

UNIVERSITÉ DE PARIS

LES DEBUTS
DE LA THEORIE QUANTIQUE
DES CHAMPS
(1925 - 1948)

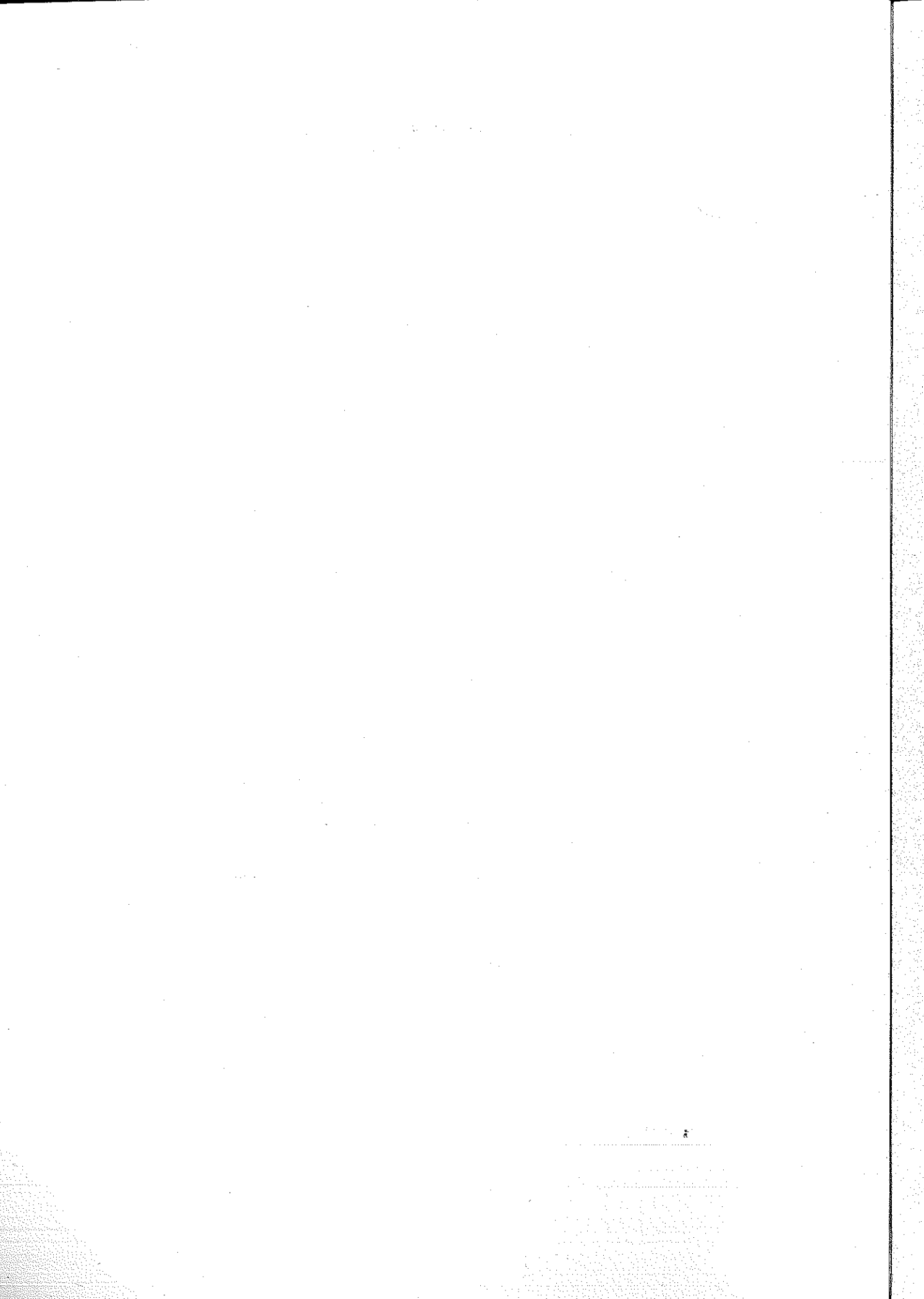
=====

THÈSE POUR LE DOCTORAT DE TROISIÈME CYCLE
PRÉSENTÉE À L'UNIVERSITÉ DE PARIS I
(PANTHÉON-SORBONNE)

PAR

OLIVIER DARRIGOL

SPÉCIALITÉ : HISTOIRE DE LA PHYSIQUE



UNIVERSITE DE PARIS

LES DEBUTS
DE LA THEORIE QUANTIQUE
DES CHAMPS
(1925 - 1948)

THESE POUR LE DOCTORAT DE TROISIEME CYCLE
PRESENTEE A L'UNIVERSITE DE PARIS I
(PANTHEON - SORBONNE)

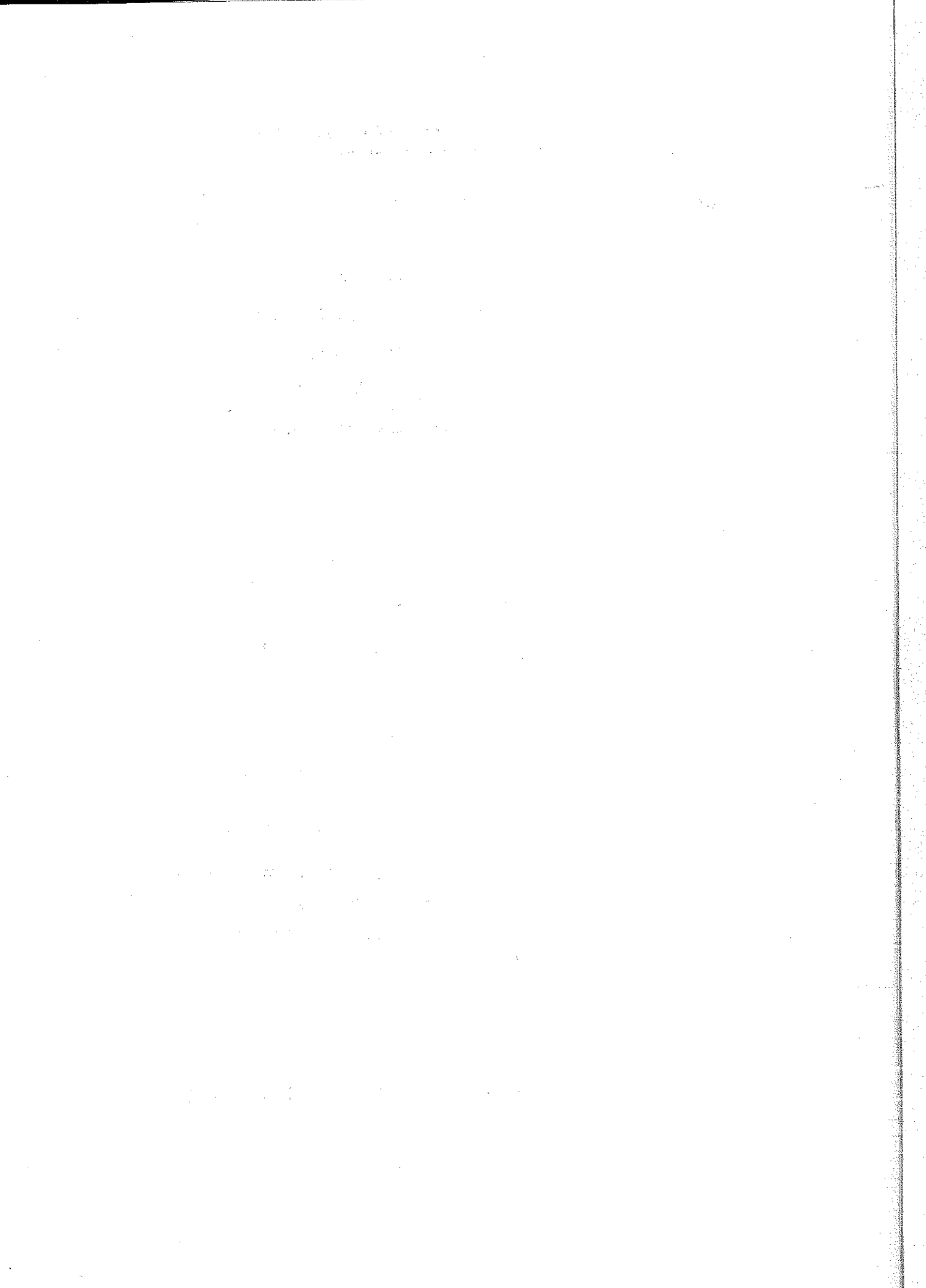
par

OLIVIER DARRIGOL

DATE DE SOUTENANCE : 27 Septembre 1982

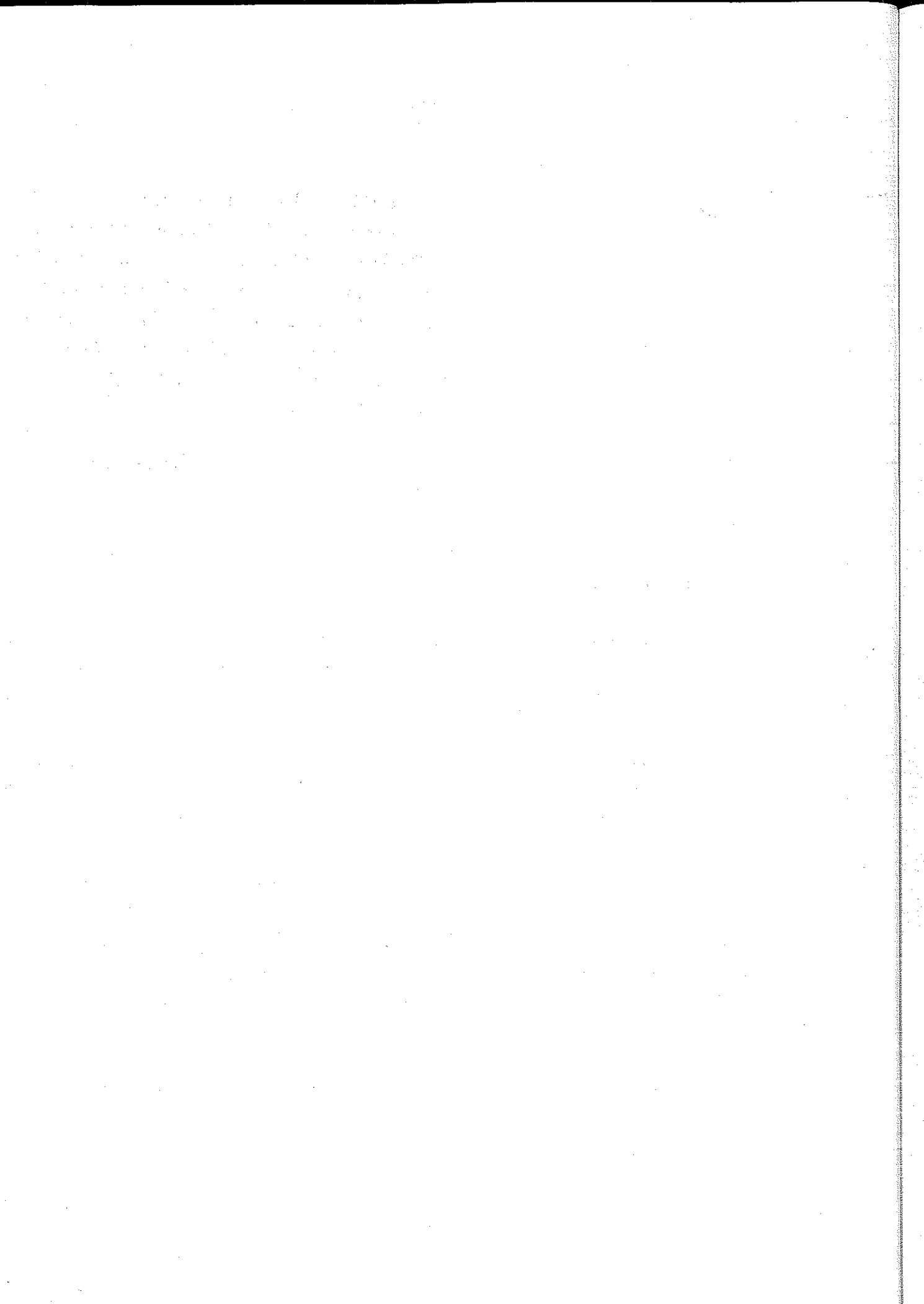
JURY : Président : M. B. D'ESPAGNAT
 Mlle S. BACHELARD
 M. C. COHEN-TANNOUDJI
 M. C. ITZYKSON

SPECIALITE : HISTOIRE DE LA PHYSIQUE



Et par là-bas s'en furent, au bruit d'élytres
de la terre, les grands Itinérants du songe et de
l'action : les Interlocuteurs avides de lointains
et les Dénonciateurs d'abîmes mugissants, grands
Interpellateurs de cimes en exil et Disputeurs de
chances aux confins, qui sur les plaines bleuis-
santes menaient un oeil lonptemps froncé par l'ar-
neau des lunettes.

SAINT-JOHN PERSE



REMERCIEMENTS

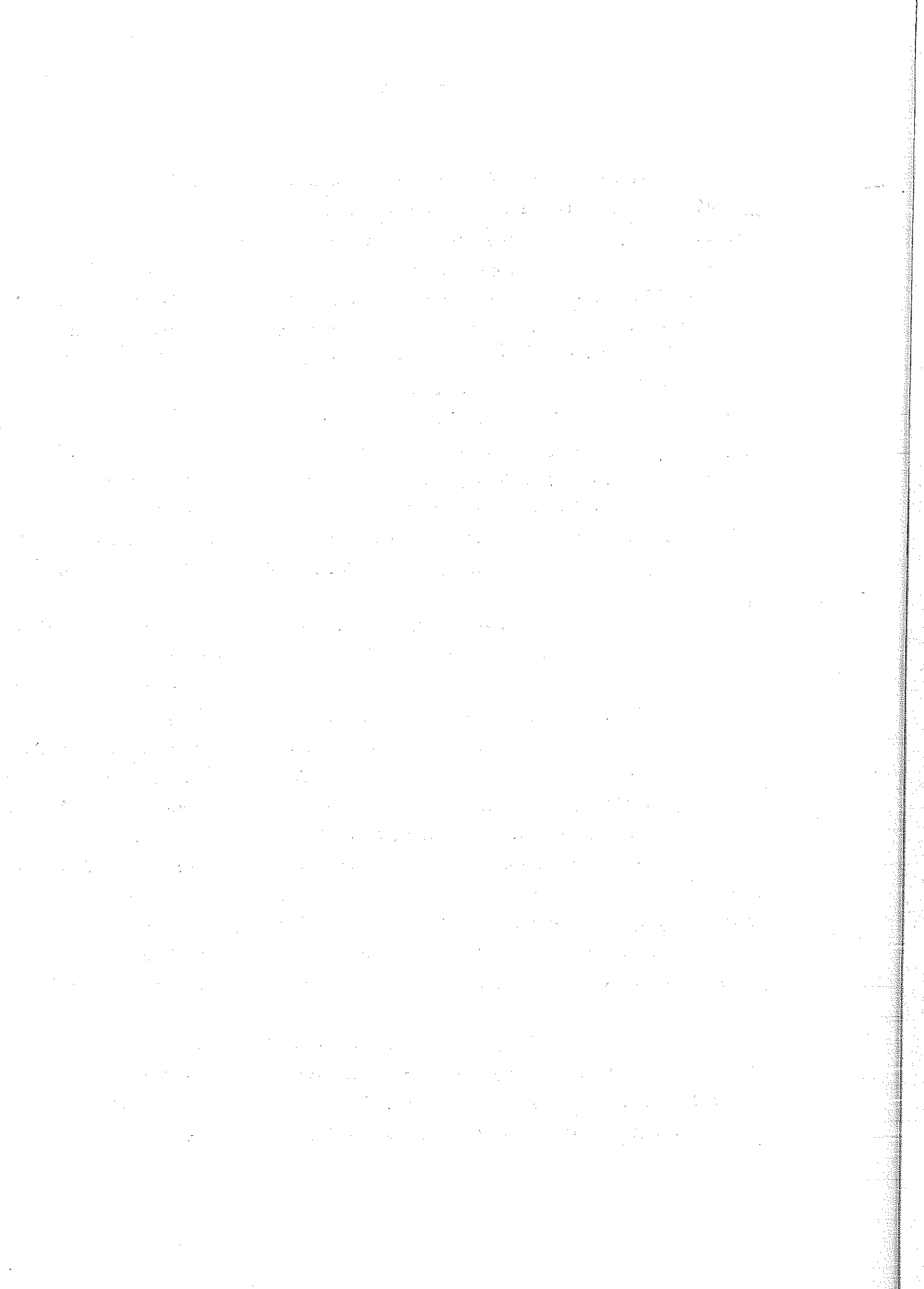
Ce travail se situe à la frontière de deux disciplines : la physique et l'histoire, aussi dois-je doublement remercier plusieurs spécialistes, plus historiens, plus physiciens ou plus philosophes que moi, de s'être intéressés à mon travail malgré mes difficultés d'expression. L'accueil chaleureux du Professeur Suzanne BACHELARD à l'Institut d'Histoire des Sciences de la rue du Four m'a doublement servi : j'ai pu à travers les cours et les séminaires m'initier à une lecture critique de textes historiques et rencontrer d'autres étudiants d'origines et d'intérêts très divers. Le Professeur Bernard d'ESPAGNAT, au tout début de mes recherches, a généreusement accepté de diriger et d'encourager ce travail de thèse. Ses critiques, souvent fondamentales, m'ont permis de rectifier quelques erreurs et d'élargir mon regard aux détours d'une histoire complexe et enchevêtrée. Bien sûr, les défauts de méthode et d'écriture m'incombent totalement : mon entreprise était trop ambitieuse et trop rapide pour que l'intégrité des conseils reçus soit respectée.

Il me fallait pour étudier vingt années d'une histoire très riche, beaucoup de temps et une activité complémentaire compatible; le Professeur Jean BROSSEL m'a fait l'honneur de sa confiance et m'a accueilli comme "caïman" dans son laboratoire. Je dois aussi ce privilège au soutien de Monsieur Jean-Louis VIALLE sans lequel ce travail n'aurait pas même été envisageable.

Avant que j'étudie l'histoire des champs quantiques, d'éminents professeurs m'ont enseigné la version moderne de ces théories. Mes principaux maîtres en la matière furent Claude COHEN-TANNOUJJI et Claude ITZYKSON.

Pour les recherches de sources historiques, je remercie les chercheurs historiens qui m'ont donné des éléments d'information ou qui m'ont communiqué leur travail, notamment David CASSIDY, Armin HERMANN, Karl von MEYENN, Michel PATY, Silvan SCHWEBER, Bruce WHEATON. Certains des "happy few" : Les Professeurs Hans BETHE et Rudolf PEIERLS m'ont fait l'honneur de répondre à mes questions.

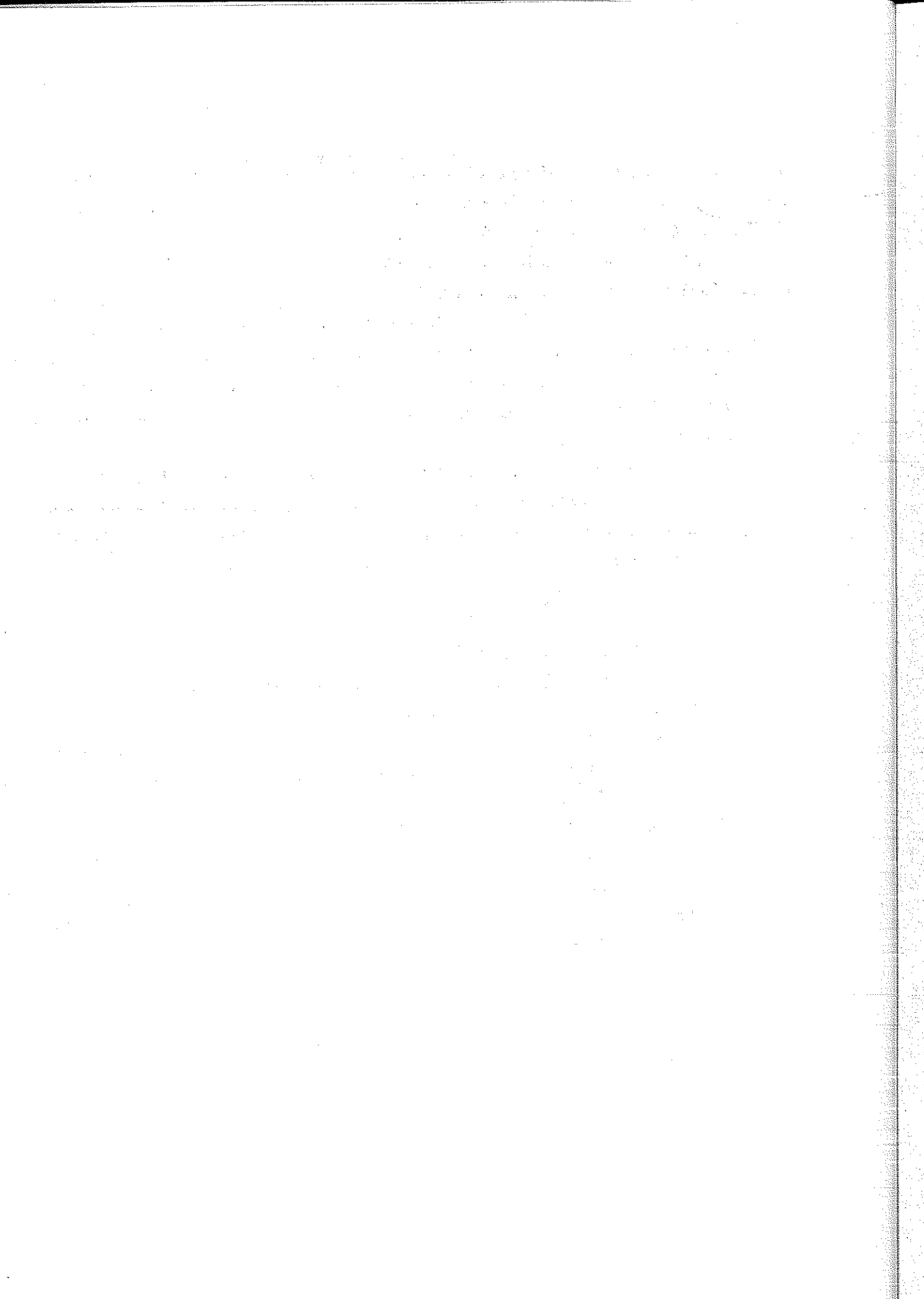
Je ne sousestime pas l'importance de la fréquentation de séminaires apparemment éloignés de mes préoccupations premières ; j'y trouverai, je l'espère la méthode historique utile à mes nouvelles recherches. Le Professeur Ernest COUMET, au Centre d'Histoire des Sciences Alexandre Koyré, m'a montré



quels jugements d'historien ou d'épistémologue (notamment Koyré et Duhem) transcendaient une période précise. Dans le cercle d'Histoire des Sciences de l'Ecole Normale, mon travail fut régulièrement stimulé et discuté par Alain ASPECT, Pierre CHAVEL, Claude FABRES, Marie FARGES, Philippe GRANGIER, Claude GODRECHE et Jean-Michel RAIMOND.

Ce dernier et Pierre DROSSART m'ont courageusement offert de décrypter le manuscrit de cette thèse. Malgré les difficultés diverses, les secrétaires responsables de la frappe : Mesdames D'AUNAY, BERNABE, CALVET, EMO, PILLET m'ont donné par leur ardeur à la tâche une preuve de leur sympathie.

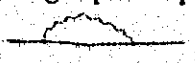
Madame AUDOUIN, bibliothécaire de l'E.N.S., doit être deux fois gratifiée de ma reconnaissance : elle m'a beaucoup facilité l'accès aux ouvrages anciens, et elle a de bonne humeur assuré le long tirage de cette thèse.



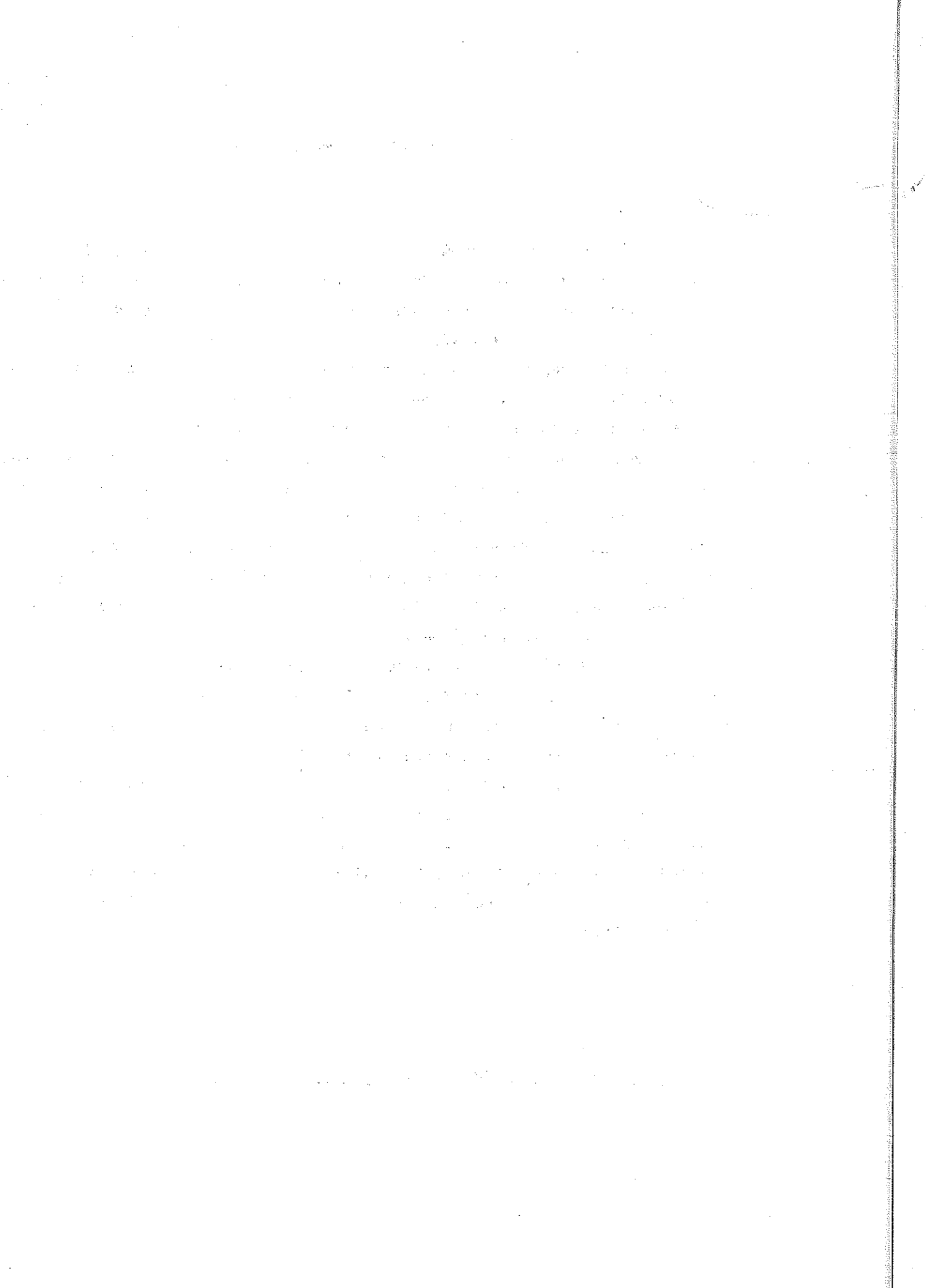
Notations, conventions, formules

Celles des notations utilisées qui ne sont pas précisées dans le texte, sont d'usage conventionnel dans la théorie moderne des champs. Le lecteur les trouvera dans le cours d'Itzykson et Zuber, cité comme Réf. I*.

Sauf exception précisée dans le texte, nous avons transcrit toutes les formules d'époque en notations modernes uniformisées, tant pour éviter de longues définitions de symboles que pour aider la compréhension. Nous pensons, comme la plupart des historiens des sciences, que cette pratique est généralement dangereuse. Si nous l'avons adoptée, c'est en raison de la proximité de la période étudiée : les changements introduits se limitent le plus souvent à une permutation des lettres de l'alphabet ou à des abréviations (par exemple $p_\mu \gamma^\mu = \not{p}$), et dans le cas contraire nous avons pris la précaution d'indiquer la portée de ces changements (significative par exemple quand Fermi utilisa des notations non relativistes dans son électrodynamique quantique).

Dans un souci de légèreté nous avons plusieurs fois fait usage de "graphes de Feynman" pour préciser l'analogie moderne de certains processus. Qu'il soit dit nettement que la seule fonction de ces graphes que nous utilisons est purement abrégative : par exemple  est synonyme de "un photon virtuel est émis puis réabsorbé par un électron". Les premiers calculs de théorie des champs se faisaient dans l'ancienne théorie des perturbations dont la "diagrammatique" précise diffère sensiblement de celle de Feynman. L'incidence historique des graphes de Feynman comme visualisation de calculs complexes est postérieure à la période que nous étudions.

* "Quantum Field Theory" Itzykson et Zuber McGraw Hill 1980



CHAPITRE 0

L'EVOLUTION DES CONCEPTS ET DES METHODES DE LA THEORIE QUANTIQUE

DES CHAMPS DE 1927 à 1948.

Introduction. Résumé-commentaire. Conclusion.

Un jeune spécialiste français de théorie des champs me confiait récemment que la période des années trente-quarante lui semblait complexe et obscure. Il avait lu, comme les mieux informés de ses collègues, Schwinger Feynman et Dyson dans le texte mais n'avait pas remonté le temps davantage, puisque la théorie des champs moderne est l'héritière directe de ces travaux datant de 1947-1949. La physique théorique d'aujourd'hui est devenue si vaste et si complexe que par économie de temps et de pensée, on en arrive à ne plus s'interroger sur les fondements historiques mêmes d'une théorie considérée par beaucoup comme la plus fondamentale de toutes. Profitant du transfert récent de revues anciennes des caves sombres de l'E.N.S. en des lieux plus propices, j'ai cherché à comprendre la genèse de la théorie quantique des champs moderne. Je n'ai pu à Paris consulter qu'une très faible partie des correspondances utiles à cette étude, et j'ai remis à plus tard les voyages lointains qui me permettraient d'interroger les acteurs vivants de cette époque. Ce travail n'est donc qu'une ébauche d'investigations plus complètes qui tiendraient compte plus finement des échanges entre physiciens et des courants d'ordre psychologique ou sociologique. Malgré l'ambition excessive de suivre sur une période aussi longue (1926 à 1947) l'évolution conceptuelle d'une théorie, je n'ai pas cherché à limiter cette étude à la seule électrodynamique quantique, ce qui conduirait à une vision complètement distordue. On verra plutôt s'entremêler de manière complexe des problèmes d'électrodynamique pure, d'autres champs quantiques, de physique nucléaire, de physique des rayons cosmiques, et un débat philosophique sur la mesure.

Le problème de la quantification du champ de rayonnement peut être considéré comme le plus vieux problème de physique quantique, si l'on démarre son histoire avec les travaux de Planck de 1900 sur le corps noir : en effet un corps noir est une enceinte contenant du rayonnement, et il se trouve que l'étude de ce rayonnement à l'équilibre thermodynamique révèle des propriétés quantiques. La théorie de Maxwell du rayonnement était relativiste avant la relativité, et le rayonnement du corps noir était quantique avant la théorie des quanta lumineux. Le problème de la synthèse de la relativité et de la mécanique quantique était implicitement posé, et il fut

abordé dès qu'en 1925-27 la mécanique quantique fut constituée. Les interactions entre particules devaient se propager à une vitesse inférieure à celle de la lumière, d'où la notion de champ satisfaisant à des équations de propagation invariantes relativistes, et il fallait aussi expliquer l'aspect corpusculaire de ces champs comme les photons associés au champ électromagnétique, d'où la notion de champ quantique. Les applications de la mécanique quantique à la physique atomique (où l'on pouvait négliger certains effets relativistes) se poursuivirent sans entrave. En revanche l'unification des idées quantiques et relativistes est difficile, à tel point qu'une solution en partie satisfaisante de ce problème dut attendre 1947. Heureusement nous verrons que les concepts les plus importants de la théorie quantique des champs furent disponibles bien avant cette solution, et ils permirent un progrès rapide de la physique des particules naissante.

Dans ce chapitre nous tenterons de donner une vue d'ensemble non technique de l'histoire conceptuelle des débuts de la théorie quantique des champs et nous tenterons de dégager les présupposés, les motivations et les craintes de ses créateurs ainsi que les traits spécifiques de la crise des infinis.

0.1. DANS L'EUPHORIE DE LA NOUVELLE MECANIQUE QUANTIQUE :

La quantification du rayonnement libre par Jordan (1925)

En 1925-26, la mécanique des matrices de HEISENBERG, BORN et JORDAN se développa suffisamment pour s'appliquer à des systèmes physiques simples. Sans entrer dans les détails de ses règles de calcul, disons qu'elle permettait de faire correspondre à certains problèmes de la mécanique classique de nouveaux problèmes de mécanique quantique grâce à l'introduction de la constante de Planck et en calculant des tableaux de nombres, les matrices, à la place de variables classiques comme les coordonnées d'une particule ou son impulsion. Le système mécanique le plus simple qu'on puisse imaginer après la particule libre est par exemple un pendule effectuant des oscillations de faible amplitude autour de sa position d'équilibre. Dans la théorie classique l'énergie du pendule peut prendre n'importe quelle valeur, imposée par exemple en lâchant le pendule après l'avoir écarté de sa position d'équilibre. Dans la théorie quantique, on démontre qu'une mesure de l'énergie d'un tel oscillateur ne peut donner que des valeurs discrètes $\frac{1}{2} h\nu, \frac{3}{2} h\nu, \frac{5}{2} h\nu, \text{etc.}$ où ν représente la fréquence de l'oscillateur et h la constante de Planck.

Grâce à la mécanique quantique une nouvelle lumière éclaircit toutes sortes de vieux problèmes que l'ancienne théorie quantique ne pouvait traiter qu'avec "la baguette magique" du principe de correspondance. Pascual JORDAN, à la fin d'un des articles fondateurs de la mécanique des matrices (la "Dreimännerarbeit" de Born, Heisenberg et Jordan)⁽¹⁾ pensa à étudier la quantification du rayonnement libre (en l'absence de toute charge). Le champ électromagnétique libre, c'était bien connu, était formellement équivalent à un ensemble d'oscillateurs du type que nous venons de décrire, et la quantification devait pouvoir se faire sans grand effort. L'intérêt de ce problème était évident : la première manifestation de propriétés quantiques du rayonnement datait de l'étude du corps noir. PLANCK⁽²⁾ pensa en 1900 que le quantum d'action h devait être introduit au niveau des particules chargées interagissant avec le rayonnement. EINSTEIN⁽³⁾ en 1905 eut l'audace de postuler une structure quantique du rayonnement-même : un rayonnement de fréquence ν devait être constitué de "quanta" d'énergie

$h\nu$. EHRENFEST⁽⁴⁾ en 1906 et à nouveau DEBYE⁽⁵⁾ en 1910 montrèrent que la distribution d'équilibre de l'énergie du corps noir pouvait être simplement obtenue en supposant que les oscillateurs virtuels équivalents au champ électromagnétique ne pouvaient prendre que des énergies discrètes. La nouvelle mécanique quantique, en établissant cette dernière propriété redonnait immédiatement la loi de Planck. Jordan vit aussi qu'elle permettait d'aller plus loin en donnant un nouveau calcul des fluctuations du rayonnement noir. De même que le nombre de particules fluctue dans un petit volume de gaz à l'équilibre thermodynamique, de même l'énergie du rayonnement contenu à l'intérieur d'un petit volume partiel d'un corps noir devait fluctuer en raison de l'entrée et de la sortie des quanta lumineux. Et EINSTEIN⁽⁶⁾ utilisa en 1909 un calcul statistique des fluctuations dérivées de la loi de Planck pour justifier son hypothèse des quanta. Les mêmes fluctuations devaient pouvoir être calculées par une autre méthode : en suivant les fluctuations temporelles lors de l'évolution libre d'un état du rayonnement. Ni la théorie classique du rayonnement, ni l'ancienne théorie quantique, ne pouvaient donner la valeur correcte de ces fluctuations. Jordan, grâce à ses matrices, put retrouver la formule d'Einstein et montrer qu'elle devait effectivement contenir un terme purement quantique lié à la structure corpusculaire du rayonnement.

Trois programmes de quantification des champs

En 1927, grâce aux travaux de HEISENBERG, SCHRÖDINGER et DIRAC, la mécanique quantique était devenue en l'espace de deux années un édifice théorique solide et puissant. La formulation de Dirac, la plus générale et la plus proche de la version moderne enseignée aujourd'hui, ouvrait un vaste champ d'applications. L'interprétation physique du formalisme, après le congrès Solvay de 1927, semblait débarrassée de toute ambiguïté, et le débat philosophique sur la complétude de la mécanique quantique se séparait déjà de l'usage pratique de la théorie. Dans la même année et au cours de la suivante furent commencés trois programmes de quantification des champs, suivant trois stratégies différentes.

- Le premier à quantifier un champ en interaction : le champ électromagnétique interagissant avec des atomes, fut Paul Adrien Maurice DIRAC.

Le but lointain de Dirac était d'unifier théorie quantique et relativité. Un tel programme, pensait Dirac, ne pouvait aboutir d'un seul coup ; il fallait, comme l'ingénieur qu'il allait devenir, résoudre pas à pas les difficultés. Le premier pas devait être la quantification du champ électromagnétique, puisque la relativité exigeait la propagation des interactions via un champ. Le programme de Dirac, riche de nombreuses découvertes parmi les plus grandes de ce siècle, continua même après guerre et après la théorie de la renormalisation (dont Dirac ne fut pas pleinement satisfait).

- Au début de 1927 Werner HEISENBERG et Wolfgang PAULI, par un échange de lettres, s'engagèrent dans la construction d'une théorie quantique des champs. Le projet était plus vaste que celui de Dirac, l'électrodynamique quantique n'étant qu'un cas particulier, et si Heisenberg, Pauli (et les autres théoriciens engagés dans ce programme : JORDAN, Oskar KLEIN, et Robert OPPENHEIMER) procédèrent finalement comme Dirac, par étapes, ce ne fut qu'en raison de la difficulté imprévue du sujet. Pauli fut même tenté un moment (en 1928) d'abandonner et d'écrire un roman utopique.... La collaboration étroite de Heisenberg et Pauli autour de ce programme cessa en 1930 avec le problème des infinis.

- L'ambition de Enrico Fermi se limita momentanément à l'électrodynamique quantique. De 1928 à 1930 il reformula la théorie de Dirac de 1927 sous une forme "plus familière" et plus générale. Fermi se souciait bien plus de la familiarité des mathématiques utilisées et de leur facilité d'application à des problèmes concrets que d'exigences formelles comme l'invariance relativiste. Pourtant il donna la solution moderne la plus élégante de l'obstacle majeur que Heisenberg et Pauli rencontraient.

La base formelle de l'électrodynamique moderne est née pendant ces trois années de l'interaction de ces trois programmes.

La théorie de Dirac de 1927

En février 1927 parut la première application des règles de la mécanique quantique à un système de charges interagissant avec le rayonnement. Ce fut "La théorie quantique de l'émission et de l'absorption du rayonnement" de Dirac⁽⁷⁾. La quantification du rayonnement libre, nous l'avons vu avec

Jordan, donnait une énergie somme de contributions $(n + \frac{1}{2}) h\nu$ avec n entier, et ν fréquence des oscillateurs virtuels variant de zéro à l'infini. Dans les problèmes de statistique, dire que l'énergie d'un oscillateur ne peut varier que de quantités $nh\nu$ revient à dire que l'on "remplit" à la manière de Bose⁽⁸⁾ l'oscillateur avec des quanta indiscernables de valeur $h\nu$. La quantification donnait donc à la fois la structure quantique du rayonnement et l'interprétation de la mystérieuse indiscernabilité des quanta lumineux. La première tâche de Dirac fut de préciser le lien entre champ quantifié et ensemble de particules de Bose (bosons). La dynamique d'une onde quantifiée était mathématiquement équivalente à celle d'un ensemble de bosons. L'utilisation des ondes quantifiées pour décrire un tel ensemble fut appelée par Dirac 2° quantification en pensant au cas où l'onde quantifiée était l'onde de Schrödinger.

Dans le cas particulier du champ de rayonnement traité dans la suite de l'article, les bosons étaient les quanta lumineux. L'effet de l'interaction avec les atomes devait être traité comme une petite perturbation ; c'est tout ce qu'on sait faire même aujourd'hui et le procédé est en partie justifié par la petitesse de la constante de couplage $\propto \frac{1}{137}$. L'effet de la perturbation est de faire varier le nombre de quanta lumineux et corrélativement d'opérer des transitions de l'atome d'un niveau à un autre. Ainsi le formalisme contenait la notion de création et d'annihilation de photon*. Pour l'immédiat cette nouvelle théorie de Dirac n'apportait pas de nouvelles prédictions expérimentales, mais elle intégrait dans un système hypothético-déductif rigoureux, les résultats qu'avait devinés plus que démontrés l'ancienne théorie des quanta. Une des contradictions logiques les plus gênantes de cette ancienne théorie était liée au calcul de l'émission spontanée : les états excités des atomes, stables dans le modèle de Bohr, devaient pourtant pouvoir se désintégrer en émettant de la lumière. Le calcul de leur durée de vie nécessitait le recours au principe de correspondance. En 1927 grâce à Dirac la même formule était cette fois démontrée. La théorie classique donnait elle aussi une durée de vie finie pour les atomes car l'électron devait rayonner au cours de son mouvement autour du noyau,

* Le terme de photon, introduit par la chimiste G.N. Lewis en 1926, (qui proposait que le nombre de photons soit conservé lors de tout processus atomique) apparut dans le titre du conseil SOLVAY de 1927: "électrons et photons".

mais elle n'expliquait pas l'évidence spectroscopique des niveaux discrets. La théorie de Dirac permettait de concilier ces deux aspects.

Préliminaires à la théorie de Heisenberg et Pauli :

Aussitôt parue, la théorie de Dirac de 1927 suscita d'importantes généralisations en Allemagne. Jordan participa aux quatre articles qui marquent le point de départ du programme de Heisenberg et Pauli. Une première généralisation, due à JORDAN⁽¹¹⁾ puis sous une forme plus précise et plus correcte à Jordan et WIGNER⁽¹²⁾, concernait l'application de la 2^o quantification à des fermions. La statistique de Fermi-Dirac interdisait à deux particules indiscernables de se trouver dans le même état quantique. Grâce à une définition convenable des opérateurs de création et d'annihilation des fermions, la structure granulaire de la matière (constituée essentiellement de fermions : électrons et protons en 1927) s'expliquait par la quantification d'ondes de matière, tout comme les quanta lumineux résultaient de la quantification des ondes électromagnétiques. Jordan se réjouissait d'obtenir ainsi une description symétrique de la matière et du rayonnement et de dépasser le dualisme onde-corpuscule à l'aide d'un formalisme contenant les deux aspects.

Revenant à la 2^o quantification des bosons, JORDAN et KLEIN⁽¹²⁾ démontrèrent qu'elle s'appliquait dans un contexte plus général que celui étudié par Dirac : on pouvait traiter des interactions "à plusieurs corps" comme l'interaction coulombienne entre deux bosons chargés. Les ondes quantifiées semblaient toutes indiquées pour une théorie de la matière en interaction.

Malheureusement les règles, dites de quantification canonique, permettant de passer d'une théorie classique à la théorie quantique correspondante, s'appliquaient mal au cas de l'électromagnétisme car il semblait impossible de mettre cette théorie sous la forme canonique. JORDAN et PAULI durent se contenter, dans une première étape de quantifier le champ électromagnétique libre grâce à l'analogie des oscillateurs. Par rapport au premier travail de Jordan sur ce sujet plusieurs progrès importants étaient réalisés : l'expression des champs quantiques libres était donnée dans tout l'espace-temps, ainsi que celle de leurs commutateurs si bien

que l'invariance relativiste de la théorie obtenue était évidente. L'introduction du calcul fonctionnel de Volterra donnait une formulation plus élégante des relations entre les différents champs.

L'électron de Dirac :

Revenons en 1925 à Cambridge, où Dirac franchissait la deuxième étape de son projet de conciliation des idées quantiques et relativistes. L'équation de Schrödinger pour l'électron était non relativiste. Une autre équation, simple et relativiste, avait déjà été écrite par Schrödinger lui-même, Klein, Gordon ... mais elle ne satisfaisait pas Dirac car elle ne s'adaptait pas à la formulation générale de la mécanique quantique par Dirac (la théorie des "transformations") intimement liée à l'interprétation probabiliste. L'équation de Klein-Gordon ne permettait pas de calculer des probabilités positives pour toutes les mesures envisageables de propriétés de l'électron. En outre Pauli⁽¹⁴⁾ et Darwin⁽¹⁵⁾ avaient, pour tenir compte du spin de l'électron, écrit une nouvelle équation de Schrödinger pour des fonctions d'onde à deux composantes ; l'accord avec l'expérience était très bon, mais le spin et la forme même des équations ne recevaient aucune justification profonde.

Au lieu de chercher une généralisation relativiste de la meilleure équation existante, celle de Pauli-Darwin, DIRAC⁽¹⁶⁾ préféra dériver sa nouvelle équation des principes mathématiques généraux de la théorie des transformations et de l'invariance relativiste. L'équation de Dirac ainsi obtenue était bien étrange : la fonction d'onde ψ avait quatre composantes et l'énergie de l'électron pouvait prendre des valeurs négatives. Pourtant l'équation de Pauli-Darwin en était une approximation et les premières corrections relativistes au spectre de l'atome d'hydrogène étaient conformes aux données empiriques. Dirac s'empessa de publier de peur que de nouvelles difficultés ne surgissent, et il laissa à DARWIN⁽¹⁷⁾ et à GORDON⁽¹⁸⁾ le calcul analytique exact de l'atome d'hydrogène par l'équation de Dirac.

Il restait la difficulté des énergies négatives. Déjà en théorie classique relativiste l'équation $E^2 = p^2 + m^2$ reliant l'énergie E à l'impulsion p et à la masse m admet deux solutions $E = \pm \sqrt{p^2 + m^2}$

Lors du mouvement classique d'un électron l'énergie varie continûment et elle conserve son signe. En revanche la théorie quantique autorise des sauts d'énergie, et en effet l'équation de Dirac permet des transitions d'un électron d'un état d'énergie positive à un état d'énergie négative. Le phénomène d'émission spontané devrait alors rendre instable tout état d'énergie positive. A ce paradoxe noté par Dirac, s'ajouta le paradoxe de KLEIN⁽¹⁹⁾ : une barrière de potentiel (c'est à dire un champ antagoniste) devenait perméable aux électrons de Dirac lorsque sa hauteur dépassait la valeur critique $2m$. A la fin de 1929 DIRAC⁽²⁰⁾ eut l'idée de remplir tous les états d'énergie négative. Le principe d'exclusion de Pauli, interdisant de mettre plus de deux électrons dans le même état, permettait ce remplissage. Une fois "la mer" pleine, les électrons d'énergie positive ne pouvaient plus tomber dedans. Comme dans la théorie de Heisenberg des couches atomiques incomplètes, un "trou" dans la mer devait être interprété comme une charge positive d'énergie positive. Comme les seules particules élémentaires disponibles à l'époque étaient l'électron, le proton et le photon, le proton était le seul candidat pour occuper la fonction de trou. Dirac se félicitait d'avoir unifié la théorie des électrons et des protons. Hélas, à la difficulté des états d'énergie négative se substituaient deux nouvelles difficultés. La matière constituée d'électrons et de protons devenait instable : la chute d'un électron dans un trou correspondait à l'annihilation réciproque d'un électron et d'un proton. La probabilité de ce processus, calculée par Dirac, Oppenheimer et Tamm, était effectivement non nulle, quoique faible. OPPENHEIMER⁽²¹⁾ décida d'abandonner l'interprétation des trous et de considérer la mer comme éternellement pleine. L'autre difficulté de l'assimilation des trous aux protons provenait de la différence de masse entre proton et électron. Dirac pensait pouvoir expliquer la lourdeur du proton par l'interaction électrostatique entre les électrons de la mer. Mais Hermann WEYL⁽²²⁾, spécialisé dans l'étude des symétries, l'en dissuada. La théorie des trous pouvait être réécrite sous une forme invariante par conjugaison de charge, et par conséquent les masses d'un électron et d'un trou devaient être rigoureusement identiques. En mai 1931 après une longue période de silence (pas de publication sur ce sujet), DIRAC⁽²³⁾ approuva l'argument de Weyl et proposa d'identifier les trous à une nouvelle particule hypothétique "l'anti-électron", conjuguée de charge de l'électron. Dirac pensait que l'observation de telles particules n'était pas pour demain. Pourtant dès 1931 le positron s'était manifesté

à l'insu des expérimentateurs et des théoriciens, dans les expériences de Gray et Tarrant⁽²⁴⁾ (par exemple).

La théorie quantique des champs d'ondes de Heisenberg et Pauli.

Cette théorie publiée en deux gros articles^(25,26) en 1929-1930, opéra une synthèse des travaux précédents de Dirac, Jordan, Wigner, Klein et Pauli, à l'exclusion de l'hypothèse de l'antiélectron (de 1931). La théorie de Dirac de 1927 n'avait pas satisfait pleinement Heisenberg et Pauli. Elle introduisait une séparation du champ électromagnétique entre champ électrostatique coulombien (responsable de la liaison atomique) et champ de rayonnement, contraire à l'esprit d'une formulation relativiste de la théorie de Maxwell-Lorentz. Un autre défaut auquel Jordan, Heisenberg et Pauli furent très sensibles était la dissymétrie entre matière chargée et rayonnement. Dirac traitait la matière chargée par des ondes de Schrödinger, et le rayonnement par des ondes électromagnétiques quantifiées ; d'un côté les particules chargées indestructibles, de l'autre les photons susceptibles de création et d'annihilation. Heisenberg et Pauli, utilisant les travaux de Jordan, Wigner et Klein, mirent matière et champ électromagnétique sur le même pied. L'électrodynamique quantique devait être obtenue en quantifiant une théorie purement ondulatoire résultant du couplage de l'onde électromagnétique de Dirac (de l'équation de Dirac) au champ électromagnétique.

Avant de particulariser leur étude à l'électrodynamique quantique, Heisenberg et Pauli développèrent une "dynamique quantique générale des champs d'ondes", par quantification canonique (suivant l'une ou l'autre des deux statistiques) de champs classiques, et donnèrent dans ce cadre général une preuve très pénible d'invariance relativiste. La difficulté, purement technique, venait du rôle privilégié joué par le temps dans la formulation habituelle de la mécanique quantique. Un autre obstacle avait retardé Heisenberg et Pauli plus longtemps : l'électrodynamique classique entraînait difficilement dans ce cadre général. Disons sans plus d'explication technique que la symétrie de jauge en était responsable. Heisenberg trouva une astuce pour sortir de l'impasse. Après quelques applications perturbatives la nouvelle électrodynamique quantique de Heisenberg et Pauli se révéla équivalente à la première théorie de Dirac de 1927 (à l'introduction près de la

théorie relativiste de l'électron), et pour une étude plus commode des applications à l'interaction atome-rayonnement, ces auteurs et Oppenheimer⁽²⁶⁾ durent opérer une transformation mathématique qui ramenait à la formulation de Dirac. Le gain de la nouvelle formulation était une preuve d'invariance relativiste et l'extension considérable du cadre de la théorie quantique des champs.

La théorie de Fermi de 1929-1930.

Enrico Fermi^(27,28) lui aussi, critiqua la théorie de Dirac de 1927. La séparation a priori de l'interaction coulombienne lui semblait surtout condamnable, mais il se souciait moins que Heisenberg et Pauli d'une preuve d'invariance relativiste et d'un traitement parallèle de la matière et du rayonnement. La théorie de Fermi, dans ses prémices, ne différait de celle de Dirac que par l'idée de quantifier tout le champ électromagnétique y compris le champ coulombien. Fermi contourna si habilement l'écueil rencontré par Heisenberg et Pauli lors de cette quantification, qu'il ne l'aperçut probablement pas. Le résultat fut la formulation la plus élégante et la plus simple de l'électrodynamique quantique de 1930, avec de surcroît de nombreuses applications développées par Fermi et ses collaborateurs et enseignées à l'Institut Poincaré et aux Etats-Unis.

En 1930, avec la théorie de Heisenberg-Pauli et la mer de Dirac, se trouvait constituée la base formelle de l'électrodynamique quantique moderne. Certains faits étranges de la première théorie quantique : la dualité onde-corpuscule, l'indiscernabilité de particules identiques, trouvaient une explication unitaire et profonde. A la notion de particule il fallait substituer la notion d'état quantique du champ, certains états présentant des propriétés plutôt ondulatoires, d'autres plutôt corpusculaires, de même qu'à la notion de position d'une particule il avait fallu substituer la notion de fonction d'onde. Les "particules" en tant que degrés d'excitation du champ étaient évidemment indiscernables. Elles étaient susceptibles d'être créées ou annihilées, y compris les électrons et les protons, particules de matière, même si pour Dirac de tels processus n'étaient qu'une réinterprétation de processus conservant le nombre total de particules (mer + électrons d'énergie positive).

Hélas, à côté de tous ces succès fut reconnu dans cette même année 1930 l'ampleur d'un nouveau problème, celui des infinis. OPPENHEIMER⁽²⁹⁾, poursuivant les applications de la théorie de Heisenberg-Pauli, montra qu'elle impliquait un déplacement infini des niveaux atomiques et WALLER⁽³⁰⁾ montra que l'énergie propre d'un électron de Dirac libre divergeait. Avant d'aborder ce problème des infinis, qui hanta la théorie quantique des champs jusqu'en 1947, nous montrerons comment celle-ci survécut et même prospéra malgré (et même parfois à cause de) la conscience des divergences.

0.2. MALGRE LES INFINIS :

Les tests de l'électrodynamique pure.

De manière générale les calculs de processus physiques en théorie quantique des champs s'effectuent perturbativement, c'est-à-dire dans un développement suivant les puissances de la constante d'interaction qui couple les différents champs fondamentaux. Le premier terme non nul (approximation de Born) est en général fini, mais le terme correctif suivant diverge. Avant qu'en 1930 soient faits les premiers calculs de ces termes correctifs, il était déjà évident que le premier terme donnait dans bien des problèmes une approximation suffisante. La plupart des théoriciens pensaient que les corrections, dans le cas de l'électrodynamique, seraient petites (de l'ordre de $\alpha = \frac{e^2}{4\pi\hbar c} \approx \frac{1}{137}$) et plusieurs d'entre eux (Heitler, Oppenheimer, ...) continuèrent de penser après la découverte des divergences que la théorie future débarrassée des divergences n'apporterait que de faibles corrections à l'approximation de Born.

A basse énergie (de l'ordre des énergies de liaison atomique), les vérifications les plus fines de la théorie du rayonnement et de l'électron de Dirac eurent lieu en physique atomique. Il fallut attendre 1947 pour que des écarts indiscutables aux prédictions de l'équation de Dirac apparaissent dans l'étude des spectres atomiques. La théorie de WIGNER et WEISSKOPF de la largeur naturelle des raies atomiques ne fut pas non plus contredite par l'expérience.

Les formules donnant les probabilités de diffusion dans la théorie de Dirac, formule de MÖLLER⁽³²⁾ (1931) pour la diffusion électron-électron, formules de MOTT⁽³³⁾ (1929) pour la diffusion coulombienne et pour la polarisation des électrons par une cible, formule de KLEIN-NISHIMA (1929)⁽³⁴⁾ pour la diffusion électron-photon, formule de BETHE et HEITLER (1933)⁽³⁵⁾ pour le rayonnement de freinage, formule de HEITLER et SAUTER (1933)⁽³⁶⁾ pour la production de paires, furent soumises au verdict de l'expérience dans le domaine des énergies relativistes, grâce à la source naturelle de particules très énergétiques que constituaient les rayons cosmiques. L'histoire de la confrontation des calculs d'électrodynamique et des données

de la physique des rayons cosmiques est complexe en raison des difficultés expérimentales diverses et des problèmes d'identification des particules en présence.

En 1932, ANDERSON⁽³⁷⁾, en Californie, découvrit le positron et une confirmation fut obtenue en Angleterre par BLACKETT et OCCHIALINI⁽³⁸⁾. Cette découverte ranima la théorie des trous de Dirac, et changea l'évaluation théorique de l'absorption par les métaux lourds des rayons cosmiques : il fallait tenir compte de la création possible de paires électrons-positrons. Les gerbes observées par Blackett et Occhialini correspondaient comme OPPENHEIMER⁽³⁹⁾, CARLSON⁽⁴¹⁾, BHABHA et HEITLER⁽⁴⁰⁾ le comprirent en 1936, à l'émission en cascade de telles paires, et non pas, comme le croyait HEISENBERG⁽⁴²⁾, à des "explosions" des particules nucléaires. Mais pour éliminer tout désaccord entre les formules relativistes d'absorption et les mesures expérimentales, il fallut d'abord éviter la confusion des mésons avec des électrons très énergétiques.

La théorie des champs en physique nucléaire.

Certains noyaux atomiques instables se désintègrent en émettant des électrons ou rayons β . D'après la loi de conservation de l'énergie on s'attendait à ce que l'énergie de ces particules β fût approximativement la différence entre la masse du noyau initial et la masse du noyau résultant. Au contraire on voyait un spectre d'énergie continu. Quand il ne fut plus possible après les expériences de Ellis et Wooster⁽⁴⁴⁾ et de Meitner⁽⁴⁴⁾ d'expliquer cette continuité par l'émission de rayons γ secondaires inobservés, PAULI proposa à la fin de 1930⁽⁴⁵⁾ d'admettre qu'une nouvelle particule neutre difficilement détectable, d'abord appelée neutron par Pauli puis neutrino par Fermi, était émise lors du processus. En 1932 à Cambridge James CHADWICK⁽⁴⁶⁾ réinterpréta certaines expériences récentes de JOLIOT et CURIE⁽⁴⁷⁾ et réalisa lui-même de nouvelles expériences prouvant l'existence d'une particule dont Rutherford avait fait depuis longtemps l'hypothèse : le neutron. Bientôt HEISENBERG⁽⁴⁸⁾ et MAJORANA⁽⁴⁹⁾ donnèrent des théories du noyau constitué de protons et de neutrons. Au congrès Solvay de 1933 le catalogue des particules élémentaires s'était enrichi de deux nouvelles espèces : le neutron et le positron, si bien que Pauli défendit plus courageusement son hypothèse du neutrino.

A son retour en Italie, FERMI, séduit par l'hypothèse du neutrino, élaborera la première application de la théorie quantique des champs à la physique nucléaire (déc. 1933). Le processus élémentaire responsable de la désintégration β , devait être neutron \rightarrow proton + électron + neutrino, l'électron et le neutron étant créés pendant l'interaction. Fermi parlait d'une analogie avec la création des photons lors des processus d'électrodynamique. Pour des fermions comme l'électron et le neutrino, une théorie ne conservant ni le nombre ni l'identité des particules était possible grâce à la quantification de Wigner-Jordan. Avec cette audacieuse théorie de Fermi⁽⁵⁰⁾, bientôt corroborée par les faits, l'image de la matière indestructible devenait insoutenable*. Le neutrino de Pauli, maintenant intégré dans un cadre théorique solide, fut adopté par ses anciens ennemis (Bohr) bien avant que sa difficile découverte (1959) eût lieu. La théorie de Fermi se maintint très tard sans altération profonde dans l'histoire de la physique nucléaire.

