uniformité exacte, amenés par quelques-uns des états d'énergie négative inoccupés.*

Examinons les propriétés des états vacants ou "trous". Le problème est analogue à celui des niveaux de rayons X dans un atome avec plusieurs électrons. Selon la théorie habituelle des niveaux de rayons X, le trou qui est formé lorsque l'un des électrons internes de l'atome est supprimé peut être décrit comme une orbite et est dessiné comme l'orbite de l'électron avant qu'il ait été supprimé. Cette description peut être justifiée par la mécanique quantique, en supposant que l'orbite est regardée, non pas au sens de Bohr, mais comme quelque chose qui est représentable, en dehors du spin, par une fonction d'onde tri-dimensionnelle. Par conséquent, le trou ou la vacance dans une région qui est sinon saturée d'électrons est à peu près la même chose qu'un seul électron dans une région qui leur est normalement dévolue.

Dans le cas du rayon X, les trous devraient être comptés comme des choses d'énergie négative, puisque pour en faire disparaître un (i.e., pour le remplir), on doit juste lui ajouter quelque chose. C'est juste le contraire qui est vérifié, pour les trous dans notre distribution d'électrons d'énergie négative. Ces trous seront des choses d'énergie positive et seront à cet égard comme des particules ordinaires. De plus, le mouvement de l'un de ces trous dans un champ électromagnétique externe sera le même que celui de l'électron d'énergie négative qui le remplira, et correspondra ainsi à sa possession d'une charge $+e$. Nous sommes alors amenés à la supposition que *les trous dans la distribution d'électrons d'énergie négative sont les protons*. Quand un électron d'énergie positive tombe dans un trou et le remplit, nous avons un électron et un proton qui disparaissent ensemble avec émission d'une radiation.

Une difficulté surgit quand nous considérons le champ produit par la distribution des électrons d'énergie négative. Il y a une densité infinie d'électricité qui, selon l'équation de Maxwell

$$(4) \quad \operatorname{div} \mathbf{E} = -4\pi\rho,$$

devrait produire un champ électrique de divergence infinie. Il semble naturel, pourtant, d'interpréter le ρ dans l'équation de Maxwell (4) comme le départ à partir de l'état normal d'électrification du monde, dont l'état normal d'électrification est, selon la théorie présente, celui où tout état électronique d'énergie négative et aucun état d'énergie positive n'est occupé. Ce ρ consistera alors en une charge $-e$ provenant de chaque état d'énergie positive qui est occupé, ainsi qu'à une charge $+e$ provenant de tout état d'énergie négative qui est inoccupé. Ainsi, le champ produit par un proton correspondra au fait qu'il ait une charge $+e$.

De cette manière, nous pouvons passer outre les trois difficultés mentionnées à la fin de la section précédente. Nous avons juste besoin de postuler une seule sorte de particule fondamentale, au lieu de deux, les électrons et les protons, qui étaient précédemment nécessaires.

La simple tendance de toutes les particules à aller dans des états d'énergie la plus faible a pour conséquence toutes les choses distinctes dans la nature qui ont une énergie positive.

La théorie présente peut-elle rendre compte de la grande dissymétrie entre les électrons et les protons, qui se manifeste à travers les différentes masses et le pouvoir qu'ont les électrons de se combiner pour former des noyaux atomiques plus lourds ? Il est évident que la théorie fournit, dans une large mesure, la symétrie entre les électrons et les protons. Nous pouvons échanger leurs rôles et affirmer que les protons sont les particules réelles et que les électrons sont simplement les trous dans la distribution des protons d'énergie négative. La symétrie n'est pas, pourtant, mathématiquement parfaite quand on prend en compte l'interaction entre les électrons. Si l'on néglige l'interaction, l'Hamiltonien décrivant le système global sera de la forme $\sum H_\alpha$, où H_α est l'Hamiltonien ou l'énergie d'un électron dans l'état α et la somme est effectuée sur tous les états occupés. Cela diffère seulement par une constante (i.e., par quelque chose d'indépendant de quels sont les états occupés) à partir de la somme $\sum (-H_\alpha)$ calculée sur tous les états inoccupés. Ainsi, nous obtenons formellement le même système dynamique si nous considérons les états inoccupés ou les protons, chacun contribuant par un terme $-H_\alpha$ dans l'Hamiltonien. D'un autre côté, si nous prenons en compte l'interaction entre les électrons, nous obtenons un terme supplémentaire de la forme $\sum V_{ab}$, dans l'Hamiltonien, la somme étant calculée sur toutes les paires d'états occupés (a, b) , et cela n'est pas équivalent à n'importe quelle somme prise sur des paires d'états inoccupés. L'interaction devrait ainsi donner un Hamiltonien profondément différent si nous regardons les protons comme les particules réelles qui occupent les états.