Avec elle naissait aussi la possibilité (Tamm et Iwanenko, 1934) d'expliquer l'interaction nucléon-nucléon par un échange virtuel de particules. Cette idée, à la base de l'image moderne des interactions entre particules, mérite quelques explications. L'exemple le plus simple est celui de l'interaction électron-électron dans la première approximation. Le calcul de la probabilité de ce processus à une structure analogue à celle d'une combinaison de deux processus : le premier électron émet un photon, puis le photon est réabsorbé par le deuxième électron, autrement dit le photon a été échangé par les deux électrons. Il ne faut pas confondre la structure du calcul et l'image qui en résulte, avec le déroulement réel des phénomènes entre les instants initial et final d'observation ; c'est pourquoi l'échange est dit virtuel. Le théoricien japonais Hideki YUKAWA⁽⁵³⁾ substitua en 1935 l'échange d'une nouvelle particule scalaire chargée à l'image électron-neutrino pour obtenir une théorie plus similaire à l'électrodynamique quantique et une interaction neutron-proton suffisamment forte. Le succès de la théorie de Yukawa ne fut pas immédiat. On n'était enclin à introduire de nouvelles particules que quand la nécessité théorique et expérimentale s'en faisait vraiment sentir ; de plus la théorie

* Dans la théorie des trous de Dirac, la création de paire était expliquée par l'éjection d'un électron de la mer, si bien que le nombre total d'électrons était conservé.

de Yukawa ne décrivait pas simplement l'interaction de deux nucléons de charge identique (elle n'utilisait pas de particule d'échange neutre par analogie avec la théorie de Heisenberg des forces nucléaires où l'interaction neutron-proton était comparée à la liaison H_2^+ par échange d'électron : $H + H^* \longleftrightarrow H^* + H$). Quand les expérimentateurs, STREET et STEVENSON reconnurent en 1937 la présence d'une nouvelle particule dans les rayons cosmiques, plusieurs théoriciens dont OPPENHEIMER⁽⁵⁵⁾ et STUECKELBERG⁽⁵⁶⁾ pensèrent que la particule de Yukawa avait été découverte. Il s'agissait en réalité d'un autre méson, le muon, interagissant faiblement avec les particules nucléaires. Cette "plaisanterie de dix ans" (Oppenheimer) eut le mérite de promouvoir la théorie de Yukawa, et même d'intéressantes théories nucléaires visant à reproduire le couplage anormalement faible entre le faux méson et le nucléon. Enfin la découverte du méson et l'explication qu'elle apporta du pouvoir pénétrant d'une partie des rayons cosmiques (la composante "dure") levèrent les soupçons à l'égard de la validité de l'électrodynamique pure dans le domaine des hautes énergies. Entre les anomalies d'absorption constatées en 1934 par Anderson et Neddermeyer⁽⁵⁷⁾ et la découverte du méson, expérimentateurs et théoriciens avaient dû tâtonner pour découvrir les failles d'un édifice théorique complexe : détecteurs biaisés ? électrodynamique violée ? confusion sur la nature de particules absorbées ?

Progrès formels.

Nous venons d'avoir avec la physique des rayons cosmiques un exemple de progression achevée de l'expérience et de la théorie quantique des champs. Mais de nombreux progrès de cette théorie se sont réalisés avant guerre sans la pression directe de l'expérience. Là où celle-ci ne pouvait guider efficacement, une physique mathématique autonome s'est développée à travers les embûches des infinis. Extensions formelles et reformulations, loin de se limiter à un pur jeu de l'esprit, préparèrent le terrain plus tard fécondé par le principe de renormalisation, de même que la mécanique analytique de Lagrange, Hamilton, Jacobi et Poisson constitua la base de la nouvelle mécanique quantique.

La mécanique quantique naquit sous deux formes différentes, mais équivalentes : les matrices de Heisenberg et les ondes de Schrödinger.

Schrödinger et Dirac montrèrent que l'on passait de l'une à l'autre en changeant de "point de vue" : l'évolution temporelle pouvait être alternativement reportée sur les opérateurs (Heisenberg) où sur les états quantiques (Schrödinger). Sans entrer dans les détails techniques disons que suivant le "point de vue" (ou "représentation") adopté en théorie quantique des champs, les propriétés des champs quantiques apparaissent différemment. A tel point que HEISENBERG⁽⁵⁸⁾ en 1931 et DIRAC⁽⁵⁹⁾ en 1932 eurent un instant l'illusion d'avoir créé une nouvelle électrodynamique quantique alors qu'ils n'avaient fait qu'introduire de nouvelles représentations. Après la preuve de l'équivalence de la théorie de Dirac de 1932 et de celle de Heisenberg-Pauli (DIRAC, FOCK, PODOLSKY⁽⁶⁰⁾ et ROSENFELD⁽⁶¹⁾), il resta que la nouvelle formulation était explicitement covariante relativiste. Elle utilisait la représentation d'interaction d'aujourd'hui (ainsi nommée par Schwinger), intermédiaire entre celles de Schrödinger et de Heisenberg et elle fut généralisée par le théoricien japonais Sin-Itiro TOMONAGA en 1943 au cas où la matière était décrite par des ondes quantifiées. Un autre article de DIRAC⁽⁶³⁾ avait inspiré Tomonaga (et Yukawa): "Le lagrangien en mécanique quantique"; toujours dans le souci de mieux respecter l'invariance relativiste, Dirac avait établi en 1933 une correspondance entre l'action classique et l'évolution quantique. Richard Feynman s'en souvint quand il introduisit les fameuses intégrales de chemin (publication en 1948).

L'invariance relativiste, bien que ses exigences restreignent de beaucoup les théories quantiques des champs possibles, laisse le choix de plusieurs possibilités de transformation des champs fondamentaux lors d'une transformation de Lorentz, et corrélativement le choix du spin des particules associées. Le champ électromagnétique (spin 1) et le champ d'électrons (spin 1/2) furent deux premiers exemples. PAULI et WEISSKOPF⁽⁶⁴⁾, en 1934, étudièrent le champ scalaire (spin 0) chargé, surtout pour combattre l'idée de la mer de Dirac (papier "anti-Dirac"). L'idée d'une infinité d'électrons dans la mer déplaisait à Pauli et pouvait être abandonnée (FOCK⁽⁶⁵⁾ l'avait déjà remarqué en 1933) au profit d'une nouvelle méthode de quantification des ondes de matière, sans rien perdre de la physique du positron. La théorie de Pauli et Weisskopf passa de la fiction

à la réalité grâce aux travaux de Yukawa et à la découverte du méson. D'autres possibilités furent développées dans de nombreux travaux de DIRAC, PROCA, KEMMER, BHABHA, STUECKELBERG, FIERZ, BELINFANTE ... soit pour raffiner la théorie mésique, soit dans un but plus spéculatif en prévoyant l'accueil de particules futures. De toutes ces théories ressortait une règle stricte : les champs de spin demi-entier ($1/2$, $3/2$, $5/2$, ...) devaient être quantifiés suivant la statistique de Fermi, et ceux de spin entier, suivant la statistique de Bose. PAULI s'était préoccupé très tôt d'éclaircir ce lien mystérieux entre le spin et la statistique et il en donna en 1940⁽⁶⁶⁾ une démonstration générale. Il avait réglé plus tôt les cas empiriquement bien connus de l'électron et du photon : l'électron de Dirac devait impérativement satisfaire à la statistique de Fermi pour que les états d'énergie négative puissent être remplis. Les théories de spin entier devaient être quantifiées suivant la statistique de Bose pour que deux mesures de champ n'interfèrent qu'à des distances à portée de lumière. Malgré les entraves des divergences, le mariage des idées quantiques et relativistes devait être, à n'en pas douter, célébré pour que des lois de la nature aussi universelles que la correspondance spin \longleftrightarrow statistique trouvent une justification profonde.

Au fur et à mesure que la théorie quantique des champs se grandissait des nouveaux succès de la physique des hautes énergies, et qu'elle prenait une forme plus "élégante" (Dirac), le problème des infinis changeait de visage. Du côté de l'électrodynamique pure les mesures d'absorption d'électrons cosmiques rendaient académique la considération des termes divergents, du côté de la physique nucléaire il fallait au contraire absolument calculer ces termes ; sans quoi plus de moment magnétique des nucléons, plus d'interactions proton-proton ...

0.3 ATTITUDES FACE AUX INFINIS :

Quand on parle aujourd'hui de divergences sans rien préciser, on pense aussitôt aux divergences ultraviolettes de la théorie quantique des champs telles qu'elles se montrèrent dans toute leur horreur en 1930. Mais le problème était bien plus ancien, et il convient de préciser son statut dans la théorie classique pour comprendre ce qu'il devint et quelles réactions il suscita après quantification.

Les divergences en théorie classique

Les systèmes simples de la mécanique classique sont caractérisés par un nombre fini et petit de variables indépendantes (les variables dynamiques) telles les coordonnées des centres de gravité des planètes et leurs angles de rotation propre. D'autres systèmes nécessitent un nombre infini de variables dynamiques, par exemple les coordonnées de l'infinité des éléments d'un fluide, ou les valeurs des champs électrique et magnétique en tout point de l'espace. Dans de nombreux cas on sait démontrer grâce à l'analyse mathématique l'existence de solutions aux problèmes de mécanique correspondants. Mais l'électrodynamique classique de Maxwell-Lorentz appliquée à des charges rigoureusement ponctuelles n'admet pas de solutions, ne serait-ce que parce que le champ créé par une charge ponctuelle diverge au lieu de cette charge, lui conférant une énergie propre infinie. Lorentz échappa à cette difficulté à l'aide d'un modèle d'électron étendu sur une distance de l'ordre du "rayon classique de l'électron", choisi pour que la masse de l'électron s'identifie avec son énergie propre électromagnétique ($\frac{e^2}{r_0} \approx mc^2$). Grâce à l'extrême petitesse de r_0 (10^{-15} m \approx dimensions nucléaires) l'électrodynamique macroscopique des électrons devait très peu dépendre de leur structure et le problème des divergences n'était pas la difficulté centrale de la théorie classique.

L'incompatibilité du modèle avec la relativité et plus tard en 1926 la découverte du spin, la menacèrent plus réellement. Enfin avec l'avènement de la théorie quantique il devint clair que les problèmes de structure de l'électron, s'ils admettaient une solution, devaient être résolus en tenant compte des effets quantiques puisque les corrections quantiques à la diffusion rayonnement-électron apparaissaient pour des longueurs d'onde $\lambda \approx \frac{h}{mc}$ très supérieures à r_0 .

A cet égard les théoriciens étaient plutôt optimistes ; l'introduction du quantum d'action h avait déjà éliminé une divergence : celle de l'énergie totale d'un corps noir classique à l'équilibre.

Les divergences en théorie quantique :

En raison de son caractère fondamental, l'électron de la théorie à quantifier fut choisi ponctuel. C'était aussi une manière d'éviter la difficulté de traiter des objets étendus en relativité. Il fallut attendre 1930 pour que s'évanouît l'espoir de guérir par quantification les divergences de l'électrodynamique. En 1927 Dirac s'était gardé de calculer les deuxièmes termes dans sa théorie du rayonnement. D'ailleurs des effets du 2^o ordre comme les processus où deux photons interagissent simultanément avec l'atome lui semblaient incongrus. Du côté de Heisenberg Pauli et Jordan, des infinis étaient apparus comme l'énergie de zéro des oscillateurs (la somme de $\frac{1}{2} h\nu$ sur toutes les fréquences) et l'énergie propre coulombienne, mais on pouvait au moins provisoirement les soustraire sans inconvénient. Les infinis se dramatisèrent en 1930 lorsque OPPENHEIMER⁽²⁹⁾ calcula un effet physique absurde de déplacement infini des niveaux atomiques, et lorsque WALLER⁽³⁰⁾ démontra que l'énergie propre de l'électron de Dirac divergeait tout autant que dans la théorie classique. En 1934 WEISSKOPF⁽⁶⁷⁾ et FURRY, grâce à la théorie du positron de Dirac, abaissèrent le degré de la divergence pour la rendre logarithmique. Cette amélioration ne donnait pas à la théorie plus de cohérence mathématique, mais elle la rendait moins sensible à l'introduction d'un petit rayon de l'électron et plus tard elle permit le succès de la renormalisation. Puis les théories de Fermi et de Yukawa donnèrent des divergences bien pires, sapant à nouveau les bases de la théorie des champs. Plusieurs théoriciens, Waller le premier, suspectèrent la méthode des perturbations. Mais cette méthode semblait conforme à "l'esprit" de la théorie pour des raisons philosophiques (idée de Bohr et Rosenfeld d'une progression par étapes) et aussi mathématiques (possibilité de faire tendre la constante de couplage vers zéro). Dirac pensait qu'une solution non perturbative, si jamais elle existait, serait trop compliquée pour être utilisée. Les divergences furent plutôt considérées comme une difficulté fondamentale, nécessitant des changements révolutionnaires.

Les infinis et la mesure

En 1927 le congrès Solvay avait permis un consensus d'interprétation du formalisme de la mécanique quantique. Le fonctionnement des appareils de mesure, pensait Bohr, devait s'exprimer dans le langage de la physique classique pour faire partie de l'expérience communicable des êtres macroscopiques que nous sommes. Il est loin d'être évident que cette conception s'applique dans le cas de la mécanique quantique relativiste. La relation entre la théorie quantique et l'expérience est plus subtile pour la mécanique quantique que pour la mécanique classique puisque tout réalisme naïf doit être éliminé. Dès l'apparition des infinis en 1930 la base observationnelle de la théorie fut suspectée ; à raison, semblera-t-il après l'application du principe de renormalisation. Bohr, Heisenberg, et Dirac tentèrent tous trois d'aborder ce problème d'interprétation du formalisme par des arguments philosophiques ou méthodologiques a priori, avec, nous allons voir, quelques succès importants, mais sans trouver pour autant la clé du problème ultraviolet.

LANDAU et PEIERLS, en 1931, donnèrent des limites de précision infranchissables aux mesures instantanées d'impulsion d'une particule et aux mesures de champs électrique et magnétique. Ces limites minaient selon eux la théorie quantique relativiste. L'argument pour les mesures de champ, tout copenhaguien qu'il était, fut énergiquement repoussé par BOHR et ROSENFELD. Ceux-ci corrigèrent une erreur de Landau et Peierls pour rétablir une harmonie parfaite entre la conception de Bohr de la mesure et la théorie quantique du champ électromagnétique libre. Cette longue discussion (achevée en 1933) ⁽⁶⁹⁾ n'avait rien changé à la théorie, mais ses bases semblaient plus solides.

De toute façon, pour Bohr, le principe de correspondance donnant les règles de quantification de la théorie classique, devait être caduque à l'échelle des dimensions nucléaires, car celles-ci étaient du même ordre que le rayon classique r_0 de l'électron, limite de la théorie classique de l'électron. En dessous de r_0 il fallait s'attendre au pire : aux divergences, et même à une violation du principe de conservation de l'énergie. En témoignait, selon Bohr, la continuité des spectres β . Heureusement Pauli sut résister à la proposition de Bohr. Pour lui il n'y avait pas de raison

d'abandonner la conservation de l'énergie plutôt que celle de la charge, que Bohr gardait. Les énergies négatives de l'équation de Dirac devaient aussi traduire, pensait Bohr, un échec de la méthode de correspondance et ce n'est qu'après le succès de la théorie du positron, qu'il vit dans la proposition de la mer de Dirac une vraie solution de la difficulté des énergies négatives et un raffinement de la méthode de correspondance. Dirac, lui, suivait une stratégie toute différente : il fallait selon lui construire d'abord la théorie mathématique la plus belle possible et se soucier en second lieu de son interprétation physique. Le succès de cette démarche culmina avec la découverte du positron. Elle avait le mérite d'être constructive, au lieu d'être critique comme celle de Bohr. Dans le domaine de la physique nucléaire Bohr continua à penser que la distance r_0 devait jouer un rôle singulier. Au-delà, seules des expériences d'un type entièrement nouveau pourraient guider le théoricien privé d'une base de correspondance classique.

HEISENBERG fut sensible aux arguments de Bohr, et après que Blackett et Occhialini eurent photographié (à partir de 1932) des gerbes de particules, il crut à l'apparition de phénomènes explosifs au-delà d'une longueur critique ⁽⁴²⁾ de l'ordre de r_0 . Cette longueur devait être (1936) une constante universelle comme c et h , caractéristique d'une théorie future plus générale, et elle devait (1938) ⁽⁷⁰⁾ fixer une limite au-delà de laquelle nos concepts d'espace et de temps perdraient leur signification. En 1943 Heisenberg tenta de renouveler l'approche "observationniste" qui avait donné la mécanique des matrices. Les seules quantités observables (ou presque) devaient être la probabilité de diffusion et il fallait construire une théorie introduisant le minimum d'inobservables. Ainsi démarra le programme de la matrice S (comme "Streuung"), reliant les états des particules incidentes d'un processus de diffusion aux états finaux. Sans champs quantiques en interaction, il n'y avait plus de divergence, mais comment construire la matrice S ? Jusqu'à ce jour une détermination complète de la matrice S n'a pu être obtenue que par l'intermédiaire de théories des champs renormalisées. Bien que l'affirmation de Heisenberg, limitant les observables aux probabilités de diffusion, ait été largement critiquée, le programme de matrice S a porté de nombreux fruits (relations de dispersion, bornes de

Froissart, dualité ...) tout en restant ennemi du programme de théorie des champs.

Pour clore les analyses du rôle du mesurable dans la théorie quantique relativiste nous présenterons à la fin de ce chapitre celles de Kramers et de Weisskopf qui vinrent à bout des infinis.

0.4 A LA RECHERCHE DE THEORIES ALTERNATIVES :

La remise en question des fondements observationnels de la théorie quantique relativiste conduit à elle seule à parler d'une véritable situation de crise. Les nombreuses tentatives pour échapper à la formulation normale de la théorie quantique des champs sont autant de symptômes de cette crise des divergences. Les alternatives dont nous parlerons ici retinrent une grande partie des postulats de la relativité et de la théorie quantique, malgré l'abandon explicite ou implicite de quelques grands principes comme la causalité relativiste (impossibilité de transmettre des signaux à une vitesse supérieure à celle de la lumière) (Dirac 1938), l'existence d'une fonction d'onde (Heisenberg), la conservation et la positivité des probabilités (Dirac 1942) ... D'autres tentatives comme celles d'Einstein visaient à substituer une théorie entièrement nouvelle à la théorie quantique ordinaire jugée incomplète, et elles avaient trouvé leur origine dans le débat sur l'interprétation de la mécanique quantique plutôt que dans la crise des divergences.

La découverte de l'anomalie des infinis ne dit en rien où et comment la théorie quantique des champs devait être corrigée. Un problème formel engendra la crise sans en donner le remède ; de même les expériences négatives à l'origine de crises plus anciennes ne furent jamais des "expériences cruciales". Les expériences nouvelles de la physique des hautes énergies ou de la physique nucléaire, loin de préciser une direction alternative de recherche, servirent plutôt à étendre la portée de la théorie normale, malgré ses infirmités formelles. Alors quel guide trouver dans la recherche de théories nouvelles ?

Les leçons de la théorie classique :

Un regard vers le passé montrait que le problème des infinis préexistait à la théorie quantique. Certains théoriciens ⁽⁷²⁾, Wentzel (1933), Dirac (1938), Kramers (1937), Stueckelberg (1941), Wheeler et Feynman (1945), Peierls et Mac Manus (1947) ... pensèrent qu'il fallait guérir la théorie classique de ses divergences avant de quantifier. Cette guérison fut possible et la théorie de l'électrodynamique classique obtint enfin un statut mathématique satisfaisant par un retour dans le passé du problème des infinis.

Hélas la quantification introduisait de nouvelles divergences spécifiquement quantiques et le seul vrai bénéfice fut une meilleure compréhension de la théorie classique. Certaines idées de Kramers, Stueckelberg et de Wheeler et Feynman devaient prendre corps plus tard, au niveau quantique, grâce au renouveau expérimental et théorique de l'électrodynamique quantique. Dirac, armé de sa nouvelle théorie classique ⁽⁷³⁾ (1938), tenta de poursuivre sur la voie royale de ses précédentes découvertes : chercher les bonnes équations, et les faire parler après. Mais la théorie de 1942 ⁽⁷⁴⁾, débarrassée de toute divergence grâce à l'introduction formelle de probabilités négatives, posa de graves problèmes d'interprétation. Pauli, pour cette théorie de Dirac, mais aussi pour d'autres théories nouvelles, joua un rôle clarificateur et critique important.

Les intuitions philosophiques :

D'autres théoriciens à la recherche d'idées nouvelles eurent recours à des présupposés d'ordre philosophique sur la nature de la matière et du rayonnement. En 1932 DIRAC ⁽⁵⁹⁾ avait voulu donner au rayonnement un statut différent de celui de la matière puisque le rayonnement intervenait nécessairement comme "moyen d'observation" ; HEISENBERG ⁽⁵⁸⁾ en 1931 avait voulu une électrodynamique quantique plus conforme à l'idée de correspondance, où les opérateurs-champs vérifiaient les mêmes équations d'évolution que les champs classiques. Ces deux tentatives, nous l'avons dit, conduisirent à des théories équivalentes à la théorie de Heisenberg-Pauli. Les tentatives "monistes" ⁽⁷²⁾, issues de la théorie de Mie de 1912 apportèrent un changement plus radical. Gustav MIE avait proposé de décrire matière et rayonnement comme des manifestations d'une même entité : le champ électromagnétique. Bien sûr ce champ devait obéir à des équations plus complexes que celles de Maxwell pour que certaines solutions (du type "soliton" dirait-on aujourd'hui) soient interprétables en terme de matière chargée. Max BORN pensait que la conception dualiste usuelle : matière versus rayonnement était à l'origine de la crise des divergences. Avec son étudiant INFELD il élaborait en 1933-34 une théorie électromagnétique classique non linéaire à la Mie, qui contenait dans une certaine approximation la théorie de Maxwell-Lorentz. Il n'y avait plus de divergence mais la quantification était vouée à l'échec en raison de la non linéarité des équations. Fritz BOPP, en 1940 essaya une théorie du même type. En réalité cette théorie

était équivalente à une théorie dualiste mixte où la matière se trouvait simultanément couplée au rayonnement usuel et à un champ de photons massifs. Les mêmes équations toléraient une interprétation moniste ou une interprétation dualiste, selon le goût du théoricien.

Les théories dualistes mixtes, où la matière chargée interagissait simultanément avec d'autres champs que le champ électromagnétique, par exemple avec le champ de Yukawa, constituèrent pour STUECKELBERG, PAIS et SAKATA un nouvel espoir ⁽⁷²⁾. Les divergences étaient peut-être dues à la séparation artificielle de l'interaction électromagnétique des autres interactions. De fait Pais (1945) ⁽⁷⁵⁾, et indépendamment Sakata et ses collègues japonais (1947) ⁽⁷⁶⁾ montrèrent que dans certains cas bien choisis les divergences se compensaient dans les théories mixtes. Peu de temps après (à la fin de 1947) il apparut que ces compensations ne suffisaient pas à elles seules à donner une théorie correcte, il fallait encore "renormaliser". Pauli et Villars ⁽⁷⁷⁾ avaient en tête ces travaux et particulièrement les "masses auxiliaires" de Stueckelberg quand ils inventèrent en 1949 la méthode de "régularisation" (= une manière de rendre finies les expressions divergentes en préservant les symétries importantes) qui porte leur nom.

De toutes les alternatives présentées, la plus révolutionnaire fut peut-être la matrice S de HEISENBERG, car elle abandonnait la notion de champ défini en tout point de l'espace-temps. Nous avons dit comment Heisenberg justifia son introduction par des présupposés "observationnistes". STUECKELBERG ⁽⁷²⁾ contribua lui aussi à la théorie de la matrice S, après en avoir donné une autre justification : les théories des champs classiques sans divergence (comme celle de Dirac de 1938) faisaient intervenir une certaine "prémonition" du système empêchant la description temporelle causale et locale usuelle en mécanique classique. Les travaux de Stueckelberg sur la matrice S inspirèrent plus tard Freeman Dyson, en ce qu'ils permettaient une construction de la matrice S à partir de l'électrodynamique usuelle. Mais pour Dyson la matrice S ne fut qu'un outil commode pour l'étude d'une théorie quantique des champs, alors que Stueckelberg comme Heisenberg rejetaient l'idée-même de champ quantique.

Les attitudes pragmatiques :

Les premières approximations, non divergentes, de l'électrodynamique quantique finirent par s'accorder (après la découverte du méson) excellemment avec les expériences à haute énergie. En revanche la faible interaction mesurée entre le nucléon et la faux méson n'était pas compatible avec les approximations du même type effectuées dans les théories mésiques. Il fallait expliquer pourquoi les termes suivants (termes "radiatifs") apportaient de faibles corrections en électrodynamique et de fortes dans les théories mésiques. HEITLER (1941)⁽⁷²⁾ découvrit une méthode de soustraction satisfaisante à cet égard. Plusieurs méthodes de "coupure" ad hoc des intégrales divergentes virent aussi le jour. Aucune n'apporta d'éclairage nouveau sur les difficultés fondamentales de ces théories.

- - - - -

Après la guerre plusieurs grands théoriciens Wentzel, Kramers, Peierls, Sakata, Pauli ... se livrèrent à des études comparatives des principales théories récentes. Aucune ne s'imposait vraiment, et toutes présentaient des défauts évidents. En réalité elles étaient toutes trop hardies, trop révolutionnaires. L'importance des divergences avait été exagérée. Quand une issue fut découverte, la plupart de ces directions novatrices de recherche furent abandonnées. Mais de nombreux sous-produits furent intégrés dans la nouvelle version de la théorie quantique des champs, et le programme de matrice S se maintint comme espoir de résolution des difficultés persistant en physique nucléaire.

0.5 L'ISSUE

La polarisation du vide

Revenons au congrès Solvay de 1933 où DIRAC se présenta glorifié par la découverte récente du positron. La mer des électrons d'énergie négative devait être prise au sérieux puisque le positron existait. La charge et l'énergie de la mer au repos étaient déclarées inobservables, mais les écarts à cette distribution, perturbée par exemple par la présence d'un électron d'énergie positive, devaient être observables et correspondre à une certaine polarisabilité du vide. Autrement dit un électron devait s'entourer d'un nuage virtuel de paires électron-positron masquant une partie de sa charge, de même que des charges plongées dans un diélectrique s'attirent moins que dans le vide. Dirac présenta le calcul de cet effet dans un cas simple. Une nouvelle divergence surgissait du calcul, et de plus le résultat était ambigu puisqu'on devait soustraire deux quantités infinies : la distribution de charge de la mer perturbée et celle de la mer au repos pour obtenir l'effet physique recherché. FURRY et OPPENHEIMER⁽⁷⁸⁾ à la fin de la même année publièrent une étude plus générale et plus détaillée du phénomène de polarisation du vide. Leur conclusion était très pessimiste : les nouveaux infinis, selon eux, ne pourraient être éliminés que par un "*changement profond de nos notions d'espace et de temps*". DIRAC⁽⁷⁹⁾, lui, ne se découragea pas immédiatement et il tenta d'établir une prescription non ambiguë de soustraction des deux quantités infinies, grâce à l'utilisation des invariances relativiste et de jauge. L'idée importante (essentielle dans les travaux ultérieurs de Tomonaga et Schwinger) de respecter les symétries fondamentales dans le calcul d'intermédiaires divergents faisait son apparition. En résultèrent des "règles" explicites de soustraction de Dirac et de HEISENBERG⁽⁸⁰⁾. Cette "physique de la soustraction" (ainsi nommée par Pauli) avait pour mérite essentiel d'aboutir à des expressions convergentes. Mais les calculs étaient extrêmement lourds et seul le critère de simplicité permettait de fixer complètement les règles de soustraction. Heisenberg était lui aussi conscient de ces défauts. Du côté californien (sous l'égide d'Oppenheimer) UEHLING⁽⁸¹⁾ et SERBER⁽⁸²⁾ calculèrent quelques effets physiques en principe observables à l'aide de cette théorie un peu monstrueuse, et même un dépla-

cement des termes spectraux (du mauvais signe par rapport au déplacement expérimentalement soupçonné). Mais la précision des expériences de l'époque était insuffisante pour mettre en évidence de si petits effets, si bien que la théorie de la polarisation du vide se présenta comme une nouvelle excroissance formelle de la théorie de l'électron, totalement coupée de l'expérience. Un autre effet encore plus faible de la polarisabilité du vide était la possibilité d'interaction de deux photons, et la modification correspondante des équations de Maxwell calculée par Heisenberg, EULER et KOCKEL⁽⁸³⁾ (1935). Les calculs n'allèrent pas beaucoup plus loin jusqu'en 1947, mais une simplification et un changement d'interprétation essentiels furent introduits par Weisskopf en 1936.

Victor WEISSKOPF⁽⁸⁴⁾ montra qu'on pouvait s'affranchir des règles de soustraction compliquées de Dirac-Heisenberg, en déclarant inobservable "une polarisabilité du vide constante et uniforme et indépendante des champs". Autrement dit un tel effet revenait à multiplier toutes les charges de la nature par une même constante. Et cette redéfinition de la charge, montrait Weisskopf, suffisait à éliminer toutes les divergences de la polarisation du vide. Weisskopf venait de découvrir rien moins que le principe de renormalisation de charge.

Les divergences infrarouges :

Quand un électron est diffusé par une charge fixe, il peut éventuellement émettre des photons, de même qu'une charge classique accélérée rayonne. La probabilité d'émettre un photon pendant la diffusion était infinie pour les photons "mous" de basse fréquence, comme il ressortait des calculs de Mott et de Sommerfeld de 1931. Ces nouvelles divergences étaient indépendantes des divergences "ultraviolettes" déjà reconnues : les unes étaient associées à l'émission de photons réels de basse fréquence, les autres à l'émission suivie de l'absorption de photons virtuels de haute fréquence. Enfin les divergences infrarouges, bien qu'elles rendent la théorie apparemment absurde par l'introduction de probabilités infinies, n'empêchaient pas de trouver des résultats finis et corrects (comme on le sut a posteriori) pour les pertes d'énergie des électrons traversant la matière. La solution du problème infrarouge attendit 1937, année de la publication d'une "Note sur le champ de

rayonnement de l'électron" par Felix BLOCH et Arnold NORDSIECK. Ces deux théoriciens remarquèrent que l'affirmation qu'un seul photon mou avait été émis lors de la déflexion d'un électron n'avait aucune signification expérimentale. L'expérimentateur ne distingue pas des états finaux qui ne diffèrent que par la présence de photons très mous indétectables. A la question : quelle est la probabilité d'avoir dans l'état final telle ou telle particule et une infinité de photons mous inobservables ?, la théorie donnait une réponse finie. La théorie de Bloch et Nordsieck et l'analyse ultérieure de PAULI et FIERZ⁽⁸⁶⁾ montrèrent que l'idée d'électron nu n'avait pas de sens expérimental. Dans la théorie quantique comme dans la théorie classique on ne pouvait séparer l'électron de son champ propre l'accompagnant dans ses déplacements. En 1947-48 Schwinger et Tomonaga utilisèrent une transformation mathématique généralisant celle introduite par Pauli et Fierz en 1938 pour intégrer le champ propre dans les états quantiques à un électron.

Les idées de Kramers :

En 1937, le physicien hollandais Hendrick Anthony KRAMERS écrivit un ouvrage⁽⁸⁷⁾ sur la théorie de l'électron et du rayonnement, assez critique vis-à-vis de la théorie de Dirac de 1927 du rayonnement. Kramers, depuis longtemps, ne comprenait pas pourquoi la théorie de Dirac marchait si bien, alors qu'elle n'introduisait pas de distinction entre champ extérieur et champ propre de l'électron, comme la théorie classique l'exigeait. Il montra d'abord qu'on pouvait tirer de l'électrodynamique classique de Maxwell-Lorentz des équations indépendantes de la structure de l'électron et sans divergence, à condition de distinguer champ extérieur et champ propre entraîné par les charges et de ne faire intervenir dans les équations que la masse physique observable (comportant une contribution d'origine électromagnétique). Seule une telle théorie pouvait selon Kramers constituer une base de correspondance correcte pour la quantification. L'idée que l'introduction de la seule masse observable dans les formules de l'électrodynamique absorbe les divergences est aujourd'hui nommée principe de renormalisation de masse. L'utilisation d'une approximation non relativiste et la difficulté de mettre les équations classiques sous une forme propice à la quantification limitèrent les premières applications de ces idées. Toutefois Kramers démontra aussitôt que la théorie

de Dirac de 1927 pouvait être réinterprétée de manière conforme à ses idées, à condition de la considérer comme une approximation, et plus tard (en 1940) les disciples de Kramers montrèrent que la quantification de sa théorie devait donner une valeur finie pour le déplacement des raies spectrales. Le calcul précis de ces déplacements ne fut exposé par Kramers qu'en 1948 (congrès Solvay), après le calcul de Bethe.

L'expérience de Lamb et la résolution de la crise :

En juin 1947 Willis LAMB exposa au congrès de l'île de Shelter (près de New York) les résultats d'une expérience effectuée dans le laboratoire de Rabi avec Retherford. Les progrès divers de la physique atomique expérimentale, et l'amélioration des techniques de radiofréquences héritées de la technologie du radar développée pendant la guerre, permettaient enfin de mesurer précisément un effet depuis longtemps soupçonné : certains niveaux de l'atome d'hydrogène, confondus dans la théorie de l'électron de Dirac, étaient en réalité séparés par un petit intervalle d'énergie (de l'ordre du milliardième en valeur relative). Au même congrès Kramers exposa ses idées déjà vieilles de dix ans, et Hans BETHE⁽⁸⁸⁾ au retour de son voyage sur l'île de Shelter fut en mesure de donner une bonne estimation du déplacement de Lamb en appliquant directement au niveau de la théorie quantique le principe de renormalisation de masse de Kramers. La majeure partie du déplacement divergent calculé par Oppenheimer en 1930 devait être absorbée par une redéfinition de la masse. Bethe pensa, à juste titre, avoir trouvé la clé du problème des divergences et il prédit que le calcul relativiste complet donnerait un résultat fini du même ordre de grandeur.

Cette intuition de Bethe fut bientôt vérifiée par ceux qui crurent aussitôt aux vertus du principe de renormalisation : Weisskopf, Feynman, Schwinger, Tomonaga, Lamb et Kroll. Les premiers calculs relativistes furent effectués dans le cadre de l'ancienne théorie des perturbations telle qu'elle était disponible dès 1934 (date des calculs d'énergie propre par Weisskopf dans la théorie du positron), non sans mal car la soustraction de la contribution infinie associée à la renormalisation de masse était ambiguë. Sin-Itiro TOMONAGA et Julian SCHWINGER avaient à leur disposition plusieurs techniques mathématiques développées au moins en partie avant 1947 comme

la théorie des perturbations covariantes de Tomonaga, la transformation canonique de Pauli et Fierz, et ils eurent bientôt la conviction que l'application conjuguée des principes de renormalisation de masse et de charge et un calcul covariant dans toutes ses étapes conduisaient à des calculs de corrections radiatives finis et non ambigus. La démarche de Richard FEYNMAN était plus profondément originale et par là-même elle fut largement incomprise avant sa réexposition par Freeman Dyson, mais elle conduisait aux mêmes résultats que la méthode de Tomonaga-Schwinger, et ce par des calculs beaucoup plus courts. En 1949-50 l'objectif était atteint, DYSON après avoir unifié et clarifié les approches de Feynman et de Tomonaga-Schwinger put démontrer l'élimination des divergences à tous les ordres de la série des perturbations.

L'électrodynamique quantique moderne ainsi constituée devint et est encore aujourd'hui un modèle de théorie quantique des champs. Pour elle le principe de renormalisation avait fait merveille. Il en alla tout autrement des théories mésiques ou de la théorie de Fermi des interactions faibles, qui, elles, n'étaient pas renormalisables. Là où le problème des divergences était le plus urgent, il restait hélas sans solution.

0.6 CONCLUSION

Plusieurs commentateurs ont souligné que tous les ingrédients mathématiques nécessaires au calcul de Bethe du déplacement des niveaux atomiques étaient disponibles dès 1930, et que même la généralisation relativiste de ce calcul était faisable disons dès 1934. On peut ajouter que certaines théories qui contenaient implicitement une renormalisation de masse comme la théorie de Wentzel-Dirac de la " λ -limite" auraient donné, dans une approximation non relativiste, exactement le calcul de Bethe sans aucune hypothèse supplémentaire. Pourquoi a-t-il fallu vingt ans de théorie quantique des champs (de 1927 à 1947) pour qu'un calcul analogue à celui de Bethe soit produit, et pourquoi mit-on tant de temps pour prendre au sérieux l'idée de renormalisation de masse ?

- Dans la réponse la plus répandue à cette question, la nécessité d'un stimulus expérimental, en l'occurrence l'expérience de Lamb-Retherford est invoquée. Cette réponse n'est que très partielle. S'il est vrai que la nécessité de calculer des effets radiatifs était peu pressante en électrodynamique pure où ils sont si faibles, il n'en était rien en physique nucléaire où les effets analogues comme le moment magnétique anormal des nucléons étaient énormes.

- Les divergences furent prises trop au sérieux. Les changements nécessaires pour les éliminer devaient être, selon la plupart des physiciens, révolutionnaires. Et une critique des fondements assortie de propositions alternatives se développa. La dramatisation des divergences compliqua la tâche déjà devenue difficile en théorie quantique d'établir les règles d'interprétation du formalisme mathématique, en favorisant des positions extrêmes inefficaces. Ni la stratégie de Bohr - utiliser des arguments philosophiques a priori pour déterminer ce qui est observable - ni celle de Dirac - développer d'abord le formalisme et deviner ensuite les règles d'interprétations - n'apportèrent une solution au problème des divergences. La philosophie de la complémentarité pouvait éclaircir le contenu observationnel de l'électrodynamique quantique mais seulement après que celle-ci soit complétée par le principe de renormalisation ; et en effet en 1950 Bohr et Rosenfeld montrèrent que les mesures de champ et de charge dans cette théorie s'harmonisaient avec la conception de Bohr de la mesure. Quant

à Dirac, il ne pouvait développer isolément le formalisme de l'électrodynamique quantique sans savoir quelles grandeurs (les grandeurs renormalisées) calculer. Il fallait pour aboutir, comme Kramers, à un principe de renormalisation considérer globalement le formalisme et ses règles d'interprétation et croire que le bon choix des règles d'interprétation permettrait d'absorber les divergences. Si l'on pense comme Léon Rosenfeld⁽⁸⁹⁾ qu'une théorie physique évoluée est nécessairement "ouverte" c'est-à-dire qu'elle n'est pas la réunion d'une théorie mathématique rigoureuse et unitaire (petit nombre d'axiomes) et de règles précises d'interprétation, alors il est vain de chercher à fermer le formalisme mathématique et à stabiliser les règles de connection du formalisme à l'expérience et il faut se contenter d'avoir seulement une partie calculable du formalisme en rapport avec des expériences réalisables. L'électrodynamique classique est un exemple de théorie ouverte dans le passé de la théorie quantique des champs : le formalisme mathématique était entâché de l'absurdité des divergences ou de problèmes de structure de l'électron, et aucune confiance ne devait être attribuée aux règles d'interprétation des charges et des champs pour des distances de l'ordre du rayon classique de l'électron. De même les divergences de l'électrodynamique quantique non renormalisée peuvent être vues comme une manifestation du caractère ouvert de la théorie, et la renormalisation ne serait qu'une manière d'extraire des résultats observables d'une théorie incomplète. D'ailleurs, même en 1982, l'électrodynamique quantique a un statut mathématique boiteux, puisqu'il est probable qu'il n'existe aucun moyen cohérent de resommer la série des perturbations (dont les termes sont finis par renormalisation, mais dont la somme diverge fortement). Cela n'enlève ni l'espoir que cette difficulté s'évanouisse un jour dans une théorie plus vaste dont l'électrodynamique pure ne serait qu'un secteur, ni les succès expérimentaux de cette théorie boiteuse (6 chiffres significatifs vérifiés par l'expérience pour le moment magnétique anormal de l'électron). Dans les années trente-quarante le refus des divergences peut être assimilé au refus d'une théorie ouverte ; les théories futures dont Dirac et Heisenberg rêvaient devaient être mathématiquement propres et définir clairement les grandeurs observables, à défaut d'être définitives.

Une autre explication de la durée de la crise des divergences tient au rôle de frein joué par le principe de correspondance. Avec l'avènement de la nouvelle mécanique quantique en 1925-27, ce principe était passé du

stade des recettes magiques à celui de règles précises de quantification. Ces règles appliquées stricto sensu donnèrent lieu à plusieurs difficultés dans la théorie de l'électron, et même le problème des divergences peut être vu comme une conséquence d'une application naïve de ces règles. Parmi les difficultés rencontrées dans le programme de quantification relativiste citons les énergies négatives de l'équation de Dirac, et les divergences de la théorie de la mer de Dirac. Il fallut beaucoup d'ingéniosité à Dirac pour dépasser ces difficultés malgré un a priori encore tiré de l'idée de correspondance : Dirac pensait que les électrons étaient indestructibles et il refusa longtemps d'utiliser la 2^o quantification de Wigner-Jordan de l'onde de Dirac et encore plus la version moderne symétrique par conjugaison de charge où le concept de mer disparaît complètement en même temps que l'idée de conservation du nombre de particules. L'obstacle essentiel rencontré dans la théorie quantique relativiste de l'électron peut être ainsi formulé : la théorie "classique" donnant cette théorie par application directe des règles de quantification usuelles (c'est-à-dire la théorie de l'onde ψ de Dirac) diffère de la théorie classique obtenue en prenant la limite classique ($\hbar \rightarrow 0$) de la version quantique (c'est-à-dire la dynamique relativiste de l'électron ponctuel). Corollaire : l'équation de Dirac prise comme équation d'une seule particule était absurde. Finalement la base de correspondance de l'électrodynamique moderne est purement ondulatoire (onde de Dirac couplée au champ de Maxwell) et dépourvue de singularités ponctuelles et de divergences, ce qui explique que le principe de renormalisation de Kramers (à l'encontre de l'opinion de son inventeur) doit être appliqué au niveau quantique et qu'il puisse être interprété comme une altération des règles naïves de correspondance.

Doit-on regretter que l'idée de renormalisation n'ait porté tous ses fruits qu'à partir de 1947 ? Probablement non. La théorie quantique des champs a pu vivre sans : les effets radiatifs de l'électrodynamique pure étaient très faibles, et les divergences de la physique nucléaire d'avant-guerre n'auraient pu de toute façon être toutes éliminées par renormalisation. De plus l'accumulation de progrès formels et de recherches de théories alternatives pendant cette période, ont permis une croissance explosive de la théorie des champs en 1947-49. En ce sens, la crise des divergences, loin d'avoir ralenti le progrès des théories concernées, y

a participé.

Enfin quelles furent la nature et l'importance du changement intervenu en 1947 en physique quantique ? Nous pensons, comme Steven Weinberg⁽⁹⁰⁾ qu'il ne peut s'agir de révolution : nous sommes en présence d'un cas (sans précédent en histoire des sciences ?) de crise aiguë dont la résolution se passa de révolution. Aucun des principes fondamentaux de la relativité ou de la théorie quantique ne fut renversé, il fallut seulement rajouter au cadre formel préexistant de la théorie quantique des champs un lien précis entre les grandeurs calculées et les observations expérimentales donné par le principe de renormalisation. Ce principe n'avait de nouveau que sa vertu à éliminer les infinis puisque Kramers, pour la masse, l'avait tiré de la théorie électromagnétique classique, et Weisskopf, pour la charge, s'était inspiré de l'analogie avec les diélectriques classiques. On ne peut même pas dire que l'année 1947 marque une discontinuité dans l'évolution des problèmes les plus urgents qui étaient ceux de la physique nucléaire. Il s'avéra même que le principe de renormalisation ne suffisait pas à éliminer les divergences des théories de Fermi et de Yukawa.

Mais justement l'évolution future de ce problème montre que l'idée de renormalisation fut tout aussi riche qu'une idée révolutionnaire. Loin d'être rejetée pour les interactions fortes et faibles de la physique nucléaire, elle a évolué en "critère de renormalisabilité" : on chercha de nouvelles théories renormalisables des forces nucléaires, modelées sur l'exemple de la théorie la plus précise qui ait jamais existé : l'électrodynamique quantique. Et ce furent les théories de jauge.

BIBLIOGRAPHIE GENERALE

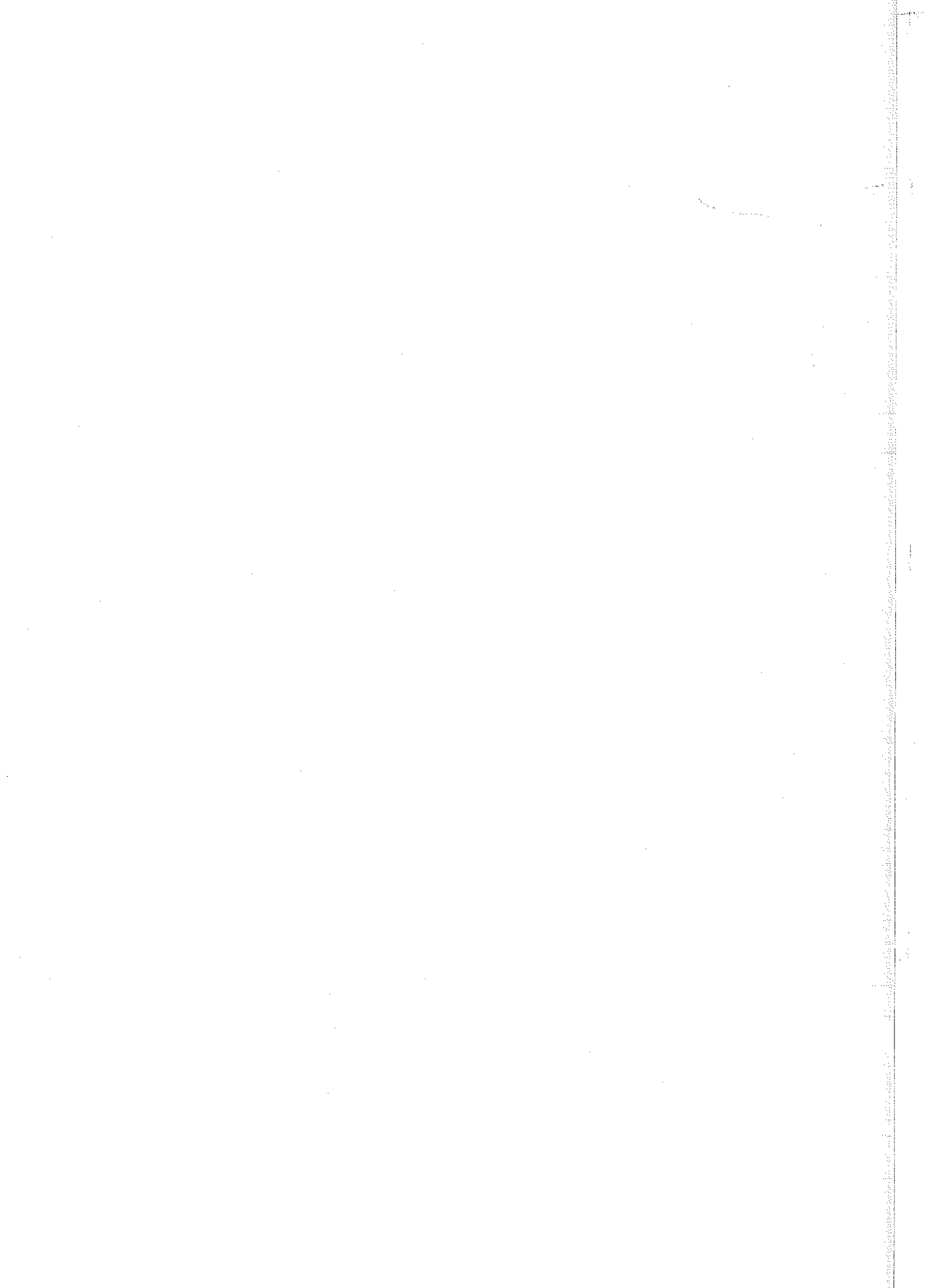
- "Theoretical Physics in the 20th century" (Pauli Memorial)
Fierz and Weisskopf Ed. Interscience (1960)
En particulier l'article de Wentzel : "Quantum Theory of Fields" (until 1947)
- "Aspects of quantum theory" Salam and Wigner Ed. Cambridge (1972)
contient une bibliographie de Dirac et plusieurs études historiques des contributions de Dirac à la physique théorique (Mehra, Jost, Pais ...)
- "The Physicists' conception of Nature" Mehra Ed. Reidel (1973)
En particulier les articles de Rohrlich, Peierls, Wentzel et Tomonaga
- "Selected Papers on Q.E.D." Schwinger Ed. Dover (1958)
- V. Weisskopf Selected Essays MIT Press (1972)
- V. Weisskopf Gregory Lectures 1979
"The Beginning of Field Theory", A personal recollection. CERN 30/10/79
- Steven Weinberg "The Search for Unity : Notes for a History of Q.F.T".
Daedalus 1977, 106
- D. Cassidy "Cosmic ray showers, high energy physics and Q.F.T.
Programmic interaction in the 1930s" HSP512 : 1 (1981)
- Joan Bromberg : "The concept of particle creation before and after
Quantum Mechanics" HSPS 7 (1976) 161
- G. Holton : "Heisenberg, Oppenheimer et l'émergence de la physique
moderne. La Recherche 130 (1982) 190
- D.F. Moyer : "Origin of Dirac's electron, 1925-28 Am. J. of Phys.
49 n° 10
"Evaluation of Dirac's electron, 1928-1932 Am. J. of Phys. 49 n° 11
"Vindications of Dirac's electron", 1932-34 ibid. n° 12

Les deux manuels qui suivent contiennent de nombreuses références d'articles originaux.

- Heitler : "The quantum theory of radiation" Oxford 1st Ed 1936
2nd Ed 1944
3rd Ed 1954
- Schweber, Bethe, de Hoffmann vol I "Fields"
Bethe, de Hoffmann vol II "Mesons"
Row, Peterson (1955)

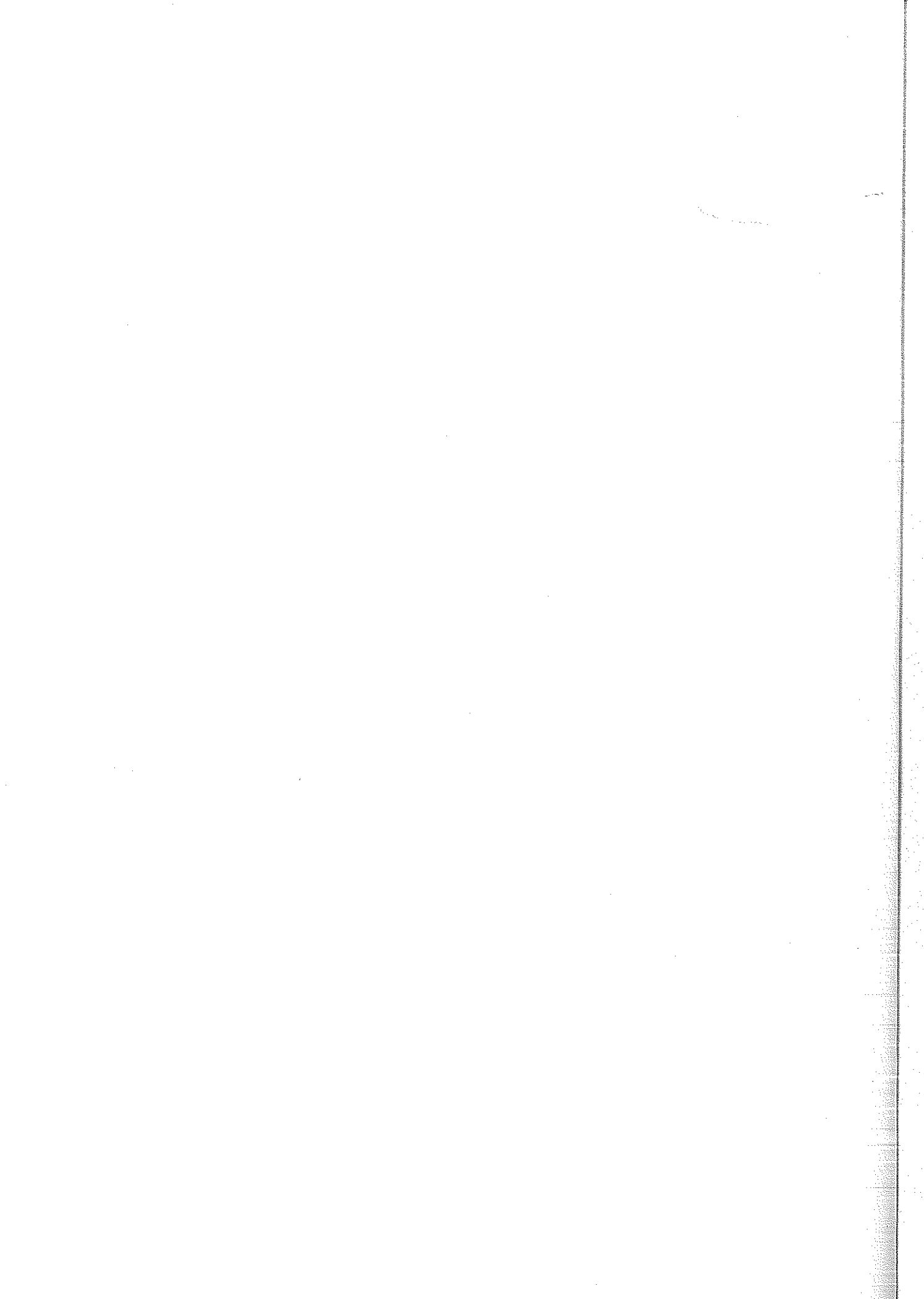
- W. Pauli Wissenschaftlicher Briefwechsel Band I Springer 1979
(Hermann, V. Meyenn, Weisskopf Ed.)

Des sources bibliographiques plus détaillées seront données par chapitre.



CHAPITRE I

LA QUANTIFICATION DU CHAMP ELECTROMAGNETIQUE



1.1 LES QUANTA DE RAYONNEMENT AU DEBUT DE LA THEORIE QUANTIQUE :

La plupart des historiens s'accordent à voir une des origines de la théorie quantique dans l'étude du rayonnement du corps noir. Il est donc naturel de chercher dans les travaux de cette période les premières hypothèses sur la nature quantique du rayonnement électromagnétique. De l'histoire maintenant bien connue de la genèse de la théorie quantique ^{*}, nous extraierons ce qui ressemble le plus à une quantification du champ électromagnétique.

La première démonstration de la loi de Planck, par PLANCK (1) lui-même (1900), ne contient pas d'hypothèse quantique pour le rayonnement : égalisant l'émission et l'absorption des oscillateurs ("résonateurs") liés à la paroi du corps noir, Planck établit le lien entre la distribution spectrale u_ν du corps noir et l'énergie moyenne U d'un oscillateur de fréquence ν , $u_\nu = \frac{8\pi\nu^2}{c^3} U$, et le problème de thermodynamique est rapporté aux oscillateurs atomiques. Non sans une certaine répugnance, Planck finit par utiliser la méthode statistique de Boltzmann (qu'il condamnait un peu avant) pour calculer l'entropie $S(U)$ d'une assemblée de N oscillateurs. Pour que la probabilité statistique associée à la répartition de l'énergie NU entre les N oscillateurs de fréquence ν soit finie, il fallait que l'énergie fût divisée en éléments ("Energieelemente") ϵ . L'entropie $S(U)$ ainsi calculée permet de déterminer $U(T)$ puis u_ν . Le résultat final de ce calcul aurait dû être obtenu en faisant tendre vers zéro les paquets ϵ . Mais une telle limite aurait donnée la loi de Rayleigh-Jeans. Pour obtenir la loi de Planck, Planck dut garder ϵ fini et de valeur $h\nu$, mais il n'attribua pas de signification essentielle à ce découpage discret de l'énergie des résonateurs.

Dans les célèbres "Vorlesungen" (3) Planck exposa aussi (§ 165) la méthode suivie indépendamment par RAYLEIGH (2) (4) et JEANS (5), dont l'intérêt comme le souligna Planck, était de ne faire intervenir que le système thermodynamique que constitue le rayonnement à lui seul. Le champ de Maxwell dans une cavité est donné par une superposition de modes d'oscillation, qui peuvent être considérés comme autant de degrés de liberté indépendants du système.

* Voir par exemple les travaux de Armin Hermann (III) , Max Jammer (IV) Thomas Kuhn (V).

La forme quadratique du hamiltonien total du rayonnement permet l'application du théorème d'équipartition de l'énergie (kT par mode). Le choix de la forme cubique pour la cavité permet de calculer la densité de modes $\frac{8\pi\nu^2}{c^3}$ *, d'où $U_\nu = \frac{8\pi\nu^2}{c^3} kT$. Cette expression n'est valable que pour les hautes températures et les basses fréquences. Planck conclut (§ 166) : "*de cette limitation résulte une difficulté certaine pour l'application de la mécanique statistique aux processus de rayonnement.... je pense que cette difficulté doit être attribuée à une application injustifiée de la loi d'équipartition de l'énergie*". Comme il est tout aussi incorrect d'appliquer le théorème d'équipartition aux résonateurs de Planck, on peut se demander pourquoi Planck ne proposa pas de découper l'énergie des modes, tout comme il avait découpé celle des résonateurs.

Dans l'année même (1906) de la publication des "Vorlesungen", EHRENFEST (6) soutint modestement ce dernier point de vue. A propos du livre de Planck il fit une première remarque : formellement les oscillateurs matériels de Planck et les modes propres de la cavité interviennent de la même façon "*pour autant que les oscillateurs ne sont définis par rien d'autre que leurs équations linéaires homogènes*". Bien sûr l'idée de considérer le champ électromagnétique dans une cavité comme un ensemble d'oscillateurs harmoniques n'avait rien de nouveau, elle découlait directement de l'analyse de Fourier du champ, mais elle devint très intéressante une fois appliquée aux problèmes posés par Planck :

- il devint évident que des résonateurs sans frottement ne pouvaient "thermaliser" le rayonnement : la distribution spectrale initiale de l'énergie de l'ensemble de tous les oscillateurs (réels et fictifs) se conservait au cours du temps et donc ne tendait pas vers celle du corps noir.

- la méthode des éléments d'énergie devait pouvoir s'appliquer directement aux oscillateurs fictifs, modes du champ et c'est ce que montra Ehrenfest. Il était ainsi établi que le problème de l'équilibre thermodynamique du rayonnement pouvait être traité isolément, sans faire intervenir les propriétés des parois, à condition de supposer que "*les modes de fréquence ν ne peuvent avoir que des*

* L'erreur de Rayleigh dans cette détermination a été corrigée par Jeans.

énergies ϵ_ν , multiples entiers du quantum d'énergie $\epsilon_\nu^0 = h\nu \dots$, comme si pour chaque fréquence ν , l'énergie de l'oscillateur se composait d' "atomes d'énergie"....." (6).

Néanmoins, la plupart des physiciens ne cherchèrent pas à donner une signification autre que formelle aux éléments d'énergie. Planck était aussi conservateur que possible : il accusait l'application de la loi d'équipartition au problème de l'équilibre du rayonnement mais pensait qu'une étude statistique classique correcte devait donner la loi de Planck. Pour Rayleigh et pour Jeans, la distribution d'équilibre thermodynamique était donnée par la loi de Jeans, (en fait il n'y avait pas d'équilibre possible puisque l'énergie totale d'équilibre donnée par cette loi diverge), et la distribution de Planck observée expérimentalement devait être considérée comme un état transitoire. Seul EINSTEIN eut l'audace dès 1905, de donner une interprétation corpusculaire à des quanta d'énergie. C'est dans le fameux volume 17 de la revue "Annalen der Physik" (7) qu'Einstein formula l'hypothèse des quanta lumineux sous le titre : "D'un point de vue heuristique sur l'émission et l'absorption de la lumière". Critiquant la différence formelle qui existe entre la théorie des gaz et la théorie de Maxwell - respectivement corpusculaire et ondulatoire- Einstein proposa de considérer que l'énergie lumineuse était elle aussi distribuée de façon discontinue dans l'espace. A l'appui de ce point de vue vint le calcul de l'entropie du rayonnement du corps noir, dans la limite de faible densité (loi de Wien). En fonction du colume $S_\nu(\nu) - S_\nu(\nu_0) = k \ln \left(\frac{\nu}{\nu_0} \right)^{\frac{\nu u_\nu}{h\nu}}$, expression identique à celle de l'entropie d'un gaz parfait d'atomes ponctuels : $S_\nu(\nu) - S_\nu(\nu_0) = k \ln \left(\frac{\nu}{\nu_0} \right)^N$ à condition d'intentifier $\frac{\nu u_\nu}{h\nu}$ au nombre de particules. Le rayonnement du corps noir pouvait être assimilé dans ce cas à un gaz parfait de quanta lumineux. Einstein exploita aussitôt cette hypothèse corpusculaire pour expliquer des caractéristiques expérimentales connues des phénomènes d'émission et d'absorption :

- La règle de Stokes suivant laquelle la fréquence de la lumière de fluorescence est toujours inférieure à la fréquence de la radiation monochromatique incidente ($h\nu_2 \leq h\nu_1$)
- L'effet photoélectrique, les lois expérimentales de Lenard : existence d'un seuil, pente universelle de la tension d'arrêt en fonction de la fréquence -reçoivent l'explication simple et percutante qu'on enseigne aujourd'hui dans tous les cours de physique atomique. Une autre justification de l'hypothèse des quanta lumineux fut donnée par Einstein en 1909 (5) :

l'expression de fluctuations d'énergie dans un petit volume obtenue à partir de la loi de Planck est composée de deux termes :

- l'un correspondant aux fluctuations d'un gaz parfait de quanta
- l'autre aux fluctuations "classiques" obtenues par la loi de Jeans.

Malgré son pouvoir heuristique l'hypothèse des quanta fut longtemps rejetée par les théoriciens influents : elle ne connut pas un succès aussi rapide que la théorie de la relativité. L'étude de son implantation progressive dans le cadre nouveau de la première mécanique quantique sort du cadre de cet exposé. Nous citerons seulement l'article ⁽⁸⁾ de DEBYE de 1910, où se trouve exposée la base de la quantification du rayonnement dans une cavité, à la manière d'Ehrenfest (bien que ce dernier ne soit pas cité). Fort des critiques d'Einstein, Debye remit en question la déduction de Planck de la loi de Planck : la preuve de Planck faisait appel d'une part à une représentation purement classique des oscillateurs matériels quand il s'agissait d'établir la formule $u_\nu = \frac{8\pi\nu^2}{c^3} U$, puis quantifiait l'énergie des résonateurs pour établir l'expression de U. Pour lever cette contradiction Debye utilisa l'analogie formelle entre le champ électromagnétique dans une cavité, et un ensemble d'oscillateurs harmoniques, et quantifia ces oscillateurs, exactement comme Ehrenfest, afin de retrouver la loi de Planck. " *Finally les considérations précédentes ne permettent pas de trancher la question : doit-on regarder l'existence de quanta comme une propriété intrinsèque de l'éther, ou, ainsi qu'on l'imaginait plus volontiers jusqu'à ce jour, comme une propriété de la matière ?* " conclut prudemment Debye.

Le problème du corps noir demeurait obscur car Planck et Einstein avaient introduit deux types différents de quanta $h\nu$: ceux de Planck concernaient les échanges d'énergie des "résonateurs" avec le rayonnement, ceux d'Einstein étaient des particules de rayonnement. Une autre différence entre les deux types de quanta fut éclaircie en 1911 dans deux publications indépendantes de NATANSON ⁽¹⁷³⁾ et d' EHRENFEST ⁽¹⁷⁴⁾. Dans la terminologie moderne, les quanta d'énergie des résonateurs de Planck étaient indiscernables et ceux d'Einstein discernables. La discernabilité des quanta lumineux d'Einstein conduisait nécessairement à la loi de Wien. Pour obtenir la loi de Planck, il fallait, comme BOSE ⁽⁷⁰⁾ en 1924, dénombrer les répartitions possibles de quanta indiscernables. Grâce à Bose, le calcul d'Ehrenfest-Debye obtenu en donnant à l'énergie totale du rayonnement la structure $U = \sum N_\nu h \nu$ (N_ν entier) s'harmonisait parfaitement avec l'idée d'Einstein d'un gaz de quanta lumineux.

Il devenait clair que la donnée d'une suite de N_ν spécifiait un état unique du rayonnement comprenant pour chaque ν , N_ν quanta indiscernables. En 1926, Heisenberg, Born et Jordan n'eurent plus qu'à démontrer, à l'aide de la nouvelle mécanique quantique, la formule $U = \sum_\nu N_\nu h\nu$ pour que soit justifiée l'idée d'Einstein des quanta lumineux et aussi leur statistique.

1.2 LES FLUCTUATIONS DU CORPS NOIR A L'AUBE DE LA MECANIQUE DES MATRICES :

Nous avons déjà mentionné comment EINSTEIN, en 1909 ⁽⁹⁾ s'était servi de l'expression des fluctuations du corps noir dérivée de celle de l'entropie par inversion de la formule de Boltzmann $S = k \ln W$, pour justifier l'hypothèse des quanta. Dans le même article, Einstein mit le doigt sur une difficulté importante : on peut calculer les mêmes fluctuations par une autre méthode, c'est-à-dire en faisant la moyenne temporelle de l'écart quadratique de l'énergie lors de l'évolution temporelle d'un état initial donné du rayonnement de la cavité, puis la moyenne statistique sur ces états initiaux. Le résultat ainsi obtenu donne, pour une fréquence donnée, $\overline{\Delta^2} = \frac{\overline{E}^2}{z_{\nu} V}$, où \overline{E} désigne l'énergie moyenne dans le sous volume V et $z_{\nu} dv$ le nombre de modes dans dv alors que le résultat déduit de la loi de Planck est $\overline{\Delta^2} = hv \overline{E} + \frac{\overline{E}^2}{z_{\nu} V}$. Ce calcul classique des fluctuations fut explicité par Lorentz ⁽¹⁰⁾ en 1916 dans "les théories statistiques en thermodynamique" et commenté par EHRENFEST ⁽¹¹⁾ en 1925 lors d'un séminaire à Göttingen durant l'été 1925. Ehrenfest expliqua la divergence des deux méthodes par la non-additivité de l'entropie du volume partiel V et du volume complémentaire.

Dans la deuxième partie ⁽¹²⁾ (désignée par "Dreimännerarbeit") de la construction impressionnante de la mécanique des matrices * , BORN HEISENBERG, et JORDAN (BHJ) ** prouvèrent à la fin de la même année que c'était en fait la dynamique classique du rayonnement qu'il fallait rejeter, en appliquant pour la première fois le formalisme général de la mécanique quantique à un problème de théorie des champs. Leur attention fut attirée sur ce problème par le séminaire d'Ehrenfest. L'oscillateur harmonique étant le problème lié le plus simple de la nouvelle mécanique, il était naturel d'examiner les conséquences de la nouvelle quantification pour un ensemble d'oscillateurs couplés, comme le champ électromagnétique

* La première partie était l'oeuvre de Born et Jordan ⁽¹⁶¹⁾. On y trouvait déjà l'idée que les champs E et B , en théorie quantique devaient être des champs d'opérateurs.

** Il semble d'après les témoignages de Jordan ⁽¹⁵⁹⁾ et de Van der Waerden ⁽¹⁶⁰⁾ que seul Jordan ait participé à la partie de l'article sur les fluctuations du rayonnement.

ou comme les vibrations d'un réseau cristallin. Enfin l'obtention par cette méthode de l'expression correcte des fluctuations du corps noir était une vérification de la nouvelle mécanique dans un domaine assez éloigné de l'étude des spectres atomiques dont était parti HEISENBERG.

Les modes propres i de la cavité peuvent être considérés comme des oscillateurs harmoniques indépendants dont les états propres d'énergie prennent les valeurs discrètes $n_i \hbar \omega_i + \frac{1}{2} \hbar \omega_i$ (n_i entier). L' "énergie de zéro" $\sum \frac{1}{2} \hbar \omega_i$ diverge pour les grandes fréquences mais, remarquent B.H.J., seules les différences d'énergie sont observables. B.H.J. supposent que les états $\{n_i\}_i = n$ sont équiprobables pour une même énergie totale $E = \sum_i n_i \hbar \omega_i$ (ensemble microcanonique) et trouvent alors que la distribution spectrale du corps noir est celle donnée par la répartition aléatoire de quanta lumineux obéissant à la statistique de Bose dans chaque mode. Le calcul de Debye ⁽⁸⁾ et Ehrenfest ⁽⁶⁾ est ainsi complètement justifié. Les fluctuations temporelles sont étudiées de manière tout à fait analogue au calcul classique : l'énergie du rayonnement dans un sous-volume est exprimée en fonction des coordonnées normales des oscillateurs, ainsi que l'écart quadratique instantané, en prenant soin de ne pas commuter les coordonnées. La moyenne temporelle de la matrice ainsi obtenue est une matrice diagonale (on est, bien sûr, en représentation de Heisenberg) $\bar{\Delta}^2$ reliée à la matrice diagonale d'énergie \bar{E} par $\bar{\Delta}^2 = \hbar \nu \bar{E} + \frac{\bar{E}^2}{z \nu}$, conformément à la formule d'Einstein. *

* B.H.J. arrêtent là le raisonnement. Pour comparer plus exactement à la formule d'Einstein, il conviendrait de passer de l'identité opératorielle de B.H.J. à une moyenne sur la distribution statistique des états $|n\rangle$.

Posons : $\langle X \rangle = \sum \langle n | X | n \rangle e^{-\beta E_n}$. Alors $\langle \Delta^2 \rangle = \hbar \nu \langle \bar{E} \rangle + \frac{\langle \bar{E}^2 \rangle}{z \nu} +$
termes d'ordre supérieur en $\frac{1}{\nu}$.

En résumé B.H.J. ont, pour la première fois quantifié le champ électromagnétique libre ^{*} en se fondant sur l'analogie avec un ensemble d'oscillateurs harmoniques, ils en ont déduit la structure discrète de l'énergie du champ et son interprétation en terme de bosons (au lieu de la supposer a priori comme Ehrenfest et Debye), tout cela en vue de l'étude statistique du rayonnement du corps noir.

* Une telle affirmation est légèrement abusive ; en fait B.H.J. quantifient une corde vibrante, c'est-à-dire un champ scalaire libre unidimensionnel. Les fluctuations d'un tel champ ne diffèrent de celles du champ électromagnétique que par le dénombrement des modes.

1.3 "LA THEORIE QUANTIQUE DE L'EMISSION ET DE L'ABSORPTION DU RAYONNEMENT
DE P.A.M. DIRAC (13) (2 Février 1927)

On ne saurait trop souligner la profondeur et l'originalité de ce travail de Dirac. En quelques pages d'une clarté et d'une concision remarquables, un point de vue unitaire fut développé sur un ensemble très vaste de phénomènes ; nombre de vieilles contradictions ou suppositions arbitraires se trouvèrent dissipées, la voie féconde de la théorie quantique des champs était ouverte.

1.3 a Circonstances et motivations

Sans entrer dans trop de détails biographiques ^{*}, nous précisons seulement dans quel cadre et dans quel entourage le jeune Paul Adrien Maurice Dirac menait ses premières recherches de physique quantique.

En 1921, à l'âge de 19 ans, Dirac acheva sa formation d'ingénieur électricien à Bristol. N'ayant pu trouver d'emploi dans ce domaine, il poursuivit des études en mathématiques appliquées et put quitter Bristol pour Cambridge en automne 1923 grâce à deux bourses d'étude. Il fut accepté comme "research student" au St John's college. Plusieurs excellents cours étaient alors donnés en mathématiques et en physique théorique, deux domaines très liés administrativement, alors que la physique expérimentale relevait d'un autre département. Les méthodes hamiltoniennes de la mécanique analytique, alors très en vogue, retinrent particulièrement l'attention de Dirac. L'influence de FOWLER, son directeur de recherche, fut particulièrement bénéfique. Fowler faisait preuve d'un esprit enthousiaste et curieux pour tout ce qui touchait la théorie quantique, et ramenait fréquemment des nouvelles fraîches de Copenhague. Dès 1922, Dirac commença à publier sur des problèmes de mécanique statistique et sur l'utilisation des invariants adiabatiques en théorie quantique, puis ce fut la série de papiers publiés en 1925 et en 1926 sur la mécanique quantique, sous l'impulsion du premier travail de Heisenberg sur la mécanique des matrices. Les jeunes chercheurs de Cambridge étaient alors assez isolés les uns des autres, en raison de l'organisation de la vie des collègues. A tel point que Mott rapporte qu'il passa presque toute l'année 1927 à lire les papiers de Dirac sur la théorie du rayonnement sans avoir eu aucune discussion avec lui sur ce sujet ⁽¹⁵⁾. L'organisation des cours était très

* J. Mehra⁽¹⁴⁾ a publié, à l'occasion du 70ème anniversaire de Dirac, une excellente biographie s'arrêtant en 1933, année d'attribution du Prix Nobel à Dirac et à Schrödinger pour leur "découverte de nouvelles formes fertiles de la théorie atomique et de ses applications".

formelle si l'on en croit Lady Jeffreys⁽¹⁵⁾ : " ces leçons étaient toujours suivies par le même petit groupe. Il y a trop longtemps pour que je sois sûre, mais je pense que Hartree, Whittaker, Wilson, Gaunt, Mott et Oppenheimer étaient présents[...] Fowler avait un don remarquable pour assimiler rapidement de nouveaux travaux et pour communiquer son enthousiasme. DIRAC nous présenta ce qu'il avait fait récemment, seulement en partie déjà publié, je pense. Nous ne constituions pas, il est vrai, un groupe très sociable, mais pour quiconque était présent, il est impossible d'oublier l'émotion ressentie. J'étais intimidée par Dirac, mais si j'avais le courage de lui poser une question j'obtenais toujours une réponse directe et utile, sans qu'il fît cas de mes erreurs". L'atmosphère était sans doute très différente à Copenhague, où des physiciens de tous pays se rencontraient régulièrement sous l'influence de la haute personnalité de BOHR, et intervenaient librement à tout moment lors des colloques. Dirac, sur les conseils de Fowler se rendit à Copenhague pour six mois à l'automne 1926. Dirac était alors armé de sa "théorie des transformations" c'est-à-dire de la formulation la plus abstraite et la plus générale de la mécanique quantique et ainsi la plus adaptée aux applications ultérieures. "La perspective plus large que Dirac obtint à Copenhague l'aïda non seulement à finir le travail d'interprétation de la théorie quantique, mais encore à commencer à étendre les applications de la mécanique quantique"⁽¹⁴⁾ (Mehra). Le texte de l'article sur la théorie de l'émission et de l'absorption du rayonnement, notons le, fut communiqué aux Proc. Roy. Soc. par Niels Bohr lui-même. Dirac eut alors l'ambition d'élaborer une généralisation relativiste de la théorie *."Ce problème de l'invariance relativiste est bien sûr relié aux autres (interactions instantanées à distance, émission de rayonnement par un électron...) et il sera impossible de répondre complètement à l'une quelconque de ces questions, sans les résoudre toutes en même temps. Cependant, il est possible de construire une théorie très satisfaisante de l'émission de rayonnement et de la réaction du champ de rayonnement sur le système émetteur sur la base d'une cinématique et d'une dynamique qui ne soient pas strictement relativistes"(introduction de l'article⁽¹³⁾). Mehra⁽¹⁴⁾ fait remarquer que cette façon d'aborder les problèmes est propre à Dirac. Elle s'oppose en particulier à celle de Heisenberg qui pensait que l'on devait traiter simultanément tout un groupe de problèmes . Dirac

* En réalité cette préoccupation était déjà apparue dans plusieurs articles antérieurs de Dirac : par exemple avec le formalisme "Energie-angle" et l'équation hamiltonienne. Nous ne traiterons pas ici de ces travaux de Dirac antérieurs à 1927 .

commentait lors d'une conversation avec Mehra en 1969 : "Beaucoup de gens ont le point de vue de Heisenberg, mais c'est seulement trop difficile de résoudre tout un ensemble de problèmes en même temps. Je me suis rendu compte assez tôt (déjà à Bristol) qu'il n'y avait rien dans la nature que d'approximatif, et que la science se développerait pour donner des approximations de plus en plus précises, mais qu'elle n'atteindrait jamais l'exactitude complète. J'ai tiré ce point de vue de ma formation d'ingénieur, qui je pense, m'a beaucoup influencé...." (14).

Dirac résume ainsi son article : "Nous traitons le problème d'un ensemble de systèmes satisfaisant la mécanique statistique d'Einstein - Bose interagissant avec un autre système, et un hamiltonien est obtenu pour décrire l'évolution. La théorie est appliquée à l'interaction d'un ensemble de quanta lumineux avec un atome ordinaire, et il est montré que cela donne les lois d'Einstein de l'émission et de l'absorption du rayonnement. Puis nous considérons l'interaction d'un atome avec des ondes électromagnétiques et montrons que si l'on considère les énergies et les phases des ondes comme des q -nombres satisfaisant les conditions de quantification et non pas comme des c -nombres, le hamiltonien prend la même forme que dans la méthode des quanta lumineux. La théorie donne les expressions correctes pour les coefficients d'Einstein A et B".

Nous reprendrons ici les points importants des déductions de Dirac et avec les notations modernes (bien peu différentes de celles de Dirac, puisque les siennes se sont imposées par leur élégance) *

1.3.b "Seconde quantification" et évolution d'un ensemble de bosons
(§§ 2,3,4) (13).

(i) Soit $H = H_0 + V$ la hamiltonien d'une particule, $|\alpha\rangle$ les états propres de H_0 . L'équation de Schrödinger projetée sur les $|\alpha\rangle$ donne :

$$\begin{cases} i\hbar \dot{a}_\alpha = \sum_\beta H_{\alpha\beta} a_\beta \\ -i\hbar \dot{a}_\alpha^* = \sum_\beta H_{\beta\alpha} a_\beta^* \end{cases} \quad (H_{\alpha\beta} = H_{\beta\alpha}^*) \quad , \text{ les } a_\alpha \text{ étant les coefficients de la fonction d'onde dans la base } |\alpha\rangle$$

* La notation bra-ket a été introduite par Dirac en 1939. Il utilisait auparavant une notation très voisine.

Ce système peut être considéré comme les équations canoniques associées au hamiltonien $F = \sum_{\alpha\beta} a_{\beta}^* H_{\alpha\beta} a_{\alpha}$, a_{α}^* étant conjugué de a_{α} . Considérons maintenant ce hamiltonien comme un hamiltonien quantique (2ème quantification). Les relations canoniques sont $[a_{\alpha}, a_{\beta}^+] = \delta_{\alpha\beta}$ identiques à celles des créateurs - annihilateurs d'un ensemble d'oscillateurs harmoniques. La base naturelle d'états est $|n_1 \dots n_{\alpha} \dots\rangle$ et $a_{\alpha} |n_1 \dots n_{\alpha} \dots\rangle = \sqrt{n_{\alpha}} |n_1 \dots n_{\alpha}-1 \dots\rangle$, $a_{\alpha}^+ |n_1 \dots n_{\alpha} \dots\rangle = \sqrt{n_{\alpha}+1} |n_1 \dots n_{\alpha}+1 \dots\rangle$. Un état quelconque du nouveau système s'écrit $|\psi\rangle = \sum_{\{n_{\alpha}\}} c_{n_1 \dots n_{\alpha} \dots} |n_1 \dots n_{\alpha} \dots\rangle$ et l'équation de Schrödinger devient :

$$i\hbar \frac{\partial}{\partial t} c_{n_1 \dots n_{\alpha} \dots} = \sum_{\alpha\beta} H_{\alpha\beta} \sqrt{n_{\alpha}} \sqrt{n_{\beta}+1} \delta_{\alpha\beta} c_{n_1 \dots n_{\alpha}-1 \dots n_{\beta}+1 \dots} \quad (1)$$

Remarquons que $\sum_{\alpha} N_{\alpha} = N$ est conservé lors de l'évolution. ($[N, F] = 0$)

ii) Soit maintenant N bosons sans interaction, et H_i leurs hamiltoniens individuels (exemplaires de H). Le hamiltonien total est $\mathcal{H} = \sum_{i=1}^N H_i$. Au lieu d'utiliser la base $|\alpha_1\rangle \dots \dots \dots |\alpha_N\rangle$ et des coefficients symétriques, on peut symétriser cette base suivant $\frac{1}{N!} \sum_{\sigma \text{ permutations}} |\alpha_{\sigma(1)}\rangle \dots \dots \dots |\alpha_{\sigma(N)}\rangle = |n_1, \dots, n_{\alpha}, \dots\rangle$, les nouveaux états étant seulement caractérisés par les nombres d'occupation des α . La racine normalise à 1 ces états.

L'équation de Schrodinger $i\hbar \frac{\partial}{\partial t} |\phi\rangle = \mathcal{H}|\phi\rangle$ exprimée dans cette nouvelle base est identique à (1).

La 2ème quantification du (i) donne donc l'évolution d'un ensemble de bosons, dans les variables naturelles $\{n_{\alpha}\}$, nombres d'occupation des états propres de H.

iii) On suppose maintenant que l'interaction décrite par V dans $H = H_0 + V$ se fait avec un autre système P de hamiltonien H_P . La méthode de la 2ème quantification s'applique aussi et le hamiltonien résultant est

$$F = H_P + \sum_{\alpha} E_{\alpha} a_{\alpha}^+ a_{\alpha} + \sum_{\alpha\beta} v_{\alpha\beta} a_{\beta}^+ a_{\alpha}$$

agissant sur les variables de P et sur les $\{n_{\alpha}\}$.

1.3.c Application de la 2ème quantification à l'interaction d'un atome avec un ensemble de photons (§§ 5,6) (13)

La théorie de la diffusion de Born (1926 - états de diffusion stationnaires) ne s'appliquant pas dans ce cas, Dirac traite le problème général des transitions

d'un état initial dans un continuum, à l'ordre le plus bas dans le potentiel d'interaction, et relie la section efficace de diffusion aux éléments de matrice du potentiel d'interaction (relation appelée aujourd'hui 2ème règle d'or de Fermi).

Le hamiltonien F de (iii) représente maintenant un atome couplé à un ensemble de photons. H_p est le hamiltonien de l'atome et les α indicent les photons suivant leur fréquence et leur polarisation. F conserve le nombre total de photons et ne semble donc pas pouvoir rendre compte des phénomènes d'absorption ou d'émission d'un photon. Dirac fait alors remarquer qu'un photon de fréquence nulle n'a ni impulsion ni énergie, il cesse donc d'exister. On doit donc considérer qu'il y a une infinité de photons dans le mode de fréquence nulle (le rôle particulier attribué aux photons mous présage de leur rôle futur dans le traitement du problème infra-rouge), disons a_0 et $a_0^+ \rightarrow \infty$ et $v_{0\alpha}$ et $v_{\alpha 0} \rightarrow 0$ pour que $v_{0\alpha} a_0$ et $v_{\alpha 0} a_0^+$ aient les limites finies respectives v_α et v_α^* .

Alors
$$\sum_{\alpha\beta} v_{\alpha\beta} a_\beta^+ a_\alpha \rightarrow \sum_{\alpha \neq 0} (v_\alpha a_\alpha + v_\alpha^* a_\alpha^+) + \sum_{\substack{\alpha \neq 0 \\ \beta \neq 0}} v_{\alpha\beta} a_\beta^+ a_\alpha \quad (2)$$

L'émission spontanée d'un photon correspond à une transition de l'état de fréquence nulle à un état $\alpha \neq 0$ et inversement pour l'absorption. La probabilité d'absorption est proportionnelle à $|\langle n_\alpha - 1 | v_\alpha a_\alpha | n_\alpha \rangle|^2 = n_\alpha |v_\alpha|^2$

d'émission à $|\langle n_\alpha + 1 | v_\alpha^* a_\alpha^+ | n_\alpha \rangle|^2 = (n_\alpha + 1) |v_\alpha|^2$

Elles sont dans le rapport $\frac{n_\alpha}{n_\alpha + 1}$, qui donne la relation d'Einstein entre les coefficients d'absorption et d'émission du rayonnement. *

* Appelons A_{12} le coefficient d'Einstein d'émission spontanée

B_{12} le coefficient d'Einstein d'émission induite

B_{21} le coefficient d'Einstein d'absorption

D'après Einstein (1916) $B_{12} = B_{21}$ et $A_{12} = \frac{8\pi h \nu^3}{c^3} B_{12}$
pour les transitions $1 \leftrightarrow 2$ de fréquence ν .

Le rapport entre absorption et émission pour une densité spectrale ρ_ν de photons est $\frac{ab_s}{\epsilon m} = \frac{\rho_\nu B_{21}}{\rho_\nu B_{12} + A_{12}} = \frac{\rho_\nu}{\rho_\nu + 8\pi h \nu^3 / c^3}$

Or le nombre n_α de photons dans le mode α de fréquence ν_α telle que $\nu < \nu_\alpha < \nu + d\nu$ est le rapport du nombre $\rho_\nu d\nu / h\nu$ de photons dans $d\nu$ au nombre $\frac{8\pi \nu^2}{c^3} d\nu$ de modes dans $d\nu$.

d'où
$$\frac{ab_s}{\epsilon m} = \frac{n_\alpha}{n_\alpha + 1}$$

1.3.d . Quantification canonique du champ de rayonnement

Dans la jauge où le potentiel scalaire est nul, le champ de radiation est donné par le potentiel vecteur \vec{A} , développable en série de Fourier suivant * : $\vec{A}(\vec{r}) = \sum_{k\lambda} C_k (a_{k\lambda} \vec{\epsilon}_{k\lambda} e^{i\vec{k}\cdot\vec{r}} + a_{k\lambda}^* \vec{\epsilon}_{k\lambda}^* e^{-i\vec{k}\cdot\vec{r}})$

Chaque terme représentant un mode de vecteur d'onde \vec{k} et de polarisation λ transverse ($\vec{k}\cdot\vec{\epsilon}_{k\lambda} = 0$; on ne considère que la partie transverse du champ telle que $\text{div } \vec{A} = 0$) ** et C_k étant un facteur de normalisation.

Les variables $a_{k\lambda}$ et $a_{k\lambda}^*$ sont conjuguées par rapport au hamiltonien du rayonnement. La quantification canonique s'exprime par $[a_{k\lambda}, a_{k'\lambda'}^+] = \delta_{k\lambda, k'\lambda'}$. (En fait Dirac prend comme variables canoniques conjuguées la phase et l'énergie d'un mode). Dirac préfère ici le point de vue formel de la quantification canonique d'un système dynamique quelconque par réinterprétation de son algèbre de Poisson, à l'analogie des oscillateurs ***

Le coefficient c_k peut être explicité en comparant l'expression de $\int \frac{1}{2} (E^2 + B^2) d^3r$ déduite de celle de \vec{A} et celle déduite du nombre de photons, $k^2 C_k^2 2a_{k\lambda}^+ a_{k\lambda} v = N_{k\lambda} k \Rightarrow C_k = \frac{1}{\sqrt{2vk}}$ ($\hbar = c = \epsilon_0 = 1$), pour une boîte de volume v .

Dans l'approximation dipolaire électrique, le hamiltonien total est $H = H_P + \sum_{k\lambda} N_{k\lambda} k_0 - \sum_i \frac{e_i}{m_i} \vec{p}_i \cdot \vec{A}(\vec{R}_i)$, e_i, m_i, \vec{p}_i étant respectivement la charge, la masse et l'impulsion des particules constitutives de l'atome.

* En réalité, Dirac se place directement dans l'approximation $e^{i\vec{k}\cdot\vec{r}} \sim 1$. L'utilisation de cette approximation n'est spécifiée que dans l'article suivant sur la théorie de la dispersion (voir § 4)

** Nous expliquerons au chap. 2 § 5 pourquoi Dirac avait raison de choisir cette jauge.

*** Bien qu'il revienne à ce dernier point de vue au § 63 de son cours, "the principles of quantum mechanics", sans doute pour des raisons pédagogiques.

Ce hamiltonien a la même forme que F avec la modification (2). Il conduit par conséquent à des résultats physiques équivalents. Les v_α étant maintenant connu explicitement, on peut retrouver l'expression du coefficient d'Einstein d'émission induite B_{12} donnée par la mécanique des matrices (le champ étant traité classiquement) et prévoir et calculer le phénomène d'émission spontanée.

Au delà des résultats acquis par la seconde quantification d'une part et par la quantification canonique du champ de radiation d'autre part, la théorie de Dirac appelle plusieurs généralisations dont la plupart sont signalées dans cet article et fournissent le thème de recherches ultérieures :

- l'étude des phénomènes d'ordre supérieur : diffusion, largeurs naturelles, interactions à plusieurs photons, réactions de rayonnement,
- la généralisation relativiste de la théorie : l'usage d'une formulation hamiltonienne est critiquée par Dirac. Une théorie relativiste concevable de l'électron fait défaut. En outre le champ coulombien et le champ de radiation ne sont pas traités sur le même pied (seul le champ de radiation est quantifié) comme en électrodynamique classique et comme semblerait l'exiger un traitement relativiste.

Nous verrons dans la suite de ce chapitre comment la plupart de ces problèmes furent résolus dans les quelques années qui suivirent, ainsi que d'autres généralisations comme la quantification canonique du champ de fermions, et la seconde quantification appliquée aux fermions.

Mais déjà des phénomènes qui, comme l'émission spontanée, ne pouvaient être prévus qu'à l'aide de règles de correspondance ad hoc, se trouvaient expliqués par Dirac grâce à l'explicitation des propriétés quantiques du champ électromagnétique. Une fois les ondes quantifiées "*il y a une complète harmonie entre les descriptions corpusculaire et ondulatoire de l'interaction*" (13).

I.4 PREMIERES APPLICATIONS DE LA THEORIE DE DIRAC

La théorie du rayonnement de Dirac , avec ou sans l'incorporation de la théorie relativiste de l'électron et de la théorie du positron, n'a cessé jusqu'à nos jours, de se montrer féconde. Elle aboutit à une phénoménologie très fine de l'atome ; récemment la découverte du laser a permis d'étendre considérablement le champ de ses vérifications expérimentales et a suscité de nouveaux travaux théoriques. * La richesse du contenu empirique de cette théorie fut reconnue dès ses premières années. D'anciens traitements semi-classiques logiquement imparfaits , se trouvèrent justifiés, mais aussi de nouvelles prédictions furent établies et vérifiées après un délai plus ou moins long. Le célèbre ouvrage de HEITLER ⁽³³⁾ , "The quantum theory of radiation" , dans sa première édition de 1936, donne une idée très juste de ces premiers développements. En raison même de leur abondance, il est exclu de les traiter in extenso dans cet exposé, et d'autant qu'ils représentent le domaine privilégié d'application de la théorie des champs où l'on peut faire abstraction des difficultés essentielles qui rendirent nécessaire l'évolution ultérieure de la théorie. Il faut exclure de ce dernier commentaire l'étude des phénomènes mettant en jeu des particules relativistes, sur lesquels nous reviendrons aux chapitres 2 et 6.

Le hamiltonien d'un système de charges non relativistes et du rayonnement s'écrit, selon Dirac :

$$H = \sum_i \frac{p_i^2}{2m_i} + \sum_{i < j} \frac{1}{8\pi} \frac{e_i e_j}{|\vec{r}_i - \vec{r}_j|} + \sum_{\vec{k}, \lambda} N_{k\lambda} \hbar \omega_k + H_I + H_{II}$$

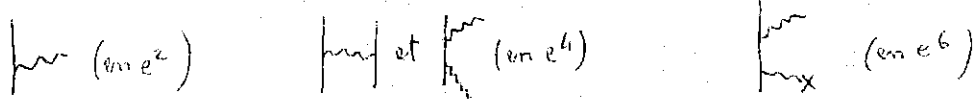
$$H_I = - \sum_i \frac{e_i}{m_i} \vec{p}_i \cdot \vec{A}(\vec{r}_i) \quad H_{II} = \sum_i \frac{e_i^2}{2m_i} \vec{A}^2(\vec{r}_i) \quad \vec{A}(\vec{r}) \text{ étant ,}$$

suivant la quantification canonique, une superposition de créateurs et d'annihilateurs de photons : $\vec{A}(\vec{r}) = \sum_{\vec{k}, \lambda} c_{\vec{k}} \left(a_{\vec{k}\lambda} \vec{\epsilon}_{\vec{k}\lambda} e^{i\vec{k}\cdot\vec{r}} + a_{\vec{k}\lambda}^+ \vec{\epsilon}_{\vec{k}\lambda}^* e^{-i\vec{k}\cdot\vec{r}} \right)$

* Voir par exemple les cours de C. COHEN-TANNOUDJI au Collège de France en 1973-1974-1975-1976 pour une étude systématique des conséquences de la théorie de Dirac.

** Dans son premier article sur l'émission et l'absorption, Dirac ⁽¹³⁾ ne considère que des processus d'amplitude linéaire en e et néglige donc H_{II} .

H_I et H_{II} dans les applications seront traités en perturbation. Il est donc naturel de classer les différents processus suivant la puissance de e^2 qui intervient dans leur probabilité. On les représenterait aujourd'hui par des graphes du type :



Notons que plusieurs de ces applications de la théorie de Dirac font aussi appel à la théorie relativiste de l'électron de Dirac et à des généralisations de la théorie de Dirac (Heisenberg-Pauli.....) dont nous ne parlerons qu'ultérieurement.

1.4.a Processus en e^2 :

Calcul de l'intensité des raies d'émission et d'absorption :

DIRAC ⁽¹³⁾ avait pu immédiatement calculer les coefficients d'Einstein d'émission et d'absorption du rayonnement, en calculant l'effet de la perturbation H_I pour des temps courts. Il restait à préciser l'évolution aux temps longs et à calculer les formes de raies qui en résultent. En Juillet 1927 LANDAU ⁽³⁴⁾ apporte une première généralisation du calcul de Dirac en étudiant l'évolution d'un atome couplé au rayonnement, partant d'un état quantique atomique quelconque. Landau remarque qu'une fonction d'onde initialement factorisée $\Psi = \Psi_{at} \cdot \Psi_{ray}$ ne le reste pas et que la valeur moyenne d'une observable atomique f s'exprime aux instants ultérieurs, après sommation sur les variables du champ, sous la forme $\sum_{nm} \rho_{nm} f_{nm}$ introduisant ainsi sans la nommer la matrice densité ρ_{nm} . Le but de Landau est maintenant de calculer l'évolution de ρ si l'on part d'abord, pour simplifier, d'un état du type $|n_0; 0\rangle$ de niveau atomique n_0 fixé et vide de rayonnement. Sans preuve rigoureuse, l'équation linéaire reliant $\frac{d\rho_{nm}}{dt}$ à $\rho_{n'm'}$ est établie (on s'attend bien sûr à retrouver le comportement exponentiel d'un oscillateur classique amorti). Les ρ_{nm} (populations des différents niveaux) évoluent suivant la loi d'Einstein, les ρ_{nm} (aujourd'hui appelés cohérences) décroissent toutes avec la même constante de temps. Landau donne la signification physique des ρ_{nm} : pour $\rho_{mn} = 0$ on connaît seulement la probabilité d'être dans tel ou tel état atomique n , si $\rho_{mn} \neq 0$ on a une information sur la localisation et la phase de l'électron et donc sur la cohérence du rayonnement émis. La présence de termes oscillants de la forme $e^{i(E_n - E_m)t - t(1/\tau_n + 1/\tau_m)}$ dans la valeur moyenne des observables

atomiques ($\Gamma_n \rho_f$) lui permet d'inférer l'expression $\Delta\omega_{nm} = \frac{1}{\tau_n} + \frac{1}{\tau_m}$ pour la largeur d'une raie et de définir la largeur $\Delta\nu_n$ à terme spectral $\frac{1}{\tau_n}$, anticipant sur les travaux de Wigner et Weisskopf.

Au mois de décembre, Felix BLOCH ⁽³⁵⁾ reprend le problème et aboutit aux mêmes conclusions, sans parler toutefois de largeur de raie. Le traitement est plus rigoureux : l'équation reliant $\dot{\rho}$ à ρ est établie à $t=0$ pour un $\rho(0)$ arbitraire et "comme cet instant ne diffère en rien des autres, on peut la prendre valable à tout autre instant". Elle est résolue exactement dans le cas d'un oscillateur harmonique couplé au rayonnement. Le sens physique des termes ρ_{nm} est expliqué comme chez Landau, grâce, dit Bloch, à une remarque amicale de Heisenberg.

Les largeurs naturelles

Il appartient à WEISSKOPF et WIGNER ^{★(37) (38)} (1930) d'avoir donné un traitement plus rigoureux de l'émission spontanée pour des temps longs et d'en avoir déduit l'expression des formes de raies, et d'avoir souligné la différence qualitative essentielle des conclusions de la nouvelle théorie avec l'ancienne théorie quantique. Dans cette dernière théorie, les niveaux d'énergie sont parfaitement définis et seule l'instabilité du niveau de départ élargit la raie : sa durée de vie finie impose une largeur finie du train d'onde émis et corrélativement une largeur spectrale finie. Si au contraire comme EHRENFEST ⁽³⁶⁾ (1923) on suppose que la relation de conservation de l'énergie $E_i - E_f = h\nu$ est exactement vérifiée, il faut que les niveaux atomiques aient une largeur finie pour expliquer une largeur finie des raies. Une conséquence expérimentale vérifiable de ce point de vue est l'existence de transitions peu probables et pourtant larges (transitions d'un niveau métastable à un niveau instable).

Wigner et Weisskopf résolvent l'équation de Schrödinger du problème avec deux hypothèses simplificatrices : seuls deux niveaux atomiques 1 et 2 interviennent, et seules les transitions à un photon de fréquence $\nu \sim \frac{E_2 - E_1}{h}$ interviennent. Un "Ansatz" de forme exponentielle donne la fonction d'onde du système globale et, par suite, la distribution spectrale des photons émis. On retrouve la forme lorentzienne classique de la raie et l'expression de Bloch et de Dirac de la largeur.

Dans une 2ème étape, les auteurs généralisent leur méthode au cas de trois niveaux, ce qui permet de vérifier que la largeur d'une raie est la somme des largeurs des deux niveaux qu'elles connectent, ces largeurs étant

★ Victor Weisskopf obtint son Ph.D à Gottingen avec Wigner en 1931. Il eut l'idée de chercher des solutions exponentielles à la théorie de Dirac, mais "trop jeune et trop ignorant" il demanda l'aide de Wigner. "Les dieux me furent favorables en me faisant travailler avec Wigner, désormais, quel que soit le papier que j'écrive et quel que soit le coauteur, je m'en tiendrai, toujours à l'ordre alphabétique des auteurs" Ce vœu de Weisskopf est rapporté dans "ma vie de physicien" (158).

inversement proportionnelles aux durées de vie des niveaux.

On trouve dans une note une remarque importante : si l'on ne considère plus comme complètement nulle la contribution des photons tels que $h\nu$ diffère sensiblement de $E_2 - E_1$ dans la somme qui intervient dans l'équation de Schrödinger, alors les niveaux se trouvent déplacés. Le déplacement estimé est fini et négligeable devant $E_2 - E_1$ dans l'approximation des deux niveaux. Mais nous faisons remarquer que cette approximation est illégitime pour le calcul de déplacement, puisque, comme Oppenheimer l'avait prouvé quelques mois plus tôt, la contribution des autres niveaux au déplacement diverge. Une expression générale du déplacement avait déjà été donnée en juin 1927 par Dirac dans son étude de la diffusion ⁽³⁹⁾, mais pour le processus inverse (l'absorption d'un photon par un atome) et sans estimation de la valeur numérique du déplacement.

Les tests expérimentaux les plus décisifs de la théorie quantique de largeurs de raie doivent être opérés dans le cas spécifiquement quantique des raies faibles mais larges. Les raies H_α de l'hydrogène remplissent cette condition ^{*} et en plus sont calculables de manière entièrement analytique. Les mesures doivent être effectuées sur des gaz très dilués afin d'éviter l'élargissement par collision, en absorption sur des couches épaisses pour avoir un signal suffisant, et loin du maximum d'absorption pour que l'intensité observée ne soit plus due qu'à la largeur naturelle : il n'y avait pas à cette époque de procédé pour éliminer l'élargissement Doppler. On devait donc se servir du fait que le profil Doppler, bien qu'il soit en général beaucoup plus large que le profil naturel, décroît exponentiellement, donc beaucoup plus vite qu'une lorentzienne. Pour avoir un grand nombre d'atomes d'hydrogène dans les états $n=2$ afin de mesurer l'absorption $2 \rightarrow 3$, UNSOLD utilisa l'atmosphère solaire ⁽⁴⁰⁾ (1931). De nombreuses mesures furent aussi effectuées en laboratoire, citons seulement celles de SCHUTZ ⁽⁴¹⁾ (1927) effectuées sur des raies de l'Hélium, dans une configuration analogue à la précédente.

Les publications ultérieures à la première de Weisskopf et Wigner sur le même sujet ne font qu'étudier de nouveaux cas particuliers ⁽³⁸⁾ ou déduire

* La probabilité de $3 \rightarrow 2$ est environ 10 fois plus petite que $2 \rightarrow 1$



des résultats analogues par d'autres méthodes : ROSENFELD (42) et CASIMIR (43) utilisèrent à cet effet la théorie de Heisenberg (44) qui diffère de la théorie de Dirac par le passage à la représentation de Heisenberg (voir chap. §4)

1.4.3. Processus en e^4 :

- interactions entre deux particules chargées :

Elles sont décrites dans la théorie de Dirac par le hamiltonien coulombien et indirectement par l'échange de photons virtuels transverses grâce au terme $-\vec{j} \cdot \vec{A}$. Cette dernière interaction est responsable du phénomène de retard. Sous cette forme de la théorie de rayonnement (représentation de Schrödinger) le champ \vec{A} ne dépend pas du temps et donc les phénomènes de propagations n'apparaissent pas explicitement dans la théorie. Plusieurs théoriciens trouvèrent utile de s'assurer que les phénomènes ondulatoires classiques étaient tout de même contenus dans la théorie de Dirac : Citons FERMI (45) (1929) pour les franges de Lippman *, RACAH (46) (1930) pour les interférences lumineuses et Fermi (21) pour une revue de cette classe de problèmes. La méthode générale utilisée consiste à traiter simultanément l'évolution d'atomes émetteurs et d'atomes récepteurs explorant le champ. L'atome récepteur n'est perturbé qu'après que la lumière a eu le temps de se propager depuis les émetteurs. **

La section efficace de diffusion de deux particules chargées, a été étudiée par de nombreux auteurs (avec le traitement relativiste de l'électron de Dirac) : BREIT (47,45) et MØLLER (45,50) (qui a donné son nom à l'expression relativiste de la section efficace de diffusion électron-électron) ont apporté les contributions les plus célèbres. Plus précisément Breit (47) (1929) a donné à l'ordre $\frac{v^2}{c^2}$ le hamiltonien effectif (après élimination des petites composantes de Dirac) d'interaction de deux électrons : pour ce faire il prend le hamiltonien de Dirac pour les deux électrons libres et un hamiltonien d'interaction de structure identique à celui de Darwin (théorie classique avec élimination des potentiels retardés) ; et il démontre que la théorie de Heisenberg Pauli (17) conduit au même hamiltonien effectif. MOTT (57) (1930) a explicité l'effet de l'antisymétrisation des fonctions d'onde (cas de la diffusion électron-électron) sur les amplitudes de diffusion non relativistes et en première quantification. Møller (49) (1931) a calculé la section efficace de diffusion

* Ondes stationnaires produites par réflexion d'une onde plane sur un miroir.

** Les facteurs $e^{i\vec{k} \cdot \vec{r}}$ dans $\vec{A}(\vec{R})$ sont responsables du retard. Le passage à la représentation de Heisenberg (voir chap. §4) substitue $e^{i(\vec{k} \cdot \vec{r} - \omega t)}$ à $e^{i\vec{k} \cdot \vec{r}}$: le lien est ainsi plus explicite.

de deux particules chargées via la détermination du potentiel retardé (quantique) créé par les ondes de Dirac libres des particules incidentes. ROSENFELD ⁽⁵⁴⁾ (1931) a repris le calcul de Møller, avec en plus quantification du champ él.-magn. par la méthode de Heisenberg (§ 4.4). Breit a repris en 1932 son premier calcul, cette fois en utilisant la théorie de Dirac proprement dite, c'est à-dire avec un traitement non perturbatif de l'interaction coulombienne. BETHE et FERMI ⁽⁵⁵⁾ (1932) ont opéré la synthèse des différents travaux en démontrant qu'avec les approximations convenables, on pouvait opérer les déductions : théorie de Dirac → formule de Møller → hamiltonien effectif de Breit.

Diffusion du rayonnement

Le hamiltonien donné par Dirac dans sa théorie de l'émission et de l'absorption du rayonnement correspondait à l'approximation "non relativiste" où l'on retient seulement le terme $-\vec{j} \cdot \vec{A}$ pour le couplage rayonnement-matière chargée : "il semblerait alors qu'il n'y a pas de processus de diffusion directe, mais cela peut être dû au caractère incomplet de la théorie présente (13). Dirac lui-même compléta sous cet aspect la théorie deux mois plus tard dans un article intitulé "the quantum theory of dispersion" ⁽¹³⁾. Le nouveau hamiltonien utilisé par Dirac est celui donné au début de ce paragraphe, il est obtenu en prenant l'approximation au 2ème ordre en $\frac{1}{c}$ de $\sqrt{(\vec{p}-e\vec{A})^2 + m^2} - e\phi$: le fait de partir d'une expression relativiste est bien significatif de ses projets à plus long terme. Sans doute à regret, Dirac constate : "*on doit utiliser une méthode de perturbation, comme on ne peut pas résoudre l'équation de Schrödinger directement*". Il utilise encore une autre approximation qui consiste à remplacer $A(\vec{R})$ par $A(\vec{0})$ l'origine étant au centre de l'atome, si la longueur d'onde de la lumière incidente est grande devant les dimensions atomiques. De cette approximation (dipolaire), il résulte une divergence de l'intégrale de la section efficace sur la fréquence du photon diffusé calculée pour une valeur finie du temps t par la théorie des perturbations dépendant du temps. Cette divergence bénigne n'a rien à voir * avec "les divergences UV" et Dirac trouve tout de suite

* Contrairement à l'affirmation de Møyer ⁽¹⁶⁸⁾ (p. 947).

la manière de l'éliminer : dans l'expression modifiée convergente de l'intégrale où l'on a tenu compte de la dépendance $\vec{A}(\vec{R})$ seules contribuent, pour des temps longs devant la période atomique, les fréquences ν_f telle que $2\pi(\nu_f - \nu_i) \approx E_i - E_f$ (règle d'or n°2), et la section efficace dans cette limite ne présente plus de divergence si l'on prend maintenant $\vec{R} \approx \vec{O}$. En bref la limite des temps longs (diffusion) doit précéder l'approximation dipolaire (Le calcul des ordres multipolaires supérieurs justifie ce procédé).

Au même ordre d'approximation (en e^2 pour les amplitudes) il faut distinguer deux types de processus : ceux que Dirac appelle "vraie diffusion", induits par les termes en a^+ et aa^+ de \vec{A}^2 , et ceux obtenus par l'effet de \vec{A} au 2ème ordre. La somme des deux permet de retrouver la formule de dispersion de Kramers et Heisenberg ⁽⁵²⁾. Cette formule avait été le point de départ des travaux de Heisenberg sur la mécanique des matrices, se trouva démontrée par celle-ci ⁽¹²⁾ dans un traitement semi-classique (rayonnement non quantifié) et enfin par la théorie entièrement quantique de Dirac : la nouvelle mécanique quantique put ainsi justifier pleinement ses origines. Pour les courtes longueurs d'ondes, on ne peut plus approximer $A(\vec{R})$ par $A(\vec{O})$, WALLER a examiné ⁽⁶⁵⁾ l'effet des facteurs $e^{i\vec{k}\cdot\vec{r}}$ dans $A(\vec{R})$ (appelés facteurs de retard) sous la formule de dispersion. Il a aussi repris ⁽⁶⁶⁾ le calcul avec le traitement relativiste de l'électron de Dirac.

Différents cas particuliers de la théorie de la diffusion de Dirac ont été étudiés par d'autres auteurs (entre autres AMALDI ⁽⁵⁸⁾, SMEKAL ⁽¹⁷⁶⁾ et HEISENBERG avaient déjà prédit l'effet Raman, correspondant au cas où l'atome change de niveau lors de la diffusion et cet effet fut effectivement observé en 1928 par RAMAN et KRISHNAN ⁽⁵⁶⁾ dans les solutions liquides et par LANDSBERG et MANDELSTAMM dans les solides ⁽⁵⁷⁾.

Le cas de la diffusion résonnante, pour lequel le rayonnement incident a une fréquence voisine de celle d'une transition atomique, doit donner (si $\Delta\nu_{\text{incident}} \gg$ largeur naturelle) une intensité de rayonnement diffusé identique à celles que donneraient une absorption et une émission successives. WEISSKOPF ^(55,60), par une méthode analogue à celle utilisée pour l'émission spontanée (Wigner et Weisskopf ⁽³⁷⁾), a donné la forme de la raie de fluorescence dans les deux cas de l'excitation par une raie étroite ("sharp line") et par une raie large ("broad line").

Le cas de l'effet Compton (où l'électron peut être considéré comme quasi libre) sera étudié au chap 2 § 2

1.4.c Processus en e^6 :

Tel est le cas de la "Bremsstrahlung" : rayonnement d'une charge accélérée, par exemple lors de la diffusion par un potentiel coulombien. Bien sûr le phénomène existe déjà en électrodynamique classique, mais celle-ci conduit à une perte d'énergie divergente. Utilisant la théorie de Dirac et l'approximation de Born où le potentiel de Coulomb et le terme d'interaction $-\vec{j}\cdot\vec{A}$ sont chacun traités au premier ordre, on obtient après sommation sur l'énergie du photon émis une perte totale d'énergie finie. Une première estimation à haute énergie de cette perte fut donnée par HEITLER ⁽⁶¹⁾ en 1933, et la "formule de BETHE et HEITLER" ⁽⁶²⁾ valable pour toute énergie incidente comptable avec les approximations précédentes, date de 1934. La démonstration utilise la théorie de l'électron relativiste de Dirac . L'étude expérimentale de l'absorption des rayons cosmiques - nous y reviendrons - donna à la formule de Bethe et Heitler un rôle important. MOTT ⁽⁶³⁾ et SOMMERFELD ⁽⁶⁴⁾ avaient donné par d'autres méthodes des expressions valables seulement à basse énergie. Toutes ces formules présentent la même difficulté pour l'émission de photons mous : la probabilité d'émettre un photon diverge à basse fréquence. Nous reviendrons sur cette difficulté (la catastrophe infra-rouge) au chapitre 5 . L'essentiel pour les premières applications de la formule de Bethe et Heitler était que l'énergie perdue, elle, convergeât à basse fréquence du photon émis.

Dans les exemples cités, la théorie de Dirac permit soit de retrouver dans un nouveau cadre des résultats expérimentaux connus, soit de faire des prédictions directement vérifiables. Il reste celles de ses conséquences qui sortaient du domaine alors accessible à l'expérience, et celles qui vinrent ébranler la structure formelle de l'édifice théorique.

Examinant la structure du terme en \vec{A}^2 du hamiltonien, DIRAC ⁽⁵¹⁾ avait noté la présence de terme en aa et en a^+a^+ "correspondant à des processus dans lesquels deux quanta lumineux sont émis ou absorbés simultanément. Ils ne peuvent apparaître dans une théorie des quanta lumineux où il n'y a pas de forces entre les quanta. L'effet de ces termes se trouve être négligeable, si bien que le désaccord avec la théorie des quanta lumineux n'est pas sérieux ". Il est bien vrai que la probabilité de ces processus multiples ne peut être calculée par application du principe de correspondance, mais cette difficulté ne justifie pas la méfiance de Dirac à l'égard de sa propre théorie : non seulement de tels processus existent, mais encore ils constituent des phénomènes spécifiquement quantiques, dont l'application nécessite la quanti-

fication de rayonnement. EINSTEIN et EHRENFEST⁽⁶⁷⁾ avaient envisagé leur existence en 1923. Le calcul de leur probabilité (1929) dans la théorie de Dirac^{*} est dû à Maria MAYER-GÖPPERT^(68,69). Dans la région optique, la probabilité d'une émission double est très faible. L'observation de phénomènes d'absorption simultanée de deux photons nécessite un accord très précis de la somme de leur fréquence sur la transition atomique ; grâce à la technologie du laser, le phénomène est couramment observé aujourd'hui. Dans les années trente, cet aspect de la théorie de Dirac restait purement spéculatif.

En somme il était prudent de s'en tenir aux prédictions rapidement accessibles à l'expérience puisque le calcul des termes d'ordres supérieurs en α conduisait, comme on le sut dès 1930 (voir chap 3), à des résultats divergents.

* Ils résultent des termes en aa et a^+a^+ de A^2 mais aussi de ceux linéaires en \vec{A} pris au 2ème ordre.

CHAPITRE 2

VERS UNE THEORIE QUANTIQUE GENERALE DES CHAMPS D'ONDES

2.1 LA SECONDE QUANTIFICATION :

De 1924 à 1926 parmi les nombreuses découvertes préluant à la nouvelle mécanique quantique furent reconnus les deux types de statistiques auxquelles obéissent les particules quantiques : d'abord celle de BOSE (70) - EINSTEIN (71) puis celle de FERMI (72) - DIRAC (73) *. Dans la nouvelle mécanique les unes devaient être décrites par des fonctions d'onde symétriques les autres par des fonctions d'onde antisymétriques. On opérait ainsi une restriction de l'espace des solutions de l'équation de SCHRÖDINGER d'un système de particules sans en comprendre bien l'origine. A cet égard BORN, HEISENBERG et JORDAN (12) (B.H.J.) apportèrent en 1926 une première explication :

La quantification du champ électromagnétique libre suivant la nouvelle mécanique des matrices conduisait directement à des quanta lumineux obéissant à la statistique de BOSE. Certes il existait déjà, dans l'ancienne théorie des quanta, un lien entre quantification de l'énergie des modes du champ électromagnétique à la EHRENFEST - DEBYE (voir chap. 1 § 1) et statistique de BOSE - EINSTEIN des quanta lumineux : Ces deux approches conduisaient toutes deux à la loi de Planck du corps noir, car il revenait au même de dire qu'un oscillateur (mode du champ) ne pouvait prendre que des énergies $n h \nu$ (n entier), ou qu'on y attachait n quanta indiscernables d'énergie $h \nu$. Mais avant B.H.J. la structure discrète de l'énergie d'un oscillateur harmonique ne pouvait être tenue comme une explication de la statistique des quanta lumineux puisqu'elle était postulée par EHRENFEST et DEBYE. Seule la nouvelle mécanique quantique donna une preuve de cette structure discrète à partir d'un petit nombre d'hypothèses générales.

2.1.a Le cas des bosons :

Le lien entre onde quantifiée et ensemble de bosons ne se trouvait établi par B.H.J. que pour le problème particulier de l'équilibre du corps noir. DIRAC, en 1927, (13) montra qu'il valait encore pour l'interaction atome-rayonnement et aussi pour l'évolution de SCHRÖDINGER d'un ensemble quelconque de bosons (soumis à des interactions à un corps).

* Voir l'article de PEIERLS (164).

Nous avons rapporté en détail les considérations de DIRAC au chap. I § 3. L'onde quantifiée dans le second cas est l'onde Ψ de Schrödinger, d'où la dénomination de "deuxième quantification", encore en vigueur aujourd'hui et celle plus désuète de "méthode des ondes quantifiées".

Dans le problème d'électromagnétique dont l'étude était l'objet du papier de Dirac, les bosons en question (les photons) n'interagissaient pas directement entre eux et seule la 2° quantification d'un hamiltonien "à un corps" du type $H = \sum_{i=1}^N h_i$, h_i n'agissant que sur les variables de la i° particule était envisagée. En octobre 1927 JORDAN et KLEIN (74) et (76) généralisèrent la méthode dans l'étude : "du problème à plusieurs corps dans la théorie quantique".

Leur but lointain (réalisé par Heisenberg et Pauli en 1929) était de traiter sur le même pied la quantification de la matière et celle du rayonnement. La description "classique" de la matière doit alors être ondulatoire, et il faut quantifier l'onde de Schrödinger (ultérieurement de Dirac) pour décrire un ensemble de particules de matière. Pour justifier cette démarche les auteurs commencent par établir sa validité pour un ensemble de bosons soumis à des interactions à deux corps : $H = \sum_i h_i + \sum_{i,j} h_{ij}$

les exemples pris sont ceux de l'interaction coulombienne $h_{ij} = e^2/4\pi|\vec{r}_i - \vec{r}_j|$ et du terme de Darwin d'interaction magnétique* $h_{ij} = -\frac{e^2}{4\pi m^2} \vec{p}_i \cdot \vec{p}_j / |\vec{r}_i - \vec{r}_j|$. L'opérateur champ $\Psi(\vec{r})$ tel que $[\Psi(\vec{r}), \Psi^\dagger(\vec{r}')] = \delta^3(\vec{r} - \vec{r}')$ est introduit pour construire le hamiltonien F en 2° quantification. Dans le cas coulombien par exemple** $F = \int d^3r \Psi^\dagger(\vec{r}) \left(-\frac{\hbar^2}{2m} \Delta\right) \Psi(\vec{r}) + \frac{e^2}{8\pi} \int \Psi^\dagger(\vec{r}) \Psi^\dagger(\vec{r}') \frac{1}{|\vec{r} - \vec{r}'|} \Psi(\vec{r}) \Psi(\vec{r}') d^3r d^3r'$

* Voir par exemple Landau et Lifchitz tome 2 § 65 - Jordan et Klein résolvent le problème de l'ambiguïté de l'ordre des opérateurs \vec{p} et \vec{r} .

** Dans le terme coulombien l'ordre normal (tous les Ψ^\dagger à gauche des Ψ) permet d'éliminer les termes divergents d'auto-interaction en $1/|\vec{r}_i - \vec{r}_i|$. Comme le remarquaient Klein et Jordan, le seul fait de prendre soin de l'ordre des créateurs et des annihilateurs suffisait dans ce problème d'une onde de Schrödinger quantifiée en auto-interaction à éliminer toute divergence. Heisenberg crut sans doute (d'après la correspondance de Pauli (162)) un moment (quand il élaborait avec Pauli l'électrodynamique quantique) qu'une élimination analogue, par simple permutation, pouvait être opérée dans le cas de l'électrodynamique quantique.

La démonstration d'équivalence de l'évolution par F dans l'espace des nombres d'occupations et de celle par H dans l'espace des configurations symétrisé est analogue à celle de Dirac. L'expression de F est aussi trouvée comme chez Dirac par quantification du problème à une particule donné par

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \psi - \frac{e^2}{4\pi} \int \frac{\psi^*(\vec{r}') \psi(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r' \psi(\vec{r}). \quad *$$

FOCK (77) en 1928 développa un formalisme analogue et précisa l'évolution de la valeur moyenne d'un opérateur dans ce cadre. Il introduisit surtout la "représentation de Fock" : tout état quantique peut être écrit sous la forme d'une série entière de créateurs a^+ agissant sur le vide et la donnée de cette série détermine complètement l'état. L'avantage (purement formel) d'une telle représentation est de permettre une résolution explicite des équations du mouvement. La seconde partie du papier de Fock est une tentative de généralisation à d'autres statistiques. Comme Fock le reconnut lui-même, les conclusions auxquelles il aboutit sont fausses car les relations de commutation du type $[a, a^+] = 1$ ne peuvent conduire qu'à la statistique de B.E. Au moment où il écrivit, à Léninegrad, cet article Fock ne connaissait sans doute pas les travaux antérieurs de Jordan et Wigner sur la 2° quantification des fermions.

2.1.b Le cas des fermions :

La première généralisation de la 2° quantification à la statistique de Fermi-Dirac est due à JORDAN (78). Datant de Juillet 1927 elle précède les autres types de généralisation de la théorie de Dirac de la 2° quantification. Dans son article (79) "De nouveaux principes pour la mécanique quantique" Jordan avait généralisé la notion d'opérateur conjugué. Une telle généralisation se trouvait nécessaire car des opérateurs hermitiques positifs à spectre discret, sans point d'accumulation comme l'opérateur N nombre de particules, ne peuvent avoir de conjugué θ au sens ordinaire $[N, \theta] = -i$.

* Schrödinger (75) s'était inquiété quelques mois plus tôt des conséquences d'une théorie de l'électrodynamique obtenue en prenant l'équation de Schrödinger et les équations de Maxwell couplées et en tenant compte du potentiel électromagnétique créé par la distribution de charge et de courant associée à la fonction d'onde $\psi(\vec{r})$. Bien sûr il aboutissait à des conséquences absurdes, ce genre d'équation ne prenant un sens, comme le montrent Jordan et Klein, qu'en 2° quantification.