Les conséquences de cette dissymétrie ne sont pas très faciles à calculer pour les résultats relativistes, mais nous espérons qu'elles amèneront finalement à une explication des masses différentes du proton et de l'électron. Une théorie possiblement plus parfaite de l'interaction, basée peut-être sur les calculs d'Eddington² de la constante de structure fine e^2/hc , est nécessaire avant qu'un tel résultat puisse être obtenu.

2. Eddington, Roy. Soc. Proc., A, vol. 122, p. 358 (1929).

3. Application à la diffraction.

Comme application élémentaire des idées précédentes, nous pouvons considérer le problème de la diffraction d'une radiation par un électron, libre ou limité. Un processus de diffraction, selon la théorie, devrait être considéré comme un double processus de transition, consistant d'abord en une absorption d'un photon simultanée au saut de l'électron dans un certain état, suivie d'une émission avec l'électron sautant dans son état final, ou bien autrement, par d'abord l'émission et ensuite l'absorption.

Nous devons donc considérer à la fois trois états du système global, l'*état initial* avec un photon incident et l'électron dans son état initial, un *état intermédiaire* avec soit les deux ou bien aucun photon existant et l'électron dans n'importe quel état, et l'*état final* avec le photon diffracté et l'électron dans son état final. Les états initial et final du système global doivent avoir la même énergie totale, mais l'état intermédiaire, qui dure seulement une très courte durée, peut avoir une énergie considérablement différente.

La question qui se pose maintenant est celle de savoir comment interpréter la manière dont cette diffraction s'effectue lorsque l'état intermédiaire est un état d'énergie négative de l'électron.

Selon les idées précédentes, ces états intermédiaires n'ont pas de sens physique réel, et il serait donc douteux que le processus de diffraction qui advient selon leur agencement devrait être inclus dans la formule du coefficient de diffraction. Cela présente une sérieuse difficulté, puisque dans des cas pratiques importants, presque toute la diffraction provient des états intermédiaires à énergie négative pour l'électron³. En fait, pour un électron libre et une radiation de basse fréquence, quand la formule classique est vérifiée, l'entière de la diffraction provient de tels états intermédiaires.

Selon la théorie du présent article, il est absolument interdit, par le principe d'exclusion, pour un électron, de sauter dans un état d'énergie négative, de telle façon que le processus de double transition avec états intermédiaires d'énergie négative pour l'électron doit être exclu. Nous avons maintenant, cependant, une autre sorte de double transition qui peut avoir lieu, notamment, dans laquelle d'abord un électron d'énergie négative saute dans l'état final de l'électron avec absorption (ou émission) d'un photon, et alors, l'électron d'énergie positive originale saute dans le trou formé par la première transition avec émission (ou absorption) d'un photon. De tels processus aboutissent à un état final du système global indistinguable de l'état final obtenu par des processus plus directs, dans lesquels le même électron fait deux sauts successifs. Ces nouveaux processus compensent ceux des processus plus directs qui sont exclus sur la base de l'état intermédiaire de charge négative pour l'électron, puisque les éléments de la matrice qui déterminent les probabilités de transition sont juste les mêmes dans les deux cas, bien qu'ils entrent en jeu dans l'ordre inverse. De cette façon, les anciennes formules de diffraction, dans lesquelles les états intermédiaires sont exclus, peuvent être justifiées.

3. Je suis reconnaissant à I. Waller d'avoir attiré mon attention sur cette difficulté.