Dirac avait déjà utilisé les opérateurs θ par exemple dans (13), mais, ainsi qu'il l'expliqua oralement à Jordan, en utilisant seulement la relation $[N, e^{i\theta}] = e^{i\theta}$; $e^{i\theta} = a^+$. Dans le cas de la statistique de B.E., N prend toutes les valeurs entières positives et les annihilateurs et créateurs introduits par Dirac avec $N_\alpha = a_\alpha^\dagger a_\alpha$ vérifient bien $[N_\alpha, a_\alpha^\dagger] = a_\alpha^\dagger$. Si maintenant N ne peut prendre que les valeurs 0 et 1 - c'est le cas de la statistique de F.D. - il peut être représenté par la matrice $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et a^+ par la matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Le hamiltonien $F = \sum_{\alpha\beta} a_\beta^\dagger H_{\alpha\beta} a_\beta$ conserve le nombre total de fermions, et donne l'évolution de Schrödinger d'un ensemble de fermions. Cependant les démonstrations et définitions de Jordan sont incomplètes car elles ne tiennent pas compte du signe introduit par l'action d'un créateur sur un état. Mais Jordan est maintenant convaincu de la possibilité de décrire des fermions identiques par une onde quantifiée dans l'espace à 3 dimensions, dans un traitement analogue au cas du rayonnement. *"Le fait fondamental de la théorie des électrons, l'existence de particules électriques discrètes, apparaît comme un phénomène typiquement quantique, c'est-à-dire que les ondes de matière n'interviennent que dans des états quantiques discrets"*. Une première application justifiant le titre de l'article : *"La mécanique quantique des gaz dégénérés"*, est un calcul unifié des fluctuations d'un gaz parfait de Bose ou de Fermi.

L'introduction correcte des signes intervenant lors de l'action des a et des a^+ et l'écriture explicite des relations d'anticommutation $a_\alpha a_\beta + a_\beta a_\alpha = 0$ $a_\alpha a_\beta^\dagger + a_\beta^\dagger a_\alpha = \delta_{\alpha\beta}$ sont dues à WIGNER et JORDAN ⁽⁸⁰⁾. A la fin de l'article fut rajoutée la preuve de l'unicité (à une transformation unitaire près) de la représentation matricielle irréductible de ces relations. Il fut aussi établi que la méthode de la 2^o quantification des fermions s'appliquait aux interactions à deux corps.

2.1.c Espace des configurations ou espace des nombres d'occupation?

Nous avons déjà rapporté le point de vue de JORDAN, un convaincu des avantages de la 2^o quantification. L'article de DIRAC ⁽¹³⁾ prouvait déjà suffisamment la nécessité de quantifier les ondes électromagnétiques. Pour d'autres particules que les photons la 2^o quantification permet aussi de justifier la structure quantique de la matière. Pour un système dans lequel le nombre

total de particules est conservé, la description habituelle dans l'espace des configurations symétrisé ou antisymétrisé est toujours possible et les avantages de la méthode alternative sont alors surtout formels : les propriétés statistiques des gaz de Fermi et de Bose peuvent être traitées sur le même pied, il suffit de choisir dans les formules finales si les annihilateurs - créateurs commutent ou anticommulent; bien des calculs se trouvent simplifiés en raison de l'harmonie entre l'utilisation de l'ensemble grand canonique et celle d'un nombre de particules opératoirel.

Pourtant le poids de ces arguments n'empêcha pas certains théoriciens de préférer, occasionnellement, le retour à l'espace des configurations et même pour les photons ! Il est toujours possible de réexprimer l'action d'un opérateur 2° - quantifié dans l'espace des configurations. FOCK ⁽⁸¹⁾ donna une démonstration claire et générale de cette possibilité formelle lors d'un séminaire à l'université de Léninegrad en Janvier 1931. La première partie de l'article contient la démonstration moderne (telle qu'on la trouve par exemple dans le cours de Landau et Lifchitz ^{*}) de l'équivalence des deux méthodes lorsque le nombre de particules est conservé. Puis l'action d'opérateurs changeant le nombre de particules comme $\Psi(\vec{r})$ est exprimée dans l'espace des configurations décrit par une suite de fonctions d'onde d'un nombre variable de particules ^{**}

$$\Psi(\vec{r}) \begin{pmatrix} \text{const.} \\ \psi(\vec{r}_1) \\ \psi(\vec{r}_1, \vec{r}_2) \\ \psi(\vec{r}_1, \vec{r}_2, \vec{r}_3) \\ \vdots \end{pmatrix} = \begin{pmatrix} \psi(\vec{r}) \\ \sqrt{2} \psi(\vec{r}, \vec{r}_1) \\ \sqrt{3} \psi(\vec{r}, \vec{r}_1, \vec{r}_2) \\ \vdots \end{pmatrix}$$

* tome 3 §§ 64 et 65

** Une telle suite fut introduite pour la première fois par Landau et Peierls dans l'article ⁽⁸²⁾ cité un peu plus loin.

Même dans une théorie où le nombre et l'identité des particules ne sont pas conservés (ce que permet la théorie de la relativité) on peut exprimer l'évolution du système dans l'espace des configurations à condition de considérer simultanément des fonctions d'ondes correspondant à des nombres différents de particules qui se trouvent couplées entre elles par l'équation d'évolution ; $|\psi|^2(\vec{r}_1, \vec{r}_2, \vec{r}_3)$ par exemple représente la densité de probabilité d'avoir trois particules occupant les positions $\vec{r}_1, \vec{r}_2, \vec{r}_3$.

Le choix du meilleur type de description se pose différemment dans le cas des photons et dans le cas des électrons. Pour les électrons et avant la théorie du positron tout processus électrodynamique conservait le nombre d'électrons. Conformément au point de vue de Jordan, HEISENBERG et PAULI⁽¹⁷⁾ (dans "la dynamique quantique des champs d'ondes") utilisèrent la 2^e quantification des électrons pour plus d'homogénéité formelle - mais dans la plupart des applications le nombre de particules chargées est faible et il est plus simple de revenir à l'espace de configuration, c'est ce que firent HEISENBERG et PAULI (et OPPENHEIMER) dans la deuxième partie⁽¹⁸⁾ de cette étude, et c'est le point de vue adopté dans presque toutes les premières applications de l'électrodynamique quantique (Dirac, Fermi, Waller, Oppenheimer...). Après la théorie du positron (DIRAC 1930) la situation fut inversée. Dans la présentation initiale de Dirac il y avait une infinité d'électrons occupant les états d'énergie négative : on avait tout intérêt à abandonner l'espace des configurations, comme le firent plusieurs auteurs travaillant sur la polarisation du vide, ou sur l'énergie propre de l'électron (voir les chapitres correspondants), même si le nombre (infini) d'électrons ne variait au cours d'aucun processus : une création de paire était interprétée comme un saut d'électron d'un état d'énergie négative dans un état d'énergie positive. Le point de vue moderne, qui consiste à développer le champ de Dirac sur des annihilateurs d'électrons et sur des créateurs de positrons, tous d'énergie positive, fut introduit par FOCK en nov 1933 et par HEISENBERG⁽⁸³⁾ (voir chap. 5 § 1) en Juin 1934 dans son étude de la polarisation du vide^{*} ; cette fois le processus de création de paires fait

* Un traitement symétrique des électrons positifs et négatifs, par introduction de nombres d'occupations d'états d'électrons positifs fut donné par H. WEYL dès 1930⁽¹²⁰⁾. (Voir chap. 2 § 1)

varier le nombre et la nature des fermions présents, la 2^o quantification s'impose. La théorie de Fermi de 1933 de l'émission β nécessita, dans sa formulation même, la 2^o quantification (Voir chap. 6 § 1), car le neutrino et l'électron étaient créés pendant le processus. Les phénomènes de création et d'absorption de particules observés en physique nucléaire furent sans doute l'argument le plus massif en faveur de la quantification de Wigner - Jordan. Il est étonnant que Dirac ait longtemps (jusqu'après la guerre) résisté à son usage en électrodynamique quantique. La théorie de PAULI et WEISSKOPF de 1934 (papier "*anti-Dirac*" ; voir chap. 5 § 3) du champ scalaire chargé montrait pourtant bien l'inutilité de la mer de Dirac ^{*}.

Pour les photons, l'introduction de l'espace des configurations ne peut se faire qu'a posteriori, contrairement au cas des électrons pour lesquels on sait définir une fonction d'onde avant même de parler de 2^o quantification. En effet l'onde que quantifie Dirac ⁽¹³⁾ dans la théorie électromagnétique quantique, n'est pas la fonction d'onde du photon, mais le champ électromagnétique classique. DIRAC ⁽¹³⁾ souligne d'emblée (P. 247) que ce dernier ne saurait être interprété comme une fonction d'onde puisque la densité de photons associée à une onde monochromatique n'est pas son module au carré, mais sa densité d'énergie divisée par $h\nu$. En fait on ne peut même pas définir de fonction d'onde pour un photon, suivant l'argument de LANDAU et PEIERLS ⁽⁸⁴⁾ (voir chap. 4 § 2). Avant de prendre conscience de cette difficulté les mêmes auteurs proposèrent une électrodynamique quantique dans l'espace des configurations ⁽⁸⁵⁾ (Février 1930). pour ce faire ils introduisirent une suite de "fonctions d'onde" correspondant à un électron et N photons, N variant. L'identification de la fonction d'onde libre pour $N = 1$ au produit du champ électrique par la fonction d'onde de Dirac permet de deviner l'expression de la densité de probabilité — qui a le mauvais goût de ne pas être positive — en fonction des fonctions d'onde et d'exprimer la condition de normalisation. Cette condition et quelques "Ansätze" simples sur la forme du couplage électron - champ, permettent d'établir un système d'équations de Schrödinger couplées auxquelles doivent satisfaire ces fonctions d'onde. Ce système est équivalent aux équations de Heisenberg - Pauli II ⁽¹⁸⁾ obtenues par la méthode des "ondes quantifiées". Par rapport à cette méthode de quantification canonique des champs, le nouveau formalisme de Landau et Peierls

* Une mer saturée de particules d'énergie négative n'est envisageable qu'en statistique de Fermi.

était à bien des égards un régression : le point de vue de Heisenberg et Pauli était plus unitaire et plus systématique, puisqu'il résultait d'une application aveugle de la quantification canonique au système dynamique "classique" constitué par le champ électromagnétique et le ψ de Dirac. La difficulté de l'infinité des énergies propres subsistait en raison de l'équivalence des deux formalismes. La démonstration de l'invariance relativiste était encore bien plus dure ^{*}. Et surtout "c'était une idée stupide", dit Peierls ^{*}, "car l'idée de coordonnées du photon est dépourvu de sens. A l'origine de la mécanique quantique ressortait l'analogie entre matière et lumière ; une analyse plus détaillée fit apparaître les limites mêmes de ce point de vue." ^{**}

Introduite par Dirac dans un problème particulier, qu'est celui de l'émission et de l'absorption de rayonnement par un atome, la seconde quantification a une portée bien plus vaste.

— elle permet par une substitution de langage de ne plus parler de la présence de telle particule à tel endroit, mais de nombres d'occupation d'états quantiques, éliminant ainsi les reliquats du langage classique des particules discernables, et elle justifie cet abandon de la discernabilité.

— elle constitue une méthode puissante de calcul en mécanique statistique

— son utilisation en physique des particules s'impose en raison du changement d'identité et de nombre des particules au cours d'une réaction inélastique.

* d'après une conversation avec M. Peierls.

** Le formalisme de Landau et Peierls présente toutefois l'avantage de ne pas contenir d'énergie de zéro infinie. Cette propriété n'est pas explicitée par les auteurs, mais elle fut remarquée par ROSENFELD et SOLOMON (voir chap. 2 § 6).

2.2 LA THEORIE RELATIVISTE DE L'ELECTRON :

La recherche d'une équation d'onde relativiste pour l'électron n'est pas a priori un problème de théorie des champs. Pourtant son étude nous semble nécessaire dans ce cadre, car elle fut une pierre indispensable à la construction de l'édifice de l'électrodynamique quantique, et elle permit ou même exigea le passage de l'onde au champ électronique par 2^o quantification.

La nécessité d'une théorie quantique relativiste de l'électron se fit sentir avec la découverte du spin. Certes l'hypothèse d'une rotation propre de l'électron est assez vieille : on la trouve déjà chez LORENTZ en 1909 ⁽⁸⁶⁾ puis chez COMPTON ⁽⁸⁷⁾ en 1921 pour expliquer les propriétés magnétiques de la matière. Son introduction fut rendue nécessaire par la structure en doublets des spectres atomiques : on observe en général deux fois plus de termes spectraux que ne prévoit le modèle de Bohr sans spin. UHLENBECK et GOUDSMIT ⁽⁸⁸⁾ montrèrent en 1925 que l'hypothèse du spin permettait d'expliquer ce phénomène. L'introduction des spineurs à deux composantes, et ainsi l'élimination de l'image classique de l'électron tournant au profit d'une représentation abstraite des nouveaux degrés de liberté, est due à PAULI et à DARWIN. En mai 1927 Pauli ⁽⁸⁹⁾, partant de l'hypothèse qu'un électron doit être décrit par un couple de fonctions d'onde, introduisit une représentation bidimensionnelle des relations de commutation de l'opérateur de moment cinétique de spin ($[S_x, S_y] = i S_z$) avec les fameuses matrices de Pauli et retrouva ainsi le spectre de l'atome d'hydrogène et l'effet Zeeman anormal. Les équations de Pauli, comme il le signala lui-même, avaient déjà été obtenues par Darwin ⁽⁹¹⁾ * en utilisant une méthode inverse : Darwin déduisit du spectre expérimental de l'hydrogène le système d'équations que devait satisfaire un couple de fonctions d'onde.

* Sir Charles Galton DARWIN, petit-fils de l'inventeur de la théorie de l'évolution est surtout connu par ses travaux en mécanique statistique (théorie de Darwin - Fowler) et par son application de l'équation de Dirac à l'atome d'hydrogène (terme de Darwin...).

Il construisit par combinaisons linéaire des deux fonctions, quatre autres fonctions devant se transformer comme un scalaire et un vecteur, expliquant ainsi le spin par la "nature vectorielle" de la fonction d'onde. Dans un article ⁽⁹⁰⁾ plus détaillé, Darwin réexposa sa théorie. Elle ne différait pas de celle de Pauli par le contenu empirique mais plutôt par le refus de Darwin d'utiliser des opérateurs (comme \vec{S}), et accessoirement par l'étrange idée d'une nature vectorielle de la fonction d'onde. L'application par Darwin de son formalisme à l'effet Zeeman (reproduisant entre autres la théorie de Heisenberg - Jordan de l'effet Zeeman anormal de mars 1926) suscita une polémique assez vive avec Pauli. Prenant le parti de Pauli, Heisenberg écrivit à Pauli [163] ⁽¹⁶²⁾ : *"Je suis toujours irrité quand, même aujourd'hui, les gens essaient d'écrire une théorie ondulatoire conjuguée pour chaque théorie matricielle ; je pense que les gens feraient mieux d'apprendre les deux."* L'équation de Dirac dépassa cette querelle d'interprétations. La théorie de Darwin et Pauli permet d'éliminer la structure classique inobservable de l'électron tournant et les hypothèses ad hoc pour réduire à deux le nombre infini des degrés de liberté de rotation propre ; l'origine même du spin restait inexplicée, et la valeur deux du facteur de Landé de l'électron restait mystérieuse.

L'équation de Klein - Gordon $[(i\partial_\mu - eA_\mu)^2 - m^2]\psi = 0$ est la première équation invariante relativiste qui fut envisagée pour décrire l'électron. C'est SCHRÖDINGER * qui le premier tenta d'utiliser cette équation pour calculer le spectre de l'atome d'hydrogène. Découragé par le mauvais résultat, il abandonna quelques mois cette équation, puis il eut l'idée de recalculer le même spectre en prenant la limite non relativiste de cette équation c'est-à-dire l'équation de Schrödinger. Cette dernière seulement fut alors publiée puisque son accord avec l'expérience était excellent. KLEIN ⁽⁹²⁾ et GORDON ⁽⁹³⁾ ** redécouvrirent indépendamment en 1926 l'équation qui porte leurs noms, et publièrent leur résultat sans trop se soucier des difficultés d'application à l'atome d'hydrogène.

* d'après Dirac, discours de réception du prix de l' "Oppenheimer Memorial" ⁽⁹⁴⁾.

** Gordon voulait plus particulièrement calculer l'effet Compton par la méthode de Schrödinger. Klein, indépendamment, cherchait une généralisation relativiste de l'équation de Schrödinger et de l'expression du courant et de la charge. (Klein a été informé du travail de Gordon seulement après l'envoi de son article à la Zeits. für Phys. (sept. 1926)). Dans son quatrième article sur "la quantification comme problème de valeurs propres" Schrödinger ⁽¹¹⁷⁾ (§ 6) avait déjà écrit l'équation de Klein - Gordon avec "une extrême réserve" et se refusait à en développer les conséquences, mentionnant seulement l'échec pour le spectre de l'atome d'hydrogène.

2.2.a L'équation de Dirac :

"Je me souviens que lors de mon séjour à Copenhague, Bohr me demanda sur quoi je travaillais, et je lui répondis que j'essayais de trouver une théorie relativiste satisfaisante de l'électron. Et Bohr dit : " Mais Klein et Gordon ont déjà fait cela." Ma première réaction fut de me troubler ; Bohr semblait pleinement satisfait par la solution de Klein, mais je ne l'étais pas à cause des probabilités négatives auxquelles elle conduisait. Je continuai mon travail, à la recherche d'une théorie qui ne présentât que des probabilités positives". (Conversation de Dirac avec Mehra, à Miami, le 28 mars 1969). (95)

Darwin et Pauli avaient cherché "l'idée physique juste" : le spineur à deux composantes. Dirac, lui "joua avec les équations" plutôt que d'essayer d'introduire l'idée physique juste. "Une grande partie de mon travail consiste seulement à jouer avec les équations et à chercher ce qu'elles donnent. La deuxième quantification sortit d'un tel jeu... j'aime bien jouer avec les équations, juste pour chercher des relations mathématiques qui peut-être n'ont pas du tout de signification physique". (Réf. précédente).

La "théorie quantique de l'électron" (96), communiquée le 2 Janvier 1928 aux "Proc. Roy. Soc.", est un des produits les plus significatifs de ce jeu. Le "modèle" de Pauli et Darwin est d'abord critiqué : "il reste à savoir pourquoi la Nature aurait choisi ce modèle particulier de l'électron au lieu de se satisfaire de l'électron ponctuel. On aimerait trouver que les méthodes antérieures de la mécanique quantique de l'électron ponctuel sont incomplètes et qu'en les complétant le phénomène de dédoublement suit sans hypothèse arbitraire." La théorie relativiste de Klein - Gordon, continue Dirac, est incomplète : du fait qu'elle contient l'opérateur $\nabla_{\partial t}$ on ne peut lui appliquer la théorie habituelle des transformations et calculer la probabilité des différentes valeurs d'une observable quelconque.* De plus les solutions d'énergie négative déjà présentes

* Corrélativement il n'existe pas dans ce cas de courant conservé dont la première composante soit positive. Il est clair aujourd'hui qu'une telle exigence ne peut être satisfaite dans la théorie scalaire que si l'on opère la quantification de ce champ et partant renonce à l'a priori implicite de Dirac : les électrons doivent être décrits par une théorie ondulatoire à une particule. Comme cet a priori s'est révélé intenable, tant pour des raisons fondamentales de limite de mesurabilité des variables cinématiques de l'électron (voir chap. 4 § 2) que pour des raisons formelles - même l'équation de Dirac doit être soumise à une deuxième quantification pour éliminer la difficulté des états d'énergie négative — on peut dire que la découverte de l'équation de Dirac reposa sur une argumentation en partie fautive.

dans la théorie classique, deviennent très gênantes dans la théorie quantique puisque les transitions (par effet Tunnel) d'un état d'énergie positive dans un état d'énergie négative deviennent possibles.

Dirac se propose de résoudre la première difficulté ; la seconde subsiste dans sa théorie. L'équation recherchée doit être du premier ordre en $\frac{\partial}{\partial t}$ et comme en relativité les différentes coordonnées doivent jouer des rôles symétriques, elle doit être linéaire en $\frac{\partial}{\partial x^\mu}$: $(-p_0 + \vec{\alpha} \cdot \vec{p} + \beta) \psi = 0$ avec $p_\mu = i \partial_\mu$. Multipliant à gauche par $(p_0 + \vec{\alpha} \cdot \vec{p} + \beta)$ on retrouvera l'équation de Klein - Gordon à condition que les constantes α_i et β vérifient $\alpha_i \alpha_j + \alpha_j \alpha_i = 2\delta_{ij}$, $\alpha_i^2 = 1, \beta^2 = 1$ et $\alpha_i \beta + \beta \alpha_i = 0$ avec $\beta = m\beta$. Ces relations sont réalisables par des matrices 4 x 4 construites à partir des matrices de Pauli.

- L'équation ainsi obtenue, qui s'écrit encore $(p_\mu \gamma^\mu - m) \psi = 0$ avec $\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}$ est bien invariante relativiste : une transformation de Lorentz sur les coordonnées et une transformation unitaire simultanée sur ψ (qui a maintenant quatre composantes) conservent la forme de l'équation.

- En présence d'un champ électromagnétique on utilise la règle du couplage minimal : $p_\mu \rightarrow p_\mu - e A_\mu$. L'équation quadratique dérivée de l'équation de Dirac contient alors un terme en $2 \frac{e}{mc} \vec{S} \cdot \vec{B}$ correspondant à un moment magnétique intrinsèque de l'électron $2 \frac{e}{mc} \vec{S}$. La valeur 2 du facteur de Landé est ainsi justifiée.

- La correction relativiste à l'ordre le plus bas apportée au mouvement d'un électron dans un champ central redonne le hamiltonien de Darwin - Pauli, et donc la valeur correcte du couplage spin - orbite.

Pour que la démonstration d'invariance relativiste soit complète il reste à prouver que $\psi^\dagger \psi$ peut bien être interprété comme une densité de probabilité c'est-à-dire qu'il existe un quadricourant conservé j^μ tel que $j^0 = \psi^\dagger \psi$. C'est ce que montre Dirac dans son article suivant (97) ; on a $\partial_\mu j^\mu = 0$ avec $j^\mu = \bar{\psi} \gamma^\mu \psi$ ($\bar{\psi} = \psi^\dagger \gamma^0$). Le même article contient le calcul des effets Zeeman et Paschenback (en conformité avec la théorie de Pauli) et la dérivation des règles de sélection sur les transitions optiques.

L'équation de Dirac suscita un intérêt immédiat dans les cercles de théoriciens. En témoigne ce commentaire de Darwin : "Dans un papier récent

Dirac a brillamment éliminé les défauts préexistants de la mécanique de l'électron, et a montré comment le phénomène habituellement appelé l'électron tournant prend place dans la théorie complète." , ainsi que l'inclusion immédiate de l'équation de Dirac dans tous les problèmes d'électrodynamique relativiste comme nous le verrons au cours de l'étude des applications.

La connaissance des matrices de Pauli aida sans doute Dirac dans la découverte de son équation. L'introduction de l'algèbre anticommutative des γ^{μ} n'en est pas moins originale, et celle de bispineurs dont deux des composantes pouvaient sembler superflues avant la découverte du positron était très audacieuse. Dirac fut plus timide dans l'exploitation immédiate des conséquences de son équation : il calcula les corrections relativistes du spectre de l'hydrogène seulement au premier ordre et publia "hâtivement" ce résultat avant d'aller plus loin parce qu'"il avait peur". "on craint beaucoup le besoin d'un succès consolidé dans de telles circonstances, et c'est en grande hâte que je fis publier cette première approximation avant que n'arrive ce qui ferait tout s'écrouler." (94)

Le calcul exact du spectre de l'atome d'hydrogène (il est remarquable que l'atome d'hydrogène de Dirac admette une solution analytique rigoureuse) fut accompli dans le mois suivant par GORDON (100) à Hambourg. Les niveaux d'énergie obtenus sont ceux de l'atome relativiste de Sommerfeld (101) , obtenus simplement en imposant les conditions de quantification de Bohr au mouvement de Kepler relativiste. L'interprétation des nombres quantiques est bien sûr différente mais la coïncidence numérique est presque miraculeuse, si l'on se souvient que d'autres théories quantiques relativistes, plus simples que celle de Dirac, comme la théorie de Schrödinger - Klein - Gordon donnent des résultats différents pour le spectre de l'atome d'hydrogène.

A Cambridge, au mois de mars, DARWIN (98) publia la solution du même problème, ainsi qu'une étude des règles de sélection et de l'effet Zeeman. Au lieu d'utiliser comme Dirac un traitement algébrique de l'équation de Dirac, Darwin préfère la considérer comme un système de quatre équations différentielles couplées, se ramenant ainsi à des notions plus familières : "il y a probablement des lecteurs qui partageront mon opinion sur les méthodes de l'algèbre non commutative, plus difficile à suivre, et certainement plus difficile à inventer que des opérations familières de l'analyse" (98). L'introduction de l'algèbre en

mécanique quantique est devenue aujourd'hui si commune, qu'on a peine à se représenter sa nouveauté et sa singularité pour les théoriciens des années trente. Dans son article suivant Darwin déterminait le courant et le rayonnement émis par un électron de Dirac en mouvement, composé d'une part du rayonnement dû au mouvement de convection, d'autre part de celui associé au moment magnétique intrinsèque de l'électron.

Plusieurs exemples d'applications conjuguées de la théorie du rayonnement de Dirac et de l'équation de Dirac ont été donnés au chapitre précédent : théories de BREIT et MØLLER pour la diffusion électron - électron 1931-32, diffusion de MOTT (166) par un potentiel coulombien, théorie de BETHE et HEITLER de la "Bremsstrahlung" en 1934... Le problème de la diffusion du rayonnement par des électrons liés ne nécessitait pas de traitement relativiste de l'électron (en première approximation tout au moins. Waller (66) a calculé les corrections relativistes). Mais pour l'effet Compton dans les domaines X et γ (pour lesquels l'énergie de liaison des électrons périphériques peut être négligée) les corrections relativistes deviennent importantes. Le problème était particulièrement important pour expliquer les longueurs d'absorption des rayons γ dans la matière et son investigation par KLEIN et NISHIMA (102, 103) fut publiée dès le mois de décembre 1928. Comme la diffusion Compton était le problème le plus important dans lequel la mécanique quantique pouvait donner des prédictions différant de celles de l'ancienne théorie quantique nous allons retracer brièvement l'histoire de ses théories. L'étude expérimentale de l'effet Compton par des rayons γ et X de 0.01 Å à 1 Å de longueur d'onde avait été effectuée par COMPTON en 1921 et 1922. En 1923 le même physicien publia une théorie en excellent accord avec ces résultats (105). Il remarquait que la relation $\lambda_d - \lambda_0 = \lambda_c (1 - \cos \theta)$ entre la longueur du photon diffusé et l'angle de diffusion ($\lambda_c = h/mc$: longueur d'onde Compton) peut être obtenue en supposant que le rayonnement est diffusé classiquement par un électron de vitesse effective $\beta = \frac{\lambda_0}{\lambda_0 + \lambda_c}$; sans justification il supposait que la distribution angulaire de l'intensité de rayonnement pouvait se calculer par la même hypothèse. Le calcul de BREIT (106) d'avril 1926 nécessitait lui aussi une hypothèse de *correspondance* avec la théorie classique assez arbitraire. Presqu'en même temps DIRAC (107) proposa le premier calcul dans la nouvelle mécanique quantique en utilisant la représentation de Heisenberg et la relation relativiste impulsion - énergie $E^2 = p^2 + m^2$. La même formule pour l'effet Compton fut obtenue par GORDON (93) en représentation de Schrödinger, c'est-à-dire en utilisant l'équation de Klein - Gordon. L'accord avec l'expérience

dans le domaine d'énergie étudié par Compton était très bon. A ce stade la nécessité de refaire le calcul en utilisant l'équation de Dirac était plus théorique qu'expérimentale. KLEIN et NISHIMA (102) l'effectuèrent en traitant classiquement le rayonnement ^{*}. La formule de Klein et Nishima diffère de la diffusion Thomson classique par des termes en $\frac{h\nu}{m c^2}$ et de la formule relativiste de Dirac (107) et de Gordon (93) par des termes en $\left(\frac{h\nu}{m c^2}\right)^2$. La comparaison avec l'expérience (108, 165) ne permit pas d'exclure immédiatement ces autres formules relativistes. Les déviations entre ces deux types de formules apparaissent nettement pour la section efficace totale de diffusion de γ durs présents dans les rayons cosmiques ($1 \lesssim h\nu/m c^2 \lesssim 10$; au delà on ne peut plus négliger la création de paires). Heitler (33) (p. 159) donna la référence d'expériences exécutées dans ce domaine en 1930 et en 1934 et conclua : "Comme d'autres équations d'onde relativistes dévient notamment dans la région considérée de la formule de Klein - Nishima, nous pouvons considérer les résultats expérimentaux comme une évidence nette en faveur de l'équation de Dirac et contre, par exemple, l'équation de Klein - Gordon." Nous verrons au chapitre 6 § 2 comment la confiance dans les calculs de diffusion et de freinage effectués dans la théorie de l'électron de Dirac oscilla avec les résultats des expériences sur les rayons cosmiques.

L'équation de Dirac, par delà ses succès expérimentaux, présente à première vue un certain nombre de propriétés étranges : l'introduction de spineurs à 4 composantes, l'utilisation d'une algèbre anticommutative, mais aussi le phénomène de "Zitterbewegung" (= mouvement de tremblement) mis en évidence par SCHRÖDINGER lui-même en 1930 (108). En représentation de Heisenberg le hamiltonien de Dirac donne pour la vitesse de l'électron $\dot{\vec{r}} = [\vec{r}, H] = c \vec{\alpha}$ d'où $|\dot{\vec{r}}| = c$. ^{**} Une mesure instantanée de la vitesse devrait donc toujours donner la valeur c .

* Tamm (111) a retrouvé en 1930 la formule de Klein - Nishima en utilisant le rayonnement quantifié. La quantification du rayonnement n'introduirait de différences qu'aux ordres supérieurs.

** Cette propriété avait déjà intrigué Breit (112) et Fock (113) qui donnèrent différentes interprétations de la vitesse dans la théorie de Dirac

Cette bizarrerie s'explique en résolvant explicitement le mouvement libre :

$\vec{x} = -\frac{1}{4} c \hbar^2 \vec{\alpha}_0 \cdot \vec{e}^{-2iHt/\hbar} H^{-2} + c^2 \vec{p} H^{-1} t + \vec{x}_0$. Le premier terme (Zitterbewegung) correspond à des oscillations de fréquence très élevée, (au moins $\frac{2mc^2}{\hbar}$) et le second à la translation de vitesse $\frac{\vec{p}}{E} c^2$; le mouvement moyen, le seul observable d'après les considérations développées par Landau et Peierls (voir chap.4 § 2), est ainsi conforme à la dynamique relativiste classique; la "Zitterbewegung" permet d'assurer que la vitesse instantanée reste égale en module à c.

Sur cet exemple de la "Zitterbewegung" apparaissent les difficultés d'interprétation de l'équation de Dirac. Une application naïve des postulats de la mesure quantique (valeur propre d'une observable \rightarrow valeur mesurée) conduit à des absurdités. BOHR lui-même se méprit en affirmant que le moment magnétique et même le spin d'un électron libre n'avaient pas de sens. Certes l'équation de Dirac prédisait correctement les facteurs de Landé atomiques, mais selon Bohr le spin de l'électron pouvait perdre son sens pour les électrons nucléaires (voir chap.4 § 5) ou pour l'électron libre. MOTT ⁽¹⁶⁶⁾ publia en 1929 deux arguments de Bohr utilisant le principe d'incertitude pour démontrer qu'une mesure directe (par un magnétomètre) du moment magnétique de l'électron n'était possible et qu'on ne pouvait en outre polariser à l'aide d'un Stern et Gerlach un faisceau d'électron. Mott répliqua que l'on pouvait polariser le faisceau par diffusion sur une cible nucléaire. En 1953, Louisell, Pidd et Crane purent utiliser la polarisation de Mott pour donner la première mesure directe du facteur de Landé de l'électron. (voir (167)) La première mise en évidence qualitative de la polarisation d'un faisceau d'électrons est due à DYMOND ⁽¹⁶⁹⁾ (1934). * Non seulement le moment magnétique de l'électron s'avérait observable, mais encore la vérification ultérieure de la formule de Mott apporta une nouvelle confirmation de la théorie de l'électron de Dirac.

* Les expériences de Rupp ⁽¹⁷⁰⁾ de 1932, donnant un résultat positif, étaient probablement fausses. Louis de Broglie a écrit en marge de l'exemplaire de l'E.N.S. du rapport de Davisson au congrès ⁽¹⁷¹⁾ international d'électricité de Paris de 1932 : "Mais Rupp a avoué avoir faussé les expériences et peu après il est devenu fou. Aussi un doute subsiste-t-il à ce sujet et on n'a pu refaire ses expériences, ce qui confirme leur fausseté".

2.2.b La théorie du positron :

L'équation de Dirac n'élimine pas la difficulté des états d'énergie négative et Dirac en fut d'emblée conscient. Elle ne pouvait être qu'une approximation d'une théorie plus cohérente. "La probabilité pour des transitions avec changement de signe de l'énergie est extrêmement petite (en $(\frac{v}{c})^4$). En conséquence la théorie présente est une approximation. Il semble que cette difficulté ne peut être levée que par un changement fondamental de nos représentations (Vorstellungen) et peut-être est-elle liée à la différence entre le passé et le futur." * Ainsi s'exprimait Dirac lors d'un séminaire aux "semaines universitaires de Leipzig" ⁽¹¹⁰⁾ (18-23 Juin 1928). La résolution de cette difficulté, et ses prolongements ultérieurs dans l'étude de la polarisation du vide peuvent être considérés comme l'exploration des ultimes conséquences de l'équation de Dirac "plus intelligente que son auteur" (Dirac) qui sut néanmoins exploiter lui-même cette "intelligence".

Les transitions vers les états d'énergie négative dont Dirac estime la probabilité (en $(\frac{v}{c})^4$) correspondent dans la théorie future à l'annihilation $e^+ e^-$. Une autre difficulté de l'équation de Dirac correspond au processus inverse : un champ électrique suffisamment fort peut créer des paires $e^+ e^-$ et on s'attend à ce que l'équation de Dirac conduise à des absurdités dans ce cas. C'est ce que montra KLEIN en 1928 en étudiant la réflexion d'électrons sur une marche de potentiel. ⁽¹⁷⁴⁾. Pour de grandes valeurs du saut de potentiel, malheureusement atteintes à l'intérieur du noyau, les électrons passent la marche et se retrouvent avec une énergie cinétique négative. ** Klein conclut "la difficulté de la mécanique quantique relativiste soulignée par Dirac peut intervenir même dans des problèmes purement mécaniques, où il n'y a pas de processus de rayonnement".

* On ne peut s'empêcher de penser au point de vue que Feynman développa bien plus tard : "The theory of positrons" (Phys. Rev. Vol. 76, 749, (1949)).

** Le calcul fut repris en 1931 par SAUTER (après la théorie des trous ; mais Sauter n'interpréta pas ses résultats dans ce nouveau cadre) avec une barrière de la forme $\sqrt{\Delta x}$ au lieu de $\sqrt{\quad}$ pour montrer conformément à une hypothèse de Bohr que si $\Delta x \gg h/mc$ la probabilité de passer dans les états d'énergie négative pour une valeur élevée de l'accroissement de potentiel est bien plus faible que celle calculée par Klein.

La théorie des trous

Pendant presque deux ans Dirac ne publia rien sur la question. D'une part il était occupé par la rédaction de son cours de mécanique quantique, et par l'étude de la mécanique statistique quantique, mais aussi l'issue n'était pas facile : "Je sentais qu'écrire un papier sur l'électron n'était pas aussi difficile que d'en écrire un sur l'interprétation physique." Le problème des états d'énergie négative "était une imperfection de la théorie et je ne voyais pas ce qu'on pouvait y faire. Ce ne fut que plus tard que j'eus l'idée de remplir tous les états d'énergie négative." (95) En décembre 1929 DIRAC publia une "théorie des électrons et des protons". Dans le premier paragraphe il reprend l'exposé des difficultés théoriques : les états d'énergie négative ne peuvent être éliminés de la théorie quantique relativiste puisque l'énergie peut varier de manière discontinue. Un électron d'énergie négative se meut dans un champ électrique comme s'il avait la charge $+|e|$. Hermann WEYL (116) avait proposé l'identification de deux des composantes de Dirac à la fonction d'onde du proton. A cette époque les seules particules fondamentales étaient en effet l'électron et le proton (l'hypothèse des neutrons de Rutherford était une pure spéculation) et ce proton devait lui aussi être décrit par un spineur. Weyl, ayant remarqué que pour $m = 0$, l'équation de Dirac se découple en deux équations de Pauli pensait que cette séparation, assimilée à la distinction électron - proton devait se conserver si l'on donnait à la masse une origine gravitationnelle. L'article de Weyl contient une autre remarque très intéressante : dans l'équation de Dirac le spineur associé à l'électron et celui "associé au proton" sont pris au même point d'espace - temps, contrairement au traitement quantique habituel du problème à deux corps, car en fait il faut quantifier l'onde de Dirac avant de comparer la théorie à l'expérience. Cette remarque contenait la solution du problème des énergies négatives ; mais curieusement Dirac la découvrit par une autre voie. *

Revenons à l'article de Dirac où il critique l'identification des états d'énergie négative à des états de proton. Cette hypothèse conduit à trois paradoxes :

* Nous reparlerons de la 2^o quantification symétrique de l'onde de Dirac à propos de Heisenberg - Pauli I, de la théorie de Pauli et Weisskopf des scalaires chargés, et prochainement à propos de la symétrie par conjugaison de charge étudiée par Weyl.

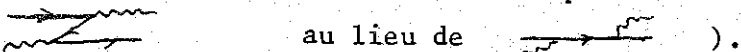
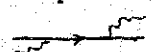
— Les transitions d'un état d'énergie positive à un état d'énergie négative ($e \rightarrow p$) violent la conservation de la charge.

— Le champ créé par un proton est celui d'une charge négative.

— L'énergie d'un proton diminue quand il va plus vite et par là même, il se meut dans la direction opposée à celle de la force appliquée : ce sont les "électrons - mules" ("donkey - electrons") de Gamow.

Le point de départ de Dirac est le problème de l'instabilité des états d'énergie positive : ils se désintègrent spontanément. Pour que cela ne soit plus possible il convient que les états d'énergie négatives soient entièrement occupés ou presque, ce qui est possible en vertu de principe d'exclusion de Pauli. Un trou dans la distribution des états d'énergie négative doit être interprété comme une particule d'énergie positive et de charge $+|e|$. Dans cette identification Dirac est guidé par la théorie des spectres des atomes lourds (voir aussi (94)). Le trou formé par l'éjection d'un électron d'une couche profonde peut être interprété comme un objet d'énergie négative.

Que faire de la distribution de charge associée à la mer infinie ? Dirac la déclare inobservable. Seules les différences par rapport à cette distribution sont observables. L'identification des trous aux protons pose un problème, remarque Dirac : on s'attend à ce que les masses soient égales — mais l'interaction électrostatique entre les particules de la mer peut éventuellement rendre compte de cette différence.

Une première justification de la théorie des trous est l'interprétation du rôle des états d'énergie négative comme états intermédiaires virtuels dans un processus de diffusion. WALLER avait fait remarquer à Dirac qu'on ne pouvait négliger dans le calcul de la diffusion du rayonnement les états intermédiaires d'énergie négatives.* La publication du calcul de Waller (66) est postérieure au papier de Dirac). L'explication donnée par Dirac est la suivante : le photon incident éjecte un électron de la mer dans un état d'énergie positive puis l'électron initial tombe dans le trou créé pour donner le photon émergent (autrement dit  au lieu de ).

La transition d'électrons de la mer dans un état d'énergie positive et le processus inverse donnèrent pour la première fois une image et une formalisation possible de la création et de l'annihilation de matière. L'idée de transformation de l'énergie électromagnétique en matière, ou celle de la transformation récipro-

* au contraire dans le cas de la diffusion Thomson, ils sont pratiquement les seuls à intervenir.

que n'étaient certes pas nouvelles et sans doute l'influence d'Eddington facilita-t-elle la tâche de Dirac. EDDINGTON, JEANS, NERNST... avaient introduit la notion d'annihilation de particules de charges opposées pour expliquer les données d'observation concernant la magnitude et la température des étoiles (diagramme d'Ertzsprung - Russel), en se fondant soit sur une théorie de l'éther soit sur la relation $E = mc^2$ d'Einstein.* La théorie de Dirac était révolutionnaire en ce qu'elle donnait de ce phénomène une interprétation mathématique précise, et sans modèle spéculatif de l'éther et qu'elle s'appuyait sur l'équation de Dirac dont le contenu empirique était déjà riche. Toutefois Dirac était resté aussi conservateur que possible en considérant le phénomène de création comme une réinterprétation d'un processus où le nombre total de particules matérielles (les électrons) était en réalité conservé. Et Dirac préféra longtemps ce point de vue à la quantification formelle symétrique de l'onde de Dirac aujourd'hui enseignée.**

Pourquoi Dirac interpréta-t-il les trous comme des protons ? L'idée d'une unification de la théorie de la matière — les seules particules matérielles simples connues étaient alors l'électron et le proton — eut sans doute quelque poids dans le choix de Dirac. Mais surtout le proton était la seule particule simple positive connue, et il eut été prématurément audacieux d'en imaginer une nouvelle après tant d'années vécues avec un nombre invariable de constituants fondamentaux.

Avant la publication de son article Dirac avait exposé par lettre son idée à Bohr. Bohr trouva cette idée "fascinante", mais avant tout pour lui le paradoxe de Klein et les états d'énergie négative montraient les limites du concept d'énergie dans le domaine ultrarelativiste, et les arguments de Landau, Peierls et Bohr lui-même rendaient inobservables certaines propriétés de l'électron de Dirac. Bohr écrivit : *"J'espère que toutes les applications réussies de votre équation d'onde sont défendables, mais j'ai le sentiment que la limitation naturelle de ces applications interdit une extrapolation comme celle décrite dans votre lettre."* ***

* Joan BROMBERG (163) a donné un excellent compte rendu de l'évolution du concept de création de particule.

** Nous reviendrons plus tard sur l'origine de la formulation moderne.

*** Ces lettres de Bohr et de Dirac ainsi qu'une analyse pertinente de leurs positions ont été publiées dans une série d'articles de D.F. Moyer (168).

L'article de Dirac, nous allons le voir, fut étudié par plusieurs théoriciens. L'idée de la mer, mutatis mutandis, fut adoptée en particulier par Oppenheimer.

Proton ou anti-électron ?

DIRAC avait noté sans commentaire dans le même article : "Quand un électron d'énergie positive tombe dans un trou et le remplit, l'électron et le proton disparaissent simultanément en émettant du rayonnement". La probabilité de ce processus est simple à calculer et elle le fut indépendamment par DIRAC (117), TAMM (111) et OPPENHEIMER (118) au début de l'année suivante (1930). Correspondant avec Dirac, Tamm se demanda quelle masse il fallait prendre dans la formule (durée de vie de l'ordre de $L^3/\lambda_0^3 c$ où $\lambda_0 = \frac{e^2}{4\pi m c^2}$ et $L =$ distance e - p) et il convint de prendre la moyenne arithmétique des deux masses. Oppenheimer tomba sur la même difficulté et prit la même moyenne car l'énergie rayonnée était $(m_e + m_p)c^2$ au lieu de $2m_e c^2$. Malheureusement la durée de vie de la matière ainsi obtenue variait de 10^{-3} s (Tamm) à 10^{-10} s (Oppenheimer) suivant la valeur prise pour L. Selon Dirac et Tamm il était possible que cette catastrophe fût évitable en tenant compte de l'interaction coulombienne entre tous les électrons, mais il ne voyaient pas très bien comment.

OPPENHEIMER en février 1930 (119) était déjà conscient de cet ordre de grandeur. Une autre difficulté le chagrinait : la diffusion Thomson sur les protons doit donner le même résultat que sur les électrons si l'on applique les règles de calcul de Dirac. De plus dans les deux cas le processus virtuel comprend la création d'une paire électron - trou par le photon incident (nécessairement dans le cas du proton ; dans le cas de l'électron on peut négliger à basse énergie les processus de diffusion directe sur l'électron par rapport au processus avec création de paire) et donc les corrections apportées par un effet éventuel de l'interaction coulombienne entre électrons seraient les mêmes. Mais la diffusion Thomson fait intervenir le carré de la masse de la particule diffusante : contradiction. Oppenheimer conclua qu'il fallait considérer que les états d'énergie négative étaient totalement remplis (mis à part les états intermédiaires virtuels nécessaires au calcul de diffusion) et que le proton devait être décrit par une équation de Dirac indépendante.

Hermann WEYL (120) reprit le problème de l'interprétation des états d'énergie négative, à la lumière de la nouvelle hypothèse de Dirac. De la préface de la deuxième édition de la "théorie des groupes et la mécanique quantique", datée de novembre 1930 nous extrayons le commentaire : "Le problème fondamental du proton et de l'électron a été discuté en relation avec les propriétés de symétrie des lois quantiques par rapport à l'échange de la gauche et de la droite, du passé et du futur, et des électricités positive et négative. A présent aucune solution du problème ne semble être en vue ; je crains que les nuages qui obscurcissent cette partie du sujet ne se regroupent pour former une nouvelle crise de la physique quantique". Cette dernière opinion est excessivement dramatique puisque la remarque de Weyl de la nécessité de l'identité de la masse du proton et de l'électron contenait la solution du problème. "Rien naturellement, ne nous interdit de quantifier les ondes de matière d'une manière analogue à celle appliquée aux ondes électromagnétiques" et "rien ne nous interdit de remplacer les nombres d'occupation $N_{p\Delta}$ par $N_{p\Delta}^- = 1 - N_{p\Delta}$ si $p_0 < 0$ et de garder $N_{p\Delta}^+ = N_{p\Delta}$ si $p_0 > 0$ " (p = impulsion, s spin). Alors N^+ et N^- ne diffèrent de 0 que pour un nombre fini d'états. Le hamiltonien en 2° quantification est invariant par conjugaison de charge, et la description de l'espace des états par N^+ et N^- est-elle aussi symétrique par cette opération. Conclusion. " Avec l'interprétation de Dirac du proton cela signifie que les électricités positives et négative ont essentiellement les mêmes propriétés, en ce sens que les lois qui les gouvernent sont invariantes par une certaine substitution qui permute les nombres quantiques de l'électron et ceux du proton. La dissimilarité des deux espèces d'électricité semble ainsi cacher un secret de la nature..."

Les arguments de Weyl convainquirent Dirac. Ce dernier dans l'introduction à son étude des singularités quantifiées du champ électromagnétique (111) publiée en mai 1931 abandonna l'identification des trous au proton. D'Oppenheimer, il retint l'idée que les états d'énergie négative étaient normalement tous occupés, mais il n'exclut pas la possibilité de créer artificiellement un trou (par exemple par collision) qu'il faudrait alors interpréter comme un "anti-électron" de charge opposée et de masse identique à l'électron. Les protons devaient alors faire l'objet d'une théorie indépendante analogue à celle de l'électron, ce qui suggéra la notion d'"antiproton" à Dirac.

Avec l'idée de l'antiélectron et de l'antiproton s'achevait le premier volet du projet initial de Dirac : trouver une théorie quantique invariante relativiste. Le cadre théorique de l'électrodynamique quantique, avec d'un côté la quantification du rayonnement, de l'autre l'électron relativiste, était prêt pour explorer les difficultés du problème couplé. La théorie de l'électron s'était constituée suivant la méthodologie définie par Dirac (121) : *"La méthode la plus puissante qui puisse être suggérée actuellement est d'employer toutes les ressources des mathématiques pures pour tenter de parfaire et de généraliser le formalisme mathématique qui forme la base existante de la physique théorique, et après chaque succès dans cette direction, d'essayer d'interpréter les nouveaux faits mathématiques en termes d'entités physiques."*

La prédiction ^{*} de Dirac du positron et de l'antiproton, à une époque où le nombre de particules connues était très limité, fut certainement beaucoup plus surprenante aux yeux de ses contemporains, que ne l'est aujourd'hui celle de tel ou tel nouveau quark. En témoigne le jugement de Pauli : *"Le défaut expérimental de telles particules est attribué à un état initial particulier. Cela paraît discutable, car les lois de la nature dans cette théorie sont exactement symétriques par rapport à l'électron et à l'antiélectron. Enfin des photons (...) devraient pouvoir se transformer d'eux-mêmes en un électron et un antiélectron. Nous ne croyons donc pas que cette issue puisse être sérieusement prise en considération."* (122) SCHRÖDINGER (123), lui aussi non satisfait par la théorie des trous de Dirac, tenta de modifier l'expression des opérateurs pour qu'ils ne connectent plus les états d'énergie positive à ceux d'énergie négative.

^{*} Moyer (168) affirme que Dirac *"n'a pas prédit que les antiélectrons (ou les monopoles magnétiques)"* seraient observés et qu'il a donné au contraire de bonnes raisons pour qu'ils ne soient pas observés facilement. Il est vrai que Dirac ne s'attendait certainement pas à ce que les antiélectrons se manifestent aussi tôt en physique expérimentale. (On sait même a posteriori que dès 1931 des expériences comme celles de Gray et Tarrant devaient s'expliquer par la production de positrons). Pourtant il me semble légitime d'employer le mot prédiction dans le sens qu'on lui attribue normalement en physique et qui diffère bien sûr de celui de la prophétie : *"Si les antiélectrons, écrivit Dirac (121), pouvaient être produits expérimentalement dans un bon vide, ils pourraient être tout à fait stables et observables."*

La découverte expérimentale du positron :

"Dans un groupe de 1300 photos de traces de rayons cosmiques dans une chambre de Wilson verticale, 15 traces appartenaient à des particules positives qui ne pouvaient avoir une masse aussi grande que celle du proton. De l'examen de la perte d'énergie et de l'ionisation produite, il résulte que la charge est moins de deux fois celle du proton et qu'elle lui est probablement égale. Si ces particules portent l'unité de charge positive les courbures et les ionisations produites requièrent une masse inférieure à 20 fois la masse de l'électron. Ces particules seront appelées positrons..." L'éditeur de la Phys. Rev. (124).

Les premiers clichés furent pris le 2 août 1932 par ANDERSON au California Institute of technology. Un bref compte rendu fut publié au bout de quelques semaines (125) "avec une juste réserve pour l'interprétation, compte tenu de l'importance et de la nature frappante de l'information" (124). Anderson ne connaissait par la prédiction de Dirac si bien qu'aucune mention n'en fut faite.

Des clichés analogues furent obtenus presque en même temps au Cavendish Laboratory par BLACKETT et OCCHIALINI (126). Blackett approuva la conclusion d'Anderson et identifia la particule découverte à l'électron positif de Dirac. Il reconnut que certains phénomènes d'absorption anormale observés par Gray et Tarrant, Meitner et Hupfeld et Chao dès 1931 pouvaient s'interpréter en terme de création-annihilation de positrons. Il soupçonna aussi la possibilité de créer des positrons à partir de rayons γ au laboratoire, ou à partir de neutrons, ce qui expliquait la découverte de CURIE et JOLIOT (127) d'électrons secondaires se mouvant en sens inverse. Curie et Joliot (128) confirmèrent un peu plus tard cette interprétation.

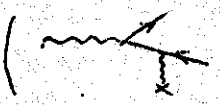
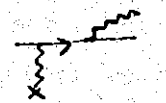
Selon Dirac (94) de tels électrons animés d'un mouvement rétrograde, avaient été observés au Cavendish Laboratory même un peu avant la "théorie des électrons et des protons" mais on interprétait le phénomène comme une "coïncidence". "Blackett fut réellement le premier à obtenir une évidence nette en faveur de l'existence du positron, mais il eut peur de publier. Il voulait une confirmation, il était réellement trop prudent."

En 1933 de nombreux résultats furent publiés sur les propriétés des positrons émis lors de l'absorption de rayons γ de sources radioactives dans la matière. Citons Anderson, Neddermeyer en Californie, Curie et Joliot en France, Meitner

et Philipp en Allemagne, Chadwick, Blackett et Occhialini en Angleterre. BLACKETT, dans un article de synthèse (129), résuma les conséquences de ces travaux et conclua que la particule observée avait bien les propriétés attendues d'un positron (production simultanée de e^- lors de $\gamma \rightarrow e^+ e^-$, seuil à $2 mc^2$, spin $\frac{1}{2}$, absorption par la matière).

L'abondance des données expérimentales sur la production de paires par collision de photons γ sur un noyau ou par les rayons cosmiques (voir chap. 6 §1), stimula les calculs théoriques de tels processus.

Une première estimation fut effectuée par OPPENHEIMER et PLESSET (130). Leur résultat était en accord avec les expériences d'absorption des rayons γ issus de sources radioactives, mais pas avec celles d'absorption des rayons cosmiques très énergétiques, HEITLER et SAUTER publièrent la formule exacte (dans l'approximation de Born) et enfin BETHE et HEITLER, dans un article (62) que nous avons déjà cité à propos de la Bremsstrahlung évaluèrent simultanément les différentes causes d'absorption des particules rapides dans la matière, en comparant soigneusement leurs résultats à ceux de l'expérience. Le calcul est très voisin de celui de la Bremsstrahlung

( au lieu de ). L'accord avec l'expérience était excellent en 1934 à part pour les électrons très rapides d'Anderson de 300 Mev.*

Le processus d'annihilation $e^+ e^- \rightarrow \gamma\gamma$ avait déjà été calculé, nous l'avons vu, par DIRAC, OPPENHEIMER et TAMM, alors qu'il était encore interprété comme $e^+ e^- \rightarrow \gamma\gamma$. Une simple substitution de noms donne des prédictions pour l'absorption des positrons dans la matière.

Bilan et nouvelles difficultés :

Au congrès Solvay de 1933 (132) Dirac présenta sa théorie du positron :
"La découverte récente de l'électron positif ou positron a rémené l'attention vers une théorie déjà ancienne sur les états d'énergie négative de l'électron,"

* Nous reviendrons, à propos de l'étude des rayons cosmiques, sur les violations apparentes de la formule de Bethe et Heitler. La première réaction des théoriciens des rayons cosmiques face à l'existence de particules très énergétiques faiblement absorbées fut de soupçonner la théorie.

les résultats expérimentaux obtenus jusqu'ici se trouvant en accord avec les prévisions de cette théorie." Puis il effectua une première évaluation de la polarisation du vide, mettant ainsi le doigt sur un nouveau phénomène physique et sur un nouveau type de divergence. Dans la discussion qui suivit transparut l'importance accordée à la théorie du positron par d'autres théoriciens. Pauli naguère hostile à l'hypothèse de l'électron positif, introduisit ainsi son intervention: "La théorie des lacunes m'a toujours paru très intéressante, à cause du rôle essentiel qu'y joue le principe d'exclusion," puis s'attarda sur les difficultés liées à la polarisation du vide. BOHR se réjouit de voir s'étendre sa méthode de correspondance: "La merveilleuse confirmation apportée à la théorie de l'électron de Dirac par la découverte des positrons à éclairé d'un jour nouveau les paradoxes qui avaient semblé un moment limiter d'une manière décisive l'application de la méthode de correspondance dans la mécanique quantique relativiste.", et s'inquiéta de l'harmonie entre cette nouvelle théorie et les limites de mesurabilité des coordonnées de l'électron (Voir chap, 4 § 5).

En 1933 Dirac reçut le prix Nobel conjointement avec Schrödinger pour "la découverte de nouvelles formes productives de théorie atomique." La conférence de Dirac fut centrée sur l'Equation de Dirac, et finit par la prédiction générale de l'antimatière: "il doit être possible d'inverser la charge de n'importe quelle particule... Si nous acceptons le point de vue d'une complète symétrie entre les charges négative et positive comme une loi fondamentale de la nature, nous devons considérer comme un accident que la terre... contienne surtout des électrons négatifs et des protons positifs. Il est tout à fait possible que pour certaines étoiles ce soit le contraire..." Dans un contexte très différent, celui de la théorie de Weyl de la gravitation et de l'électromagnétisme (Voir § 5) complétée d'un Ansatz sur le lagrangien d'interaction, Pauli (155) nota en 1919: "une conséquence générale de la théorie est que dans un champ statique les lois de la nature sont symétriques en ce qui concerne les électricités positive et négative", mais loin d'y voir la prédiction d'antiparticules Pauli continuait ainsi: "cette circonstance est une difficulté bien grande, voire insurmontable, pour la théorie de Weyl". Lors de ses tentatives de théorie unitaire des champs EINSTEIN déplora lui aussi en 1925 (156) que les hypothétiques solutions particulières de ses équations aient leur pendant de charge opposée,

il en douta qu'une unification de l'électromagnétisme et de la gravitation fût possible. L'hypothèse de Dirac, d'existence de l'antimatière était vraiment très audacieuse. *"On peut voir combien l'idée d'antimatière semblait déraisonnable à cette époque au fait que beaucoup d'entre nous ne crurent pas à l'existence d'une antiparticule du proton à cause de son moment magnétique anormal."*, se rappelait récemment WEISSKOPF (157).

La découverte du positron en 1932 est le dernier exemple avant 1947 d'une interaction spectaculaire entre la théorie de l'électrodynamique quantique et l'expérience. D'un côté le problème des divergences UV allait persister, malgré le progrès dans l'étude de la polarisation du vide (Voir chap.5), de l'autre les techniques expérimentales s'affinèrent jusqu'à ce que les corrections radiatives soient décelables.

2.3 GENERALISATIONS DE LA THEORIE DU RAYONNEMENT DE DIRAC :

Il fallut d'abord mettre l'électrodynamique pure sous une forme qui permette une vision et une exploitation plus claires de certaines propriétés générales attendues comme l'invariance relativiste et l'invariance de jauge. Mise à part l'introduction d'éléments nouveaux comme l'équation de Dirac pour l'électron, les nouvelles théories de Fermi et de Heisenberg et Pauli étaient en fait équivalentes à la théorie de Dirac. Cependant elles la justifiaient plus profondément, et ouvraient la porte à un champ plus vaste d'applications, en physique nucléaire et en physique des solides.

Peu après avoir pris connaissance de l'article de Dirac du 2 février 1927 sur la théorie quantique du rayonnement, PAULI s'engagea avec HEISENBERG dans un programme de construction de l'électrodynamique quantique. Ce programme se déroula sur environ trois années. Le 23 février 1927 Heisenberg adressa une lettre ⁽¹⁶²⁾ à Pauli où il présentait son interprétation physique du formalisme de la mécanique quantique. En 1927 l'élaboration de l'interprétation copenhagienne fut l'activité prépondérante de théoriciens comme Bohr, Heisenberg, Pauli et Kramers. La recherche de nouvelles applications, bien qu'elle fût déjà commencée par Dirac avec l'étude de l'émission et de l'absorption du rayonnement vint en général un peu plus tard. En tête de la lettre citée Heisenberg dit son enthousiasme pour le programme de Pauli : *"J'approuve tout à fait votre programme concernant l'électrodynamique... Que l'on doive quantifier les équations de Maxwell pour obtenir les quanta lumineux etc., à la Dirac, je le crois déjà ; mais on doit alors peut-être quantifier aussi les ondes de de Broglie, pour obtenir la charge, la masse et la statistique [!] des électrons et des noyaux."* avant même que la notion de 2^o quantification ait été étendue aux ondes de matière et aux fermions par Jordan et Klein et par Jordan (resp.), Heisenberg prophétisait l'existence d'une théorie expliquant la structure quantique de la charge et de la masse. Dans une lettre à Bohr (29 mars 1927, [160] ⁽¹⁶²⁾) Pauli ne cachait pas son enthousiasme : *"Actuellement je m'occupe beaucoup d'électrodynamique quantique. Je suis vraiment de bonne humeur et je juge plutôt bonne la situation actuelle de la physique. Il semble presque que la solution de l'énigme quantique ne soit plus très loin"*. La première étape du programme concerne seulement la quantification du champ électromagnétique libre par PAULI et

Pascual JORDAN. Une lettre de Heisenberg à Pauli du 3 juin 1927 ([165] (162)) montre que Jordan et Pauli avaient déjà établis leur expression du commutateur du champ électromagnétique libre ; ils avaient rencontré la difficulté que présente la quantification en variables A_μ , (Voir les articles de Heisenberg-Pauli). Ce qui explique que leur article (133) du 7 décembre 1927 utilise l'analogie des oscillateurs : l'écriture des équations de Maxwell en transformée de Fourier donne un choix de variables indépendantes pour le système.

2.3.a La fonction Δ :

En 1927, une théorie relativiste correcte de l'électron faisait défaut ; il était donc prématuré d'aborder le problème complet d'une électrodynamique quantique invariante relativiste. Seul le cas du champ électromagnétique libre fut traité à la fin de l'année 1927 par JORDAN et PAULI (133). Dans la représentation de Schrödinger, utilisée initialement par Dirac dans la théorie du rayonnement, les opérateurs-champ ne dépendent pas du temps et les relations de commutation non plus, si bien que l'invariance relativiste du procédé de quantification est cachée. Bien qu'ils ne mettent pas l'accent sur ce point, Jordan et Pauli utilisent la représentation de Heisenberg, ce qui revient à remplacer $e^{i\vec{k}\cdot\vec{r}}$ par $e^{i(\vec{k}\cdot\vec{r} - \omega t)}$ dans le développement de Fourier de $F_{\mu\nu}$. Les relations de commutation des coefficients de Fourier donnent alors :

$$[F_{\mu\nu}(x), F_{\rho\lambda}(y)] = i \left[(g_{\nu\lambda} \frac{\partial}{\partial y^\rho} - g_{\nu\rho} \frac{\partial}{\partial y^\lambda}) \frac{\partial}{\partial x^\mu} - (g_{\mu\lambda} \frac{\partial}{\partial y^\rho} - g_{\mu\rho} \frac{\partial}{\partial y^\lambda}) \frac{\partial}{\partial x^\nu} \right] \Delta(x-y)$$

avec $\Delta(x-y) = -i \int \frac{d^4k}{(2\pi)^3} \delta(k^2) \varepsilon(k^0) e^{-ik(x-y)}$ ($\varepsilon(k^0)$ fonction signe de k^0).

La "fonction" Δ est introduite, comme limite de fonctions, avec toutes les précautions utiles à une époque où la théorie des distributions n'existait pas encore. L'invariance relativiste est maintenant manifeste. Cependant ces auteurs soulignent que le passage à la limite continue (avec introduction de fonctions

* Jordan et Pauli utilisent directement les composantes de Fourier de $F_{\mu\nu}$ et non pas A_μ si bien que les difficultés liées à l'ambiguïté de jauge n'apparaissent pas.

de Dirac et de dérivations fonctionnelles de Volterra)* présente une difficulté : l'infinité de l'énergie de zéro des oscillateurs. Le caractère purement formel de cette difficulté ne devait être reconnu que plus tard (Voir § 6) . La fonction Δ (plus généralement sa version dans le cas d'une masse non nulle), qui apparaît plus simplement dans les règles de commutation du champ scalaire libre : ($[\varphi(x), \varphi(y)] = i \Delta(x-y)$) est omniprésente dans la théorie des champs moderne. L'idée d'utiliser une représentation (ici la représentation de Heisenberg; pour le problème couplé ce sera la représentation d'interaction) où le champ d'opérateurs électromagnétique satisfait les équations de Maxwell libres et les relations de commutation explicites et covariantes de Jordan-Pauli fera aussi fortune en théorie des champs.

2.3.b L'électrodynamique quantique de FERMI :

Selon Fermi, le travail de Jordan et Pauli mettait le problème ⁽¹⁹⁾ "de l'électrodynamique en absence de charge sous une forme presque définitive". La théorie de Dirac, elle, le satisfait moins, et il s'employa à la reprendre sous une forme mathématique plus familière. Il élaborait sa propre théorie pendant l'hiver 1928-1929, c'est-à-dire pratiquement en même temps que celle de Heisenberg-Pauli, mais tout à fait indépendamment.** Selon Amaldi ⁽¹³⁶⁾, il n'était pas encore très familier avec le formalisme de la deuxième quantification. Le formalisme de Fermi, que nous allons maintenant résumer, est exposé en deux articles essentiellement. FERMI lui-même et ses collègues AMALDI et RACAH développèrent nombre d'applications dont nous avons déjà dit un mot

* Nous reviendrons sur cet aspect des premiers travaux de Heisenberg Pauli et Jordan à propos du premier article de Heisenberg-Pauli. La mathématique de Volterra intervenait dans l'article de Jordan et Pauli pour écrire directement à l'aide du champ d'opérateurs $F_{\mu\nu}$ et non pas à l'aide de ses composantes de Fourier l'équation de conservation de l'énergie en représentation de Heisenberg.

** "... est paru un travail de Fermi (mai 1929) sur l'électrodynamique quantique que j'ai étudié de près. Il est entièrement indépendant du travail de Heisenberg et moi-même; la méthode est intéressante (bien qu'elle n'apporte pas de nouveaux résultats)" (lettre de Pauli à Jordan du 30 nov. 1929 [238] (162)).

à propos des applications de la théorie de Dirac. Tous ces résultats furent exposés en détails lors des cours donnés en Avril 1929 à l'institut Poincaré à Paris et à l'école d'été de physique théorique de Ann Arbor (Michigan) pendant l'été 1930.

Le premier article de Fermi sur l'électrodynamique quantique fut présenté en mai 1929 aux "Rendiconti della accademia nazionale dei lincei" Fermi n'était pas satisfait par la théorie de Dirac. "L'électrodynamique de Dirac est incomplète parce qu'elle considère seulement le champ de radiation[.] en réalité le champ électromagnétique qui se produit au voisinage d'un corpuscule électrique en mouvement est autrement plus compliqué[.] on ne peut traiter le problème de l'interaction de deux électrons[.] tenant compte du champ qui se propage de l'un à l'autre par les potentiels retardés..." Ces reproches ne sont qu'en partie justifiés : la théorie de Dirac, tient compte de l'interaction coulombienne, contenue dans le hamiltonien atomique ; les effets de retard peuvent être introduits dans la théorie de Dirac, et Dirac lui-même dans sa théorie de la dispersion ⁽⁵¹⁾ souligne que l'expression exacte du potentiel vecteur \vec{A} doit contenir la dépendance $e^{i\vec{k}\cdot\vec{r}}$. Il reste que la description séparée de l'interaction coulombienne par un potentiel paraît moins satisfaisante qu'un traitement homogène du champ électromagnétique complet tel que Fermi le propose.

Le point de départ de Fermi ⁽¹⁹⁾ est l'équation de D'Alembert pour le quadri-potentiel \star : $\square A_\mu = j_\mu$ qui donne une équation linéaire du 2° ordre en $\frac{\partial}{\partial t}$ pour les coefficients de Fourier de A_μ . De telles équations peuvent être considérées comme les équations de Hamilton d'un hamiltonien que Fermi écrit explicitement (ce hamiltonien est celui déduit du lagrangien $L = -\frac{1}{4} F^2 - \frac{1}{4} (\partial A)^2 - \vec{j}\cdot\vec{A}$ à une 4-divergence près). Ayant trouvé le hamiltonien on peut appliquer les règles de quantification canonique et l'équation du mouvement est celle de Schrödinger. Avant la quantification Fermi note que $\square(\partial A) = \partial j = 0$ et donc $\partial A = 0$ à tout instant si $\partial A = 0$ et $\partial A = 0$ à l'instant origine ; bien sûr pour que $\square A_\mu = j_\mu$

* Nous employons les notations modernes pour simplifier l'exposé mais il est important de noter que Fermi utilisait les composantes de Fourier du champ et qu'il se souciait peu d'avoir des notations covariantes.

soit équivalente aux équations de Maxwell il convient que cette condition soit vérifiée. Fermi n'explique pas comment elle se traduit au niveau quantique dans ce premier article.

Ce point important est explicité dans l'article ⁽²⁰⁾ de septembre 1930, Fermi avait alors pris connaissance des deux papiers de Heisenberg et Pauli : *"mais les méthodes suivies par ces auteurs sont essentiellement différentes des miennes, et je ne crois pas inutile de publier mes résultats"*. Les électrons sont maintenant des électrons de Dirac ; à part cette différence le hamiltonien est le même qu'auparavant. La condition $\partial_t A = 0$ et $\nabla \cdot \vec{A} = 0$ doit être considérée comme une restriction dans l'espace des états : $(\nabla \cdot \vec{A})|\psi\rangle = 0$ et $(\partial_t A)|\psi\rangle = 0$. La dépendance de la fonction d'onde dans les degrés de liberté longitudinaux du champ peut alors être factorisée, et l'autre facteur vérifie une équation de Schrödinger effective qui n'est autre que celle donnée par le hamiltonien de la théorie du rayonnement de Dirac. L'apparition de l'interaction coulombienne se trouve ainsi justifiée. Le lecteur intéressé par l'aspect technique du problème se reportera au paragraphe sur l'invariance de jauge, où nous montrerons les avantages de la méthode de Fermi par rapport aux autres. Après l'avoir critiquée sérieusement, Fermi retombe sur la théorie du rayonnement : *"De là nous pouvons conclure que sous cette forme le problème de l'électrodynamique quantique n'est nullement plus compliqué qu'un problème ordinaire de la théorie du rayonnement"*. Mais au départ tous les degrés de liberté des champs sont traités symétriquement ce qui permet entre autres avantages formels une preuve d'invariance relativiste analogue à celle de Heisenberg-Pauli, bien que Fermi ne la donne pas (il s'intéressait plutôt aux applications). L'astuce essentielle de Fermi est d'être parti des équations du mouvement et non pas du lagrangien ($L = -\frac{1}{4} F^2$) pour trouver le hamiltonien ; il contourne ainsi les difficultés rencontrées par Heisenberg et Pauli.

Beaucoup de physiciens des années trente ont appris l'électrodynamique quantique grâce aux cours de Fermi ⁽²¹⁾, réputés pour leur clarté et leur souci des applications concrètes.

D'après Amaldi ⁽¹³⁶⁾, le problème de la divergence de l'énergie propre de l'électron préoccupa Fermi. Il effectua plusieurs tentatives pour trouver une expression du hamiltonien pour laquelle ces divergences n'apparaî-

traient pas. Il publia ⁽¹³⁷⁾ un calcul de la masse électromagnétique de l'électron pour une coupure finie et compara la divergence obtenue à celle de la théorie classique.

2.3.c La dynamique quantique des champs d'onde de HEISENBERG et PAULI :

Ce travail de Heisenberg et Pauli procède d'un esprit assez différent de celui de Fermi. Fermi trouva une méthode directe et élégante de quantification de l'électrodynamique aboutissant à des équations susceptibles d'applications nombreuses et immédiates. Heisenberg et Pauli élaborèrent un cadre plus général et plus formel, donnant des justifications a priori, et aussi une élaboration systématique des propriétés formelles de leur théorie. Bien sûr la première application en vue était toujours l'électrodynamique, mais les preuves données pour la quantification canonique des champs sont beaucoup plus générales. L'aspect détaillé et pédagogique des démonstrations donne aux deux célèbres articles sur la dynamique des champs d'onde ^(17 et 18) de mars et septembre 1929 un volume considérable : 61 pages pour le premier, 22 pour le second. Ces deux articles et celui d'Oppenheimer constituent l'aboutissement du programme lancé par Heisenberg et Pauli au début de 1927. Au cours de cette première année Jordan et Pauli réussirent, nous l'avons vu, à quantifier le champ électromagnétique libre sous une forme explicitement covariante. (Grâce à l'usage de la transformée de Fourier ils avaient alors pu éviter la difficulté dans la définition des variables canoniques conjuguées dans le formalisme lagrangien conventionnel : si $L(A_\mu, \dot{A}_\mu) = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$, $\frac{\delta L}{\delta A_0} = F^{00} = 0$ et on ne peut définir de conjuguée de A_0 . (Voir § 5 pour l'invariance de Jauge). La difficulté n'était pas si grave dans le problème du champ de Maxwell libre puisque la transformée de Fourier permettait d'identifier des variables canoniques indépendantes ; l'expression obtenue pour le hamiltonien était finie ; l'évolution temporelle étant très simple, on pouvait expliciter la forme des relations de commutation en représentation de Heisenberg et prouver leur covariance. Une lettre de Pauli à Dirac du 17 février 1928 ([187] ⁽¹⁶²⁾) rend compte des difficultés auxquelles Heisenberg et Pauli se heurtèrent dans le cas du problème couplé. A défaut de pouvoir mettre les équations de l'électrodynamique sous une

forme hamiltonienne simple, Heisenberg et Pauli se plaçaient en représentation de Heisenberg : l'évolution des champs peut alors être obtenue par les équations de Maxwell-Lorentz, sans connaître le hamiltonien du système. Mais alors on ne pouvait plus contrairement au cas de Jordan-Pauli, expliciter les relations de commutation des divers champs, et l'expression de l'énergie-impulsion totale divergeait. La situation était doublement désastreuse : on n'avait pas de preuve formelle de covariance et de conservation des grandeurs physiques intéressantes (une telle preuve nécessitait d'abord de mettre les équations sous forme hamiltonienne) et ces grandeurs divergeaient. Pauli demanda à Dirac son diagnostic. Mais Dirac avait déjà fort à faire avec les états d'énergie négative de l'équation qu'il venait de découvrir. Momentanément à cours d'idées, Heisenberg et Pauli se tournèrent vers d'autres problèmes : Heisenberg s'intéressa au ferromagnétisme, et Pauli au théorème H. A la fin de 1928 l'intérêt de Pauli pour la physique en vint à décroître : *"Au début de ce semestre (automne 1928) la physique était assez éloignée de mes préoccupations, j'étais très paresseux..., mais encore très alerte et en bonne forme. Pour mon propre "amusement" j'esquissai alors un roman utopique qui devait s'intituler "Le voyage de Gulliver chez Uranie" et donner dans le style de Swift une satire contre la démocratisation d'aujourd'hui, c'est-à-dire de tout ce qui ne fait que ramper vers de lointains parlements, élections et majorités ! Comme j'étais plongé dans ces rêves, une nouvelle de Heisenberg me surprit : il avait pu, grâce à un artifice[.], éliminer les difficultés qui s'opposaient à la confection de notre électrodynamique quantique.[.] Je partis bientôt pour Leipzig où nous convinmes Heisenberg et moi d'un programme de travail pour compléter les calculs..."* L'astuce de Heisenberg, que nous exposerons au § 5, permettait de mettre les équations de l'électrodynamique sous forme canonique.

Dans l'introduction du premier article les auteurs analysent les difficultés d'une théorie invariante relativiste des interactions entre particules chargées : à ses débuts la mécanique quantique ne considérait que des interactions obtenues par un terme ajouté dans le hamiltonien libre des particules, par exemple l'interaction coulombienne. Classiquement une telle description

hamiltonienne où seraient éliminés les degrés de liberté du champ électromagnétique n'est pas possible (on ne peut les éliminer que jusqu'à l'ordre $(\frac{v}{c})^2$, ce qui donne le hamiltonien de Darwin) il faut donc abandonner l'espoir de construire la théorie quantique correspondante. *

La base classique de correspondance doit faire intervenir le champ électromagnétique et l'onde de Dirac couplés, concluent Heisenberg et Pauli. L'introduction d'un champ se propageant à vitesse finie pour décrire les interactions paraît naturelle, voire indispensable, dans une théorie invariante relativiste.

En revanche, on peut se demander pourquoi Heisenberg et Pauli choisissent d'emblée de décrire les particules chargées par l'onde de Dirac quantifiée et non par l'ensemble de leurs positions, à un moment où la justification la plus tranchée de ce choix — l'existence du positron, et la possibilité de créer ou d'annihiler un électron — n'était pas connue. Heisenberg et Pauli connaissaient parfaitement les travaux de JORDAN, WIGNER et KLEIN sur la quantification des ondes de matières **: pour ces derniers il ne faisait guère de doute que les ondes de matières devaient comme les ondes électromagnétiques subir une quantification, et à cet égard la nature fermionique des électrons ne posait pas de problème. Des théories des ondes de matière réciproquement couplées aux ondes électromagnétiques avaient déjà été formulées, d'abord par SCHRÖDINGER (77) en 1927 (équation de Schrodinger + équations de Maxwell couplées) avec des prédictions expérimentales fausses (voir § 1); TETRODE (139) en 1928 avait montré en étudiant le tenseur énergie-impulsion associé à l'équation de Dirac, qu'on ne pouvait obtenir d'équation de conservation qu'en rajoutant la contribution du champ électromagnétique qui ne pouvait plus ainsi être considéré comme imposé de l'extérieur; MÜGLICH (140) alla plus loin en écrivant le lagrangien du système couplé : onde de Dirac + champ de Maxwell, et en préconisant de quan-

* Eddington (138) essaya de construire une théorie de ce type, avec une interaction fonction des intervalles $(x_i - x_j)^2$, mais une telle théorie ne donne pas de prédictions correctes à la limite classique.

** Comme le prouve sa correspondance, Pauli manifesta aussitôt un intérêt très vif pour la quantification des ondes de matières dont Heisenberg avait pressenti la possibilité tout au début du programme commun de l'électrodynamique quantique (Voir l'introduction de ce §)

tifier le tout. Heisenberg et Pauli n'avaient plus qu'à opérer cette fusion de deux idées : l'idée de quantifier ondes de matière et ondes électromagnétiques, et l'idée d'une base de correspondance où les ondes de matière sont couplées aux ondes électromagnétiques.*

Tous les travaux précédents sont cités explicitement par Heisenberg et Pauli. Un autre antécédent important est la théorie de MIE ⁽¹⁴¹⁾ de 1928, où le champ électromagnétique est quantifié sous une forme invariante relativiste en partant de l'équation de Hamilton-Jacobi. La forme du lagrangien, non spécifiée, devait rendre compte, d'après la théorie de l'électron de Mie de 1912, (Voir chap. 7 §1) de l'existence d'états particuliers du champ électromagnétique identifiables à des électrons. Pour Heisenberg et Pauli cette théorie ne pouvait être qu'un simple "schéma formel".

Ils ont surtout retenu du travail de Mie, la généralisation du formalisme canonique au cas des champs : la valeur d'un champ $Q_\alpha(\vec{x})$ en chaque point constitue un degré de liberté indépendant. Pour définir les variables conjuguées il faut utiliser la dérivée fonctionnelle de Volterra : **

$P_\alpha(\vec{x}) = \delta L / \delta \dot{Q}_\alpha(\vec{x})$. Les relations de commutation canoniques s'écrivent à l'aide de la fonction δ de Dirac : $[P_\alpha(\vec{x}; t), Q_\beta(\vec{y}; t)] = i \delta_3(\vec{x} - \vec{y}) \delta_{\alpha\beta}$

La preuve d'invariance relativiste de ces relations constitue le morceau de bravoure du premier article ⁽¹⁷⁾. En représentation de Heisenberg les P_α et Q_α dépendent du temps, et les relations de commutation sont écrites pour des temps égaux dans un certain référentiel. Il s'agit de savoir si elles sont équivalentes à celles que l'on écrirait, à temps égaux, dans un autre référentiel. Une transformation de Lorentz a pour effet :

1°) - de changer les coordonnées $x \rightarrow x'$ $y \rightarrow y'$

2°) - d'agir sur les indices spinoriels ou tensoriels des champs

(notés α) $Q_\alpha \rightarrow \sum_\beta S_{\alpha\beta} Q_\beta$ etc.

Pour une transformation de Lorentz infinitésimale les auteurs examinent séparément les deux effets; le calcul du premier est particulièrement laborieux.

— "Ich warne Neugierige !" (gare aux curieux !) disait Pauli, et les deux

* Remarquons, bien que Heisenberg et Pauli n'en parlent pas, qu'une électrodynamique classique cohérente de particules chargées ponctuelles n'existait de toute manière pas à cette époque.

** Dès le début de 1927, au moment où Pauli s'engageait avec Heisenberg dans un programme d'élaboration d'une électrodynamique quantique, Pauli s'intéressa à la mathématique de Volterra nécessaire à l'extension de formalisme canonique aux systèmes continus. En témoignent les lettres [150] [157] et [162] de la correspondance de Pauli.

effets se compensent pour donner les relations de commutation attendues. Remarquons que cette preuve est très générale, qu'elle ne fait intervenir ni la nature des champs, ni la dynamique particulière : il suffit que le lagrangien soit invariant.

L'application du formalisme précédent au cas de l'électrodynamique se heurte à une première difficulté : il n'existe pas de hamiltonien associé au lagrangien usuel de l'électrodynamique. Nous étudierons ce problème au § 5. L'équation de Schrödinger pour les états stationnaires du système est écrite dans l'espace de Fock, dans les deux cas où les électrons (de Dirac) sont des fermions ou des bosons. Le deuxième cas irréaliste est tout de même traité, car les auteurs n'y voyaient pas de contradiction formelle (Voir chap. 5 § 3) et croyaient alors que les deux choix étaient possibles.

Le calcul des états et énergies stationnaires au 2° ordre en e et en négligeant la contribution en $\frac{v^2}{c^2}$ redonne l'énergie attendue au 2° ordre pour un ensemble d'électrons interagissant par le potentiel coulombien. Vue la forme de leurs équations les auteurs ne doutent pas que leur théorie redonne les prédictions de Dirac sur l'absorption et l'émission du rayonnement. Mais ils pensent (à tort) que leurs prédictions pour l'ionisation d'un atome avec émission simultanée d'un photon diffèrent de ce que donnerait la théorie de Dirac, qui, elle, "ne quantifie qu'une partie du champ".*

Le DEUXIEME ARTICLE ⁽¹⁸⁾ (7 sept 1929) est centré sur l'étude et l'exploitation des propriétés de symétrie du lagrangien de l'électrodynamique. Nous verrons au § 5 comment est exploitée l'invariance de jauge. La preuve d'invariance de Lorentz est considérablement simplifiée et très proche de la preuve moderne : il faut démontrer qu'à une transformation de Lorentz infinitésimale $x_\mu \rightarrow x'_\mu = x_\mu + \delta\omega_{\mu\nu} x^\nu$ correspond une transformation unitaire de Lorentz infinitésimale $U = 1 - \frac{i}{2} \delta\omega_{\mu\nu} M^{\mu\nu}$ telle que $Q'_\mu(x') = U Q_\mu(x) U^{-1}$ et $|\psi\rangle' = U |\psi\rangle$. La solution est donnée par $M^{\mu\nu} = \int d^3x (x^\mu \theta^{\nu\lambda} - x^\nu \theta^{\lambda\mu})$ ($\theta^{\mu\nu}$ étant le tenseur d'énergie-impulsion exprimé en fonction des champs quantiques. Les prédictions

* Ce calcul est effectué pour cette raison précise, mais aussi parce qu'il permet d'expliquer l'existence d'un continuum d'énergie pour l'électron éjecté. Les auteurs proposent aussi des analogies avec l'effet Auger et avec la continuité des Spectres β ; l'ennui, dans ce dernier cas, c'est qu'on ne voit pas les rayons γ qui devraient être simultanément émis, remarquent Heisenberg et Pauli (puisque ce sont des neutrinos !) (Voir chap. 6 § 1).

sur les états stationnaires sont identiques à celles du premier papier. Pour la commodité des applications ultérieures (commencées par Oppenheimer ⁽²⁶⁾) l'espace des configurations des électrons est introduit ^{*} à la fin de l'article, ce qui ne présente aucune difficulté puisque le nombre des électrons est conservé par la dynamique.

Heisenberg et Pauli étaient conscients que plusieurs difficultés entâchaient leur théorie. Toutefois l'importance relative de ces difficultés était mal perçue ; la difficulté des solutions d'énergie négative de l'équation de Dirac était considérée comme la plus grave. L'énergie de zéro des "oscillateurs", et le problème de l'énergie propre coulombienne des électrons ^{**} étaient placés sur le même pied.

Malgré tout, une classe très large de problèmes physiques — ceux que nous avons répertoriés dans les applications de la théorie de Dirac — pouvaient être traités par la théorie de Heisenberg et Pauli. Le gain apporté par rapport à la théorie du rayonnement de Dirac était une preuve d'invariance relativiste dans le formalisme général de la théorie des champs, et un traitement plus symétrique de la matière et du rayonnement. Ce travail impressionnant fit autorité pendant de nombreuses années et nombre de publications traitant le problème de théorie quantique des champs dans la décennie suivante s'y réfèrent.

* C'est à cette partie qu'Oppenheimer apporta la contribution la plus importante. *"Oppenheimer ne toussote ici [à Zurich] presque plus, et mentalement il est tout-à-fait épanoui. Nous faisons maintenant ensemble l'électrodynamique quantique."* écrivit Pauli à Kronig en Juin 1929. ([226] ⁽¹⁶²⁾).

** Cette contribution à l'énergie propre peut être considérée comme une constante additive infinie dans le hamiltonien ; elle n'affecte pas par exemple les différences de niveaux atomiques. La difficulté la plus grave était contenue dans les autres contributions calculées par Oppenheimer ⁽²⁶⁾. (Voir Chap.3 § 2).

2.4 LA QUERELLE DES REPRESENTATIONS :

La mécanique des matrices de Heisenberg et la mécanique ondulatoire de Schrödinger résultèrent de deux approches très différentes, développées indépendamment. Appliquées à des problèmes concrets, elles conduisaient aux mêmes résultats et pourtant semblaient utiliser des concepts assez étrangers les uns aux autres. La démonstration d'équivalence donnée par Schrödinger, était vraiment indispensable. Suivant la représentation, ce sont les états (Schrödinger) ou les opérateurs (Heisenberg) qui évoluent dans le temps. Cette terminologie fut introduite par Dirac dans son cours de mécanique quantique *. La représentation de Heisenberg présente l'avantage de se fondre mieux avec l'idée de correspondance : les opérateurs évoluent alors comme les grandeurs classiques correspondantes. Pour cette raison HEISENBERG, puis DIRAC, FOCK et PODOLSKY utilisèrent des méthodes de quantification du champ électromagnétique qui en fait diffèrent de la théorie initiale de Dirac par un changement de représentation. L'équivalence des différentes méthodes, là encore, ne parut pas évidente à leurs auteurs, plus sensibles aux différences dans l'utilisation pratique. Une telle situation se répètera une troisième fois avec les théories de Schwinger et de Feynman qui ne diffèrent, elles aussi, comme le montra Dyson, que par un changement de représentation. Nous avons vu que Heisenberg et Pauli, pour démontrer l'invariance relativiste de leur théorie se plaçaient en représentation de Heisenberg, mais ils n'examinèrent pas alors davantage comment se traduisait dans les applications l'utilisation de cette représentation.

2.4. a La méthode de Heisenberg (44) :

En février 1931 Heisenberg opposa le point de vue de Heisenberg à la théorie du rayonnement de Dirac (qui utilisait celui de Schrödinger) "*Dans ce qui suit nous décrivons une méthode de traitement du problème du rayonne-*

* Il serait plus correct de traduire "picture" par "point de vue", en tout cas par un mot différent de "représentation", qui est déjà utilisé pour spécifier un choix de base dans l'espace des états (représentations p et q par exemple).

ment qui se raccorde beaucoup plus étroitement aux représentations intuitives ("anschaulich") de la théorie classique et de la mécanique ondulatoire et qui par conséquent donne sans détour dans la plupart des cas le résultat attendu de l'application du principe de correspondance."

Heisenberg retrouve les phénomènes d'émission et d'absorption et les conclusions de la théorie de la diffusion de Dirac. Comme le champ électromagnétique quantifié vérifie les équations de Maxwell, sa propagation à la vitesse c apparaît plus immédiatement que dans la théorie de Dirac. Enfin le concept d'énergie n'apparaissant pas — seule interviennent les équations classiques du mouvement — le problème de l'énergie de zéro infinie est inexistant.

Heisenberg mentionne le travail de KLEIN ⁽⁹²⁾ (cité à propos de l'équation de Klein-Gordon) de 1926 intitulé "Electrodynamique et mécanique ondulatoire du point de vue du principe de correspondance" où celui-ci cherchait lui aussi à traiter le rayonnement calculé par la formule des potentiels retardés en fonction du courant de probabilité par une méthode de correspondance. Mais les règles de Klein n'étaient que des recettes alors que Heisenberg appliquait la mécanique quantique. *

La méthode de Heisenberg, en tant que rivale de la théorie de Dirac, eut un certain succès auprès d'autres théoriciens. ROSENFELD ⁽¹⁴²⁾, à l'issue de son étude de la divergence de l'énergie propre, accorda sa confiance à "un raffinement de la méthode de correspondance... à une première approximation d'une électrodynamique quantique qu'on ne peut développer rigoureusement" puis il publia deux nouvelles applications de la théorie de Heisenberg d'abord au calcul des largeurs naturelles ⁽⁴²⁾ puis à la diffusion électron-électron ⁽⁵⁴⁾. Dans le dernier cas il suffit d'utiliser comme Møller les potentiels retardés, en supposant en plus qu'ils ne commutent pas. Bien sûr Rosenfeld retrouva les prédictions de la théorie de Dirac. CASIMIR ⁽⁴³⁾, en 1933, reprit

* Pauli, dans son article de 1933 dans le "Handbuch der Physik" ⁽¹²²⁾ (p.201 à 210) prit la peine d'exposer en détail les règles de Klein (et leur justification par la méthode de Heisenberg) car il n'y avait pas après tout de raison de préférer une électrodynamique quantique sans existence mathématique à des règles de correspondance arbitraires.

le calcul de Rosenfeld des largeurs naturelles et montra qu'il devait aussi y avoir dans la théorie de Dirac un déplacement infini des raies et que par conséquent la méthode de Heisenberg présentait les mêmes difficultés.

WEISSKOPF ⁽¹⁴³⁾ utilisa la méthode de Heisenberg pour calculer (1934) l'énergie propre de l'électron (Voir chap. suivant) parce qu'il avait ainsi une interprétation physique plus claire des différentes contributions ; mais il affirma que les deux méthodes étaient équivalentes puisqu'elles conduisaient aux mêmes prédictions et aux mêmes difficultés. On peut se demander pourquoi Heisenberg ne fut pas immédiatement conscient de cette équivalence. Sans doute parce que l'équivalence de deux théories sans existence mathématique n'a qu'un sens limité, elle est purement formelle ; déjà une divergence comme celle de l'énergie de zéro n'apparaît que dans l'un des deux points de vue ; on pouvait espérer un instant qu'il en serait de même pour les autres.

FRENKEL ⁽¹⁴⁵⁾, juste après la publication de la méthode de Heisenberg, la considéra comme un cas particulier de son traitement des systèmes incomplets (qu'on appellerait aujourd'hui passage à la représentation d'interaction), et malgré cela pensa qu'elle pouvait différer de la théorie de Dirac : *"l'avantage de cette théorie sur celle de Dirac réside dans sa simplicité relative, et dans le fait qu'elle reste dépourvue des difficultés fondamentales de la théorie de Dirac quand il s'agit seulement du champ de rayonnement."*

La représentation d'interaction, plus exactement, fut introduite par Dirac en électrodynamique, bien qu'il ne reconnût l'analogie avec le travail de Frenkel que plus tard ⁽¹⁴⁶⁾.

2.4.b La nouvelle théorie de Dirac (Mars 1932) :

Ce nouveau programme de Dirac fut présenté dans un premier article dans Proc. Roy. Soc. de mars 1932 rédigé dans un style assez différent des précédents : les calculs sont rares et les considérations d'ordre général, physiques, historiques, voire même philosophiques abondent. Comme Heisenberg, Dirac veut revenir à une application plus stricte du principe de correspondance de Bohr. Néanmoins Dirac ne fait aucune allusion au travail précédent de Heisenberg, et insiste contrairement à ce dernier sur la nécessité de

traiter dissymétriquement matière et rayonnement. La théorie de Heisenberg-Pauli est critiquée à ce propos : "Ainsi le rôle du champ est de fournir un moyen d'observation. La nature même d'une observation nécessite une interaction entre le champ et les particules. Nous ne pouvons donc pas considérer le champ comme un système dynamique sur le même pied que les particules et ainsi comme quelque chose d'observable de la même façon que les particules". * Les seules quantités observables sont les états d'un rayonnement incident et du rayonnement diffusé correspondant. Revenant au point de vue "observationaliste" de Heisenberg en 1925 il faudrait n'introduire pour décrire le rayonnement, que des amplitudes de diffusion de rayonnement **. Comment alors, établir la correspondance avec la théorie classique ? Celle-ci souffre de la divergence de la réaction de rayonnement sur des charges ponctuelles (Voir chap. 3 § 1) il faut donc espérer que la quantification lèvera cette difficulté (Dirac soutiendra en 1938 le point de vue contraire : guérir d'abord la théorie classique de ses divergences) et faire en sorte qu'elle ne fasse intervenir qu'un champ libre imposé. Comme une première illustration de ces idées Dirac envisage le cas du spin zéro. La fonction d'onde d'une particule chargée devra satisfaire $[(i\partial_\mu - eA_\mu)^2 - m^2]\psi = 0$ où A_μ est un champ imposé libre, vérifiant les relations de commutation canoniques ; pour N particules d'indice i Dirac introduit N temps distincts ; ψ dépend alors de (x_0, x_1, \dots, x_N) avec $x_i = (t_i; \vec{x}_i)$ et $[(i\partial_\mu^i - e_i A_\mu(x_i))^2 - m_i^2]\psi = 0$. Enfin Dirac prouve que cette théorie, dans le cas à une dimension d'espace, contient implicitement l'interaction coulombienne. Une solution ψ de ce système dépend de A_μ , elle peut donc être considérée comme un opérateur connectant états "in" et "out" de rayonnement, comme souhaité.

Un mois plus tard ROSENFELD (147) démontra en utilisant la représentation

* L'idée de Dirac de traiter la matière différemment du rayonnement persistera dans ses travaux ultérieurs : par exemple dans la théorie du positron (Voir chap. 5 § 1) il évitera la quantification de l'onde de Dirac.

** Dirac soutient ainsi momentanément un point de vue voisin de celui de Heisenberg initiateur du programme de la matrice S (Voir chap. 7 § 5).

d'interaction que la théorie de Heisenberg-Pauli satisfaisait aux exigences de Dirac : l'équation de Schrödinger donnée par la théorie de Heisenberg-

Pauli II : $(H_R + H_D)|\psi\rangle = i \frac{\partial |\psi\rangle}{\partial t}$ devient $H'_D |\psi\rangle' = i \frac{\partial |\psi\rangle'}{\partial t}$ si l'on pose

$$|\psi\rangle = e^{-iH_R t} |\psi\rangle' \quad \text{et} \quad H'_D = e^{+iH_R t} H_D e^{-iH_R t} \quad \text{ou encore}$$

$H'_D = \sum_i (-i\vec{\alpha}_i \cdot \vec{\nabla}_i + \beta_i m_i - e_i \vec{\alpha}_i \cdot \vec{A}'_i + e_i \vec{A}'_{0i})$ où A'_μ suit l'évolution libre. On a ainsi la nouvelle théorie de Dirac, à l'introduction près des temps distincts.

Sur ce dernier point la démonstration de Rosenfeld fut qualifiée par Dirac d'obscur (146) et DIRAC éprouva le besoin avec FOCK et PODOLSKY (146) de compléter son formalisme, en mettant l'accent cette fois sur l'introduction des temps multiples qui permettait une preuve simplifiée de l'invariance relativiste. Fock et Podolsky avant de collaborer à cet article avec Dirac, s'étaient intéressés à sa théorie et à son application à l'interaction de deux électrons (148). Ce faisant ils avaient développé une méthode de quantification du champ électromagnétique analogue à celle de Fermi (adoptée par Dirac). Les trois auteurs aboutirent au système :

$$\begin{cases} (i\vec{\nabla}_i - e_i \vec{\alpha}_i - m_i) \psi = 0 \\ \square A_\mu = 0 \end{cases} \quad \text{et} \quad \begin{cases} [\partial_\alpha A_\mu - \sum_i e_i \Delta(\alpha - \alpha_i)] \psi = 0 \\ [A_\mu(\alpha), A_\nu(\beta)] = i g_{\mu\nu} \Delta(\alpha - \beta) \end{cases}$$

où ψ dépend des coordonnées x_i d'espace-temps et des variables du champ A_μ (la fonction Δ de Jordan-Pauli s'introduit dans la condition supplémentaire de Fermi pour assurer sa compatibilité avec l'équation de Dirac).

Un tel système est manifestement invariant relativiste et de plus il redonne la théorie de Heisenberg-Pauli en égalisant tous les temps (et par transformation unitaire).

BLOCH (149) donna l'interprétation physique des fonctions d'onde ψ à temps multiples : elles sont associées à des mesures des positions des différentes particules à des temps distincts. (On ne saurait éviter d'envisager de telles mesures dans une théorie invariante relativiste, puisque la simultanéité n'est pas invariante). De plus pour que les différentes équations de Dirac soient compatibles, il faut que les x_i soient séparés par des intervalles du genre espace (ce qui est aussi nécessaire quand il existe un temps commun).

Une fois l'équivalence avec la théorie de Heisenberg-Pauli prouvée, le problème de l'application de la théorie de Dirac à l'interaction électron-électron,

réétudié par Fock-Podolsky (150) et par Nikolsky (151), devenait académique.

Tentation avortée de créer une nouvelle électrodynamique quantique, la nouvelle théorie de Dirac garde cependant une importance historique à cause de son invariance relativiste manifeste. La représentation d'interaction et l'introduction des temps multiples fut généralisée par TOMONAGA (152) en 1946. (Voir chap. 5 § 4).

Dans ce paragraphe nous avons vu apparaître trois représentations différentes et la prise de conscience plus ou moins rapide de leur équivalence formelle ; ce sont celles de Schrödinger, de Heisenberg et la représentation d'interaction de Dirac qui est en fait une représentation d'interaction mixte, où seule l'évolution libre du champ électromagnétique est éliminée des vecteurs d'état. La vraie représentation d'interaction utilisée le plus fréquemment aujourd'hui a été employée en 1938 par Heisenberg (153) * (Voir aussi chap. 4 § 6) pour établir le célèbre formule : $U(t, t_0) = T \exp -i \int H_I(x) d^4x$ (H_I désigne le hamiltonien d'interaction en représentation d'interaction ; les notations ne sont pas d'époque) dont la covariance relativiste est plus simple à prouver que celle de l'équation de Schrödinger.

* Elle prit plus tard toute son importance dans les travaux de Stueckelberg, de Tomonaga et de Dyson.

2.5. L'INVARIANCE DE JAUGE ET LA QUANTIFICATION DE L'ELECTRODYNAMIQUE :

Pour étudier l'émission et l'absorption de rayonnement par un atome DIRAC (13) avait scindé le champ électromagnétique en deux parties :

- le potentiel coulombien statique traité comme interaction instantanée entre les différents constituants chargés de l'atome.
- le champ de radiation satisfaisant aux équations de Maxwell. Cette séparation peut se justifier formellement (Voir l'appendice) par l'élimination, au niveau de la théorie-classique, des degrés de liberté longitudinaux du champ dans la Jauge de Coulomb, ce qui explique qu'elle conduise à des résultats équivalents à ceux des théories plus élaborées. Mais elle est contraire à "l'esprit" (Fermi) de la théorie de Maxwell, dans lequel les sources de champ déterminent complètement le champ, et elle n'est pas invariante par changement de référentiel puisque champ statique et champ de radiation se mélangent par transformation de Lorentz.

Si l'on traite tous les degrés de liberté sur le même pied, des difficultés apparaissent lors de la quantification canonique et l'invariance de jauge de l'électrodynamique fut bientôt reconnue comme responsable.

La difficulté formelle de base est bien connue: si $L = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu}$ est le lagrangien du champ électromagnétique, alors $\pi^0 = -F^{00}$ impulsion conjuguée de A^0 est nulle. On ne peut donc pas avoir $[A_0(\vec{x}), \pi_0(\vec{y})] = i \delta_3(\vec{x}-\vec{y})$

2.5.a FERMI : "Sopra l'elettrodinamica quantistica" (19)

Nous avons déjà décrit les grandes lignes de la méthode de Fermi (§.3). La difficulté citée n'apparaît pas alors, puisqu'en fait Fermi utilise un hamiltonien qui se déduirait du lagrangien $L = -\frac{1}{4} F^2 - \frac{1}{2} (\partial_0 A)^2$ * pour lequel $\pi^0 = -\partial_0 A \neq 0$. Les équations du mouvement $\square A_\mu = 0$ dans le cas libre

* à une dérivée par rapport au temps près : en réalité le lagrangien associé à la méthode de Fermi originale est $-\frac{1}{4} F^2 - \frac{1}{2} (\partial_0 A)^2 + \partial_0 (\vec{A} \cdot \vec{\nabla} A_0)$

(c'est d'elles que Fermi part) ne sont identiques aux équations de Maxwell que si la condition de jauge de Lorentz $\partial A = 0$ est vérifiée. Cette condition sera vérifiée à tout instant si à l'instant origine $\partial A = 0$ et $\overline{\partial A} = 0$; en effet $\square A_\mu = 0 \Rightarrow \square(\partial A) = 0$ et cette dernière équation n'admet que la solution $\partial A = 0$ pour les conditions initiales données. Après quantification on a $[A_\alpha(\vec{x}), \pi_\beta(\vec{y})] = i\delta_{\alpha\beta}(\vec{x}-\vec{y})$, donc $\partial A = \pi_0$ ne saurait être un champ d'opérateurs nul.

Les "conditions supplémentaires" à prendre concernent les états quantiques du système : les seules états physiques sont ceux tels que $\partial A |\psi\rangle = 0$ et $\overline{\partial A} |\psi\rangle = 0$.

2.5.b Heisenberg-Pauli I (17) :

Comme ces auteurs partent de l'expression du lagrangien habituel de l'électrodynamique $L = -\frac{1}{4} F^2$ et non pas des équations du mouvement, ils mettent directement le doigt sur la difficulté $\pi_0 = 0$. La difficulté est levée en ajoutant au lagrangien un terme $-\xi (\partial A)^2$, alors $\pi_0 = -\xi (\partial A) \neq 0$, en calculant les effets physiques recherchés (probabilités d'absorption, d'émission, perturbation des niveaux d'énergie...) avec $\xi \neq 0$ et en faisant tendre in fine ξ vers zéro.

2.5.c Heisenberg-Pauli II (18) et l'invariance de jauge de WEYL :

la différence de point de vue avec la première étude de quantification des champs d'onde est exprimée dans l'introduction : *"dans la mécanique quantique du point, des progrès essentiels ont été atteints par l'examen des propriétés d'invariance du hamiltonien... De même par les considérations suivantes nous exploiterons les propriétés d'invariance des équations d'onde"*. Les auteurs en viennent vite à l'invariance de jauge en se référant à Hermann WEYL (22), un des premiers à avoir vu toute l'importance des propriétés de symétrie et de la théorie des groupes en mécanique quantique. L'invariance de jauge a pris une telle importance dans les théories modernes que nous ouvrirons une brève parenthèse sur son histoire.

La dénomination bizarre d'invariance de jauge ("Eichinvarianz") prend sa

source dans les travaux de Weyl sur la relativité générale. En 1918 (23) WEYL proposa une généralisation de la théorie d'Einstein invariante non seulement par changement de coordonnées généralisées, mais encore par changement local des étalons d'espace-temps. Dans une telle théorie il n'y a pas d'étalon universel, il faut donc savoir comment varie la convention d'étalonnage quand on passe d'un point de l'espace-temps à un autre : $d\ell = -\ell A_\mu dx^\mu$ représente cette variation quand on passe de x_μ à $x_{\mu'} dx_{\mu'}$. Une transformation de jauge, c'est-à-dire un changement local dans un rapport $\lambda(x)$ des étalons implique $A_\mu \rightarrow A_\mu - \frac{1}{\lambda} \frac{\partial \lambda}{\partial x^\mu}$. Weyl identifia A_μ au potentiel électromagnétique, réalisant ainsi le "rêve de Descartes" d'une physique toute géométrique. EINSTEIN critiqua les conséquences physiques de cette théorie par exemple sur le comportement des horloges ($\oint \frac{d\ell}{\ell} \neq 0$) lors de leur déplacement et Weyl finit par se rallier plus tard au point de vue moderne sur l'interprétation de la symétrie de jauge. Ce point de vue fait intervenir la mécanique quantique : Schrödinger, en 1922, (24) remarqua que $\oint \frac{d\ell}{\ell}$, variation de l'étalonnage sur un circuit fermé, s'annule, comme on pouvait le souhaiter suivant Einstein, à condition que $\frac{d\ell}{\ell}$ soit imaginaire pur, et seulement sur les orbites de Bohr de systèmes liés simples. LONDON en 1927 (25) (donc après la découverte de l'équation de Schrödinger) identifia $\frac{d\ell}{\ell}$ à la variation de phase de la fonction d'onde. En 1928 dans "Gruppentheorie und Quantenmechanik" de Weyl on trouve (P. 100 de l'éd. anglaise) l'expression moderne de l'invariance de jauge : "les équations du champ pour les potentiels ψ et A_μ de la matière et des ondes électromagnétiques sont invariantes par le remplacement simultané de ψ par $e^{i\ell} \psi$ et A_μ par $A_\mu - \frac{\hbar}{e} \frac{\partial \ell}{\partial x^\mu}$ ".

Ainsi en 1929 Heisenberg et Pauli pouvaient reconnaître toute l'importance de cette symétrie. Le terme supplémentaire $-\frac{1}{2} \epsilon (\partial \cdot A)^2$ violait l'invariance de jauge. Plutôt que de l'introduire, ou de se rallier à la méthode de Fermi jugée intéressante mais introduisant elle aussi un terme $-(\partial \cdot A)^2$ non invariant de jauge, ils ont préféré réduire l'arbitraire de jauge en fixant $A_0 = 0$ et en considérant seulement les \vec{A} comme des variables dynamiques. Comme seuls les \vec{A} sont variés lors de l'application du principe de moindre action, $\text{div } \vec{E} - \rho = 0$

n'est plus une équation du mouvement. Cependant la liberté de jauge qui subsiste ($\vec{A} \rightarrow \vec{A} - \vec{\nabla}\xi$, $\dot{\xi} = 0$) impose que $\text{div } \vec{E} - \rho$ soit une constante du mouvement. Il suffit de prendre $\text{div } \vec{E} - \rho = 0$ à l'instant origine pour que les équations de Maxwell soient toutes satisfaites.

Quantiquement on restreint les états quantiques aux $|\psi\rangle$ tels que $(\text{div } \vec{E} - \rho)|\psi\rangle = 0$. Heisenberg et Pauli vérifient que cette nouvelle méthode conduit aux mêmes perturbations de l'énergie d'un système de charges au 2^o ordre.

2.5.d Equivalence des différentes méthodes :

En novembre 1930 FERMI publia (20) une généralisation relativiste de sa méthode (il prit des électrons de Dirac au lieu des électrons classiques) ainsi qu'une démonstration de l'équivalence de sa théorie et de celle de Dirac (à l'énergie propre coulombienne $\frac{1}{|r_1 - r_2|} = \infty$ près) au moyen de l'élimination des conditions supplémentaires $\nabla \cdot \vec{A} |\psi\rangle = 0$ et $\nabla \cdot \vec{A} |\psi\rangle = 0$. Fermi avait alors eu connaissance des deux papiers de Heisenberg et Pauli : "j'ai noté que récemment W. Heisenberg et W. Pauli ont traité du problème de l'électrodynamique quantique. Mais comme je crois que les méthodes suivies par ces auteurs diffèrent essentiellement des miennes, je ne crois pas inutile de publier mes résultats.*" Nous présentons les démonstrations d'équivalence avec la théorie de Dirac en appendice.

Oppenheimer collabora à la dernière partie du deuxième article de Heisenberg et Pauli : le passage à l'espace des configurations des électrons, puis dans son étude** de l'interaction champ-matière (26) (novembre 1929), il donna aussi une démonstration de la légitimité de la méthode de Dirac, toujours au terme additif infini de l'énergie propre coulombienne près.

Le fait que l'on puisse se ramener à la théorie de Dirac donne non seulement une preuve détournée de l'équivalence des différentes méthodes, mais encore

** En réalité le travail fut commencé en collaboration avec Pauli à Zürich. Mais Oppenheimer était déjà revenu aux Etats-Unis quand il adressa l'article à Phys.Rev..

* Déjà cité plus haut en partie.

le procédé le plus naturel pour traiter les interactions atome-rayonnement. En effet le terme d'interaction coulombienne $\frac{1}{3n} \sum_j \frac{e_i e_j}{|r_i - r_j|}$ présent dans la théorie de Dirac est traité en première approximation sans le couplage avec le champ de radiation pour donner les états stationnaires atomiques, et on prend ces états comme base d'états dans l'étude des perturbations par le rayonnement.

Dans toutes les méthodes présentées, on peut voir que la présence dans le hamiltonien (réel ou effectif) du terme coulombien résulte de l'élimination des degrés de liberté longitudinaux du champ électromagnétique ($\text{div } \vec{A} \neq 0$). Pour Dirac ces degrés de liberté sont éliminés au niveau de la théorie classique par le choix de jauge $\text{div } \vec{A} = 0$, pour Fermi et Heisenberg-Pauli II, par la résolution des conditions supplémentaires qui relie la distribution de charge aux degrés de liberté longitudinaux : $(\text{div } \vec{E} - \rho) |\psi\rangle = 0$. (I.e. $\vec{E} = -\frac{\partial \vec{A}}{\partial t}$)

2.5.e Avantages et inconvénients des différentes méthodes :

Comme nous l'avons dit, la méthode de Dirac, bien qu'elle soit naturellement privilégiée dans l'étude des interactions atome-rayonnement, présente le défaut de traiter dissymétriquement les degrés de liberté transversaux et longitudinaux du champ et elle ne permet pas d'établir une preuve d'invariance relativiste.

A cet égard la méthode de Heisenberg-Pauli I est bien préférable puisqu'elle procède d'un lagrangien invariant relativiste et permet une démonstration de cette invariance pour la théorie quantifiée. Cependant le terme supplémentaire $-\frac{1}{2} E (\partial.A)^2$ complique substantiellement les équations du mouvement et l'expression des modes propres de la cavité utilisée pour "discrétiser" les degrés de liberté, si bien que les auteurs eux-mêmes jugèrent utile de mettre en place une autre méthode. Celle-ci conduit effectivement à des calculs plus simples, mais rend plus tortueuse la preuve d'invariance relativiste puisque le potentiel A_0 , pris nul, joue un rôle particulier.

La méthode de Fermi, elle, conjugue les avantages des deux méthodes précédentes : formulation invariante relativiste de la théorie classique, et simplicité des équations du mouvement. Elle fut adoptée dans la plupart des

premiers cours d'électrodynamique quantique par exemple par Dirac dans "The principles of Quantum Mechanics" * et par Wentzel dans "Einführung in die Quantentheorie der Wellenfelder". Aujourd'hui encore elle reste la plus utilisée (Réf I par ex.) et elle conduit à l'expression de Feynman du propagateur du photon. Entièrement satisfaisante pour la dérivation formelle des règles de Feynman, elle n'est toutefois pas exempte de difficultés : elle nécessite l'introduction d'une "métrique indéfinie" dans l'espace des états (physiques et non physiques). S.T. MA (28) puis BELINFANTE (29) notèrent en 1949 que les états physiques (tels que $(\partial.A) |\psi\rangle = 0$) ne sauraient être normalisables. La difficulté fut résolue par GUPTA (30) et BLEULER (31) dans l'année suivante en introduisant une métrique "indéfinie" (plus exactement non positive) dans l'espace des états et en affaiblissant la condition supplémentaire (on exige seulement que la partie de fréquence positive de $\partial.A$ annule $|\psi\rangle$). Ces travaux tardifs sortent de notre épure.

2.5.f Une méthode générale d'étude des symétries
(ROSENFELD (32) - 1930) :

Aucun des trois traitements mentionnés pour l'invariance de jauge, disons Fermi, HPI, HPII, n'est pleinement satisfaisant pour Rosenfeld. HPI est un "pur artifice analytique", HPII fait jouer un rôle particulier à une composante du potentiel, ce qui rend difficile la preuve d'invariance relativiste, enfin Fermi introduit un lagrangien qui a le défaut de ne pas être invariant de jauge. En fait aucune des trois méthodes ne permet d'exprimer l'invariance de jauge au niveau quantique, et la preuve de leur équivalence est indirecte : elle passe par l'équivalence avec la théorie de Dirac.

Sur une idée de Pauli, Rosenfeld propose une méthode générale qui s'applique à toute symétrie "continue" du lagrangien classique. Les cas particuliers

* - Nous avons vu au § 4 comment l'utilisation conjuguée de la méthode de Fermi, de la représentation d'interaction et du formalisme à temps multiples donne une théorie manifestement covariante relativiste. (Einstein - Fock - Podolsky)

- Dans la troisième édition, le terme d'énergie propre coulombienne infinie, après l'élimination des degrés de liberté longitudinaux n'apparaît pas, grâce à l'utilisation de la " λ - limite".

du groupe de Lorentz-Poincaré, du groupe de jauge et de la relativité générale sont examinés. Nous exposerons les idées de Rosenfeld en appendice et seulement à propos du groupe de jauge.

Dans la discussion précédente nous avons omis à dessein un autre type d'approche qui consiste à ne pas introduire du tout le potentiel A_μ et à travailler directement avec $F_{\mu\nu}$: cette méthode n'a pas eu beaucoup d'incidence historique, car elle est isolée des précédentes et surtout parce qu'elle conduit à un formalisme plutôt lourd en raison de la non localité de l'interaction j_μ^A réexprimée à l'aide de $F_{\mu\nu}$. Néanmoins elle fut adoptée par Pauli dans son célèbre cours publié dans le "Handbuch der Physik" 1933 (22) sur les principes fondamentaux de la mécanique quantique, même dans ce cas des champs couplés. Pour le champ électromagnétique libre elle ne présente pas de difficulté et elle fut utilisée dès 1927 par Jordan et Pauli (133) (Voir § 3), plus tard par Rosenfeld et Solomon (134) qui s'intéressaient alors à l'élimination de l'énergie de zéro (Voir § 6) et pour la première fois dans le problème couplé par Landau et Peierls (85) (Voir § 1), associée à l'espace des configurations pour les photons.

Arrêtant en 1930 l'histoire de l'invariance de Jauge en électrodynamique nous n'avons certainement pas fait le tour du problème. Mais pour l'essentiel les différentes voies possibles étaient déjà connues et les raffinements ultérieurs ne vinrent qu'après le regain d'intérêt qu'apporta la théorie de la renormalisation, donc bien plus tard (Voir les travaux de Gubta et Bleuler déjà cités). En 1930 les théoriciens étaient capables de dériver correctement les règles formelles de quantification du champ électromagnétique et d'aboutir à une interprétation cohérente en terme de photons. D'autres aspects de l'invariance de Jauge, comme les identités de Ward ou la réduction des divergences de l'électrodynamique à la divergence logarithmique constatée par Weisskopf, Dirac, Heisenberg... (Voir chap. suivant) et alors presque miraculeuse, ne furent reconnus que dans les années cinquante.

2.6. L'ENERGIE DE ZERO DES OSCILLATEURS :

Un champ peut être comparé à un ensemble continu d'oscillateurs. La quantification de chaque oscillateur introduit une somme divergente de termes $\frac{1}{2} h\nu$ dans l'expression du hamiltonien. Cette divergence est considérée aujourd'hui comme bénigne. Born Heisenberg et Jordan ⁽¹²⁾, quand ils la rencontrèrent pour la première fois, la soustrayèrent du hamiltonien comme on le fait encore aujourd'hui. Contrairement au cas du cristal solide où les effets liés à l'existence des demi-quanta sont mesurables (par exemple par l'étude de la chaleur latente de fusion) aucun effet physique n'est associé à l'énergie de zéro du champ électromagnétique *. Mais ce n'était pas évident au début de la théorie des champs. Le calcul de B.H.J. des fluctuations du corps noir est présenté d'une manière qui peut laisser croire que l'énergie de zéro joue un rôle, et nous verrons au chap. 3 qu'on ne fut pas immédiatement certain que la divergence des énergies propres soit étrangère à celle de l'énergie de zéro. Dans les deux cas il fut établi que les résultats étaient finalement les mêmes dans une théorie dépourvue d'énergie de zéro. MÖGLICH ⁽¹³⁵⁾ (1929) reprit le calcul de B.H.J. après introduction du produit normal (créateurs à gauche des annihilateurs) dans le hamiltonien. ROSENFELD et SOLOMON ⁽¹³⁴⁾ souhaitaient trouver une méthode moins artificielle: ils introduisent dans leur article de 1931 les champs \vec{F} et \vec{F}^* combinaisons de \vec{E} et \vec{B} tels que $H = \frac{1}{2} \int \vec{F} \cdot \vec{F}^* d^3x$ conduise aux bonnes équations de Maxwell, et ne présente pas d'énergie de zéro (\vec{F} n'est rien d'autre que la partie de fréquence négative de \vec{E}) et ils retrouvent la formule d'Einstein pour les fluctuations du corps noir. En réalité pour être compatibles avec $\text{div } \vec{F} = 0$ (cas du champ libre) les relations de commutation de \vec{F} et \vec{F}^* doivent être plus complexes que celles données par Rosenfeld et Solomon et alors le procédé perd tout son naturel. L'introduction de quantités analogues dans le problème couplé (Landau-

* Si nous anticipons jusqu'en 1948, même l'effet Casimir peut s'expliquer par analogie avec les forces de van der Waals, sans utiliser l'énergie de zéro.

Peierls (85) *, PAULI (122) n'est pas plus avantageuse. Comme PAULI (122) (p. 256) le fait remarquer, l'apparition de l'énergie de zéro est une des manifestations des limites de la méthode de correspondance dans son application à la quantification de systèmes possédant une infinité de degrés de liberté. Il est nécessaire et inévitable de modifier les règles de correspondance pour obtenir une théorie cohérente.

* L'absence d'énergie de zéro n'était pas le but de Landau et Peierls (Voir chap. 2 § 4) et ils ne l'établirent pas.

APPENDICE : Premiers traitements de l'ambiguïté de jauge en électrodynamique quantique.

1. DIRAC Le hamiltonien total s'écrit $H = H_m + H_c + H_r + H_I$
- H_m : particules chargées libres
 - H_c : Interaction coulombienne entre ces particules
 - H_r : Hamiltonien du champ de radiation \vec{A} ($\text{div } \vec{A} = 0$) libre
 - H_I : Hamiltonien d'interaction de ce champ et de la matière chargée :
- $$H_I = - \vec{j} \cdot \vec{A}$$

La justification de ce traitement n'est pas donnée explicitement par Dirac (13).
 Les équations de Maxwell dans la jauge de Coulomb ($\text{div } \vec{A} = 0$) s'écrivent :

$$\begin{cases} \Delta \varphi + \rho = 0 \\ -\Delta \vec{A} + \frac{\partial^2 \vec{A}}{\partial t^2} + \frac{\partial}{\partial t} \vec{\nabla} \varphi = 0 \end{cases}$$

Le potentiel φ , pris nul à l'infini, peut être éliminé suivant $\varphi = \frac{1}{4\pi} \rho * \frac{1}{r}$ et alors les équations du mouvement résultent du hamiltonien $H = H_m + \frac{1}{2} \int [(\eta_0 + \vec{A})^2 + (\frac{\partial \vec{A}}{\partial t})^2] d^3r + H_c - \vec{j} \cdot \vec{A}$ en prenant pour variables dynamiques les coordonnées des particules chargées et les degrés de liberté transverses de \vec{A} . La quantification est obtenue en écrivant les relations de commutation entre ces variables et leurs conjuguées canoniques.

2. HEISENBERG-PAULI I

Le Lagrangien du champ électromagnétique est $L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$
 La première composante de l' "impulsion" est $\pi_c = \frac{\partial L}{\partial \dot{A}_0} = F^{00} = 0$ rendant impossible la quantification canonique ($[A_0, \pi_c] = i\delta_3$).
 Heisenberg et Pauli rajoutent un terme $-\frac{1}{2} \epsilon (\partial_\mu A^\mu)^2$ dans L. Alors $\pi_c = -\epsilon (\partial_\mu A^\mu) \neq 0$ et la quantification devient possible. La théorie physique est obtenue en passant à la limite $\epsilon \rightarrow 0$ dans les résultats finaux.

3. HEISENBERG-PAULI II

On se place dans la Jauge $A_0 = 0$. Alors les \vec{A} ($\vec{\nabla} \cdot \vec{A} = 0$) admettent des variables canoniques conjuguées. Le hamiltonien total pour un ensemble de particules de Dirac i couplées au champ électromagnétique s'écrit

$$H = \int d^3r \left[\frac{1}{2} (\vec{E}^2 + \vec{B}^2) - \varphi \operatorname{div} \vec{E} + \rho \varphi \right] + \sum_i (\vec{\alpha}_i \cdot \vec{p}_i + \beta_i m_i) - \sum_i e_i \vec{\alpha}_i \cdot \vec{A}(\vec{r}_i)$$

Les termes contenant φ sont nuls puisque $\varphi \equiv A_0 = 0$. Les équations canoniques redonnent $\partial_\mu F^{\mu\alpha} = j^\alpha$ ($\alpha = 1, 2, 3$) mais pas $\partial_\mu F^{\mu 0} = j^0$ (ou $\operatorname{div} \vec{E} - \rho = 0$) puisque cette équation est associée à la variable canonique A_0 . Néanmoins $A_0 = 0$ ne fixe pas complètement la Jauge. Les transformations $\vec{A} \rightarrow \vec{A} - \vec{\nabla} \xi(\vec{r})$, ξ ne dépendant pas du temps, sont encore permises. A cette symétrie correspond (d'après le théorème de Noether) une constante du mouvement : $\operatorname{div} \vec{E} - \rho$. On peut aussi vérifier que

$$\partial_\mu F^{\mu\alpha} = j^\alpha \Leftrightarrow \operatorname{rot} \vec{B} - \frac{\partial \vec{E}}{\partial t} = \vec{j} \Rightarrow -\frac{\partial}{\partial t} \operatorname{div} \vec{E} = \operatorname{div} \vec{j} \Rightarrow \frac{\partial}{\partial t} (\operatorname{div} \vec{E} - \rho) = -(\operatorname{div} \vec{j} + \frac{\partial \rho}{\partial t}) = 0$$

On retrouvera donc les équations de Maxwell habituelles à condition de prendre la condition supplémentaire $\operatorname{div} \vec{E} - \rho = 0$, compatible avec l'évolution temporelle.

Quantiquement on ne peut imposer que l'opérateur $\operatorname{div} \vec{E} - \rho$ soit nul.

(en effet, dans le cas libre $\rho = 0$, $\operatorname{div} \vec{E} = -\operatorname{div} \vec{\pi}$ ($\vec{\pi}$ impulsion associée à \vec{A}) et $[\operatorname{div} \vec{E}(\vec{r}), \vec{A}(\vec{r}')] = \vec{\nabla} \delta(\vec{r} - \vec{r}') \neq 0$). Mais on ne peut restreindre l'espace des états aux $|\psi\rangle$ tels que $(\operatorname{div} \vec{E} - \rho)|\psi\rangle = 0$ et ce sous-espace est invariant lors de l'évolution.

Explicitons maintenant cette condition. On a $\vec{E} = -\dot{\vec{A}} = -\vec{\pi}$ avec

$$[A^i(\vec{r}), \pi^j(\vec{r}')] = i\delta_{ij} \delta(\vec{r} - \vec{r}'). \text{ Posant } \chi(\vec{r}) = \int \tilde{\chi}(\vec{k}) e^{i\vec{k} \cdot \vec{r}} d^3k \text{ il vient } (-i\vec{k} \cdot \vec{\pi} - \rho)|\psi\rangle = 0 \text{ et } [\tilde{A}_k^a, \tilde{\pi}_k^a] = \frac{i}{(2\pi)^3} \delta_3(\vec{k} - \vec{k}')$$

Distinguons les parties longitudinale \tilde{A}_\parallel ($\parallel \vec{k}$) et transversale ($\perp \vec{k}$) \tilde{A}_\perp

$$\text{alors } (i\vec{k} \cdot \vec{\pi}_\parallel - \rho)|\psi\rangle = 0 \text{ ou encore comme } \tilde{\pi}_\parallel = -\frac{i}{(2\pi)^3} \frac{\delta}{\delta \tilde{A}_\parallel^*} \text{ dans la}$$

$$\text{représentation } \psi(A_\parallel, \vec{A}_\perp, \vec{\pi}_\perp), \left(-\frac{\delta}{\delta \tilde{A}_\parallel^*} - (2\pi)^3 \frac{\delta}{\delta k} \right) \psi = 0 \text{ ce qui}$$

impose la forme de ψ : $\psi = \int \frac{\delta}{\delta k} \tilde{A}_\parallel^* d^3k \varphi(\vec{A}_\perp, \vec{\pi}_\perp)$. Il en résulte

$$\int \frac{1}{2} E_{\mu}^2 d^3x \psi = (2\pi)^3 \frac{1}{2} \int \tilde{\pi}_{\mu}^* \tilde{\pi}_{\mu} d^3k \psi = \frac{(2\pi)^3}{2} \int \frac{\tilde{\rho} \tilde{\rho}^*}{k^2} d^3k \psi = \int \frac{\rho(\vec{x}) \rho(\vec{x}')}{|\vec{x} - \vec{x}'|} d^3x d^3x' \psi$$

et $\vec{P}_i \psi = -i \frac{\partial}{\partial x_i} \psi = i (2\pi)^3 \int \frac{\tilde{A}_{\mu}^*}{k} \frac{\partial}{\partial x_i} \tilde{\rho} d^3k \psi + \exp[-(2\pi)^3 \int \frac{\tilde{\rho} \tilde{A}_{\mu}^*}{k} d^3k] \frac{\partial}{\partial x_i} \psi$

comme $\rho = \sum_i e_i \delta(\vec{x} - \vec{x}_i)$ et $\tilde{\rho} = \frac{1}{(2\pi)^3} \sum_i e_i e^{-i\vec{k} \cdot \vec{x}_i}$

le premier terme devient $\int \tilde{A}_{\mu}^* \frac{\vec{k}}{k} e_i e^{-i\vec{k} \cdot \vec{x}_i} d^3k \psi = e_i \vec{A}_{\mu}(\vec{R}_i) \psi$

$H \psi = i\hbar \frac{\partial \psi}{\partial t}$ est donc équivalente à $H_{eff} \psi = i\hbar \frac{\partial \psi}{\partial t}$

avec $H_{eff} = \frac{1}{2} \int d^3x [E_{\perp}^2 + B^2] + \sum_i (\vec{\alpha}_i \cdot \vec{p}_i + \beta_i m_i) - \sum_i e_i \vec{\alpha}_i \vec{A}_i(\vec{R}_i) + H_C$

et $H_C = \frac{1}{8\pi} \int \frac{\rho(\vec{x}') \rho(\vec{x})}{|\vec{x} - \vec{x}'|} d^3x d^3x'$

Remarque : Comme Oppenheimer, nous avons fait le calcul dans le cas de la première quantification des électrons (espace de configuration des électrons) ; dans la théorie des "ondes quantifiées" il est aisé de démontrer un résultat identique avec $\rho = \psi^* \psi$ (en remarquant que

$[\psi^* \vec{\alpha}_i \cdot \vec{p}_i \psi(\vec{x}), \tilde{\rho}(\vec{k})] = -\psi^* \vec{\alpha}_i \cdot \vec{k} \psi(\vec{x}) e^{-i\vec{k} \cdot \vec{x}}$ est analogue à

$[\vec{\alpha}_i \cdot \vec{p}_i, \tilde{\rho}(\vec{k})] = -\vec{\alpha}_i \cdot \vec{k} e^{-i\vec{k} \cdot \vec{x}}$ en première quantification).

4- FERMI :

Le Lagrangien du champ électromagnétique est $L_{e.m.} = -\frac{1}{4} F^2 - \frac{1}{2} (\partial \cdot A)^2$

Les équations du mouvement qui en résultent : $\square A_{\mu} = 0$ ne sont équivalentes aux équations de Maxwell déduites du lagrangien $-\frac{1}{4} F^2$ que si $\partial \cdot A = 0$

Comme $\square A_{\mu} = 0$ on a nécessairement $\square (\partial \cdot A) = 0$, donc si à l'instant t , $\partial \cdot A = 0$ et $\vec{\partial} \cdot A = 0$, alors $\partial \cdot A = 0$ à tout instant ultérieur et les équations de Maxwell sont vérifiées.

Avec ce nouveau lagrangien les impulsion généralisées sont

$$\pi^\mu = -F^{\nu\mu} - g(\partial_\nu A) \quad . \text{ On n'a plus l'ennui } \pi^0 = 0$$

Le hamiltonien total est :

$$H = \int \left[\frac{1}{2} (E^2 + B^2) - \varphi \operatorname{div} \vec{E} + \frac{1}{2} (\operatorname{div} \vec{A})^2 - \frac{1}{2} \dot{\varphi}^2 \right] d^3x + \sum_i (\alpha_i \vec{p}_i + \beta_i m_i) + \sum_i e_i (\varphi(\vec{r}_i) - \vec{A}(\vec{r}_i))$$

$\partial_\nu A = 0$ et $\dot{\vec{A}} = 0$ s'écrivent encore $\pi_0 = 0$ et $\rho + \operatorname{div} \vec{\pi} = 0$ car $\partial_\nu A = -\pi_0$

$$\text{et } \dot{\vec{A}} = \partial_0 (\partial_0 A + \vec{\partial} \cdot \vec{A}) = \partial_0^2 A + \operatorname{div} \partial_0 \vec{A}$$

$$\left. \begin{aligned} \partial_0^2 A + \vec{\partial} \cdot \partial_0 \vec{A} &= \rho \\ \vec{\pi} &= \partial_0 \vec{A} + \vec{\partial} A_0 \end{aligned} \right\} \Rightarrow \dot{\vec{A}} = \rho + \vec{\partial}^2 A_0 + \operatorname{div} (\vec{\pi} - \vec{\partial} A_0) \Rightarrow \dot{\vec{A}} = \rho + \operatorname{div} \vec{\pi}$$

Comme précédemment au niveau quantique on impose $\left. \begin{aligned} \pi_0 |\psi\rangle &= 0 \\ (\rho + \operatorname{div} \vec{\pi}) |\psi\rangle &= 0 \end{aligned} \right\}$

En représentation $\psi(A_0, A_{\parallel}, \vec{A}_{\perp}, \vec{r}_i)$ il en résulte que ψ ne dépend pas

de A_0 et a la forme donnée au 3° : $\psi = e^{-i(\pi)^3 \int \frac{\vec{p}}{k} \vec{A}_{\parallel}^* d^1k} \varphi(\vec{A}_{\perp}, \vec{r}_i)$

De plus comme $\pi_0 = -\partial_\nu A = -(\dot{\varphi} + \operatorname{div} \vec{A})$ et $[\dot{\varphi}, \operatorname{div} \vec{A}] = [-\pi_0, \operatorname{div} \vec{A}] = 0$

$$\text{on a } \left[\frac{1}{2} (\operatorname{div} \vec{A})^2 - \frac{1}{2} \dot{\varphi}^2 \right] |\psi\rangle = 0$$

Enfin comme $(\rho - \operatorname{div} \vec{E}) |\psi\rangle = 0$, $\int \varphi \operatorname{div} \vec{E} |\psi\rangle d^3x = -\int \rho \varphi |\psi\rangle d^3x = \sum_i e_i \varphi(\vec{r}_i) |\psi\rangle$

On a donc à nouveau $H_{\text{eff}} \varphi = i\hbar \frac{\partial \varphi}{\partial t}$ avec la même expression de H_{eff} que précédemment.

5- ROSENFELD :

— L'invariance du lagrangien par un groupe continu implique une relation entre les impulsions généralisées. Pour une transformation de jauge infinitésimale quelconque $A_\mu \rightarrow A_\mu + \partial_\mu \xi$, $\delta L = \frac{\partial L}{\partial(\partial_\nu A_\mu)} \partial_\nu \partial_\mu \xi = 0$ ce qui impose

que $\pi^{\mu\nu} = \frac{\partial L}{\partial(\partial_\nu A_\mu)}$ soit antisymétrique. Or $\pi^\mu = \frac{\partial L}{\partial(\partial_0 A_\mu)}$ sont les impulsions généralisées.

$$\pi^{\mu\nu} = -\pi^{\nu\mu} \Rightarrow \pi^{00} = 0 \Leftrightarrow \pi^0 = 0$$

— D'après cette relation les équations $\frac{\partial L}{\partial \dot{Q}_\alpha}(Q, \dot{Q}) = P_\alpha$

ne sont pas indépendantes et leur solution $\dot{Q}(Q, P)$ n'est pas unique.

Ici
$$\begin{cases} \vec{\nabla} A_0 + \frac{\partial \vec{A}}{\partial t} = \vec{\pi} \\ \partial_0 A_0 - \partial_0 A_0 = \pi_0 = 0 \end{cases}$$
 ont pour solution
$$\begin{cases} \dot{A}_0 = d \quad \text{avec } d(x) \text{ quelconque} \\ \vec{A} = \vec{\pi} - \vec{\nabla} A_0 \end{cases}$$

Le hamiltonien $\sum_\alpha P_\alpha \dot{Q}_\alpha - L$ construit avec l'une de ces solutions a la forme $\pi_0 \dot{d} + H_0$ où H_0 est le hamiltonien usuel (avec $\pi_0 = 0$).

Les équations du mouvement de Hamilton s'identifient aux équations de Maxwell à condition d'imposer les conditions supplémentaires $\pi_0 = 0$.

— La quantification canonique assortie de la condition supplémentaire

$\pi_0 |\psi\rangle = 0$ est un schéma invariant de jauge : en effet à toute transformation de Jauge $A_\mu \rightarrow A'_\mu$ correspond une transformation unitaire S indépendante du temps telle que $A'_\mu = S^{-1} A_\mu S$ et $\pi'_\mu = S^{-1} \pi_\mu S$, donc les équations du mouvement et les relations canoniques sont invariantes de jauge et la condition

$\pi_0 |\psi\rangle = 0$ l'est aussi.

Rosenfeld note que la deuxième méthode de Heisenberg et Pauli correspond au choix $\lambda = 0$ et que Fermi et HPI détruisent l'invariance du lagrangien initial. Mais bien que le hamiltonien ne soit pas le même, nous faisons remarquer que le système des équations du mouvement et des conditions supplémentaires de Fermi (de HPI) est équivalent au système analogue de Rosenfeld à condition de prendre $\lambda = -\text{div } \vec{A}$ (resp. $\lambda = -\epsilon \text{div } \vec{A}$) cette remarque permet de déduire immédiatement de l'étude de Rosenfeld l'équivalence Fermi \equiv HPI \equiv HP11.

CHAPITRE 3

LA CRISE DES DIVERGENCES ULTRAVIOLETES

Rosenfeld note que la deuxième méthode de Heisenberg et Pauli correspond au choix $\lambda = 0$ et que Fermi et HPI détruisent l'invariance du lagrangien initial. Mais bien que le hamiltonien ne soit pas le même, nous faisons remarquer que le système des équations du mouvement et des conditions supplémentaires de Fermi (de HPI) est équivalent au système analogue de Rosenfeld à condition de prendre $\lambda = -\text{div } \vec{A}$ (resp. $\lambda = -\epsilon \text{div } \vec{A}$) cette remarque permet de déduire immédiatement de l'étude de Rosenfeld l'équivalence Fermi \equiv HPI \equiv HPIL.

CHAPITRE 3

LA CRISE DES DIVERGENCES ULTRAVIOLETTES

Certains obstacles à la quantification des champs classiques furent - nous l'avons vu - aisément franchis. Dès les premiers développements, plusieurs solutions furent trouvées aux difficultés liées à l'invariance de jauge de l'électrodynamique. Une autre difficulté apparaissant elle aussi dès la première étape de la quantification (pour les champs libres), s'est avérée superficielle : le problème de l'énergie de zéro infinie des oscillateurs. Il en va tout autrement de la classe de problèmes désignés aujourd'hui sous le terme générique de divergences ultra-violettes. Ces divergences apparaissent seulement lorsqu'on calcule les corrections qui suivent l'ordre non nul le plus bas (dans un développement en puissance de la charge e), comme l'énergie propre de l'électron à l'ordre e^2 , ce qui explique qu'elles ne furent pas immédiatement signalées par Dirac, dans son article sur l'émission et l'absorption du rayonnement ⁽²⁾. Néanmoins EHRENFEST les devina dès cette dernière parution ; Pauli écrivit ⁽¹⁾ en 1933 après le suicide d'Ehrenfest : *"Dirac venait juste de rendre public son premier travail sur la théorie du rayonnement, dans lequel il quantifia le champ électromagnétique. Alors Ehrenfest fit aussitôt remarquer que cette théorie devait contenir la difficulté de l'énergie propre infinie de l'électron, puisqu'elle faisait intervenir de manière essentielle le potentiel du champ au lieu même de l'électron et qu'elle devait être prise comme une réinterprétation conforme au principe de correspondance, de la théorie classique d'un électron ponctuel. Une difficulté, qui s'avéra extrêmement pénible et déroutante lors de l'élaboration ultérieure de l'électrodynamique quantique et qui reste encore sans solution"*

Effectivement le problème de l'énergie propre de l'électron n'était pas nouveau et il fut une difficulté importante lors des premiers développements de la théorie classique de l'électron. Avant d'aborder le cas quantique, il nous paraît utile de résumer les derniers

tourments des théoriciens de l'électron classique, et d'établir ainsi le premier terme d'une comparaison toujours fructueuse entre les deux électrodynamiques classique et quantique : non seulement ces deux théories présentent le même type de divergence, mais encore les méthodes pour les traiter, comme on le sait maintenant, (voir Rohrlich réf 5) présentent certaines analogies : renormalisation infinie de la masse, introduction de conditions asymptotiques.....

3.1 L'AUTO-INTERACTION EN ELECTRODYNAMIQUE CLASSIQUE :

Les difficultés liées à l'interaction des particules chargées avec leur propre champ (auto-interaction) ont donné à l'histoire de l'électrodynamique classique un aspect complexe et presque désordonné. Son étude détaillée sortirait du cadre de cet exposé. Nous nous contenterons ici d'extraire ce qui touche de plus près le problème des divergences retrouvées dans la théorie quantique, et renvoyons le lecteur aux travaux de Whittaker (4), Rohrlich (5) et Pais (6) pour plus d'informations.

Avant même sa découverte de l'électron lors de l'étude des rayons cathodiques (1897), J.J. THOMSON calcula en 1891 (7) la masse électromagnétique d'une particule chargée. La théorie de Maxwell (1864) était alors à sa disposition, et l'on commençait à croire en l'existence d'atomes d'électricité, pendents des atomes de matière. En 1842 STOKES avait montré que l'effet du fluide environnant (incompressible et irrotationnel) lors du mouvement d'un solide se traduit par l'introduction d'une masse effective somme de la masse du solide et d'une fraction de la masse du liquide déplacé *. De même une particule chargée sphérique de rayon R en mouvement uniforme de vitesse v est accompagnée par un champ électromagnétique d'énergie cinétique $T_{e.m.} = f \frac{e^2}{4\pi R C^2} \frac{v^2}{2}$ (f facteur de forme dépendant de la distribution de charge à l'intérieur de la sphère). La première remarque est que cette expression diverge lorsque $R \rightarrow 0$, ce qui semble imposer une taille finie pour l'électron. Si d'après J.J. Thomson, l'électron est un atome d'électricité, sa masse doit être d'origine purement électromagnétique, elle est donnée par $m_{em} = f \frac{e^2}{4\pi R C^2}$, d'où la valeur du rayon $R = f \frac{e^2}{4\pi m C^2}$.

* voir par exemple Landau et Lifchitz Tome 6 § 11.

Le même modèle de l'électron, sphère rigide avec distribution de charge de symétrie sphérique, fut à nouveau étudié par ABRAHAM ⁽⁸⁾ et LORENTZ ⁽⁹⁾, dans le cadre de l'électrodynamique des particules de Lorentz : dans cette théorie les particules n'ont plus d'interactions directes (contrairement à celle de Clausius par exemple), mais elles interagissent via le champ électromagnétique. Maxwell avait trouvé comment calculer le champ en fonction des charges et des courants ; la formule de Lorentz $\vec{f} = \rho(\vec{E} + \vec{v} \wedge \vec{B})$ exprime l'action du champ sur les charges. Lorentz espérait ainsi construire une théorie microscopique de l'électrodynamique et de l'optique macroscopique . Le modèle de l'électron inclus dans cette théorie donnait, dans sa version de 1904 tenant compte de la contraction de Lorentz . et dans la limite $\frac{v}{c} \ll 1$, la valeur de Thomson de l'énergie cinétique :

$$T_{em} = \frac{1}{2} m_{em} v^2 \quad \text{et pour l'impulsion} \quad \vec{P}_{em} = \frac{4}{3} m_{em} \vec{v}$$

Bien sûr ce résultat n'est pas compatible avec l'invariance de Lorentz et contredit les expériences de KAUFMANN ⁽¹⁰⁾ (1906) sur la déviation d'un électron dans un champ électrique et un champ magnétique parallèles.

A ce stade POINCARÉ ⁽¹¹⁾ (1905) fit remarquer qu'un tel modèle donnait un électron instable puisque ses parties se repoussaient ; de plus, la pression (due à des forces de cohésion de nature différente) qu'il fallait rajouter pour assurer la stabilité compensait le facteur $\frac{4}{3}$ pour donner 1 .

Aujourd'hui nous savons que le facteur $\frac{4}{3}$ de Lorentz résultait d'une définition fautive de l'impulsion du champ, et donc il n'y avait pas de place pour les forces de cohésion de Poincaré. Le résultat correct $\vec{P}_{em} = m_{em} \vec{v}$ fut établi indépendamment d'abord par Fermi en 1922 puis par W. Wilson en 1939, par Kwal en 1949 et enfin par Rohrlich en 1960.

Même sans que cette critique fût connue, l'électron de Poincaré ne satisfait pas Lorentz, car il était instable par rapport aux déformations. De plus il faisait intervenir des forces d'origine extra-électromagnétiques. La découverte du spin en 1926 par UHLENBECK et GOUDSMIT ⁽¹²⁾ vint encore obscurcir la situation et la dernière communication de Lorentz à la conférence de Côme de 1927, peu avant sa mort, apporta, à l'issue de calculs complexes, plus de questions que de réponses.

La théorie de l'électron posait alors un dilemme : on avait le choix entre d'une part une théorie de rayon fini nécessitant des forces externes de cohésion et donnant des réactions de rayonnement dépendant de la structure (ce qui contredit l'élémentarité de la particule), et d'autre part une théorie de rayon nul avec une énergie propre infinie. On comprend que les espoirs se soient tournés vers la théorie quantique. Signalons toutefois que l'opinion encore aujourd'hui répandue, selon laquelle il n'existerait pas d'électrodynamique classique cohérente des particules ponctuelles est erronée : les travaux de Fokker (1929), Wentzel (1933) Dirac (1938) Gupta (1951) et Rohrlich (1960), sur lesquels nous reviendrons au Chapitre 4, ont prouvé le contraire.

3.2. MISE EN EVIDENCE DE LA DIVERGENCE DE L'ENERGIE PROPRE

"Bien des difficultés de la théorie classique de l'électron avaient disparu dans sa version quantique ; pour l'énergie propre, comme sa divergence était liée à la nature ponctuelle de l'électron, et comme par ailleurs la notion de position précise d'une particule était altérée par le principe d'incertitude, je pensais qu'il y aurait une amélioration, mais il n'en fut rien". Ainsi témoigne R. Peierls de la déception ressentie après les premiers calculs d'énergie propre de l'électron. *

HEISENBERG et PAULI (13,14) mentionnèrent une *"interaction infinie de l'électron avec lui-même, qui dans beaucoup de cas rendra impossible l'application de la théorie"* - à la suite de leur calcul de l'énergie au 2ème ordre d'un ensemble d'électrons. Les auteurs retenaient seulement les termes permettant de retrouver l'interaction coulombienne entre les électrons (qui n'est pas présente dans le hamiltonien de départ), et seul l'infini en $\frac{1}{|\vec{r}-\vec{r}'|}$ résultant de l'interaction coulombienne d'un électron avec lui-même était mis en évidence. Cet infini était relativement bénin en ce qu'il s'éliminait dans les calculs des différences d'énergies, qui elles seules sont observables. Sa non covariance relativiste (ce qui est purement coulombien dans un référentiel ne l'est pas dans un autre) était plus problématique mais Heisenberg et Pauli espéraient encore que cet infini pourrait disparaître en ordonnant convenablement les opérateurs comme chez Jordan et Klein (voir chap. 2 § 1), *"Je veux encore m'entretenir avec toi*

* Entretien avec Peierls.

de cette question toujours obscure : est-ce que cette énergie propre peut être éliminée par une simple permutation de facteurs, comme dans ton travail avec Jordan ? " écrit Pauli à Klein ([218]⁽³⁸⁾) le 16 Mars 1929 . La poursuite du travail sur l'électrodynamique quantique par Pauli et Oppenheimer à Zurich (Heisenberg était parti en voyage aux U.S.A.) apporta une réponse négative à cette question. L'article d'Oppenheimer ne parut que le 12 Novembre 1929, mais ses principaux résultats et tout le sérieux des divergences ultraviolettes était connus quelques mois plus tôt. Pauli écrivit à Bohr le 17 Juillet 1929 ([231]⁽³⁸⁾) : "Très satisfait de toute la théorie d'Heisenberg et moi-même, non je ne le suis pas (bien que je crois qu'elle présente certains traits d'une théorie future exacte). En particulier, l'énergie propre de l'électron est une difficulté bien plus grande que Heisenberg ne l'avait pensé au début. Aussi les nouveaux résultats auxquels conduit notre théorie sont-ils très suspects, et le risque est grand que toute l'affaire ne perde peu à peu le contact avec la physique et ne dégénère en pure mathématique". A la fin du second article⁽¹⁴⁾, le travail d'OPPENHEIMER⁽¹⁵⁾ était annoncé ainsi que ses conclusions sur l'énergie propre .

3.2a Le calcul d'Oppenheimer du déplacement des niveaux d'un atome * (nov. 1929).

Lors de son séjour à Zurich, OPPENHEIMER avait pu prendre connaissance des derniers travaux de Heisenberg et Pauli avant leur publication et avait même participé à l'élaboration de la dernière partie du II où les auteurs passent de l'espace des nombres d'occupation à l'espace des configurations (des positions des différents électrons) pour les électrons .

* Dans le second article de Wigner et Weisskopf sur la largeur naturelle des raies spectrales (voir chap. 1 § 4), le déplacement divergent des niveaux n'est pas calculé (de toute façon il serait fini dans l'approximation à deux niveaux où ils se placent), mais les auteurs sont conscients de la difficulté : "nous soulignerons encore que les "transitions anormales" [transitions virtuelles faisant intervenir des niveaux autres que ceux de la transition réelle étudiée] sont ici aussi négligées ; leur considération ne serait nullement simple, au contraire elle serait selon toute vraisemblance impossible, car le système d'équations de Dirac sans approximation diverge à cause du rayon infiniment petit de l'électron". (Zeit. Phys. 65 p. 21 (1930)).

Nous avons montré au chapitre 2 comment Oppenheimer put, en éliminant dans la jauge de Coulomb les degrés de liberté longitudinaux du champ e.m. se ramener à la théorie du rayonnement de Dirac ⁽²⁾, aux termes d'énergie propre électrostatique près. L'équation de Schrödinger effective est $(H_D + H_R + H_I + H_C) |\psi\rangle = i\hbar \frac{\partial |\psi\rangle}{\partial t}$ où $|\psi\rangle$ ne fait plus intervenir que les modes transverses du champ. $H_D = \sum_{i=1}^N (\vec{\alpha}_i \cdot \vec{p}_i + \beta_i m_i)$ est le hamiltonien de Dirac des particules chargées i.

$$H_R = \frac{1}{2} \int d^3x (\vec{E}^2 + \vec{B}^2) \text{ est celui du rayonnement}$$

$$H_I = \sum_{i=1}^N -e_i \vec{\alpha}_i \cdot \vec{A}(\vec{r}_i) \text{ celui d'interaction}$$

enfin $H_C = \sum_{i \neq j} \frac{e_i e_j}{4\pi r_{ij}}$ est le terme coulombien.

Pour un atome, le hamiltonien de départ est $H_0 = H_D + H_C + H_R$. L'effet de H_I au 2ème ordre, $E_n^{(2)}$, sur les niveaux atomiques E_n correspondant aux états propres de H_0 $|\psi_n; 0\rangle$ s'obtient d'après la théorie de la diffusion de Dirac ⁽¹⁶⁾ par la formule *

$$E_m^{(2)} = \text{Partie Principale} \sum_{\vec{k}, \lambda; n} \frac{|\langle \psi_m; 0 | H_I | \psi_n; \vec{k}, \lambda \rangle|^2}{E_m - E_n - \hbar k} \text{ correspondant}$$

à l'émission et à la réabsorption de photons virtuels transverses d'impulsion \vec{k} et de polarisation \vec{e}_λ .

Les termes de $\vec{A}(\vec{r})$ qui contribuent à l'élément de matrice sont de la forme $\frac{1}{\sqrt{k}} \sum_{\lambda} a_{k\lambda} \vec{e}_\lambda e^{ik \cdot \vec{r}}$; d'où

$$E_m^{(2)} \propto -\frac{1}{2} \int k dk d\Omega \sum_{\lambda, m} \frac{|\langle \psi_m | \sum_i \vec{\alpha}_i \cdot \vec{e}_\lambda e^{i\vec{k} \cdot \vec{r}_i} | \psi_m \rangle|^2}{\hbar k + E_m - E_n}$$

Nous allons maintenant étudier le comportement aux grands k de l'intégrale (Oppenheimer donne le résultat (divergence logarithmique des termes spectraux) sans démonstration). Pour $m > n$ (Oppenheimer fait l'hypothèse supplémentaire qu'il existe un état fondamental, malgré l'existence des solutions d'énergie négative de l'équation de Dirac) $0 < \frac{1}{\hbar k + E_m - E_n} < \frac{1}{\hbar k}$

* voir par ex. Réf II p. 1332

La série \sum_m^k est donc majorée par $\sum_{i \neq j} \langle \psi_n | \vec{\alpha}_i \cdot \vec{\epsilon}_\lambda \vec{\alpha}_j \cdot \vec{\epsilon}_\lambda^* e^{ik(\vec{R}_i - \vec{R}_j)} | \psi_n \rangle$
 (car $\sum_m |\psi_m\rangle \langle \psi_m| = 1$). Les termes $i \neq j$ (de retard) sont d'ordre plus élevé en $\frac{1}{k}$ quand $k \rightarrow \infty$.

Nous examinons seulement l'effet des termes $i=j$:
 $\langle \psi_n | (\vec{\alpha}_i \cdot \vec{\epsilon}_\lambda) (\vec{\alpha}_i \cdot \vec{\epsilon}_\lambda^*) | \psi_n \rangle$ ou encore comme $(\alpha_1^4)^2 = (\alpha_1^2)^2 = (\alpha_1^3)^2 = 1$, le majorant est $\langle \psi_n | \psi_n \rangle = 1$.

Le comportement aux grands k de la série majorée s'obtient en prenant la limite $k \rightarrow \infty$; elle tend donc vers le majorant que nous venons de calculer terme à terme d'où $E_n^{(2)} \int dk$ diverge linéairement. Si l'on considère maintenant l'effet de la perturbation sur un terme spectral : $E_n^{(2)} - E_0^{(2)}$ les divergences linéaires s'éliminent car toutes deux sont proportionnelles à $\langle \psi_n | \psi_n \rangle = \langle \psi_0 | \psi_0 \rangle = 1$ et il reste une divergence logarithmique. C'est la conclusion d'Oppenheimer.

Mais nous devons remarquer que le calcul précédent, sans la théorie du positron, n'est pas un calcul relativiste correct, et il serait donc plus cohérent à ce stade, de prendre $\frac{p_1^2}{2m}$ pour hamiltonien des particules et $-e_i \vec{v}_i \cdot \vec{A}(\vec{r}_i)$ comme hamiltonien d'interaction; dans ce cas (qui correspond au calcul historique de Bethe du déplacement de Lamb ⁽¹⁷⁾), il faut substituer \vec{v}_i à $\vec{\alpha}_i$ dans le raisonnement précédent, et un terme spectral diverge linéairement en $(\langle n | v^2 | n \rangle - \langle 0 | v^2 | 0 \rangle) \int dk$ et une telle divergence ne peut être éliminée que par renormalisation de masse, comme nous le verrons au chap 8 § 6.

Le calcul relativiste de l'énergie propre dans la théorie du positron, bien qu'il ait été traité par Weisskopf pour l'électron libre (voir au § suivant), n'a été entrepris ⁽¹⁸⁾ qu'après la découverte expérimentale du déplacement, dans le cas d'un électron lié. Signalons toutefois (nous y reviendrons au chap 8) que French et Weisskopf avaient commencé un tel calcul avant 1947, et qu'avec ces moyens (disponibles au début des années trente) ils auraient ^{pu} trouver la valeur relativiste exacte du déplacement de Lamb (et sans renormalisation).

3.2.b Le calcul de Waller ⁽¹⁹⁾ de l'énergie propre de l'électron libre : (mars 1930)

Bien que ce cas soit plus simple que celui envisagé par Oppenheimer, il lui succède historiquement. WALLER utilisa la théorie

de Dirac ⁽²⁾, équivalente à la théorie de Heisenberg-Pauli dans sa 2ème version, mise à part l'énergie propre électrostatique, comme le montrait l'article d'Oppenheimer.

Pour un seul électron (et le nombre d'électrons est constant avant la théorie des trous), le hamiltonien total se réduit à $H = H_D + H_R + H_I$ en utilisant les notations précédentes. Les états non perturbés sont les $|\vec{p}, \epsilon; (n_\alpha)\rangle$ où \vec{p} et ϵ sont l'impulsion et la polarisation de l'électron libre de Dirac et n_α les nombres d'occupation des modes α du champ électro-magnétique.

Dans ce cas, $E_p^{(2)}$ se calcule explicitement, et le terme le plus divergent est proportionnel à $\frac{\alpha}{E_p} \int k dk$. Les différences $E_{p'}^{(2)} - E_p^{(2)}$ divergent elles-aussi quadratiquement.

Le calcul classique, où $H = \frac{(\vec{p} - e \vec{A})^2}{2m} + H_R$, donne le même degré de divergence. Les états d'énergie négative de l'équation de Dirac ne peuvent donc être tenus pour responsables.

ROSENFELD ⁽²⁰⁾ aboutit aux mêmes conclusions que Oppenheimer et Waller. Il établit que l'énergie propre d'un oscillateur harmonique couplé au champ électromagnétique diverge elle aussi et en $k^{n+1/2}$ pour le niveau n de l'oscillateur. Suivant un conseil de Pauli, Rosenfeld donna un autre aspect des divergences : il appliqua la théorie des perturbations dépendant du temps à un état initial $|n; 0\rangle$ composé d'un électron de niveau n et de 0 photon pour calculer la probabilité d'être dans $|m; \{n_\alpha\}\rangle$ au bout du temps t , ainsi que l'énergie moyenne de l'électron et l'énergie moyenne du rayonnement au même instant. Aussi court que soit le temps t , cette dernière quantité diverge, rendant absurde la méthode de perturbation.

Avant la théorie des trous de Dirac, un photon ne pouvait créer de paires, et donc le problème de l'énergie propre du photon ne se posait pas. Sortant de l'électrodynamique pure, Rosenfeld puis Solomon calculèrent l'énergie propre gravitationnelle du photon.

3.2.c. L'énergie propre gravitationnelle du photon :

L'intérêt de ce problème peut sembler anecdotique, car il est d'usage de négliger les effets gravitationnels quand on traite de problèmes d'électromagnétisme. Pourtant il apporte une nouvelle difficulté sérieuse pour la théorie d'Heisenberg et Pauli : l'énergie du champ de gravitation créé par un photon diverge. La question fut d'abord soulevée par Heisenberg.

ROSENFELD y répondit en septembre 1930 : dans ce calcul le champ gravitationnel est traité comme faible : $g_{\mu\nu} = \delta_{\mu\nu} + \epsilon\gamma_{\mu\nu}$ ($\epsilon = \sqrt{\frac{8\pi f}{c}}$; f constante de Newton). Alors l'équation d'Einstein se réduit à $\square\gamma_{\mu\nu} = -2\epsilon T_{\mu\nu}$, où $T_{\mu\nu}$ désigne le tenseur d'énergie-impulsion du champ électromagnétique: $T_{\mu\nu} = \frac{1}{4} g^{\mu\nu} F^2 + F^{\mu\rho} F_{\rho\nu}$, et permet d'exprimer $\gamma_{\mu\nu}$ en fonction de $F_{\mu\nu}$ libre. L'expression de $\gamma_{\mu\nu}$ ainsi obtenue est reportée dans l'énergie totale H . Dans un traitement semi-quantique où le champ de Maxwell seulement est quantifié (en représentation de Heisenberg), l'énergie totale du système photon-champ de gravitation s'obtient en prenant la valeur moyenne dans un état à 1 photon de H fonction du seul champ libre. Le résultat diverge, même si l'on soustrait la partie indépendante du nombre de photons. SOLOMON ⁽²²⁾ a montré, en utilisant la méthode développée par lui-même et par Rosenfeld (voir chap 1 § 6) pour éliminer l'énergie de zéro du champ électromagnétique, que l'énergie propre diverge indépendamment de cette élimination. Ce dernier point n'était alors pas évident, comme il transparait de l'étude de HEISENBERG ⁽³⁶⁾ de l'énergie propre de l'électron (1930) : Heisenberg trouvait des solutions d'énergie propre finie à l'électrodynamique de Dirac de masse nulle, à condition de faire deux suppositions (malheureusement contradictoires). 1) l'ordre des créateurs-annihilateurs dans l'expression de H est l'ordre naïf déduit de l'expression du hamiltonien classique.

2) il existe une solution stationnaire sans énergie de zéro correspondant au vide de particules.

Ce résultat pouvait laisser croire que l'énergie de zéro jouait un rôle important dans la divergence de l'énergie propre, mais, comme Heisenberg le nota à la fin du même article, l'introduction d'un hamiltonien modifié (à la Landau-Feierls, Rosenfeld-Solomon voir chap. 2 § 6) rendait impossible sa méthode, et l'énergie propre demeurait infinie. En 1928, dans une situation plus confuse, Jordan ⁽³⁷⁾, avant même que le calcul des divergences soit explicité, s'était exprimé ainsi :
"mathématiquement, ce sont les mêmes propriétés des champs quantifiés \vec{E} et \vec{B} qui sont responsables de cette difficulté [la divergence de l'énergie propre coulombienne de l'électron] et qui expliquent la présence gênante de l'énergie de zéro". *

Les théories classiques de la gravitation et du rayonnement font intervenir des champs étendus : la divergence de l'énergie propre n'était donc pas nécessairement liée à la présence de particules ponctuelles dans la théorie classique, mais plutôt à la forme du terme de couplage dans le lagrangien qui faisait intervenir le produit de deux champs interagissant au même point de l'espace-temps.

3.3 LES DIVERGENCES DANS LA THEORIE DES TROUS DE DIRAC :

Dans la théorie de Dirac où le vide quantique est obtenu en remplissant les états d'énergie négative de l'équation de Dirac, le phénomène de création de paires électrons-trou ou électron-positron devient possible : un électron d'énergie négative peut sauter dans un état d'énergie positive en laissant un trou derrière lui. Il en résulte que le vide est polarisable (voir chap. 5 § 1) : un électron

* Alors Jordan pensait que la théorie de Heisenberg- Pauli (encore non publiée) éliminerait la difficulté de l'énergie propre coulombienne déjà rencontrée par Jordan et Klein.

s'entoure d'un nuage virtuel de paires qui écranter sa charge, comme s'il était plongé dans un milieu diélectrique. On peut alors espérer que l'énergie d'interaction de l'électron avec son propre champ va se trouver diminuée et que l'énergie propre divergera moins. Un calcul rigoureux montre effectivement que la divergence qui subsiste est seulement logarithmique.

3.3.a Les calculs de Weisskopf de l'énergie propre de l'électron :

Dans sa première investigation du problème ⁽²³⁾, suggérée par Pauli en 1934, WEISSKOPF trouva la même divergence (de Waller en k^2) ^{que} pour la vieille théorie à un seul électron ; il corrigea son erreur ⁽²⁴⁾ en juillet 1934 après une remarque de FURRY . Weisskopf rappelait encore récemment et avec beaucoup de modestie le rôle de Furry . *"Cela me remémore l'un des moments les plus sombres de ma carrière de physicien. Je fis une erreur dans la première publication, conduisant à une divergence quadratique de l'énergie propre. Puis je reçus une lettre de W. Furry qui notifiât mon erreur plutôt stupide et le caractère logarithmique de la divergence. Au lieu de publier le résultat lui-même, il m'autorisa à publier une correction citant son intervention. Voilà comment la découverte de la divergence logarithmique de l'énergie propre de l'électron est faussement attribuée à moi et non pas à Wendell Furry"*.

Pour plus de clarté nous indiquerons la méthode suivie par Weisskopf en suivant plutôt la présentation utilisée dans une publication ultérieure ⁽²⁵⁾ sur le même sujet.

Le hamiltonien total est $H = H_0 + eV$ avec :

$$H_0 = \frac{1}{2} \int (\mathbf{E}^2 + \mathbf{B}^2) d^3x - \int \psi \operatorname{div} \vec{E} d^3x + \int \psi^\dagger (\vec{\alpha} \cdot \vec{p} + \beta m) \psi d^3x$$

$$\text{et } eV = e \int \bar{\psi} A \psi d^3x = \int (\rho \varphi - \vec{j} \cdot \vec{A}) d^3x$$

comme il résulte de l'expression du lagrangien : $L = -\frac{1}{4} F^2 + \bar{\psi} (i \not{\partial} - m - eA) \psi$ - $A t \rightarrow -\infty$, on suppose que l'interaction est débranchée ($e=0$). Les expressions des champs quantiques libres $\dot{\psi}$ et \dot{A}_μ sont obtenues par quantification canonique, en jauge de coulomb $\dot{A}_0 = 0$ pour le second.

- Puis on branche adiabatiquement l'interaction. En représentation de Heisenberg (voir chap. 2 § 4) les champs ψ et A_μ évoluent suivant les équations classiques du mouvement $\square A_\mu = j_\mu$ et $(i\gamma - e\mathcal{A} - m)\psi = 0$ avec les conditions initiales $\psi(-\infty) = \psi^0$ et $A_\mu(-\infty) = A_\mu^0$.

- Si la charge $\epsilon(t)$ croît lentement de 0 à e, alors un état propre (à 1 électron) à $t = -\infty$ reste un état propre du hamiltonien en représentation de Heisenberg $H(t) |1\rangle = E(t) |1\rangle$; entre t et $t + dt$, la théorie des perturbations au 1er ordre donne $dE = \langle 1 | V d\epsilon | 1 \rangle$; V dépend implicitement de t et donc de ϵ par les champs ψ et A : $V = V_0 + \epsilon V_1 + \dots$; à l'ordre le plus bas $dE = \langle 1 | V_0 d\epsilon | 1 \rangle + \langle 1 | V_1 \epsilon d\epsilon | 1 \rangle$.

L'énergie propre d'un électron s'obtient en intégrant de 0 à e

$$E_{\text{self}} = 1/2 e^2 \langle 1 | V_1 | 1 \rangle \quad (\text{car } \langle 1 | V_0 | 1 \rangle = \langle 1 | \bar{\psi} \mathcal{A} \psi | 1 \rangle = 0)$$

Weisskopf utilise cette méthode plutôt que la théorie usuelle des perturbations utilisée par Waller, "pour mieux faire ressortir la signification physique des différents termes". E_{self} est obtenu au 2ème ordre par $E_{\text{self}} = \frac{1}{2} \int \langle \vec{j}_\mu A^\mu \rangle d^3x$ avec $A^\mu = \overset{\circ}{A}^\mu + A'^\mu$, A'^μ potentiel retardé créé par $\overset{\circ}{j}_\mu$. $E_{\text{self}} = -1/2 \int \langle \vec{j} \cdot \vec{A} \rangle d^3x + 1/2 \int \langle (E'^2 - B'^2) \rangle d^3x$. \vec{E}' est séparé en ses parties longitudinale \vec{E}'_{\parallel} et transverse \vec{E}'_{\perp} .

Alors $E_{\text{self}} = E_{\text{stat.}} + E_{\text{spin}} + E_{\text{fluct.}}$

$$E_{\text{stat.}} = \frac{1}{8\pi} \int \left\langle \frac{\rho(\vec{r}-\vec{r}')\rho(\vec{r}')}{|\vec{r}-\vec{r}'|} \right\rangle d^3x \quad E_{\text{spin}} = \frac{1}{2} \int \langle E_{\perp}'^2 - B'^2 \rangle d^3x$$

$$E_{\text{fluct.}} = -\frac{1}{2} \int \langle \vec{j} \cdot \vec{A}' \rangle d^3x$$

Dans ces expressions, on substitue les expressions des champs ψ et A_μ calculés à l'ordre convenable en e, et on soustrait enfin la contribution du vide (mer de Dirac au repos). Les résultats corrects pour ces trois parties sont dans le cas d'un électron au repos

$$E_{\text{stat.}} = \lim_{P \rightarrow +\infty} \frac{e^2}{4\pi^2} \frac{mc}{h} \ln \frac{2P}{mc} \quad E_{\text{spin}} = \lim_{P \rightarrow +\infty} \left(-\frac{e^2}{4\pi^2} \frac{P^2}{mhc} + \frac{e^4}{4\pi^2} \frac{mc}{2h} \ln \frac{2P}{mc} \right)$$

$$E_{\text{fluct.}} = \lim_{P \rightarrow +\infty} \frac{e^2}{4\pi^2} \frac{P^2}{mhc}$$

On constate que les divergences quadratiques (de Waller) s'éliminent, et il reste seulement une divergence logarithmique $E_{\text{self}} = \lim_{p \rightarrow \infty} \frac{3e^2}{8\pi^2} \frac{mc}{h} \ln \frac{2p}{mc}$. Dans son premier calcul ⁽²³⁾, Weisskopf, en raison d'une erreur de signe, trouvait une divergence quadratique pour $E_{\text{spin}} + E_{\text{fluct}}$.

Dans le premier article ⁽²³⁾, les calculs étaient présentés à l'état brut. En 1939 Weisskopf ⁽²⁵⁾ réexposa le problème en faisant précéder les calculs de justifications "physiques" du comportement des différents termes. Il ne pouvait s'agir que d'arguments a posteriori, comme le prouvait l'erreur initiale de l'auteur. De plus la discussion faisait appel aux propriétés de grandeurs inobservables, comme les mouvements de la mer d'états d'énergie négative, et à l'interprétation du spin comme "Zitterbewegung" ⁽²⁶⁾ (mouvement circulaire fluctuant de l'électron) (voir chap. 2 § 2) dont on connaît les limites. L'argument d'étalement de la charge -qui n'explique que le comportement du terme électrostatique- a déjà été présenté au début de ce paragraphe. L'ajout d'autres images colorées fera mieux de consulter le texte original ⁽²⁵⁾.

Ce même texte contenait deux généralisations importantes du problème étudié. D'abord le cas des bosons scalaires chargés dont la théorie venait d'être développée par Pauli et Weisskopf ⁽²⁷⁾ (voir chap 5 § 3), se prêtait à la même méthode de calcul :

$$E_{\text{stat.}} = \frac{e^2}{16\pi h c} \lim p^2, \quad E_{\text{spin}} = 0, \quad E_{\text{fluct}} = \frac{e^2}{4\pi^2 m h c} \lim_{p \rightarrow \infty} p^2$$

Malgré l'existence du phénomène de création de paires dans cette théorie, la divergence reste ici quadratique.

Enfin la dernière partie de l'article contient une démonstration élégante du fait que dans le cas de l'électron, la divergence est logarithmique à tous les ordres de la série des perturbations. Dans le cas de masse nulle $m = 0$, $E_{\text{self}}^{(n)} = 0$ comme il résulte de la symétrie électron-positron appliquée à la forme explicite du n° terme de la série des perturbations stationnaires. Par ailleurs l'homogénéité de l'intégrand du n° terme par rapport aux grandes valeurs des impulsions des photons virtuels permet d'établir la forme de $E_{\text{self}}^{(n)}$:

$$E_{\text{self}}^{(n)} \sim mc^2 \left(\frac{e^2}{hc}\right)^n \left(\frac{p}{mc}\right)^N \left(\ln \frac{p}{mc}\right)^n$$

$E_{\text{self}}^{(n)} = 0$ pour $m = 0$ n'est possible que si $N \leq 0$, donc

$$E_{\text{self}}^{(n)} \approx mc^2 \left(\frac{e^2}{hc} \ln \frac{P}{mc} \right)^n .$$

Weisskopf définit une longueur critique de la théorie $\frac{h}{mc} e^{-1/\alpha}$ pour laquelle l'énergie propre à tous les ordres est de l'ordre de la masse au repos. Cette longueur est 10^{-58} fois le rayon classique de l'électron. "La "longueur critique" de la théorie du positron est ainsi infiniment plus petite qu'on le suppose habituellement" concluait Weisskopf. Le terme de longueur critique était sans doute emprunté à Heisenberg (voir chap. 4 § 6) pour signifier l'existence d'une limite au delà de laquelle les fondements de la théorie correspondante ne seraient plus valides.

3.3.b La polarisation du vide

Au Congrès SOLVAY de 1933 ⁽³¹⁾ Dirac attira l'attention des physiciens sur le phénomène de polarisation du vide : un champ électromagnétique induit la présence de paires réelles ou virtuelles dans le vide, si bien que celui-ci se comporte comme un milieu diélectrique polarisable. Le champ de polarisation diverge logarithmiquement. Nous étudierons en détail au chap 5 § 1 l'histoire de ce problème. Le lecteur peut se demander pourquoi nous différons cette étude : la raison est que les théoriciens surent traiter beaucoup plutôt cette divergence que les autres, l'idée de renormalisation de charge étant apparue bien avant la théorie complète de la renormalisation. La polarisabilité infinie du vide ne fut donc en général pas considérée comme un vice aussi désastreux de la théorie (sauf sans doute par Dirac).

3.3.c. L'énergie propre du photon dans la théorie des trous de Dirac

Un photon est maintenant susceptible de créer des paires dans son voisinage, donnant lieu à une énergie propre non nulle. Le calcul de cette énergie propre suppose au moins un traitement correct du problème de la polarisation du vide dans la théorie des trous. Nous verrons au chapitre 5 § 1 comment HEISENBERG ⁽²⁸⁾ en 1934 put soustraire de manière univoque les infinis qui apparaissent alors, en exploitant l'invariance relativiste et la symétrie de jauge. Muni de l'expression correcte de la polarisation du vide au 2ème ordre en e , il put calculer dans le même article l'énergie propre du photon : elle

divergeait logarithmiquement, en $(\alpha k c \ln \frac{P}{mc})$ où k est le moment du photon. Heisenberg fit remarquer que cette divergence existait dans la théorie quantique bien qu'elle n'ait pas d'analogue classique dans la théorie des ondes électromagnétiques. Cette divergence est en quelque sorte créée par la quantification (on a pu faire la même remarque à propos de l'énergie propre gravitationnelle du photon). Comme il est souhaitable que la masse du photon soit nulle, il n'est même pas possible de donner un sens à une valeur finie de la coupure, contrairement au cas de l'électron.* D'où le commentaire de Wentzel (29) : *"Un fait plus sérieux fut que selon la théorie du positron, le photon acquiert aussi une énergie propre infinie. [...] aucun remède ne semblait disponible pour sauver l'invariance relativiste et l'invariance de jauge de la théorie"*.

Quelques mots sur l'évolution ultérieure du problème : SERBER (39) (1936) critiqua le calcul de Heisenberg pour des raisons techniques et aussi parce qu'il utilisait les prescriptions de soustraction (des divergences de la polarisation du vide) en dehors de leur domaine de validité supposée. Quoiqu'il en soit, l'énergie propre du photon ne pouvait qu'encore plus diverger. L'élimination de cette divergence fut la plus tardive de toute. C'est seulement la régularisation de Pauli-Villars (1949) qui permit une solution rigoureuse. Comme le dit Wentzel, une énergie propre du photon est très choquante en ce qu'elle viole l'invariance de jauge du lagrangien de départ. Cette violation s'introduit quelque part dans les calculs, et l'on n'était bien loin en 1936 de savoir où. Il fallait maintenir l'invariance de jauge à tous les stades du calcul grâce à une méthode invariante de régularisation et de perturbation.

3.3.d. Corrections radiatives à la diffusion d'un électron

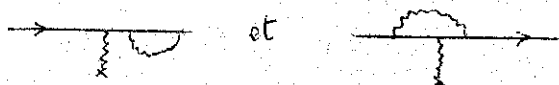
Une fois traité par Bloch, Nordsieck (voir chap. 5 § 2) le problème infra-rouge lié à la présence de photons mous dans les états intermédiaires et finals. DANCOFF entreprit de calculer le comportement des corrections radiatives pour les grandes impulsions des photons virtuels. On a là un dernier type de divergence rencontré en électrodynamique, une

* Bien sûr on sait aujourd'hui que la renormalisation du propagateur du photon (polarisation du vide) se traduit non pas par une masse propre du photon mais par une renormalisation de charge (Ref I p. 335)

nouvelle fois logarithmique pour l'électrodynamique spinorielle.

Plus précisément Dancóff calcule les corrections radiatives à la diffusion par un potentiel classique coulombien $V(r)$. La méthode utilisée est la suivante. On calcule les éléments de matrice de l'énergie d'interaction coulombienne $e \int \psi^\dagger V \psi d^3r$ (nous prenons les notations de l'électrodynamique spinorielle) entre les états incident et sortant (conformément à l'approximation de Born), états à un électron perturbés par le couplage avec le champ de rayonnement $e \bar{\psi} \mathbf{A} \psi$ calculés au 2ème ordre en e . On obtient ainsi l'analogie dans l'ancienne théorie des perturbations (non covariante) du calcul des graphes

du type



et



Suivant une remarque de Serber, il faut aussi tenir compte de

Dans le cas des spineurs la divergence est logarithmique. Le caractère logarithmique de cette divergence est exact, toutefois l'expression précise de Dancoff est fautive, en raison de l'omission des termes correspondant à la partie longitudinale du champ de rayonnement, comme le remarquent plus tard Lewis et Tomonaga (voir chap 8 § 6). Mais la correction est finie pour le champ scalaire chargé (de Pauli et Weisskopf) en raison du caractère particulier de l'interaction coulombienne (d'autres potentiels conduiraient à des divergences quadratiques comme le fait remarquer Dancoff).

Résumé des §§ 3.2. et 3.3.

Il fallut attendre la théorie des trous de Dirac pour trouver les divergences fondamentales de l'électrodynamique quantique, telles qu'on les connaît aujourd'hui. Ces divergences sont logarithmiques en raison de l'élimination, semblant miraculeuse dans les premières études, des termes les plus divergents, linéaires ou quadratiques dans la coupure ultra-violette. On sait aujourd'hui que c'est l'invariance de jauge qui provoque cet abaissement du degré de divergence. On peut classer les divergences en trois catégories.

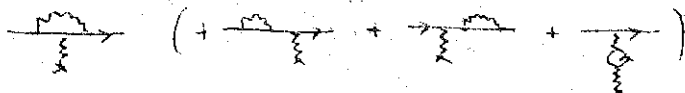
1) Celles de l'énergie propre de l'électron libre étudiées par Waller (1930) et Weisskopf (1934) associées aujourd'hui au graphe



2) Celles liées à la polarisabilité du vide étudiées par Dirac (1933) et autres,



3) Les corrections radiatives à la diffusion étudiées par Dancoff (1939)



Si l'ombre portée par les divergences du 2ème type fut vite moins noire (dès 1934), il n'en est rien des deux autres dont on ne sut que faire avant 1947. Nous allons voir maintenant comme ces difficultés furent accueillies par les théoriciens des années trente.

3.4. COMMENTAIRES SUR LA NATURE DES INFINIS ET SUR L'AVENIR DE LA THEORIE

Dans la plupart des travaux cités il semble que les auteurs aient trouvé les résultats des calculs suffisamment éloquents pour s'abstenir de longs commentaires. Cependant les quelques phrases de conclusions, résumant sans doute de plus amples conversations de cafeteria, permettent de discerner différentes sortes de critique.

3.4.a Origine mathématique des infinis

En premier lieu il convient de déterminer l'origine mathématique des infinis : à quel moment du calcul apparaissent-ils , et par quel abus ?

- La méthode des perturbations est souvent condamnée
"Ce calcul laisse naturellement ouverte la possibilité que la faute réside dans le processus d'approximation " (WALLER⁽¹⁹⁾). ROSENFELD⁽²⁰⁾, après avoir montré que l'énergie du rayonnement produit à partir d'un état à un électron après branchement de la perturbation diverge, conclut : "Ce résultat paradoxal semble enfin ne pas être conforme à la conservation de l'énergie, bien que cette dernière soit à la base de toute méthode de traitement ; il faut en déduire que l'énergie de couplage s'avère infiniment grande, si bien que cela n'a pas de sens de partir de systèmes de particules découplés. Autrement dit, il apparaît que le formalisme reposant sur l'hypothèse d'un couplage faible conduit à des résultats qui contredisent cette hypothèse". Mais que substituer à la méthode de perturbation ? "Cette méthode semble exprimer le sens de la théorie, puisque la charge en principe peut être envisagée arbitrairement petite. Une autre méthode ne semble guère possible, du point de vue général " (Waller⁽¹⁹⁾). Dans l'édition de 1947 des célèbres "Principles of Quantum Mechanics " (32) DIRAC va plus loin en doutant de l'utilité d'une éventuelle solution non perturbative : "Nous avons là une difficulté fondamentale de l'électrodynamique quantique, une difficulté qui n'a pas encore été résolue. Il se peut que l'équation d'onde ait des solutions qui ne soient pas de la forme d'une série

entière en e . De telles solutions n'ont pas encore été trouvées. Si elles existent elles sont probablement très compliquées. Enfin même si elles existent la théorie ne serait pas satisfaisante, car nous devons exiger d'une théorie satisfaisante que ses équations aient une solution simple pour tout problème physique simple, et la solution de l'équation d'onde pour le problème banal du mouvement d'une seule particule chargée en l'absence de champ de radiation incident n'a pas été trouvée".

- Corrélativement la forme du couplage est tenue pour responsable des divergences. "Il est aussi clair que les difficultés soulevées sont en rapport direct avec le fait que l'électron a été considéré comme un objet ponctuel" (Waller) ⁽¹⁹⁾. Pourtant comme PAULI ⁽³³⁾ le fait remarquer : "De manière formellement analogue au cas de l'électron, intervient une énergie infinie du champ de gravitation créé par un quantum lumineux ... bien qu'ici aucune singularité ponctuelle ne soit introduite dans la théorie classique". La même remarque peut être faite pour l'électrodynamique quantique pure si l'on considère, comme aujourd'hui, que la théorie "classique" de départ est donnée par les équations de Maxwell et l'équation de Dirac couplées qu'elle est donc purement ondulatoire et dépourvue de singularités ponctuelles. La divergence est plutôt liée à la forme du terme de couplage qui fait intervenir les champs interagissant au même point d'espace-temps, et au nombre infini de degrés de liberté : "l'application du formalisme de la mécanique ondulatoire à des systèmes comportant une infinité de degrés de liberté représente une violation des limites d'application du principe de correspondance". (Pauli ⁽³³⁾).

Ce dernier commentaire de Pauli dépasse l'aspect technique des calculs; il relève d'une analyse critique des fondements mêmes de la théorie.

3.4.b. Critique des fondements

- Pour HEISENBERG (en 1934) ⁽²⁸⁾ la théorie classique est valide, c'est la méthode de quantification qu'il faut suspecter : "La circonstance selon laquelle l'application de la mécanique quantique conduit à des divergences qui n'apparaissent pas dans la théorie naturelle des champs d'onde, laisse supposer que cette théorie naturelle donne vraiment pour l'essentiel la description exacte des faits en tant que base de correspondance, mais que surtout le passage à la théorie quantique ne peut pas être entrepris de la matière primitive, comme on a tenté de le faire

- Inversement on peut penser que la méthode de quantification est solide, mais que c'est la théorie classique qui doit être modifiée en premier lieu. Ce point de vue fut à l'origine de plusieurs travaux (Born, Infeld, Wentzel, Dirac....) dont nous traiterons au Chap 7 . HEITLER, à la suite d'une longue discussion : "*critique et limites de la théorie*" (§ 24) publiée dans la première édition (1935) de "*La théorie quantique du rayonnement*" ⁽³⁾ soutient ce point de vue et accorde sa faveur à la nouvelle théorie de Born et Infeld (1934). Dans un ouvrage essentiellement consacré aux applications de la théorie du rayonnement (en particulier à la physique des rayons cosmiques), on comprendra que Heitler cherche une origine et une mesure de leurs limites. Dans la théorie classique d'Abraham-Lorentz la réaction du rayonnement apporte des corrections au mouvement d'une charge, dépendant de la structure de la charge dès que les longueurs d'ondes du rayonnement interagissant avec cette charge sont plus petites que le rayon classique de l'électron ($\lambda < r_0$) (Heitler le démontre dans le cas particulier d'un oscillateur harmonique). "*En conséquence, il est essentiel de souligner que d'après la correspondance avec la théorie classique nous ne devons pas nous attendre à ce que ces approximations d'ordre supérieur aient un sens physique, puisque les termes analogues de la théorie classique dépendraient de la structure de l'électron*"* (p. 93) "*A de très hautes énergies ($E > 137 mc^2$) [c'est-à-dire $\lambda \sim \frac{hc}{E} < r_0$] il est hautement probable que la théorie ne s'accorde pas avec l'expérience, et qu'elle donne une valeur trop élevée de la perte d'énergie d'un électron dans la matière*" (p. 235). La prudence de Heitler : "*Une telle discussion ne peut être que provisoire. En dernière analyse les limites d'une théorie ne peuvent être comprises que si une nouvelle théorie dépassant ces limites existe déjà*" était de rigueur, car la dernière conclusion sur les pertes d'énergie (des électrons cosmiques ; voir Chap. 6 § 2) était fautive. A la fin du paragraphe (§24), la raison théorique (due à Williams et à von Weizsäcker) d'espérer la validité des calculs de Bremsstrahlung au delà de la limite $137 mc^2$ est donnée, et elle débouche sur une sage conclusion : "*Une discussion finale de ce problème n'est pas encore possible, car les données expérimentales sont encore insuffisantes*". Nous verrons au chapitre 6 § 2 comment l'expérience donnera raison à Weizsäcker et à Williams

* Cet argument fut aussi fréquemment avancé par Niels Bohr (voir chap 4 § 5)

- D'autres théoriciens remettent en question d'autres principes fondamentaux : "il n'est pas certain qu'une telle théorie sera possible sur la base de la théorie de la relativité restreinte" (Oppenheimer⁽¹⁵⁾)

A cette date (1930) Oppenheimer intègre le problème des énergies négatives dans l'ensemble des difficultés d'une électrodynamique quantique. PAULI⁽³³⁾, dès 1933, suggère de modifier le concept d'espace-temps : "Nous voudrions remarquer que non seulement le concept de champ, mais encore le concept d'espace-temps nécessite une modification fondamentale, dans le domaine de l'infiniment petit". Pauli ne précise pas de quel genre de modification il peut d'agir, * mais explicite plutôt sa critique du concept de champ libre et de la dichotomie matière rayonnement., "La théorie présente repose sur deux bases logiquement indépendantes, auxquelles correspondent dans la théorie classique la mécanique du point et la théorie de Maxwell : c'est-à-dire sur la mécanique ondulatoire de particules matérielles et sur la mécanique ondulatoire (théorie des photons) du rayonnement dans une boîte. En conséquence, elle ne permet pas de prédire la structure granulaire de l'électricité, puisqu'elle est compatible avec un nombre arbitraire de charges, de valeur arbitraire ** En résulte aussi l'essai (malheureux), de résoudre en photons tout le champ d'un électron en mouvement, au lieu de le considérer comme partie d'un tout indivisible, qui est inévitablement lié à une valeur numérique déterminée du nombre sans dimension $\frac{e^2}{4\pi\hbar c}$. Une théorie quantique relativiste complète devra apporter une profonde unification des fondements". Heisenberg,⁽²⁸⁾ en 1934, trouve dans son étude la polarisation du vide un nouvel argument en faveur d'une théorie qui donne une valeur déterminée à la constante de structure fine : on ne peut séparer, d'une matière physique pertinente et indépendante de la valeur de α , un champ électromagnétique, du champ de matière réel ou virtuel qu'il induit dans le vide polarisable.

* Pauli avait critiqué la tentative de Heisenberg (début de 1930) de remplacer l'espace continu par un réseau de pas r_0 , essentiellement parce qu'elle n'était pas conforme à la relativité, et Heisenberg⁽³⁶⁾ (p. 4 et 5) avait alors abandonné ce projet.

** Cette critique est à la base de la théorie de Dirac⁽⁴⁰⁾ de 1931 des monopoles magnétiques.

- Une révision des fondements peut aussi commencer par l'élargissement du cadre théorique, englobant maintenant les phénomènes nucléaires. OPPENHEIMER ⁽¹⁵⁾ dès 1930 exprime un lien possible avec la physique nucléaire "la théorie ne peut pas être appliquée à la discussion de la structure du nucléon. Il paraît improbable que les difficultés discutées dans ce travail soient soluble sans une théorie adéquate des masses de l'électron et du proton". Bohr et Heisenberg (voir Chap. 4 § 5 et 6) remarquent qu'il n'y a pas de raison d'espérer obtenir des prédictions correctes par l'électrodynamique quantique des phénomènes à l'échelle nucléaire, et pensent (contrairement à Dirac) qu'il est vain de chercher à résoudre isolément les difficultés de l'électrodynamique pure. Les dimensions nucléaires donnent une coupure naturelle pour les intégrales divergentes, et après tout, selon Weisskopf, celles-ci ne dépendent que peu (logarithmiquement) de la coupure. Pour Heisenberg (1938) la théorie future doit faire intervenir une longueur fondamentale si bien que là encore, notre conception de l'espace doit être révisée. Ce programme de Heisenberg sera discuté au chap. 4 § 6.

- La discussion peut être située à un niveau encore plus fondamental où la mesurabilité des variables dynamiques de la théorie est remise en question. Cette discussion commencée par Landau et Peierls sera évoquée au chap 4.

La plupart des interprétations des divergences que nous venons de relever, sont étroitement liées à des conjectures sur la théorie future. Celles-ci se préciseront au chap. 7 quand nous présenterons quelques alternatives à la théorie quantique des champs normale. On peut chercher à extraire ce qui fait le plus penser à la théorie de la renormalisation : la nécessité d'une théorie de la masse de l'électron citée par Oppenheimer, la notion d'électron "habillé" de Pauli..., mais ces propositions, parfois audacieuses, restent vagues. Les jugements rétrospectifs des théoriciens qui ont connu la solution de ces problèmes sont évidemment les plus pertinents : pour bien critiquer la théorie ancienne il faut connaître la nouvelle qui vient la généraliser ou la détruire. Les nombreuses tentatives de modification ou de refonte de la théorie (voir chap 7) et la solution moderne apportée par la théorie de la renormalisation montrent qu'il ne suffit pas de pouvoir justifier par des arguments physiques a priori les équations de base, pour

qu'elles possèdent des solutions mathématiques ; il faut préciser dès le départ, avant la résolution des équations, à quelles observations elles sont associées. C'est le point de vue de VILLARS ⁽³⁴⁾ : " *Ce qui donne au problème des divergences son aspect particulier, c'est la relation réciproque entre le problème purement physique de la formulation des équations de base et la question mathématique plus subtile soulevée par leur résolution..... le problème du physicien n'est pas seulement la solution d'un ensemble donné d'équations, mais plutôt l'élaboration d'un code, par lequel une correspondance biunivoque peut être établie entre les observations et un certain schéma mathématique.* ". Dès 1930, les théoriciens de la mécanique quantique étaient apparemment très proches de la solution, il suffisait de poser la bonne question : quelles sont les grandeurs observables et surtout en fonction de quels paramètres observables les exprimer ? Une question ainsi posée par KRAMERS dès 1937, mais dont la réponse ne fut donnée qu'en 1947 dans le cadre de l'électrodynamique quantique (voir chap 8).

La longue stagnation du problème UV (de 1930 à 1947) suffit à démontrer son importance historique ; de plus la solution vint non pas de la suppression des divergences, mais de leur compensation : l'élaboration de la théorie moderne passe par l'intermédiaire d'intégrales divergentes . Toutefois la théorie, ainsi imparfaite fut-elle, ne tomba pas dans l'abandon : " *Tout ce travail sur l'énergie propre est aujourd'hui dépassé et d'intérêt seulement historique. Malgré tous les échecs, la confiance générale en l'électrodynamique quantique comme l'outil suprême, quoiqu'encore imparfait, de la théorie atomique, resta vivante. Il est vrai que la plupart des applications particulières étaient fondées sur la théorie du rayonnement de Dirac, et les infinis, quand ils apparaissaient, étaient traités d'une manière heuristique (pure soustraction ou coupure). Aucune contradiction avec les faits expérimentaux n'était apparente (l'existence du déplacement de Lamb était soupçonnée, mais encore douteuse en ce temps-là). Néanmoins la conscience des difficultés fondamentales pesait lourdement dans nos esprits* ". (Wentzel ⁽³⁵⁾). Et le domaine d'application de la théorie quantique des champs continua de s'étendre à la physique nucléaire où affluaient les résultats expérimentaux nouveaux, avec un traitement ad hoc des divergences. Enfin certains problèmes, comme nous le verrons au chap. 5 s'avérèrent solubles, malgré les divergences UV.

CHAPITRE 4

LE PROBLEME DE LA MESURABILITE DES GRANDEURS
PHYSIQUES DANS LE DOMAINE QUANTIQUE RELATIVISTE

La théorie quantique des champs prit naissance presque en même temps que l'interprétation de Copenhague du formalisme quantique : 1927 est l'année de l'article de Dirac sur la quantification du rayonnement, et aussi celle des relations d'incertitude de Heisenberg. La théorie de Dirac en tant qu'application formelle de la méthode de quantification, devait pouvoir supporter la même interprétation physique que la mécanique quantique d'une particule. Mais les difficultés nouvelles, comme l'infinité de l'énergie propre de l'électron et les énergies négatives de l'équation de Dirac, jetèrent un doute sur les fondements observationnels de la mécanique quantique appliquée à des problèmes relativistes. Pour les fondateurs de l'interprétation de Copenhague de la théorie quantique, le problème de la mesurabilité des grandeurs physiques relativistes devint un problème central. LANDAU et PEIERLS, en 1931, obtinrent des résultats négatifs : pour eux, il n'existait pas de grandeur physique associée à l'impulsion d'une particule ou au champ électromagnétique dans la théorie quantique relativiste. BOHR et ROSENFELD en 1932-1933 réfutèrent une partie de leur argumentation et montrèrent que la base observationnelle de la théorie quantique du rayonnement était aussi ferme que celle des problèmes quantiques plus simples. L'importance historique de cette polémique entre Landau-Peierls et Bohr n'est pas dans une amélioration du formalisme, ni dans une solution du problème des divergences ; elle réside d'une part dans la plus grande confiance que purent accorder les théoriciens à la théorie quantique des champs, d'autre part dans le succès de la complémentarité de Bohr. Toutefois, il serait hasardeux de considérer comme définitives les conclusions de Bohr et Rosenfeld et celles de Landau-Peierls que ces derniers approuvèrent. Le problème de la mesure en théorie quantique relativiste est un problème complexe, et comme le montrent des études récentes, ^(38, 39) l'existence d'observables relativistes non locales, comme l'impulsion d'une particule, est une question difficile qui ne peut être tranchée par des arguments aussi simples que ceux de Landau et Peierls. Nous n'essayerons pas ici de dépasser la discussion historique telle qu'elle s'est déroulée jusqu'au second article de Bohr et Rosenfeld (1950) ; nous montrerons plutôt en quoi le type de "Gedankenexperimente" (expériences par la pensée) imaginées par ces auteurs illustrent la conception copenhagienne de la mécanique quantique et en quoi elles justifient les fondements de la

théorie quantique des champs. Cette étude, détaillée dans les quatre premiers paragraphes de ce Chapitre, sera suivie de celle de la position de Bohr contre Dirac pour la théorie de l'électron, et contre Pauli pour la désintégration β , puis de celle des limites à la mesurabilité de grandeurs locales, postulées par Heisenberg. Ces dernières limites servirent à Heisenberg pour justifier une nouvelle approche "observationiste" de la physique des particules : la matrice S.

4.1 INTRODUCTION DE LA PROBLEMATIQUE

Avant d'aborder les articles de Landau-Peierls et de Bohr-Rosenfeld, il importe de rappeler les présupposés de Bohr et de l'école de Copenhague concernant la mesure en théorie quantique, de préciser les termes utilisés dans la discussion (l'article de Landau-Peierls est précieux à cet égard) et de présenter succinctement celles des conclusions de ces auteurs qui ne furent guère contestées jusqu'à un passé récent.

4.1.a Les idées de Bohr concernant la mesure en théorie quantique

L'interprétation de Copenhague de la mécanique quantique fut une oeuvre collective et la lecture des seuls textes de Bohr* ne suffit pas à en donner un panorama complet. Citons le livre de Heisenberg sur "*Les principes physiques de la théorie quantique*"⁽⁴¹⁾, les introductions de Dirac⁽⁴²⁾ et de Landau⁽⁴³⁾ à leurs cours de mécanique quantique. Pour un résumé de la thèse de Bohr nous avons utilisé un texte de B. d'Espagnat.⁽⁴⁴⁾

Suivant Bohr une expérience de physique se décrit par un ensemble d'opérations communicables dans un langage qui est celui de la physique classique. Partant, toute mesure physique doit pouvoir être analysée en termes purement classiques. Dans le cas de la physique classique, on peut définir des entités attachées au système mesuré, par exemple la position et l'impulsion des particules, de manière non contradictoire et indépendante des appareils de mesures et des grandeurs effectivement mesurées dans une expérience donnée. Il n'en est pas ainsi dans le cas de la physique quantique : la perturbation minimale introduite par la mesure, finie en raison de l'existence du quantum d'action, interdit de

* Il existe une version anglaise de quelques-uns des plus importants ⁽⁴⁸⁾

définir des propriétés individuelles du système mesuré indépendantes du dispositif expérimental de mesure utilisé. Il n'est pas possible d'imaginer un dispositif expérimental dans lequel on effectuerait toutes les mesures possibles (par exemple celles de position et d'impulsion d'une particule) et pour lesquelles on obtiendrait un résultat certain (reproductible). L'analyse classique d'un tel dispositif montre qu'il existe toujours des relations d'incertitude du type $\Delta x \Delta p_x \geq \hbar/2$. Dans le cas du principe d'incertitude de Heisenberg on considère, par exemple, la mesure de la position d'un électron par un diaphragme suivie de celle de son impulsion; le diaphragme diffracte l'onde de Broglie associée à l'électron, si bien que l'impulsion transverse est indéterminée de $\Delta p_x \approx \frac{\hbar}{2\Delta x}$ où Δx désigne l'ouverture du diaphragme. Tout comme ce problème des corrélations de deux mesures successives, toute question physique doit être formulée en terme de dispositifs expérimentaux imaginables (peut importe la difficulté de leur réalisation).

Cette exigence (doublée de celle de la nature classique de l'appareil de mesure) doit pouvoir s'appliquer à l'une des questions les plus fondamentales de la mécanique quantique relativiste : existe-t-il des mesures "prédictibles"? c'est-à-dire, existe-t-il des dispositifs expérimentaux pour lesquels la dernière mesure effectuée conduit à des résultats certains? " Si de telles mesures n'existaient pas, la notion d'état quantique introduite dans le formalisme perdrait son sens usuel, puisqu'il n'y aurait pas moyen de savoir si le système est ou non dans cet état.

Cette question fut posée par Landau, Peierls, Bohr et Rosenfeld pour différentes mesures dans le domaine relativiste. Il importe dans la réponse de ne pas confondre indétermination (Δ) et imprécision (δ). L'imprécision d'une mesure est liée à la qualité de l'appareil utilisé; cette notion intervient aussi bien dans le cas de mesures classiques ou quantiques : par exemple, la précision d'une mesure de position (non relativiste) par un diaphragme est caractérisée par le diamètre du trou. L'indétermination d'une mesure est liée à l'ensemble du dispositif expérimental : elle est nulle (resp. ∞) pour une mesure de position si le reste du dispositif "prépare" un état propre $|\alpha\rangle$ (resp. $|p\rangle$).

* Comme il n'existe pas en général, en relativité, de séquence ordonnée d'événements dont l'ordre soit invariant, il faudra se garder de donner toujours à cet énoncé une interprétation causale.

Landau et Peierls démontrent que l'imprécision δp de mesures d'impulsion relativistes est au moins $\frac{h}{c\delta t}$ (δt durée de la mesure) d'où la non-prédictibilité. Bohr et Rosenfeld démontrent que la précision des mesures de champs est arbitrairement grande $\delta E = \delta B = 0$ et que ces mesures sont "répétables" : effectuées sur des domaines d'espace-temps presque confondues, elles donnent des résultats presque identiques. On a ainsi une généralisation relativiste de la prédictibilité (la notion d'antériorité est perdue).

La précision des appareils de mesures est indépendante de la nature quantique des objets mesurés, en vertu même du principe de Bohr selon lequel l'analyse des appareils de mesure doit se faire en termes classiques. S'il est vrai qu'en physique classique, on peut construire des appareils de mesure arbitrairement précis, comment se fait-il qu'on n'en soit pas assuré en physique quantique? Voici la raison : certaines mesures s'effectuent indirectement par la mesure préalable d'autres grandeurs. Même à supposer que ces mesures préalables soient infiniment précises, il peut se faire qu'elles concernent des grandeurs conjuguées et donc qu'elles se perturbent mutuellement, si bien qu'on ne peut les connaître simultanément avec une précision arbitraire. La grandeur finalement recherchée, fonction de ces grandeurs-là, ne peut être connue exactement.

Pour ce qui concerne l'élaboration du formalisme quantique relativiste, le choix des variables dynamiques et des observables associées ne pose pas a priori de problème : le principe de correspondance nous dit que ce sont celles de la théorie classique correspondante. Mais pour que l'interprétation de Copenhague puisse s'appliquer à ce formalisme, il faut vérifier l'exigence de prédictibilité des mesures associées aux observables.

4.1.b 1° cas étudié : mesures d'impulsion

Suivant Landau et Peierls, il n'existe pas de mesure certaine de l'impulsion. Quelque soit l'état initial de la particule mesurée, on ne peut mesurer son impulsion à mieux que $\delta p > \frac{h}{\delta t}$, pendant un temps δt .

Les mesures se font indirectement par l'étude de processus de collision. Supposons que la collision s'effectue entre la particule dans un état

* Nous nuancerons plus loin (4.1.c) cette affirmation

propre d'impulsion et d'énergie, et un "miroir" dans un état $|\vec{p}, E\rangle$. (On prend un miroir pour qu'une seule collision permette d'extraire l'impulsion normale de la particule). Le hamiltonien total est invariant par translation, et donc l'état final est une superposition linéaire d'états $|\vec{\pi}', \epsilon'\rangle |\vec{p}', E'\rangle$ tels que $\vec{p} + \vec{\pi} = \vec{p}' + \vec{\pi}'$. Par ailleurs si l'interaction a duré δt , ces états vérifient $E + E \sim E' + E' \pm \frac{\hbar}{\delta t}$ (théorie des perturbations dépendant du temps). Dans le meilleur des cas $\pi_{\perp}, \pi'_{\perp}, E, E'$ du miroir sont mesurées exactement. Le système :

$$\begin{cases} p_{\perp} - p'_{\perp} = \pi'_{\perp} - \pi_{\perp} \\ p_{\parallel} = p'_{\parallel} \quad \pi_{\parallel} = \pi'_{\parallel} \\ E - E' \sim E' - E \pm \frac{\hbar}{\delta t} \end{cases}$$

permet d'extraire p_{\perp} et p'_{\perp} avec des précisions telles que

$$\delta p_{\perp} = \delta p'_{\perp} \quad \text{et} \quad v_{\perp} \delta p_{\perp} - v'_{\perp} \delta p'_{\perp} \sim \pm \hbar / \delta t \quad \left(\text{car } \frac{\partial E}{\partial p_{\perp}} = v_{\perp} \right)$$

d'où
$$\boxed{(v_{\perp} - v'_{\perp}) \delta p_{\perp} \sim \hbar / \delta t}$$

Dans la théorie non relativiste v'_{\perp} peut être rendue arbitrairement grande et donc Δp_{\perp} arbitrairement petite.

Autrement $|v - v'| < 2c$ et
$$\boxed{\delta p_{\perp} > \frac{\hbar}{c \delta t}}$$

Dans le domaine relativiste, il n'existe pas de mesure "prédictible" d'impulsion. Plus précisément on ne pourra connaître une impulsion exactement que dans la limite asymptotique où l'impulsion ne varie pas et peut être mesurée pendant un temps δt très long.

Si l'on croit à l'interprétation de Copenhague, selon laquelle à toute observable doit correspondre un type de mesure prédictible, on doit conclure que les impulsions des particules ne peuvent être les variables dynamiques de la théorie quantique relativiste. Ce résultat n'interdit pas de prendre comme observables les valeurs aux différents points de l'espace des champs électromagnétiques ou du courant associé au champ ψ de Dirac. Les dernières mesures sont locales : elles peuvent être effectuées avec des appareils situés au voisinage d'un point de l'espace. L'impossibilité de réaliser des mesures non locales relativistes, comme celles d'impulsion ruinerait effectivement la théorie si, comme le pensèrent Landau et Peierls, il n'y avait aucune autre mesure prédictible dans la théorie quantique relativiste. Heureusement les mesures locales citées

* Heisenberg-Pauli (voir chap. 1 § 3)

sauvent la situation suivant Bohr et Rosenfeld. Aujourd'hui, le problème se pose différemment puisqu'Aharonov et Albert ⁽³⁹⁾ (1981) pensent qu'il existe des mesures non locales relativistes.

4.1.c 2ème cas étudié : mesures du champ électromagnétique

Comme $\vec{E}(\vec{x}, t)$ est une variable dynamique (du moins pour ce qui est du champ de radiation) de la théorie électrodynamique classique, il doit exister, selon Bohr, des mesures "prédictibles" de ce champ dans le cas quantique. En fait, l'opérateur-champ $\vec{E}(\vec{x}, t)$ de la théorie quantique, même dans le cas libre, n'est pas bien défini. (On dirait aujourd'hui qu'il est défini seulement comme distribution à valeur d'opérateur). Seules les moyennes sur des domaines L, T d'extension spatio-temporelle finie L, T : $\bar{E} = \frac{1}{L \cdot T} \int_{L, T} \vec{E}(\vec{x}, t) d^4x$ sont bien définies et c'est à ces "observables" là que doivent être associées des mesures prédictibles.

Il reste à savoir si l'on peut effectivement imaginer un dispositif de mesure qui permette de mesurer \vec{E} avec une précision arbitraire.

Bohr et Rosenfeld donnent une solution utilisant des corps d'épreuve constitués de distributions de charge continues et classiques, et supposent que les sources de champ sont elles aussi classiques (voir §3).

Le point le plus délicat est l'interprétation des fluctuations quantiques de \vec{E} ; bien sûr, celles-ci s'annulent pour les états propres de l'opérateur hermitique \vec{E} , mais la difficulté vient du fait que pour tout état produit par des sources classiques ("états cohérents"), ces fluctuations sont non nulles et indépendantes des sources, donc identiques aux fluctuations du vide:

$$\Delta \bar{E} = \langle \bar{E}^2 - \langle \bar{E} \rangle^2 \rangle \sim \begin{cases} \frac{\hbar c}{L \cdot T} & \text{si } L < cT \\ \frac{\hbar c}{L^2} & \text{si } L > cT \end{cases}$$

(résultat obtenu à partir de l'expression du champ libre en fonction des créateurs et des annihilateurs et de celle des états cohérents).

Par conséquent, le formalisme prédit qu'aucun des états quantiques du rayonnement considéré dans cet approche ne peut conduire à des ^{certaines} mesures de \vec{E} . Néanmoins Bohr et Rosenfeld imaginent un dispositif pour lequel la seule imprécision introduite s'identifie aux fluctuations du vide (voir §3). Ces mesures permettent donc d'extraire toutes les informations contenues dans la théorie : l'enregistrement d'un grand nombre de valeurs de \vec{E} mesurées pour des conditions initiales identiques des sources classiques donne $\langle \bar{E} \rangle$ et $\Delta \bar{E}$.

De plus l'analyse des dispositifs de mesures doubles de champ conduit à une généralisation relativiste de la "répétabilité" : si l'on effectue deux mesures de champ moyen dans deux domaines très voisins de l'espace-temps, les valeurs trouvées doivent être presque identiques puisqu'à chaque partie de la première charge d'épreuve correspond une partie de la seconde soumise presque au même champ. Le sens du mot "répétable" est ici élargi, car la cinématique relativiste interdit en général de préciser l'ordre temporel des mesures.

En résumé, dans le cadre d'étude de Bohr et Rosenfeld (sources classiques) il n'existe pas d'état quantique du champ conduisant à des prédictions certaines sur la valeur du champ. (C'est nous qui soulignons ce point). Mais l'harmonie entre le formalisme et la théorie copenhaguienne de la mesure est maintenue, parce qu'il est, en principe, possible de construire des \bar{E} -mètres satisfaisant l'exigence de "répétabilité" et suffisamment précis pour vérifier toutes les prédictions -alors statistiques- de la théorie.

Cet exposé concis de la problématique de la mesurabilité s'éloigne sensiblement de la démarche adoptée historiquement par Landau, Peierls, Bohr et Rosenfeld, mais il peut faciliter l'abord des textes originaux.

4.2 "GENERALISATION DES RELATIONS D'INCERTITUDE A LA THEORIE QUANTIQUE RELATIVISTE". LANDAU ET PEIERLS ⁽¹⁾ (Mars 1931)

Résumé-commentaire :

"Par l'examen des méthodes possibles de mesure, il est montré qu'aucune des grandeurs physiques qui entrent en jeu dans la mécanique ondulatoire ne reste définissable dans le domaine relativiste".

Deux des difficultés des théories de Dirac ^(2,5), Heisenberg et Pauli ^(3,4) sont d'abord mentionnées :

- le problème des états d'énergies négatives de l'électron de Dirac (il n'est pas question ici de la théorie de la mer de Dirac ⁽⁵⁾ conçue à la fin de 1929).

- le problème de la self-énergie d'une particule chargée.

"Cet échec complet de la théorie prête à croire que dans ce domaine les hypothèses physiques nécessaires à l'application des méthodes de mécanique ondulatoire ne sont plus vérifiées".

Suivent quelques considérations générales sur la mesure en mécanique quantique". "Le sens de toute théorie physique est le suivant : à partir du résultat d'une expérience déduire les résultats d'expériences ultérieures". Cette définition positiviste a au moins le mérite de faire ressortir les propriétés minimales de la mesure : elle doit être "répétable" ou tout au moins "prédictible". Ces deux notions ne sont pas en général équivalentes et pour l'énergie, par exemple, on doit se contenter de la seconde : en vertu de la "4° relation d'incertitude", l'énergie du système étudié après une mesure de durée Δt n'est définie qu'à $\hbar/\Delta t$ près; la répétition donnera des écarts de cette ordre.

Par un argument exposé au §1, Landau et Peierls montrent que dans le cas relativiste, les mesures d'impulsion ne satisfont pas l'exigence de prédictibilité : "le concept d'impulsion n'a de sens précis que pour des temps longs". Notons déjà que l'argument est repris dans le cours de Landau et Lifchitz^(7,8) en faveur des approches du type "matrice S", et qu'il a été approuvé par Bohr^{***}

L'étude des mesures de champ est reliée à la précédente : si la mesure est faite à l'aide d'une charge test q soumise au champ \vec{E} pendant le temps δt , à une imprécision δp sur la mesure des impulsions initiale et finale de la charge test correspondra une imprécision δE sur la mesure de E , avec $q\delta E\delta t > \delta p$. Pour une particule chargée, en plus de l'imprécision $\delta p > \frac{h}{(v' - v)\delta t}$, on a une incertitude due à la Bremsstrahlung : $\delta p > \frac{q^2}{3\delta t} (v' - v)$ car l'énergie rayonnée par unité de temps est $\frac{dE_{rad}}{dt} = -\frac{2}{3} q^2 \frac{\dot{v}^2}{4\pi c^3}$ selon Larmor⁽⁹⁾. On en déduit, en faisant le produit des deux inégalités, $\delta p > \frac{h}{c\delta t} \sqrt{\frac{q^2}{\hbar c}}$ et donc $\delta E > \frac{V\hbar c}{(c\delta t)^2}$. Conclusion : "dans le domaine de la mécanique quantique les champs ne sont pas des grandeurs mesurables". Notons que la déduction de la dernière inégalité suppose que la charge q est ponctuelle (formule de Larmor), comme Bohr le souligna plus tard.

Nous omettons la discussion sur la détection des quanta lumineux car elle fait appel au résultat précédent, réfuté par Bohr et Rosenfeld. D'ailleurs les vérifications ultérieures à haute énergie de la formule de Klein-Nishima (voir Bohr; Solvay 1933) montrèrent bien la possibilité expérimentale de dépasser ces

* dont la paternité revient pour l'essentiel à Niels Bohrs, cité plusieurs fois dans ce papier (6). L'autorité du maître de Copenhague est particulièrement évidente ici.

*** information due à Peierls

incertitudes.

La localisation d'une particule d'énergie E dans le temps et dans l'espace ne peut se faire à mieux que $\delta t > \frac{h}{E}$ et $\delta q > \frac{hc}{E}$ comme le montre l'étude d'une expérience de diffusion.

Quel est le lien entre les relations d'incertitude précédentes et le formalisme ?

- Pour l'impulsion, il est inexistant car la théorie la traite comme une grandeur physique analogue à la grandeur correspondante dans la théorie non relativiste, et elle devrait donc faire l'objet de mesures "prédictibles".

- Pour les photons, il ne peut pas exister d'expression de la densité de probabilité (l'usage de l'espace des configurations est un simple prolongement formel). En effet, la fonction d'onde devrait être un tenseur de rang deux à cause des propriétés de polarisation⁽¹⁰⁾ et à partir d'un tel tenseur, on ne peut pas construire un tenseur quadratique de rang un qui serait le courant de probabilité *.

- Pour les électrons, la théorie de Dirac montre que l'on ne peut les localiser à mieux que $\lambda_c = h/mc$ (un paquet d'ondes constitué d'états d'énergie positive de l'équation de Dirac ne peut être plus étroit que λ_c ***).

- Pour le champ électro-magnétique, les fluctuations du vide divergent. Contrairement à l'avis des auteurs, ce résultat est sans conséquence sur la mesurabilité des champs, puisque formellement il existe des états du champ qui annulent les fluctuations de \bar{E} .

"En conclusion, il n'existe pas de mesures "prédictibles" pour les grandeurs fondamentales de la mécanique ondulatoire (sauf quand elles varient peu dans le temps)". La parenthèse deviendra l'essentiel dans l'introduction du cours de théorie quantique relativiste⁽⁸⁾ (1967). Les auteurs poursuivent : "On ne doit donc pas s'étonner que le formalisme conduise à toutes sortes d'infinités". "Dans la théorie future, il n'y aura pas de "grandeur physique" au sens de la mécanique ondulatoire". La théorie donnera seulement la probabilité de trouver tel ou tel résultat dans une expérience bien définie. Les auteurs sont ici proches des conceptions "observationnistes" de Heisenberg⁽¹¹⁾ (1938).

* L'argument est raffiné dans le cours de Landau et Lifchitz tome 4 §4

*** Voir par exemple Réf I p.60

A la fin de l'article, l'interprétation du caractère continu des spectres est plutôt malheureuse: pour les auteurs, les spectres β sont continus alors que les α sont discrets parce que les particules α "dans le noyau" ne sont pas relativistes, alors que les β le sont, et qu'ils sont soumis à une incertitude supplémentaire sur l'énergie. Cette interprétation s'apparente à celle que Bohr maintint avant qu'un crédit universel ne fût accordé au neutrino de Pauli (Voir chap. 6 § 1)

Les conclusions de Landau et Peierls sont surtout négatives : il faut rejeter la théorie existante, et rien n'est dit que de très vague sur la théorie future. Elles deviendront plus positives lors de leur utilisation ultérieure pour justifier l'étude de la matrice S. En 1960, Landau écrit : "Il y a presque trente ans, Peierls et moi-même avons remarqué que dans le domaine de la théorie quantique relativiste aucune quantité concernant des particules en interaction ne peut être mesurée, et les seules quantités observables sont les quantités de mouvement et les polarisations de particules libres : en conséquence, si nous ne voulons pas introduire de quantités inobservables, nous ne pouvons introduire dans la théorie comme quantités fondamentales que les amplitudes de diffusion " (13).

Au regard des réactions de Bohr, il faut distinguer dans cet article deux parties * : les limitations sur la mesure de l'impulsion que Bohr approuva, et celles sur la mesure du champ électromagnétique ainsi que leur lien supposé avec les fluctuations du vide, que Bohr et Rosenfeld rejetèrent complètement. On peut se demander pourquoi Bohr considérait moins gênant l'abandon de mesures instantanées de l'impulsion. A défaut de commentaires précis de Bohr et Rosenfeld, on peut supposer que ces auteurs avaient l'espoir de voir naître une théorie relativiste quantique où l'impulsion et la position d'une particule ne soient pas les variables dynamiques les plus fondamentales. La théorie relativiste de l'électron était déjà boîteuse en raison de la difficulté formelle des états d'énergie négative de l'équation de Dirac; il n'était pas très étonnant que les concepts d'impulsion et de position utilisés dans cette équation soient mal définis. En revanche la théorie du champ électromagnétique quantique libre sous la forme que lui avaient donné Jordan et Pauli, semblait parfaitement cohérente, et on devait pouvoir justifier la mesurabilité de ce champ libre.

* Peierls, conversation privée.

4.3 "DU PROBLEME DE LA MESURABILITE DES CHAMPS ELECTROMAGNETIQUES".
BOHR - ROSENFELD (1933) (14,15)

Pour introduire cet article nous ne pouvons faire mieux que de citer le début de l'article de Rosenfeld "On Q.E.D." (1955)⁽¹⁶⁾

"...et discors concordia fetibus apta est."

Ovide, metam. I,433

"Quand j'arrivai à l'Institut le dernier jour de février 1931 pour mon séjour annuel, la première personne que je vis fut Gamow. Comme je lui demandais les nouvelles, il répondit à sa manière pittoresque, en me montrant un joli dessin à la plume qu'il venait juste de faire[†]. Ce dessin représentait Landau solidement ligotté à une chaise et baillonné, tandis que Bohr, debout devant lui avec l'index levé, disait : "Bitte, bitte, Landau, ^{††}muß ich nur ein Wort sagen!" J'appris que Landau et Peierls étaient venus quelques jours avant avec un nouveau papier qu'ils voulaient montrer à Bohr, "mais" (ajouta Gamow d'un air dégagé) "il ne semble pas approuver - et voilà le genre de discussion qui a continué tout le temps". Peierls était parti la veille, "dans un état d'épuisement complet", dit Gamow. Landau resta quelques semaines de plus, et j'eus l'occasion de vérifier que la représentation de Gamow de la situation n'était exagérée que dans la mesure normalement concédée à l'imagination artistique".

"Il y avait en effet raison de s'exciter, car le point soulevé par Landau et Peierls était très fondamental. Ils remettaient en question la structure logique de Q.E.D. en prétendant que le concept même de champ électromagnétique n'est susceptible en théorie quantique d'aucune détermination physique au moyen de mesures ..."

[†] Je crains que cette oeuvre d'art n'ait péri avant même que sa valeur historique pût être reconnue.

^{††} C'est là un danicisme familier de Bohr pour "darf ich".

Le caractère "fondamental" du problème justifia deux années de labeur assidu (le manuscrit fut présenté à une réunion de l'Académie danoise le 2 décembre 1932) débouchant sur quarante quatre pages de discussions et démonstrations subtiles. Doutant que beaucoup de lecteurs aient le courage de consulter le texte original, nous présentons un résumé simplifié, émaillé de quelques citations.

(1) Introduction

Suivant le principe de correspondance, la version quantique de la théorie électromagnétique ne devrait être rien d'autre qu'une "réinterprétation de la théorie classique adaptée à l'existence du quantum d'action", d'où le caractère choquant des conclusions de Landau-Peierls.

Après que Rosenfeld lui eut enseigné les bases de la quantification des champs, Bohr fit une première remarque : il est essentiel de séparer la théorie quantique du champ électromagnétique, de la théorie atomique. On doit d'abord analyser la mesurabilité du champ électro-magnétique en traitant classiquement les sources de champ ; "l'interprétation physique du formalisme quantique doit être fondée en dernière analyse sur des idéalizations classiques"⁽¹⁷⁾ (Rosenfeld).

Mettant de côté la structure atomique de la matière, on peut ne considérer que des distributions de charge continues, ce qui exclut l'effet dramatique des réactions de rayonnement sur les charges ponctuelles; et ainsi on donne un sens aux champs singuliers de la théorie en ne considérant plus que leur moyenne sur des régions étendues de l'espace-temps.

(2) Mesurabilité des champs selon la théorie quantique

Pour effectuer commodément les moyennes temporelles, Bohr et Rosenfeld utilisent la représentation de Heisenberg et les relations de commutation du champ libre correspondantes dues à Jordan et Pauli⁽¹⁸⁾.

On a par exemple $[E_x^1, E_x^2] = i\hbar (A_{xx}^{1,2} - A_{xx}^{2,1})$
avec $A_{\alpha\beta}^{1,2} = -(\frac{\partial^2}{\partial x_\alpha \partial x_\beta} - \frac{1}{c^2} \frac{\partial^2}{\partial t_1 \partial t_2}) \frac{1}{r_{12}} \mathcal{J}(t_2 - t_1 - \frac{r_{12}}{c})$, ce qui donne pour les moyennes \bar{E}_x^I et \bar{E}_x^{II} dans des domaines I et II d'espace-temps

$$\Delta \bar{E}_x^I \Delta \bar{E}_x^{II} \sim \hbar | \bar{A}_{xx}^{I,II} - \bar{A}_{xx}^{II,I} | \quad *$$

Si les domaines I et II se recouvrent partiellement sur une longueur L et sur

un temps T $\sqrt{\Delta \bar{E}_x^I \Delta \bar{E}_x^{II}} \sim \frac{\hbar c}{LcT}$ si $L < cT$
 $\sim \sqrt{\frac{\hbar}{L^3 T}}$ si $L > cT$

Et si le domaine I tend vers le domaine II alors $\bar{A}_{xx}^{I,II} \rightarrow \bar{A}_{xx}^{II,I}$

et donc $\sqrt{\Delta \bar{E}_x^I \Delta \bar{E}_x^{II}} \rightarrow 0$

ce qui est l'expression théorique de la "répétabilité" des mesures de champ.

* Notons qu'il serait plus correct d'écrire $\Delta \bar{E}_x^I \Delta \bar{E}_x^{II} \geq \frac{1}{2} \hbar | \bar{A}_{xx}^{II,I} - \bar{A}_{xx}^{I,II} |$ (inégalité rigoureuse si $\Delta \bar{E}_x^I = \langle \bar{E}_x^2 \rangle - \langle \bar{E}_x \rangle^2$ d'autant plus que pour I=II et pour les états cohérents le membre de gauche vaut \sum_c^2 et celui de droite 0.

Suit le calcul des fluctuations $\Delta \bar{E}_\alpha = \langle \bar{E}_\alpha^2 \rangle - \langle \bar{E}_\alpha \rangle^2$, d'abord pour un état correspondant à un nombre déterminé de photons de chaque espèce. Alors

$$\Delta \bar{E}_\alpha > S_0, \quad S_0 \text{ fluctuations du vide. } S_0 \sim \frac{\sqrt{\hbar c}}{LcT} \quad \text{si } L < cT$$

$$\sim \frac{\sqrt{\hbar c}}{L} \quad \text{si } L > cT$$

Pour un champ créé par des sources classiques, on trouve que les états quantiques sont des états cohérents (dans la terminologie moderne) pour lesquels la distribution des photons dans chaque mode est poissonnienne : $w(m) = \frac{m_0^m e^{-m_0}}{m!}$. La formule est ici établie en supposant que les processus d'émission sont statistiquement indépendants. D'après Rosenfeld, Bohr et Pauli furent les premiers à reconnaître cette propriété*. Pour les états cohérents, les fluctuations sont identiques aux fluctuations du vide S_0 .

(3) Présuppositions pour les mesures de champ

On utilise un corps d'épreuve de volume V, de densité de charge homogène ρ , et soumis au champ \vec{E} pendant le temps T si bien que son impulsion passe de

$$\vec{p}' \text{ à } \vec{p}'' \quad p''_\alpha - p'_\alpha = \rho \bar{E}_\alpha VT$$

Un premier type d'incertitude sur la mesure de \bar{E}_α vient de l'incertitude sur la mesure des impulsions. On a choisi le corps suffisamment lourd pour que son déplacement pendant le temps T sous l'effet de \bar{E}_α soit négligeable; le seul déplacement Δx qui subsiste est celui provoqué par les mesures d'impulsions. S'il est vrai que $\delta p_\alpha \sim \frac{\hbar}{\Delta x}$, alors il en résulte un $\delta \bar{E}_\alpha \sim \frac{\hbar}{\rho \Delta x VT}$ qui peut être rendu petit en prenant un ρ grand.

Mais il est loin d'être évident que l'on a effectivement $\delta p_\alpha \sim \frac{\hbar}{\Delta x}$ comme seule incertitude sur p_α .

Une autre incertitude sur la mesure de p_α vient de la Bremsstrahlung.

$$\delta' p_\alpha \sim \frac{q^2}{c^3} \frac{d\alpha}{(\delta t)^2} \quad (\delta t \text{ temps de mesure de } p_\alpha) \text{ si l'on suppose comme}$$

Landau et Peierls que c'est une charge ponctuelle qui rayonne.

Alors $\delta \bar{E}_\alpha = \frac{1}{qT} (\delta' p_\alpha + \delta p_\alpha) \sim \frac{\hbar}{qT \Delta x} + \frac{q \delta \alpha}{c^3 T \delta t^2} \geq \frac{\sqrt{\hbar c}}{c^2 T \delta t}$ indépendamment de $q = \rho V$, ce qui est le résultat de Landau-Peierls avec $\delta t = T$. En fait, ce résultat est incorrect pour une charge étendue : on a alors un courant réel

$$\rho \frac{\delta x}{\delta t} \quad \text{du même ordre de grandeur que le courant de déplacement } \frac{\delta \bar{E}_\alpha^R}{\delta t}$$

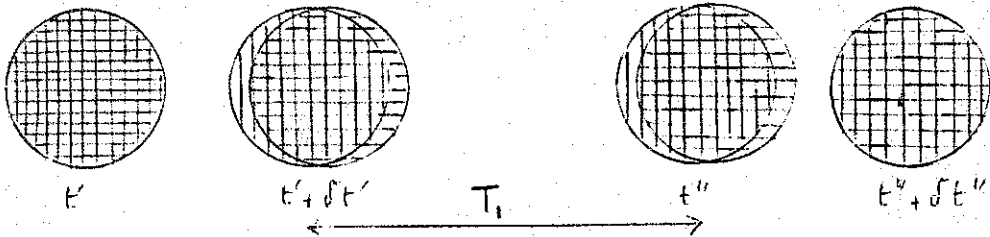
* On trouvera la démonstration moderne due à Thirring et Touschek (19) (entre autres) dans Ref I p 163 et suivantes.

(E_x^R rayonné) d'où $\delta p_x' = \rho v E_x^R \delta t = \rho^2 v \delta x \delta t$. A condition de prendre δt suffisamment petit, on aura $\delta p_x' \ll \delta p_x$ et pourra ainsi négliger l'effet de Bremsstrahlung.

A priori, une autre cause d'incertitude vient de la nécessité de décomposer le corps d'épreuve en parties de petites dimensions : le cadre relativiste de l'étude exclut de considérer un corps solide, étant donné la vitesse finie de propagation de toute action. Bohr et Rosenfeld montrent qu'il est possible de faire une mesure de p_x au cours de laquelle le déplacement δx de toutes les parties est le même (l'exemple pris est l'effet Doppler), et pour laquelle on a bien $\delta p_x \sim \frac{h}{\delta x}$

(4) Calcul des effets de rayonnement du corps d'épreuve

Il est essentiel de pouvoir calculer le champ rayonné par le corps d'épreuve puisqu'il va contribuer au champ mesuré. Avant la mesure le corps d'épreuve est attaché à un corps neutralisant. A l'instant t' , toutes les parties du corps d'épreuve sont simultanément détachées du corps neutralisant, puis une première détermination de l'impulsion est effectuée entre t' et $t' + \delta t'$: pour cela, on mesure la valeur de l'impulsion par diffusion Compton, et on ramène à zéro cette valeur par une diffusion dans l'autre sens. Lors de cette détermination, le corps d'épreuve a été globalement déplacé de δx . Pendant le temps T où agit \bar{E}_x , ce corps est suffisamment lourd pour que ce déplacement δx reste constant; puis on fait la 2ème mesure d'impulsion à t'' , et rattache le corps d'épreuve au corps neutralisant à $t'' + \delta t''$.



Dans cette situation, le calcul classique du champ rayonné est très simple et donne pour la valeur moyenne dans le corps II du champ créé par I : $\bar{E}_x^{II} = \int \alpha_1 \rho v T, \bar{A}_{\alpha\beta}^{II}$ où le $A_{\alpha\beta}^{II}$ est le même que celui du (2). Cette coïncidence (16), découverte au début de leur investigation, fit croire à Bohr et à Rosenfeld qu'ils avaient mis le doigt, par chance, sur la bonne méthode de mesure, soumise aux limitations prévues par la théorie. Aujourd'hui,

on sait bien qu'il ne s'agit pas vraiment d'une coïncidence : $E_{\alpha}^{I,II}$ rayonné par une distribution singulière de polarisation est simplement relié à la fonction de Green de l'équation de propagation, par ailleurs le commutateur $[E_{\alpha}^I, E_{\alpha}^II]$ des champs libres est aussi relié à la même fonction de Green (20).

(5) Mesure d'une seule moyenne de champ

Si l'on sépare le champ extérieur \bar{E}_{α}^I et le champ $\bar{E}_{\alpha}^{I,II}$ créé par le corps I, on a $P_{\alpha}^{I''} - P_{\alpha}^{I'} = \rho_1 V_1 T_1 (\bar{E}_{\alpha}^I + \bar{E}_{\alpha}^{I,II})$

Si l'on suppose que $\bar{E}_{\alpha}^{I,II}$ introduit une erreur incontrôlable sur la mesure de \bar{E}_{α}^I on a $\delta \bar{E}_{\alpha}^I \sim \frac{\delta P_{\alpha}^I}{\rho_1 V_1 T_1} + \rho_1 V_1 T_1 |\bar{A}_{\alpha\alpha}^{I,II}| \delta \alpha_1 \geq \sqrt{\hbar} |\bar{A}_{\alpha\alpha}^{I,II}| \sim \sqrt{\frac{\hbar}{L_1^3 T_1}}$

On aurait alors une limite absolue à la mesurabilité de \bar{E}_{α}^I et rejoindrait la conclusion de Landau et Peierls. L'astuce consiste à introduire un ressort entre le corps-test \bar{E}_{α}^I le corps neutralisant fixe, de raideur $K_{\alpha} = \rho_1^2 V_1^2 T_1 |\bar{A}_{\alpha\alpha}^{I,II}|$ qui compense exactement l'effet de $\bar{E}_{\alpha}^{I,II}$ (proportionnel à $\delta \alpha_1$). Alors \bar{E}_{α}^I peut être mesuré avec une précision arbitraire. Il reste une dernière difficulté: a-t-on le droit d'écarter les fluctuations quantiques de $\bar{E}_{\alpha}^{I,II}$? Les auteurs donnent dans ce paragraphe une justification d'ordre général, qui sera précisée en (6) : "... toute mesure de quantités physiques, par définition, doit être une question d'application de concepts classiques; et par conséquent, lors des mesures de champ, toute considération des limitations de la "stricte" applicabilité de l'électrodynamique classique serait en contradiction avec le concept de mesure même".

(6) Mesurabilité de moyennes doubles d'une composante de champ

En utilisant les mêmes ressorts que dans (5°) pour les corps test I et II on obtient $\Delta \bar{E}_{\alpha}^I \Delta \bar{E}_{\alpha}^{II} \sim \hbar [|\bar{A}_{\alpha\alpha}^{I,II}| + |\bar{A}_{\alpha\alpha}^{II,I}|]$ au lieu de la limite $\hbar |\bar{A}_{\alpha\alpha}^{I,II} - \bar{A}_{\alpha\alpha}^{II,I}|$ donnée par le formalisme. Arrivé à ce stade, Bohr alla jusqu'à suspecter les règles de commutation du champ libre et reprit plusieurs fois en détail leur déduction (16). La remarque suivante fournit la clé du problème : les deux expressions ne diffèrent que lorsque qu'il existe des intervalles du genre temps entre I et II. Cette circonstance permet de corrélérer les deux mesures à l'aide d'échanges de signaux lumineux, et de l'introduction de nouveaux ressorts, si bien que l'on obtient la valeur théorique.

Toutes ces conclusions sont soumises à une interprétation convenable des fluctuations quantiques : les fluctuations des champs créés par le corps d'épreuve devraient être considérées comme "partie intégrante du champ mesuré". Cette hypothèse n'altère nullement les prédictions statistiques de la théorie (la vérification de la valeur de $\Delta \bar{E}_\nu$ et $\langle \bar{E}_\nu \rangle$) puisque les fluctuations du champ total, toujours égales aux fluctuations du vide, ne sont pas affectées par l'existence du champ de rayonnement du corps-test. Bohr et Rosenfeld vont encore plus loin : "... nous obtenons une interprétation cohérente de toutes les conséquences de la théorie quantique des champs, si, dans une nécessaire généralisation du concept de mesure, nous interprétons les résultats de mesure obtenus par les dispositifs décrits, comme les moyennes de champ désirées". Soulignons qu'il s'agit bien d'une généralisation, puisqu'il n'existe pas, dans ce cadre, d'état du champ conduisant à des mesures certaines.

Quant aux mesures doubles de champ, les auteurs montrent qu'elles donnent des valeurs identiques quand I \rightarrow II indépendamment des corrections par des ressorts des effets du champ rayonné. En effet, les deux corps-test sont alors soumis pratiquement au même champ, quelque soit son origine. Donc "la répétabilité" -généralisée ici à une situation sans ordre temporel bien défini- n'est pas affectée par la nature, classique ou quantique, du champ rayonné par les corps-test. Curieusement (selon nous), ce type de corrélation dans une mesure double n'est pas lié à la modification de l'état quantique par la mesure, puisqu'après interaction avec une charge d'épreuve classique, on a toujours des états cohérents, qui ne sont pas des états propres ; la corrélation s'exprime ici par un argument de continuité : on trouve presque le même résultat, quand on mesure presque la même chose.

(7) Mesurabilité de moyennes doubles de composantes différentes du champ :

La discussion est très analogue au 6°.

(8) Conclusion :

Les idées de base de l'introduction sont reprises avec un nouveau commentaire (bien naturel) : "Néanmoins il semble que le caractère quelque peu compliqué des considérations utilisées pour démontrer l'accord du formalisme

et de la mesurabilité ne soit guère évitable", en raison de la forme intégrale du formalisme, et de la nécessité de donner une généralisation relativiste à la notion de complémentarité.

L'étape suivante dans l'analyse de Bohr et Rosenfeld est le problème de la mesurabilité des charges et des courants, susceptible, comme nous allons le voir, d'un traitement analogue à celui utilisé pour le champ électromagnétique.

4.4 MESURABILITE DES CHARGES ET DES COURANTS

En juillet 1934, HEISENBERG prononça une conférence sur les "fluctuations de charge liées à la création de matière à partir du rayonnement" (23). Cet article, suivant de peu une étude de la polarisation du vide par Heisenberg (25) fut cité à plusieurs reprises par Bohr et Rosenfeld dans leurs publications ultérieures, et aussi dans des études de problèmes reliés comme l'énergie propre de l'électron, ou la polarisation du vide.

La théorie de la mer de Dirac (5') rend possible la création de paires à partir du rayonnement (et le positron fut effectivement découvert en 1933 par Anderson). "Chaque fois que la mesure d'un processus physique nécessite de grands champs magnétiques, il faut tenir compte d'une perturbation imprévue de l'objet examiné par le dispositif d'observation, plus précisément de la création de matière par l'appareil de mesure". Le calcul de Heisenberg (voir l'appendice) des fluctuations de la charge à l'intérieur d'un certain volume, montre que ces fluctuations divergent si le volume est abruptement délimité (dans le cas du champ électro-magnétique, il faut en plus que le volume tende vers zéro). Une délimitation abrupte du volume de mesure correspond formellement (voir l'appendice) à l'introduction d'un potentiel électromagnétique classique discontinu, et donc à un champ associé divergent susceptible de créer une infinité de paires.

Ce phénomène représente une nouvelle limitation de l'idée de particules individuelles : plus on essaye de localiser la "particule", et plus la charge fluctue. Il donne d'un point de vue pratique, un rôle plus important aux fluctuations au cours des mesures (par la pensée) de charge, mais comme nous allons le voir, il n'altère pas fondamentalement la discussion de la mesurabilité.

Après leur étude de la mesurabilité des champs électromagnétiques, BOHR et ROSENFELD s'intéressèrent au problème analogue pour la charge électrique. Les résultats préliminaires ne furent pas publiés (en tout cas pas avant 1948 ⁽²⁶⁾) et c'est seulement en 1950, après les travaux de Tomonaga, Schwinger et autres sur la renormalisation, que parut dans Phys. Rev. ⁽²⁷⁾ un article faisant le point de la situation. Tout d'abord, les auteurs résument le cas du champ électro-magnétique. L'extension au cas des mesures de charge et de courant ne pose aucune difficulté particulière, comme nous allons le voir.

"En dernière analyse, les distributions de courant et de charge dues à la présence de particules sont définies par les champs qu'elles produisent". La moyenne $\frac{1}{R} \int_R j_\mu(x) d^4x$ est obtenue en mesurant le flux $\oint_S f_{\mu\nu} d\sigma^\nu$, S étant la surface qui limite le domaine R d'espace-temps. Dans le cas de la mesure des charges moyennes, on dispose sur S des corps d'épreuve analogues à ceux du § 3. Cette fois, l'effet des fluctuations des champs de corps d'épreuve se réduit à zéro, puisque le flux du champ électromagnétique quantique libre est nul à travers une surface fermée. Dans la première étape d'approximation, on traite classiquement les sources de champ. Le champ électromagnétique produit par la coquille de corps d'épreuve est entièrement confiné à l'intérieur de la coquille*, et donc deux mesures de charge dans des régions différentes n'interfèrent pas, en conformité avec la nature classique supposée des charges.

Le vrai problème commence quand on ne néglige plus le phénomène de polarisation du vide qui apparaît à l'ordre suivant en α . Alors la théorie prédit d'une part l'existence de fluctuations de charge ⁽²³⁾, d'autre part, une valeur non nulle de commutateur $[J_\mu, J_\nu]$ faisant intervenir les longueurs et temps caractéristiques associés à l'électron, linéairement en fonction du champ électro-magnétique perturbateur des corps d'épreuve et donc linéairement en fonction de leur déplacement. Ainsi on peut compenser ses effets sur la mesure de la charge moyenne dans un domaine de l'espace, aux fluctuations quantiques de la charge près, par une méthode analogue aux compensations de champ e.m.. Le problème des mesures doubles se ramène aussi à une limitation conforme à la valeur théorique de $[J_\mu, J_\nu]$. Les conclusions générales relatives aux mesures de champ se répètent donc pour les mesures de charge (et de courant).

* à supposer que l'on ait introduit un mécanisme qui égalise tous les déplacements normaux des corps tests. Voir Rosenfeld (16) p. 80 et 82.

La seule différence est d'ordre pratique : les fluctuations de charge divergent si l'épaisseur de la coquille de corps d'épreuve devient petite par rapport à la longueur d'onde Compton; par conséquent cette longueur fixe une échelle au-delà de laquelle les valeurs des charges deviennent de moins en moins accessibles, quoique théoriquement mesurables (Rosenfeld ⁽¹⁶⁾ p. 83).

L'article s'achève par une réflexion sur l'utilisation de ces conclusions pour la mesure en physique nucléaire : on ne peut appliquer le raisonnement précédent que si les objets étudiés se comportent conformément à une théorie des champs locale analogue à l'électrodynamique quantique. A une époque où les forces nucléaires à courte portée restaient inexplicables, on pouvait effectivement en douter.

En 1955, à l'occasion du soixante-dixième anniversaire de Bohr, Rosenfeld publia un essai ⁽¹⁶⁾ intitulé "De l'électrodynamique quantique". Dans les deux premières parties, que nous avons plusieurs fois citées, il donne un compte-rendu particulièrement clair des problèmes de mesurabilité, sans ajouter d'éléments nouveaux à l'analyse. Les deux dernières parties, sur la structure formelle de Q.E.D. et sur le couplage métrique sortent du cadre de cette étude.

Le premier reproche que l'on peut faire à Bohr et à Rosenfeld est celui de la complexité ("inévitable") des dispositifs qu'ils envisagent. "... on doit tant perturber l'espace, en mettant tant de choses dedans, qu'on ne sait pas jusqu'à quel point cela peut être considéré comme une mesure. Mais le principal objet, naturellement, était de voir s'il pouvait y avoir une difficulté pour les applications ultérieures" dit Peierls *

Une autre difficulté est liée au traitement perturbatif de l'interaction charge-rayonnement. Bien que α soit petit, l'énergie d'interaction entre matière quantique et rayonnement diverge : "l'énergie de couplage s'avère infiniment grande, si bien que cela n'a pas de sens, de partir de systèmes

* Extrait d'une conversation avec Peierls

découplés"⁽²¹⁾ (Rosenfeld). Bohr avait l'espoir (voir par exemple Congrès de Rome 1931⁽⁴⁵⁾) que la théorie future, libérée du problème des divergences, leverait ces difficultés et donnerait une valeur faible pour les corrections en α^n . "Ce n'est qu'après la mise en place du schéma de renormalisation, que l'on put donner un sens rigoureux à un développement en α " comme le remarquent Bohr et Rosenfeld dans l'article de 1950⁽²²⁾. La justification physique de ce traitement par étape est donnée en partie par les études des mesures de charge mais les champs et les charges mesurées sont alors des champs renormalisés, sans rapport direct avec les variables canoniques du problème non couplé. La renormalisation rend les quantités observables plus éloignées des ingrédients primitifs de la théorie. Finalement Bohr et Rosenfeld peuvent établir la mesurabilité de certaines observables locales, mais ces observables diffèrent considérablement de celles (liées aux champs nus) introduites par une application naïve du principe de correspondance*.

Une autre critique plus fondamentale (nous y reviendrons) s'attaquerait aux conceptions de Bohr de la mesure, et réclamerait un traitement quantique de l'appareil de mesure, en l'occurrence des charges d'épreuve; en d'autres termes, où faut-il placer la frontière entre l'objet quantique et l'observateur?

Une version faible des conclusions de la discussion serait : dans l'idéologie de Copenhague tout au moins, on peut envisager sans contradiction logique qu'il existe des observables locales (mesurables par définition) associées aux champs fondamentaux d'une théorie quantique des champs. Ce résultat n'est pas sans conséquence importante immédiatement vérifiable par l'expérience : s'il existe des observables locales, la "micro-causalité" relativiste impose qu'elles commutent lorsqu'elles sont séparées par des intervalles du genre espace, et c'est là une des hypothèses fondamentales qui conduisent au théorème spin-statistique et au théorème PCT (Réf. I p. 107). (Nous y reviendrons au chap. 5 § 3).

* Quand on introduit dans le lagrangien de départ les contre-termes de renormalisation, on peut alternativement :

- les considérer comme une partie du lagrangien total réécrit différemment; alors le principe de correspondance est violé en ce que les champs intervenant dans le lagrangien total (champs nus) ne sont pas observables.

- les considérer formellement comme une modification du principe de correspondance permettant de passer de l'expression d'un lagrangien, aux amplitudes de probabilité de la théorie quantique.

4.5 L'ELECTRON DE DIRAC ET LA PHYSIQUE DU NOYAU FACE A LA CRITIQUE DE BOHR

Bohr lui-même remit en question l'application de la mécanique quantique ordinaire à la physique subatomique. Le succès du principe de complémentarité se trouvait confirmé par l'article de Bohr - Rosenfeld montrant qu'il pouvait s'étendre aux mesures de champs quantiques. Bohr continua d'utiliser la complémentarité et la conception de la mesure qui en résultait, comme guide dans la discussion et l'élaboration de nouvelles théories. L'idée de complémentarité reposait sur une base philosophique suffisamment générale pour qu'elle pût s'accomoder de situation très diverses. Toutefois, l'approche de Bohr fut moins heureuse dans ses applications à la théorie de l'électron ou par exemple à la désintégration β . Pour la théorie de l'électron, il est certain que la conception de Dirac de la physique théorique fut plus fertile; contrairement à Bohr, Dirac préférait développer d'abord abstraitement une théorie mathématique, et essayer seulement en second lieu de donner une interprétation physique à cette théorie.

Bohr appliqua le type de raisonnement utilisé pour discuter la mesurabilité du champ électromagnétique, à la mesure du moment magnétique de l'électron de Dirac. Un "Stern et Gerlach" ne pouvait polariser un électron de Dirac, et un magnétomètre ne pouvait isoler la partie du champ due au dipôle magnétique, si bien que la notion de moment magnétique d'un électron libre perdait son sens (rapport de Bohr à Rome ⁽⁴⁵⁾). Mott, au congrès de Rome de 1931 ⁽⁴⁵⁾, approuva Bohr : *"Nous pouvons dire par conséquent qu'aucune signification ne peut être attachée à l'affirmation qu'un électron libre a un moment magnétique."* Mais, au même congrès, Mott présentait la possibilité de polariser un faisceau d'électrons par diffusion de Mott (à travers une cible nucléaire), et Rupp en apportait une preuve expérimentale (fausse? voir chap. 2 § 2). Aujourd'hui, non seulement tout théoricien accorde un sens au moment magnétique de l'électron libre, mais encore il sait que sa mesure atteint une précision fantastique. Pourtant les raisonnements de Bohr étaient justes; ce que déniait Bohr, c'était la possibilité de mettre en évidence la polarisation d'un électron libre à l'aide de dispositifs classiques. La diffusion de Mott implique un phénomène quantique d'interaction entre le faisceau électronique et la cible. Sur cet exemple, il nous semble clair que la conception de Bohr de la mesure est trop restrictive : ce que la plupart des physiciens conviennent en général d'appeller instrument de mesure (par exemple, l'ensemble du dispositif de mesure

"directe" du $g-2$) nécessite une théorie quantique. Seulement une partie de l'appareil de mesure, celle directement perçue par l'observateur, devra être considérée comme un objet classique, si l'on veut garder une conception de la mesure en harmonie avec le principe de complémentarité.

Si le spin de l'électron n'est pas susceptible de mesure par un appareil "classique", c'est parce qu'il n'a pas d'analogue de correspondance. Les difficultés d'application naïve du principe de correspondance, conduisirent Bohr à douter de la validité de la mécanique quantique ordinaire dans le domaine subatomique. Elles apparaissaient déjà, selon Bohr, avec le problème des divergences ultra-violettes. La théorie classique de l'électron, ambiguë pour les distances de l'ordre du rayon classique de l'électron, ne pouvait, dans ce domaine, constituer une base de correspondance correcte. De plus la physique nucléaire apportait, à cette échelle de longueur, toutes sortes de phénomènes nouveaux sans antécédents dans les anciennes théories classiques : *"La seule différence caractéristique entre le problème de la constitution des noyaux et la théorie de la constitution des atomes est que dans le premier cas, par opposition au dernier, nous ne pouvons a priori tirer des lois de l'électromagnétisme classique aucun renseignement sûr au sujet des forces agissant entre les particules nucléaires, mais que toutes nos déductions concernant ces forces reposent sur un ensemble d'expériences entièrement nouvelles."*

Le rayon classique de l'électron devait tracer la limite de la pertinence des concepts de la théorie quantique usuels : *"Que l'on regarde le proton comme formé d'un neutron et d'un positron, ce qui d'après les dernières expériences paraît bien être l'hypothèse la plus naturelle, ou bien qu'on le regarde comme le produit d'une dissociation de neutrons accompagnée de la libération d'un électron, il s'agit là de processus qui ne peuvent pas être décrits par les moyens actuels, et dont la possibilité doit être cherchée dans le fait que les dimensions empiriquement connues du neutron sont du même ordre de grandeur que le diamètre de l'électron, lequel exprime la limite à partir de laquelle les concepts de la théorie classique de l'électron et leur utilisation d'après la méthode de correspondance nous font complètement défaut."*
(Bohr; Solvay 1933).

Bohr avait cru voir une autre limite d'application de la méthode de correspondance dans le problème des états d'énergie négative et dans le paradoxe de Klein, et il rejeta, à la fin de 1929, dans une lettre à Dirac, la proposition

de la mer de Dirac. Après la découverte du positron, il considéra plutôt la nouvelle théorie de Dirac comme un raffinement de la méthode de correspondance, grâce auquel la longueur d'onde Compton h/mc n'était plus la limite d'application de la théorie quantique de l'électron, la seule limite subsistant étant celle, λ_c , attendue suivant l'argument de correspondance. Bien après cet épisode dont Dirac était sorti victorieux, Bohr continua à soutenir que les phénomènes nucléaires pouvaient violer les principes de la mécanique quantique ordinaire, en l'absence de base de correspondance claire. La désintégration et le problème de la statistique des noyaux (voir chap. 6 § 1) étaient, selon Bohr, ce point de vue : la conservation de l'énergie et la statistique de l'électron devaient être abandonnées pour des processus intervenant à l'intérieur des noyaux. Au congrès Solvay de 1933, malgré le nouveau poids des arguments de Pauli en faveur du neutrino, Bohr continua de croire à une violation possible du principe de conservation de l'énergie.

Heisenberg retint de ce débat l'idée que la mécanique quantique ordinaire pouvait être violée à l'échelle nucléaire, et que des concepts macroscopiques (comme ceux de longueur et de temps) perdaient alors leur sens.

4.6 LA LONGUEUR UNIVERSELLE DE HEISENBERG

Pour Bohr, le rayon classique de l'électron devait tracer la frontière d'un monde nouveau où la méthode usuelle de correspondance devenait caduque et rien n'était dit (ou très peu) sur la théorie future de cette *terra incognita* que l'atomicité de nos instruments de mesure rendait difficile à explorer. Heisenberg adopta un point de vue plus constructif en prenant dès 1930 la longueur critique comme élément constitutif de la théorie : à l'espace-temps continu, il tenta de substituer un réseau ("Gitterwelt")*. Mais Heisenberg⁽⁵⁰⁾ fut sensible à la critique de Pauli : réseau et covariance relativiste semblaient incompatibles. En 1936, cette longueur devint une "longueur universelle", avec le statut d'une constante fondamentale comme h et c .

Cette idée apparut à la fin d'un article de Heisenberg⁽³⁰⁾ (voir chap. 6 § 2) de 1936 expliquant les gerbes de Blackett et Occhialini par la théorie des

* Voir l'article de D. Cassidy⁽⁴⁹⁾

explosions. "De même que la statistique classique ne peut être appliquée de manière sensée au corps noir avec son infinité de degré de liberté qu'après l'introduction d'une action universelle (\hbar), il se peut que la théorie quantique ne soit applicable à des systèmes continus qu'après l'introduction d'une longueur universelle."

En 1938, dans deux publications successives (28,29), Heisenberg développa cette idée de l'existence d'une longueur universelle λ_0 de l'ordre de grandeur du rayon classique de l'électron $r_e = \frac{e^2}{4\pi m c^2} \approx 2,8 \cdot 10^{-15} \text{ m}$, qui est aussi celui des dimensions nucléaires.

Sous un éclairage historique, il examine la signification des constantes fondamentales, plus précisément de celles qui sont vraiment fondamentales, comme c et \hbar . (Par exemple, la constante de Boltzmann n'a plus de raison d'être si l'on décide de mesurer les températures en unité d'énergie; la masse du proton ne peut être prise comme fondamentale puisqu'il existe sûrement des masses plus petites). L'introduction de c a permis l'unification* de l'électromagnétisme et de la mécanique. De même, celle de \hbar , en donnant une interprétation cohérente du rayonnement du corps noir, a permis d'unifier l'électromagnétisme et la théorie statistique de la chaleur (tout en supprimant la divergence de l'énergie totale donnée par la loi de Rayleigh - Jeans). Dans ces deux cas les constantes fondamentales donnent les limites d'utilisation des concepts de l'ancienne théorie (temps absolu, variables dynamiques d'une particule) et les conditions d'approximation de la nouvelle théorie par les anciennes. Enfin c et \hbar sont toutes deux reliées à une invariance : la relativité exige l'invariance de Lorentz, la mécanique quantique exige les relations canoniques et l'invariance par "rotation" dans l'espace de Hilbert (unitarité). Heisenberg pense qu'une telle situation va se reproduire pour une troisième fois : l'électrodynamique quantique et la physique nucléaire théorique sont en crise : les divergences sont de plus en plus sérieuses quand la constante de couplage a la dimension d'une puissance positive d'une longueur (30), comme dans les théories de Fermi, Yukawa, Wentzel ...** (et l'on sait bien aujourd'hui que ces

* Bien sûr, Heisenberg ne se méprend pas sur le sens du mot "Vereinigen" il s'agit d'une généralisation dans laquelle les théories anciennes apparaissent comme approximations de la nouvelle théorie.

** Voir chap. 6 §2

théories ne sont pas renormalisables).

Les difficultés sont liées au comportement à courte distance de la théorie. L'argument placé en tête du premier article est la possibilité de faire "exploser" une particule en la bombardant avec des projectiles d'énergie supérieure à $\frac{hc}{r_0}$. (Voir chap. 6 § 2). Cette prédiction traduirait l'impossibilité de faire des mesures de longueurs à mieux que r_0 . Nous signalons que cette hypothèse contredit l'opinion de Pauli **: d'après Landau et Peierls (§2), la précision maximale sur la mesure de la position d'un électron est $\Delta x \sim hc/E$ (arbitrairement petite si E est grand), "on en déduit qu'une longueur universelle ne peut sûrement pas exister, comme il résulte aussi de l'invariance relativiste"⁽³¹⁾. Mais le point de vue de Heisenberg reste viable, si l'on remarque que la notion de longueur limite "propre", c'est-à-dire attachée par exemple au référentiel d'un système lié, ne contredit nullement la relativité, et l'on est en droit d'en POSTULER l'existence même si les déductions de Landau et Peierls ne fournissent pas une telle limitation.

De même que h donne les limites d'application de la mécanique classique, r_0 donnerait celles de la mécanique quantique. La théorie future, tenant compte de r_0 permettrait :

- de résoudre le problème des divergences de l'électrodynamique en tenant compte des phénomènes nouveaux qui apparaissent à des longueurs de l'ordre de r_0 .
- de donner une interprétation cohérente des phénomènes nucléaires et de prédire les masses hadroniques, la portée des interactions fortes, le phénomène d'explosion pour des énergies de collision supérieures à hc/r_0 .

A l'introduction de r_0 devrait être associée une nouvelle symétrie fondamentale, et une révision profonde des concepts de base s'impose. Sont citées la tentative de Born et Infeld (voir chap. 7 § 1) pour modifier l'électrodynamique, et celle de March pour changer la géométrie de l'espace pour les courtes distances; mais là encore "trop de concepts de la physique actuelle sont inconsciemment mis en jeu". A défaut de proposer une nouvelle théorie, Heisenberg précise dans un deuxième article ⁽²⁹⁾ "les limites d'application de la théorie

** à laquelle se rallia Heisenberg ⁽³⁰⁾ en 1930 (Zf Phys. 65. 1930 p.5 où il est souligné qu'un découpage de l'espace en cellules de taille r_0 n'est pas compatible avec l'invariance relativiste).

quantique" auxquelles conduit l'hypothèse d'une longueur fondamentale. Un sous-produit intéressant de cet article est l'établissement de l'expression de l'évolution en représentation d'interaction (voir Réf I p178) :

$$U(t, t_0) = T \exp - \frac{i}{\hbar} \int_{t_0}^t d^4x \mathcal{L}_{int}$$
, qui présente l'avantage d'être manifestement (ou presque) covariante. Une condition "plausible" (obtenue en assimilant provisoirement l'effet de r_0 à un terme correctif dans le lagrangien) pour que la mécanique quantique habituelle soit applicable, est que l'impulsion transférée $(p_1 - p_2)^2$ pendant l'intervalle de temps propre $(x_1 - x_2)^2$ d'une quelconque des particules du processus envisagé, vérifie $(p_1 - p_2)^2 \gg \left(\frac{\hbar}{r_0}\right)^2$. Cette condition permet de classer en deux catégories les phénomènes jusqu'alors étudiés ; nous ne détaillerons pas d'avantage les conclusions d'Heisenberg, puisque nous savons bien aujourd'hui que cette condition correspond en gros au seuil de création de particules hadroniques.

La seule différence réside dans l'interprétation de r_0 , comme une limite imposée sur la mesure des longueurs. Une telle hypothèse est purement spéculative: l'existence d'une longueur remarquable en physique nucléaire n'impose nullement celle d'une longueur limite. Pour Heisenberg, il est nécessaire d'abandonner un nombre croissant de concepts hérités de notre expérience à l'échelle macroscopique, si nous voulons progresser dans l'exploration de l'univers ultramicroscopique. Et pour ce faire "peut-être que dans la tentative de flairer de nouveaux concepts, on se souviendra de nouveau avec profit, qu'il ne peut s'agir en physique théorique que de la connection mathématique de grandeurs observables ..." (fin de Réf.28). Heisenberg a pris un certain recul par rapport aux théories contemporaines, aux données confuses des expériences de physique nucléaire, et l'on trouve ici en germe l'idée de la matrice S.

4.7 LA MATRICE S

Quatre ans plus tard, en 1942, dans un article intitulé "Les grandeurs observables dans la théorie des particules élémentaires"⁽³²⁾, Heisenberg précise dans quelle direction il faut chercher une théorie rendant compte de l'existence d'une longueur fondamentale.

"Dans cette situation il semble utile de se demander quels concepts de l'ancienne théorie pourront être conservés dans la future, et cette question

est à peu près équivalente à la suivante : quelles grandeurs de l'ancienne théorie sont observables ? Car la théorie future, elle-aussi, doit naturellement contenir en premier lieu des relations entre grandeurs observables. Certes, seule la théorie achevée permettra de décider quelles grandeurs sont véritablement observables ...". En somme, Heisenberg reprend le programme qui s'était avéré si fructueux dans la découverte de la mécanique des matrices; alors la constante était \hbar , et les observables les transitions atomiques. Maintenant, c'est la longueur fondamentale r_0 qui donne les nouvelles limites de mesurabilité : "il est très douteux que l'on puisse déterminer la position d'une particule à mieux que $10^{-15} m$ et un instant à mieux que $10^{-24} s$ ".

Cependant l'existence de r_0 n'affecte pas la précision des mesures d'impulsion et d'énergie de particules libres. Heisenberg fait l'hypothèse que

- les énergies discrètes des états stationnaires de systèmes fermés
- le comportement asymptotique des fonctions d'onde dans les processus de collision, émission, et absorption

sont en principe observables.

Dans cet article, il sera surtout question du second type d'observable, compte tenu du fait que le premier est relié au second : les états stationnaires de diffusion deviennent discrets si l'on enferme l'ensemble diffuseur-diffusé dans une boîte. Enfin comme Heisenberg l'a montré un an plus tard ⁽³³⁾, suivant une idée de Kramers, le prolongement analytique des valeurs propres de la matrice S considérées comme fonctions de l'énergie peut servir à déterminer les énergies des états stationnaires de systèmes liés* (voir chap. 7 § 5).

La théorie de la diffusion de Dirac ⁽³⁴⁾ est généralisée à des états finaux contenant des particules différentes de celle de l'état initial pour introduire la matrice S (comme "Streuung"), reliant la fonction d'onde sortante aux impulsions des particules incidentes (la définition de Heisenberg diffère de la définition moderne comme limite asymptotique de l'opérateur d'évolution en représentation d'interaction, par la factorisation de la relation de conservation de l'énergie-impulsion).

Les propriétés générales de la matrice S résultent de sa définition :

* Remarquons toutefois qu'une telle relation peut se justifier dans la méthode de Schrödinger, mais doit être introduite ad hoc dans une approche de matrice S.

-elle est unitaire, donc diagonalisable et de valeurs propres de module 1 (généralisation de la méthode des déphasages) et vérifie le théorème optique**.

-elle commute avec l'impulsion-énergie, ce qui donne lieu à des singularités en δ de ses éléments de matrice entre états propres d'impulsion et d'énergie.

Bien sûr ces restrictions ne suffisent pas à déterminer la matrice S : "c'est comme le cadre vide d'un tableau qui reste à peindre" dit Wentzel⁽³⁵⁾. Une théorie hamiltonienne déplaît à Heisenberg, car il n'y a pas de raison qu'un éventuel hamiltonien soit simple et surtout une telle théorie sous-tend une description exacte des déroulements spatio-temporels, contredisant l'existence de r_0 . Il n'y a pas beaucoup plus de raison que la matrice η telle que $S = e^{i\eta}$ (qui s'identifie au premier ordre seulement à un lagrangien d'interaction) soit simple; cependant Heisenberg étudiera quelques "Ansätze" simples pour η dans un article ultérieur⁽³⁶⁾. (Voir chap. 7 § 5).

Nous n'explorerions pas beaucoup plus loin l'avenir de la matrice S sans sortir de la période étudiée et aussi de notre sujet, puisqu'elle est à l'origine de méthodes concurrentes de la théorie des champs (la matrice S analytique). Mentionnons toutefois une autre justification des approches dites "de matrice S " par Landau⁽¹³⁾, la relation $\Delta p \Delta t > \frac{h}{c}$ rend impossible la description temporelle détaillée de particules en interaction. Landau continue : "les opérateurs ψ qui contiennent de l'information inobservable doivent disparaître de la théorie et, puisqu'un hamiltonien ne peut être construit qu'à partir d'opérateurs ψ , nous sommes amenés à la conclusion que la méthode hamiltonienne pour les interactions fortes est morte et doit être enterrée, mais bien sûr avec tout l'honneur qui lui est dû."

Peierls soutient aujourd'hui un point de vue opposé*** " ... cela [le programme "matrice S "] présente nombre de difficultés, en particulier parce qu'il n'y a pas que les processus de diffusion qui soient observables, les états stationnaires le sont aussi; et une autre difficulté fait que personne n'a jamais réussi à obtenir des équations fondamentales pour la matrice S qui

** Déjà établi par Breit (Phys. Rev. 1941)

*** Conversation privée

ne soient pas équivalentes à l'écriture d'une fonction d'onde... Mais je me souviens à ce propos du récit de Heisenberg d'une de ses conversations avec Einstein ⁽³⁷⁾, alors qu'il essayait de formuler le principe d'incertitude et n'en était pas loin; Einstein n'aima pas du tout son approche et dit : "c'est la théorie qui détermine ce qui est observable"* et cette remarque qu'Einstein fait à Heisenberg est vraie : nous devons trouver l'argumentation dans la théorie elle-même, et en effet le "principe" d'incertitude découle bien de la nature de l'électron. Dans la théorie de la matrice S , nous n'avons pas de situation analogue, il n'y a rien dans le formalisme qui pourrait rendre autre chose que la matrice S inobservable".

* Heisenberg pensait que l'approche "observationiste" fondant la mécanique des matrices était à rapprocher de l'utilisation par Einstein de définitions opérationnelles des mesures d'espace et de temps pour fonder la relativité en 1905. Einstein s'en défendait: "Il se peut que cela ait été ma philosophie, mais de toutes façons cela n'a aucun sens. Il n'est jamais possible d'introduire seulement des quantités observables dans une théorie. C'est la théorie qui décide ce qui peut être observé. Dans les expériences de cette complexité, nous ne pouvons tout simplement pas séparer les processus d'observation empirique des constructions mathématiques et des autres concepts théoriques." (Rapporté à G. Holton (51) par Einstein).

Appendice : Le calcul par Heisenberg des fluctuations de charge dans la théorie du positron.

Heisenberg étudie les fluctuations de charge d'un gaz d'électrons libres de densité $\frac{N}{V}$ dans un volume \mathcal{V} ($\mathcal{V} \ll V$). L'étude statistique classique ** donne $\bar{q} = eN \frac{\mathcal{V}}{V}$ pour la charge moyenne et $(\Delta q)^2 = \overline{q^2} - \bar{q}^2 = e^2 N \frac{\mathcal{V}}{V} = e^2 \times$ fluctuations du nombre d'électrons, pour les fluctuations quadratiques moyennes.

Le calcul quantique relativiste est fait à l'aide de l'expression du champ de Dirac libre $\psi(\vec{r}, t) = \sum_m b_m u_m + d_m^\dagger v_m$
 b_n : annihilateur d'électron , u_n : solution d'énergie positive de l'éq. de Dirac
 d_n : " de positron , v_n : " " négative " "

L'opérateur densité de charge est $\rho = e:\psi^\dagger\psi:$ (réf I, p 146). Se référant à Bohr-Rosenfeld, Heisenberg étudie la moyenne spatio-temporelle de ρ effectuée à l'aide d'une fonction $f(t)$ de support de largeur T telle que $\int_{-\infty}^{\infty} f(t) dt = 1$ et d'une fonction $g(\vec{r})$ de support de volume \mathcal{V} telle que $\int_{\mathcal{V}} g(\vec{r}) d^3r = 1$. Alors l'opérateur charge moyenne est : $\langle Q \rangle = \int \rho(\vec{r}, t) g(\vec{r}) f(t) d^3r dt$. Comme on va le voir, il est nécessaire d'introduire une frontière douce du domaine d'intégration pour que les fluctuations soient finies. Le calcul explicite des moyennes quantiques dans un état $|N_1, \dots, N_n, \dots\rangle$ où les nombres d'électrons sont bien définis donne $\langle Q \rangle = e \frac{\sum N_n}{V}$, résultat identique au résultat classique statistique (ergodicité), et $\Delta Q^2 = \langle Q^2 \rangle - \langle Q \rangle^2$ se compose de deux termes (pour $\mathcal{V} \ll V$).

- Le premier s'identifie à $e^2 \sum N_n \frac{\mathcal{V}}{V}$, c'est-à-dire aux fluctuations statistiques classiques si toutefois la longueur d'onde Broglieenne $\lambda_n \ll \mathcal{V}^{1/3}$ et la vitesse $v_n \ll \mathcal{V}^{1/3}/T$

- Le deuxième terme est indépendant de $(N_n)_n$ et donc existe même pour le vide. Il diverge si l'intervalle de temps et le volume d'intégration sont limités abruptement. Si par exemple on choisit de limiter \mathcal{V} par une gaussienne de largeur b et si de plus $T \ll \frac{b}{c}$, ce terme vaut

$$\frac{e^2 \mathcal{V}^{2/3}}{cTb} \quad \text{pour } T \ll \frac{b}{mc}$$

$$\frac{e^2 \mathcal{V}^{2/3}}{(cT)^2 b mc} \quad \text{pour } T \gg \frac{b}{mc}$$

* Voir par exemple Landau et Lifchitz. Tome 5 § 115

L'exposant $\frac{2}{3}$ de v et la présence de b au dénominateur montrent qu'il y a création de matière à la surface de v , et d'autant plus que sa délimitation est plus nette. Ce résultat est rapproché du calcul par l'équation de Dirac de la création de paire dans un champ fort (Klein). Le lien entre forme du profil de moyennage ($g(\vec{r}), f(t)$) et champ perturbateur n'est pas explicité. Il est pourtant simple de le faire en remarquant que dans un champ électrique perturbateur de potentiel A^0 , la probabilité P d'émettre des paires en partant du vide est $\sum_{f \neq 0} |\langle f | S | 0 \rangle|^2$ ($|f\rangle$ base d'états finaux)

A l'ordre le plus bas $S = 1 - \frac{e}{\hbar} \int d^4x \rho A^0$ où $\rho = e: \psi^\dagger \psi :$ (en représentation d'interaction). Donc $P = \sum_f \frac{1}{\hbar^2} |\langle f | \int d^4x \rho A^0 | 0 \rangle|^2$
ou encore $P = \frac{1}{\hbar^2} \langle 0 | \left[\int d^4x \rho A^0 \right]^2 | 0 \rangle$. On voit bien que si A^0 est proportionnel au profil $g(\vec{r}) f(t)$, P est proportionnel aux fluctuations de ψ : $\langle 0 | Q^2 | 0 \rangle$.

CHAPITRE 5

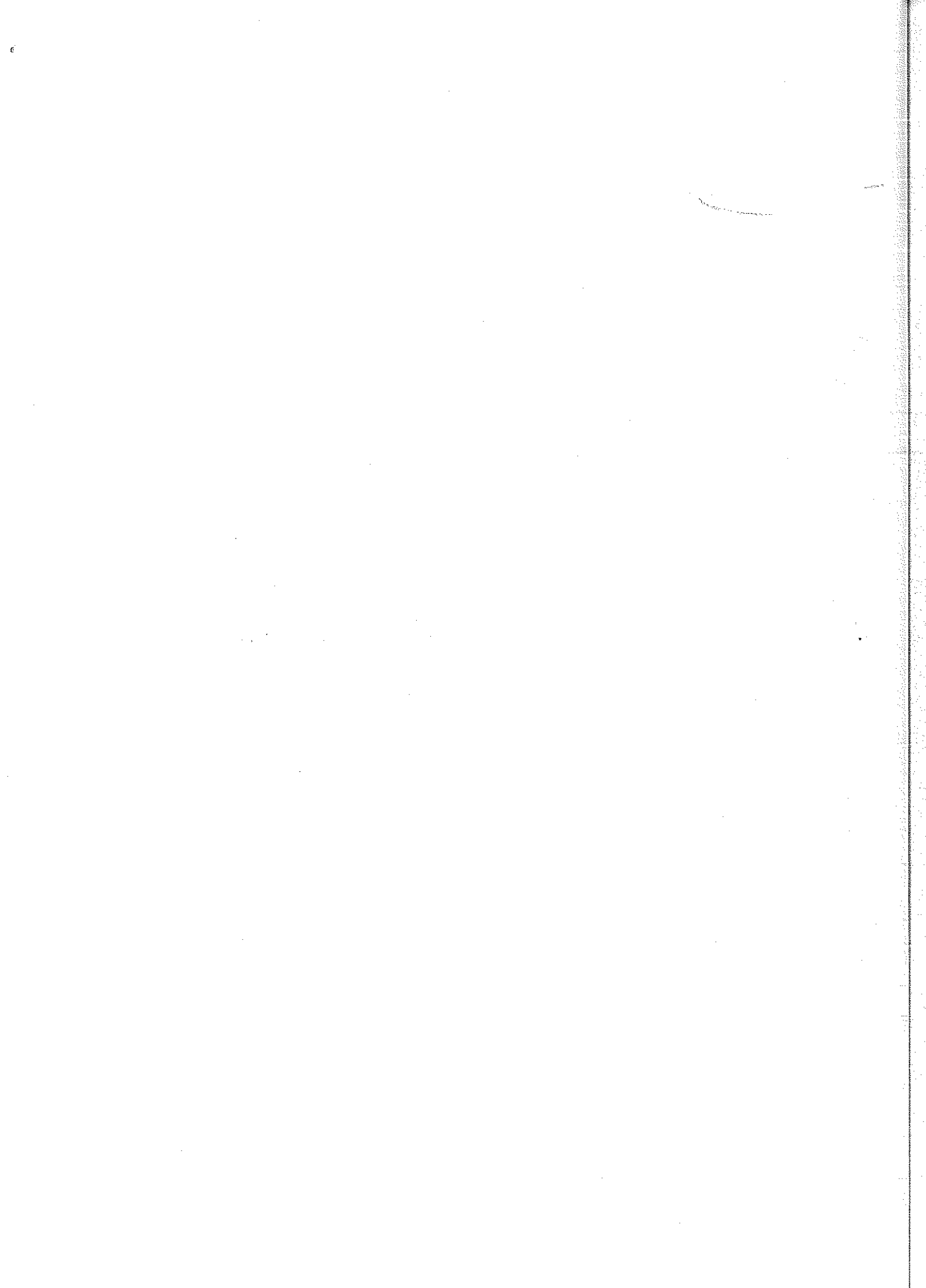
PROGRES FORMELS DE LA THEORIE QUANTIQUE DES CHAMPS PENDANT SA CRISE

L'exposant $\frac{2}{3}$ de v et la présence de b au dénominateur montrent qu'il y a création de matière à la surface de v , et d'autant plus que sa délimitation est plus nette. Ce résultat est rapproché du calcul par l'équation de Dirac de la création de paire dans un champ fort (Klein). Le lien entre forme du profil de moyennage ($g(\vec{r}), f(t)$) et champ perturbateur n'est pas explicité. Il est pourtant simple de le faire en remarquant que dans un champ électrique perturbateur de potentiel A^0 , la probabilité P d'émettre des paires en partant du vide est $\sum_{f \neq 0} |\langle f | S | 0 \rangle|^2$ ($|f\rangle$ base d'états finaux)

A l'ordre le plus bas $S = 1 - \frac{i}{\hbar} \int d^4x \rho A^0$ où $\rho = e: \psi^\dagger \psi :$ (en représentation d'interaction). Donc $P = \sum_f \frac{1}{\hbar^2} |\langle f | \int d^4x \rho A^0 | 0 \rangle|^2$
 ou encore $P = \frac{1}{\hbar^2} \langle 0 | \left[\int d^4x \rho A^0 \right]^2 | 0 \rangle$. On voit bien que si A^0 est proportionnel au profil $g(\vec{r}) f(t)$, P est proportionnel aux fluctuations de ψ :
 $\langle 0 | \psi^2 | 0 \rangle$.

CHAPITRE 5

PROGRES FORMELS DE LA THEORIE QUANTIQUE DES CHAMPS PENDANT SA CRISE



5.1. LA POLARISATION DU VIDE

La théorie des trous de Dirac (voir chap. 3.2) visait à résoudre la difficulté des états d'énergie négative de l'équation de Dirac. Dans sa première version, à travers la brume qui couvrait les propriétés de la mer infinie d'électrons d'énergie négative, elle permit de distinguer des processus élémentaires plus ou moins étranges. Parmi ceux-ci la possibilité de créer des antiélectrons fut bientôt (en 1932) vérifiée par l'expérience. Un succès aussi éclatant laissait espérer que les difficultés subsistantes liées au nombre infini des électrons de la mer s'évanouiraient bientôt. Elles auraient sans doute été immédiatement surmontées, avec les méthodes de calcul de l'époque, si les idées de base avaient été aussi claires qu'aujourd'hui : il faut s'intéresser aux seules quantités "physiques" observables, comme la charge renormalisée et les extraire de la théorie "nue" divergente par un procédé qui respecte les invariances fondamentales par transformation de Lorentz, par transformation de jauge et par conjugaison de charge. L'utilisation d'un formalisme respectant la symétrie de charge, c'est à dire de la version symétrique de la 2^o quantification des ondes de Dirac, aurait permis des calculs plus simples mais elle n'était pas indispensable et le problème fut historiquement résolu par la méthode ondulatoire* ; en revanche l'idée qu'une renormalisation infinie de la charge absorbe toutes les divergences du problème considéré est essentielle et l'histoire du problème est celle de la prise de conscience progressive de cette idée, achevée dans un article de Weisskopf de 1936.

Pour la clarté de l'exposé il importe de bien préciser quel est l'énoncé moderne du problème historique de la polarisation du vide. Ce n'est pas l'étude du problème quantique couplé au champ de Dirac et champ de Maxwell interagissant ; les divergences ultraviolettes de ce problème subsistent même dans la théorie des trous (voir chap. 3.3) et ne peuvent être éliminées sans renormalisation de masse. Le problème ici envisagé est celui de la polarisation du vide créée par un champ extérieur classique imposé : le champ électromagnétique appliqué est susceptible de créer des paires réelles ou virtuelles donnant lieu respectivement à

* Nous en donnerons une raison à propos de l'article de Heisenberg (5.1.c)

des phénomènes d'absorption ou de dispersion, comme dans tout milieu polarisable.

5.1.a La communication de Dirac au congrès de Solvay d'octobre 1933 ⁽¹⁷⁾:

Après avoir résumé les succès de sa théorie des trous - la découverte expérimentale du positron était encore fraîche - DIRAC aborda les difficultés formelles posées par la complexité du vide et en tira une prédiction de phénomènes physiques nouveaux très surprenants en regard des propriétés du vide de l'électrodynamique classique. Le vide de la théorie de Dirac est peuplé d'une infinité d'électrons d'énergie négative dont la charge et l'énergie-impulsion sont déclarées inobservables. Pour avoir les caractéristiques physiques d'un état excité il faut soustraire la contribution du vide. En absence de champ électromagnétique extérieur une telle soustraction s'opère formellement sans problème. Mais, souligne Dirac, une telle soustraction de deux infinis doit présenter dans le cas général des ambiguïtés. Un cas simple où, moyennant des hypothèses supplémentaires assez naturelles, le calcul peut être mené jusqu'au bout est celui de la polarisation du vide induite par un champ statique faible $V(\vec{r})$.

Pour décrire la mer de Dirac, Dirac utilise l'approximation de HARTREE, qui consiste à attribuer à chaque électron sa propre fonction d'onde et à décrire l'interaction entre les électrons par un champ effectif ("de Hartree") ici contenu dans le potentiel $V(\vec{r})$. Inventée par Hartree pour décrire les atomes à plusieurs électrons cette méthode avait déjà fait ses preuves et Dirac lui-même l'avait perfectionnée ⁽¹⁸⁾ en utilisant l'opérateur densité défini par $R = \sum_n |r\rangle\langle r|$ où $|r\rangle$ désigne l'état du r^{o} électron; les équations de Hartree peuvent alors se mettre sous la forme $i\dot{R} = [H, R]$ et les grandeurs physiques intéressantes se déduisant toutes de R : la charge est $+e \text{Tr } R$ l'énergie $\text{Tr } R H$ etc. Le principe d'exclusion s'exprime par $R^2 - R = 0$: R ne peut avoir que les valeurs propres 0 et 1 correspondant aux nombres d'occupation possibles d'un état quantique. Nous verrons pourquoi au § 5.1.c les calculs effectués par Dirac et ses successeurs en 1° quantification avec approximation de Hartree conduisent à des résultats équivalents à ceux donnés par la 2° quantification.

Dirac cherche la solution stationnaire R ($[H, R] = 0$) avec

$H = \vec{\alpha} \cdot \vec{p} + \beta m + eV$) qui s'identifie pour $V = 0$ à la distribution $R_0 = \sum_{\epsilon < 0} |\epsilon\rangle \langle \epsilon|$ (λ =polarisation) où tous les états d'énergie négative non perturbés sont occupés. Pour V petit $R = R_0 + R_1$ où R_1 représente la partie observable de la distribution perturbée. La charge induite est

$$\delta\rho(\alpha) = e \cdot \text{Tr}_{\Delta p_{\text{min}}} \langle \alpha | R_1 | \alpha \rangle$$

Le résultat de Dirac est $\delta\rho(\alpha) = -\frac{2\alpha}{3\pi} \left(\ln \frac{2P}{mc} - \frac{5}{6} \right) \rho - \frac{4\alpha}{15\pi} \left(\frac{\hbar}{mc} \right)^2 \Delta\rho$

P est une coupure ultraviolette que Dirac situe à $\frac{m^*}{\alpha}$.

"Comme conséquence du calcul précédent, il semblerait que les charges électriques normalement observées sur les électrons, protons ou autres particules électrisées ne sont pas les charges véritables portées par ces particules figurant dans les équations fondamentales, mais sont plus petites dans le rapport d'environ 136 à 137 (avec la valeur estimée de la coupure). Pour des processus comportant des échanges d'énergie de l'ordre de mc^2 , il n'y aurait probablement pas le temps, cependant, pour que la polarisation des électrons d'énergie négative s'établisse de manière complète, de sorte qu'on doit s'attendre à ce que les charges observées soient plus voisines des charges réelles. Il en résulterait des déviations de l'ordre de 1 % dans des expressions telles que la formule de Klein-Nishina..."**

Dirac applique ces commentaires au 1° terme de $\delta\rho$. L'analyse est particulièrement pénétrante. Il faut cependant critiquer l'estimation de l'ordre de grandeur des corrections dynamiques à la charge observée de l'électron : le 1° terme de $\delta\rho$ serait aujourd'hui considéré comme une renormalisation infinie de la charge (P infini). Le second terme "représente une correction importante seulement lorsque la densité ρ varie rapidement avec la position et change de manière appréciable sur une distance de l'ordre de $\frac{\hbar}{mc}$ ". Notons que le calcul n'est pas aussi simple qu'on pourrait le croire, il nécessite une "intégration compliquée" et Dirac s'est trompé d'un facteur 4 dans le 2° terme. (voir réf. I p. 327).

Lors de la discussion qui suivit la communication, PEIERLS fut le premier à intervenir pour mentionner son calcul dans le cas d'un champ

** Le rapport de la conférence est rédigé en (mauvais) français.

* Cette valeur correspond à des longueurs d'onde de l'ordre du rayon classique de l'électron considéré par Bohr et par Dirac comme la limite en delà de laquelle interviendrait la structure inconnue de l'électron négligée dans la théorie du positron.

magnétique et préciser qu'elle était, selon lui, la nature de la difficulté : il faut établir une correspondance entre états d' $E < 0$ sans champ et avec champ pour préciser la procédure de soustraction. Nous verrons plus loin comment Peierls développa ce point de vue.

PAULI, quant à lui, adopta une position très critique qu'il gardera tout au long de l'évolution du problème *"Toutefois l'aspect général de la théorie n'est pas satisfaisant, par suite de la manière dont elle est obligée de se servir de la conception de l'infini ..."*

Néanmoins Dirac avait eu raison, dans cette première investigation du sujet, de mettre l'accent sur les phénomènes physiques prévisibles, et dont l'existence n'est plus en doute aujourd'hui.

5.1.b A la recherche d'une méthode de soustraction :

Le terrain où devait se compléter la théorie du positron paraissait à la fin de 1933 assez mouvant. En témoigne la longue discussion de FURRY et OPPENHEIMER⁽¹⁹⁾ (déc. 1933) des problèmes de mesurabilité des variables dynamiques de l'électron.* Lors de la mesure de la position d'un électron, il est impossible de s'abstraire totalement du phénomène de création de paires. La notion de fonction d'onde d'un électron a donc une signification limitée et il vaut mieux utiliser le formalisme qui intègre naturellement des variations du nombre de particules : celui de la 2^o quantification de WIGNER et JORDAN. Malheureusement cette juste conclusion n'aida pas beaucoup FURRY et OPPENHEIMER dans leurs calculs et ne fut pas reprise dans les études suivantes. Leur procédé de soustraction n'était pas invariant de jauge et la 2^o quantification qu'ils utilisaient n'était pas symétrique (positron interprété comme trou). Les conclusions pour la polarisation du vide sont les mêmes que celles de Dirac : la notion de charge apparente, la charge réelle n'étant visible que dans des expériences à haute énergie. Le calcul de l'énergie d'interaction de deux charges au 2^o ordre de la série des perturbations montre que la polarisation du vide apporte des corrections à la loi de Coulomb : à courte distance $R \ll \lambda_c = \frac{h}{mc}$
 $e \rightarrow e + \delta e$ avec $\delta e = \frac{2ae}{3\pi} \ln \frac{\lambda_c}{R}$ ($\lambda_c = \frac{h}{mc}$) et corrélativement un déplacement des raies d'un atome. Cette dernière prédiction est

* les arguments avancés ressemblent beaucoup à ceux de Landau Peierls et de Bohr et Rosenfeld présentés au chap 4.

bien vérifiée aujourd'hui : la polarisation du vide contribue (d'environ - 27 MHz) au déplacement de Lamb (voir chap. 3). La redéfinition de la constante diélectrique du vide (c'est à dire la renormalisation de la charge) suffit à retrouver les résultats de l'électrodynamique classique, dans le cas contraire où $R \gg \lambda_c$: l'expression de l'énergie mutuelle de deux charges, l'expression de l'énergie d'un champ statique. Mais le fait que cette redéfinition soit infinie tourmente les deux auteurs, d'autant plus qu'elle ne semble pas liée à la structure ponctuelle des charges introduites dans la théorie classique (en effet l'énergie de polarisation induite par un champ statique même lentement variable diverge) : "les difficultés sont d'un tel caractère qu'il semble qu'elles ne puissent être surmontées en modifiant seulement le champ d'un électron à courte distance, mais qu'elles nécessitent un changement plus profond de nos notions d'espace et de temps sur lesquelles, en dernière analyse, reposent nos méthodes de mécanique quantique qui à leur tour demandent l'existence de particules stables pour être définies". Cette conclusion nous le verrons bientôt, était exagérément pessimiste.

FOCK (86) un peu avant (nov 1933), dans une étude purement formelle de la polarisation du vide avait abouti à des conclusions analogues : "les difficultés formelles ci-dessus mentionnées [séparation non invariante de jauge entre états d'énergie cinétique négative et positive et les divergences] montrent que même lorsque l'on fait abstraction des difficultés conceptuelles de la théorie [remplir l'espace d'un nombre infini d'électrons d'énergie négative] une mise en oeuvre des idées de Dirac n'est pas possible sans artifice particulier. C'est pourquoi la possibilité que la théorie soit fausse n'est pas exclue, malgré les succès déjà atteints". Fock utilisait, comme Furry et Oppenheimer le formalisme de la 2° quantification mais cette fois en introduisant des créateurs de positrons à la place des annihilateurs d'électrons d'énergie négative, probablement pour la première fois. Il fut ainsi le premier à bien comprendre que la divergence de la polarisation du vide n'avait rien à voir avec l'infinité des électrons de la mer, ce qui ne fut évident pour tout le monde qu'après que Pauli et Weisskopf eurent montré que la divergence existait même pour le champ scalaire chargé. (voir plus loin).

L'article de DIRAC de février 1934 : "Discussion de la distribution infinie des électrons dans la théorie du positron"⁽²⁰⁾ représente la première étape dans l'élaboration de prescriptions précises permettant d'éliminer les divergences de la distribution d'électrons.

Le but de Dirac était de trouver une extension relativiste et invariante de jauge de la méthode décrite au congrès Solvay. Pour ce faire, il faut d'abord généraliser la notion d'opérateur densité. Il agit maintenant dans l'espace des $|\vec{x}, t\rangle$ et se définit par $\langle \vec{x}'', t'' | R | \vec{x}', t' \rangle = \sum_n \psi_n^*(\vec{x}'', t'') \psi_n(\vec{x}', t')$ (nous omettons les variables de spin). R vérifie l'équation de Dirac : $(i\gamma - eA - m)R = 0$

- En premier lieu Dirac calcule R_0 dans le cas libre ($A_\mu = 0$) et introduit ainsi pour la première fois (sans utiliser cette notation) la fonction invariante Δ^* et l'expression à l'aide de fonctions de BESSEL pour préciser ses singularités sur le cône de lumière qui vont jouer un rôle crucial puisque la quantité intéressante est $\lim_{\substack{t'' \rightarrow 0 \\ \vec{x}'' \rightarrow 0}} \langle \vec{x}'', t'' | R | \vec{x}', t' \rangle$

- La soustraction pure et simple de R_0 à R suivie du passage à la limite précédent ne suffit pas à éliminer les divergences. Dirac va donc étudier plus précisément les singularités de R, à partir d'un Ansatz sur leur forme : elles doivent s'obtenir en donnant aux coefficients des singularités de R_0 une dépendance explicite en A_μ .

Ainsi R se décompose en $R_a + R_b$ où R_a contient toutes les singularités et se calcule explicitement en fonction de A_μ et R_b dépend à la fois des conditions initiales et de A_μ suivant l'équation inhomogène $(i\gamma - eA - m)R_b = -(i\gamma - eA - m)R_a$. R_b est interprété comme la partie observable de la distribution, comme il se doit elle est invariante relativiste et de jauge.

Dirac remet à plus tard l'étude de conséquences physiques de ce formalisme. La détermination de R_0 est ardue (et encore plus dans les cas spécifiques où l'on précise la forme de A_μ) et souffre de plusieurs ambiguïtés. A chaque étape Dirac a seulement pris le choix le plus simple.

Cet article est la dernière contribution de Dirac à l'étude de la polarisation du vide. Il n'était pas satisfait par la complexité croissante des calculs, et de toutes façons, même après résolution de ce problème, les autres divergences de l'électrodynamique quantique subsisteraient.

* Jordan et Pauli l'avaient introduit seulement dans le cas particulier de masse nulle. Dans la théorie moderne Δ apparaît dans :

$$\{\psi_\xi(x), \bar{\psi}_\xi(x')\} = (i\gamma_x + m)_{\xi\xi'} i \Delta(x-x')$$

Nous verrons au chapitre 7§2 quelle autre voie Dirac choisit d'explorer.

PEIERLS⁽²¹⁾ (Mars 1934) n'apprécia pas la méthode de Dirac car elle revenait à soustraire au courant $\vec{j}_\mu(x) = T_n \langle \alpha | \delta_\mu R | \alpha \rangle$ une fonction du champ singulière incomplètement déterminée. Pour lui il était plus naturel de soustraire le courant en champ nul. L'ambiguïté de la soustraction est levée en précisant la correspondance entre les états d'énergie négative perturbés et non perturbés tout en respectant l'invariance de jauge. Malheureusement le résultat contient toujours la divergence logarithmique de la permittivité du vide : "l'auteur croit qu'il faut attribuer un sens au fait que la méthode ici expliquée, qui semble être très naturelle, conduite à un résultat divergent et qu'il y a là une difficulté qui ne pourra être surmontée qu'en apportant des modifications essentielles à la théorie. De surcroît des modifications sérieuses de la théorie sont rendues nécessaires par d'autres difficultés (énergie propre de l'électron, etc) qui n'ont pas de relation immédiate avec ces problèmes". Ainsi Peierls rejoint avec Hurry et Oppenheimer le camp des pessimistes qui voyaient dans la divergence de la polarisation du vide une difficulté tout aussi insurmontable que celle des énergies propres infinies.

5.1.c. Les prescriptions de Heisenberg-Dirac

HEISENBERG⁽²²⁾ (juin 1939), lui, suivit la voie indiquée par Dirac en perfectionnant sa méthode de soustraction. En exploitant la conservation du courant et du tenseur d'énergie impulsion il put lever une partie des ambiguïtés contenues dans la procédure de Dirac*. Quelques erreurs de calcul (témoignant de sa difficulté) furent corrigées par la suite⁽²³⁾⁽²⁴⁾, grâce aux critiques d'Euler, de Kockel et de Weisskopf. Mais l'essentiel est que Heisenberg aboutit à des formules explicites pour la partie singulière à soustraire et qu'il put, ainsi que ses successeurs, passer directement aux applications. L'autorité scientifique de Dirac et de Heisenberg était sans doute la meilleure garantie de la validité de leur méthode car comme Heisenberg le reconnut plus tard⁽²⁴⁾ : "dans cette théorie les prescriptions pour la construction de la matrice S (identique à R_b de Dirac)** paraissent arbitraires. En faveur de la matrice S particulière ici choisie on ne

* Les expressions obtenues pour la matrice densité soustraite donnent des densités de charge identiques à celles de Dirac, mais des densités d'énergie-impulsion différentes.

peut guère citer que sa relative simplicité mathématique (conjointement à quelques conséquences des équations de conservation)". A défaut d'une meilleure justification de cette méthode, plusieurs théoriciens Heisenberg, Euler, Kockel, Uehling Serber développèrent son utilisation. Il se trouve que la forme des singularités du courant j_μ sur le cône de lumière proposée par Dirac et Heisenberg était exacte ainsi qu'il fut démontré dans une approche plus rigoureuse (25) (26) et surtout que les prescriptions de soustraction étaient équivalentes à une renormalisation de charge (voir l'article de Weisskopf). Heisenberg recalcula comme Dirac au congrès de Solvay, mais avec le facteur 4 corrigé la charge
$$\tilde{p} = -\frac{\alpha}{15\pi} \left(\frac{\hbar}{mc}\right)^2 \Delta \rho_0$$
 induite par une distribution extérieure ρ_0 , et aussi par un champ électrique oscillant : il n'y a donc pas, conclut-il, d'effet de polarisation statique global sur une charge (puisque cet effet divergent se trouve automatiquement soustrait ; il faut s'attendre à des effets observables seulement pour des problèmes non statiques de temps caractéristique \hbar/mc^2).

Tous les calculs précédents étaient effectués dans la théorie ondulatoire intuitive ("anschauliche") de Dirac, c'est-à-dire sans le formalisme de la 2^o quantification. Heisenberg n'introduisit des ondes de Dirac quantifiées que dans la deuxième partie de cet article à propos du problème couplé où champ électronique et champ électromagnétique interagissent. Pour lui la quantification des ondes de matière ne devenaient vraiment préférable que dans ce problème-ci, où elle devient le pendant de la quantification du champ électromagnétique. Dès l'invention de la théorie du positron, Dirac avait préféré le formalisme ondulatoire et ses utilisateurs suivirent en général la même voie, à l'exception déjà vue de Fock, Furry et Oppenheimer. Nous mentionnerons en fin de cette partie comment Heisenberg modifia le hamiltonien de l'électrodynamique quantique à partir des règles de soustraction établies dans le cas particulier du champ extérieur classique imposé. Ici nous insisterons plutôt sur l'éclairage nouveau que l'introduction du champ de Dirac quantique apporta à la méthode de Dirac, et sur la modification importante que Heisenberg utilisa pour décrire les états de ce champ.

- Heisenberg donna la correspondance entre la matrice densité de

Dirac (R) et l'opérateur ψ : $\langle \alpha'' | R | \alpha' \rangle = \langle 0 | \psi^\dagger(\alpha'') \psi(\alpha') | 0 \rangle$
 Mis à part le fait que $|0\rangle$ désigne le vide de Dirac (plein d'électrons d'énergie négative et ψ le champ de Dirac interprété en terme d'annihilation d'électrons d'énergie positive ou négative, le lien est établi entre les anciens calculs de la polarisation du vide et le formalisme moderne. L'identité des deux membres résulte immédiatement de leur égalité quand $A_\mu = 0$ et du fait qu'ils vérifient tous deux l'équation de Dirac. L'approximation de Hartree selon laquelle la mer d'électrons d'énergie négative est décrite par un opérateur densité donne en fait le comportement rigoureux de cette mer dans un champ extérieur classique. Il n'y a approximation que dans la mesure où l'interaction entre les électrons de la mer est contenue dans le champ classique (de Hartree) au lieu d'être décrite par le champ électromagnétique quantifié.

- Heisenberg décomposa enfin le champ ψ en annihilateurs d'électrons et en créateurs de positrons comme on le fait aujourd'hui. La symétrie par conjugaison de charge de la théorie devient ainsi plus explicite et le concept de mer infinie disparaît (c'est nous qui insistons sur ce dernier point). L'idée d'introduire un formalisme qui traite les trous comme des particules avait déjà été exploitée par Heisenberg en 1931 pour décrire les couches incomplètes des atomes et utilisées, répétons-le pour la théorie des trous de Dirac par Fock^{*(86)} (nov. 1933). Il est probable qu'Heisenberg n'accorda alors pas plus d'importance à cette nouveauté qu'à un simple changement de notation, et l'ancienne description continua d'être largement utilisée. Elle était par exemple encore enseignée dans le cours de Wentzel⁽²³⁾ de 1942.

Sous l'égide d'Oppenheimer, UEHLING⁽²⁹⁾ utilisa en avril 1935 le formalisme de Heisenberg dans quelques cas particuliers et aboutit à des prédictions expérimentales. La loi de Coulomb, comme l'avaient vu Furry et Oppenheimer se trouve modifiée : l'énergie d'interaction de deux charges est corrigée de $\frac{\delta E_{int}}{E_{int}} \approx -\frac{2\alpha}{3\pi} \ln \frac{r}{d_c}$ ($d_c = h/mc$)
 pour $r \ll d_c$ et $\frac{\delta E_{int}}{E_{int}} \approx \frac{e^{-r/d_c}}{r^{3/2}}$ pour $r \gg d_c$

Il en résulte des écarts à la diffusion coulombienne habituelle et un déplacement des niveaux atomiques. Par exemple pour l'hydrogène,

* Il est inexact d'attribuer comme on le fait souvent la nouvelle décomposition de ψ à Heisenberg. Fock en fut probablement le premier utilisateur.

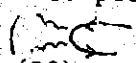
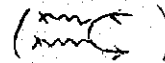
la prédiction, cette fois quantitative, est exacte mais ne représente qu'une faible partie (et avec le mauvais) signe) du déplacement observé dans l'expérience de Lamb (voir chap. 3 § 3).

SERBER⁽³⁰⁾ calcula en même temps, et aussi sous la direction d'Oppenheimer le courant induit δj_μ à l'ordre α par un courant extérieur j_μ ; on passe du cas statique à ce cas en remplaçant seulement le laplacien par le d'Alembertien : $\delta j_\mu = \alpha \Lambda * \square \square j_\mu$
(* = produit de convolution,
 Λ = noyau explicitement calculable)

La correction aux équations de Maxwell apportée par ce phénomène de polarisation ne peut être mise sous forme hamiltonienne puisqu'elle fait intervenir l'opérateur intégral $\Lambda *$. Nous verrons plus loin que cette impossibilité n'a plus cours lorsqu'on considère des champs lentement variables (pour lesquels la polarisation du vide au 1° ordre en α est négligeable devant celle en α^3).

PAULI et ROSE⁽³¹⁾ en janvier 36 publièrent le même résultat que Serber en calculant directement en régime variable les expressions formelles de Heisenberg au lieu d'utiliser comme Serber la solution statique et l'invariance relativiste de la procédure de soustraction de Heisenberg. Ainsi à défaut d'une théorie plus satisfaisante Pauli n'hésita pas à se salir les mains : "Le formalisme de la théorie du positron, aujourd'hui accepté et malheureusement pas encore remplacé par un meilleur, n'identifie pas la charge électrique physique avec $T_n \langle \alpha' | R | \alpha \rangle$ pour $\alpha' = \alpha$ " mais soustrait d'abord des termes dépendant de $x' - x$ de telle manière que la différence soit finie pour $x' = x$ ("subtraction-physics").

5.1.d. Les propriétés dispersives du vide :

Le vide peut absorber l'énergie du rayonnement incident, via la création de paires réelles : c'est le processus de diffusion photon-photon avec émission de paires  à l'ordre le plus bas) calculé par BREIT et WHEELER⁽³²⁾ en octobre 1934 ou encore la création de paires dans un champ électrique uniforme  processus calculé par SAUTER⁽³³⁾ en 1931 (voir chap. 2 § 2). Comme dans tout milieu diélectrique on doit s'attendre aussi à l'existence de phénomènes dispersifs en électrodynamique du vide et corrélativement

- 191 -

à la possibilité d'interaction entre photons^{*}. Une telle interaction, représentée par la création suivie de l'annihilation d'une paire virtuelle $(\text{---}\bigcirc\text{---})$ fut envisagée par HALPERN⁽³⁴⁾ dans une lettre à la Phys. Rev. et aussi par DEBYE lors d'une discussion avec Heisenberg. DELBRÜCK⁽³⁵⁾ lors de la discussion d'expériences de Lise MEITNER imagina la diffusion d'un photon par un champ coulombien comme processus possible d'interaction des rayons γ avec la matière. La polarisabilité du vide par une charge était déjà une chose bizarre, mais la dispersion en l'absence de toute matière était encore plus étrange quoique implicitement contenue dans la théorie des trous de Dirac. Heisenberg, après avoir développé sa méthode de soustraction, rappela la possibilité de diffusion de la lumière par la lumière et donna la correction à ce processus apporté par ses soustractions.

C'est à EULER^{**} et KOCKEL deux élèves de Heisenberg qu'il revient d'avoir calculé pour la 1^o fois la diffusion photon-photon à l'ordre le plus bas (α^4) . Leur lettre⁽³⁶⁾ à "Naturwissenschaften" de février 1935 donne seulement le résultat, sous une forme spectaculaire : l'amplitude de transition de l'état $|\gamma_1, \delta_1\rangle$ à l'état $|\gamma'_1, \gamma'_2\rangle$ est donnée par l'élément de matrice d'une correction effective au hamiltonien du rayonnement égale à

$$\delta W = -\frac{1}{90\pi} \frac{1}{\alpha E_0^2} \int d^3x \left[(\vec{B} \cdot \vec{D})^2 + 7(\vec{B} \cdot \vec{D}')^2 \right] d^3x$$

$$\left(E_0 = \frac{e}{4\pi r_0^2}, r_0 = \frac{e^2}{4\pi m c^2} \text{ et } \vec{D}' \text{ conjugué canonique de } \vec{A} \right)$$

Les deux auteurs, ainsi que leurs successeurs, ne manquèrent pas de faire la comparaison avec la récente théorie classique de Born et Infeld qui modifiait le lagrangien de l'électrodynamique classique par des termes de forme analogue. (voir chap. 7). Leur calcul utilisait

* Il peut sembler au lecteur que nous sortons maintenant du cadre défini à l'entrée de ce paragraphe : celui du champ électromagnétique classique imposé de l'extérieur. Il n'en est rien car les effets dispersifs calculés peuvent se déduire de la seule expression de la polarisabilité calculée dans ce cadre, de même qu'il suffit de connaître la polarisabilité d'un milieu matériel, pour y écrire les équations de l'électrodynamique. Certains calculs historiques comme celui d'Euler et Kockel font intervenir une quantification du champ électromagnétique, mais les processus élémentaires calculés sont tous interprétables en termes classiques pour le champ électromagnétique (les photons n'interviennent que comme pattes externes des graphes).

** Hans Euler était le petit-fils du célèbre mathématicien, il disparut pendant la guerre au cours d'un vol de reconnaissance.

les expressions de Heisenberg éventuellement corrigées*, et il fut suivi avec intérêt par Heisenberg lui-même.

L'intégralité du calcul ne fut publié que quelques mois plus tard (en juin 1935) par Hans Euler⁽³⁷⁾, La démonstration est présentée en trois étapes :

- Une correction éventuelle au lagrangien de l'électrodynamique classique est nécessairement de la forme $\int \mathcal{L} = \frac{1}{\sqrt{F_0}} \left[-A (E^2 - B^2) - A' (\vec{E} \cdot \vec{B}) \right]$ où A et A' sont des constantes numériques. En effet $\int \mathcal{L}$, pour pouvoir décrire les processus à 4 photons doit être du 4^o ordre, il ne peut dépendre que des invariants relativistes $E^2 - B^2$ et $\vec{E} \cdot \vec{B}$ et ne contient pas de dérivée dans le cas où les champs varient lentement sur λ_c et $\frac{\lambda_c}{c}$

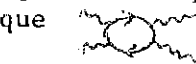
- Le calcul perturbatif en électrodynamique quantique avec les soustractions de Heisenberg donne une amplitude de diffusion photon-photon, de même forme que celle donnée par la correction du hamiltonien correspondant à $\int \mathcal{L}$ (après quantification du champ électromagnétique).

- Le calcul complet de 2 cas particuliers (pour la configuration des photons) permet de déterminer les constantes A et A'.

Il reste à dire que cette dernière étape fut particulièrement laborieuse : le théorème de FURRY⁽³⁸⁾ prévoyant la nullité des termes d'ordre impair en α n'existait pas, et surtout dans l'ancienne théorie des perturbations le nombre de termes à calculer était énorme. En témoigne l'extrait ci-joint (P. 193) d'un article de 50 pages). L'utilisation de graphes, reproduits notamment pour la première fois dans un article d'électrodynamique, ne pouvait être d'un grand secours. (Voir fig. page 194)

La valeur simple des constantes A et A' ($A = -\frac{1}{90\pi}$, $A' = -\frac{1}{90\pi}$) fit penser à Euler qu'une méthode plus simple devait exister. Une première simplification fut apportée par KEMMER et WEISSKOPF⁽⁸²⁾⁽⁸⁷⁾ (fév. 1936) :

* Certains lecteurs s'étonneront que le calcul du processus nécessite ici des soustractions infinies. Il est vrai que dans le formalisme moderne de l'électrodynamique quantique ce graphe, quoique superficiellement divergent, ne diverge pas en conséquence de l'invariance de jauge. Dans l'ancienne théorie des perturbations et sans régularisation invariante, l'invariance de jauge n'était pas respectée à tous les stades du calcul ce qui explique que ou plutôt la somme de tous les graphes ordonnés en temps correspondants dans l'ancienne théorie des perturbations) divergeait. Cependant, comme nous le verrons à propos de l'article de Kemmer et Weisskopf le processus élémentaire de diffusion d'un photon par un champ coulombien ne divergeait pas.



(8,10)

Reihenfolge	Zähler	1		$(p \cdot g)$		$ g ^2$		$(p \cdot g)^2$		$(p \cdot g)^2 g ^2$		$(p \cdot g)^2$		$ g ^4$		$(p \cdot g)^2 g ^2$		$(p \cdot g)^4$	
		$2p_\mu^2$	$8p_\mu^2 p_\nu^2$	$2p_\mu^2$	$9p_\mu^2 p_\nu^2$	$2p_\mu^2$	$8p_\mu^2 p_\nu^2$	$2p_\mu^2$	$8p_\mu^2 p_\nu^2$	$2p_\mu^2$	$8p_\mu^2 p_\nu^2$	$2p_\mu^2$	$8p_\mu^2 p_\nu^2$	$2p_\mu^2$	$8p_\mu^2 p_\nu^2$	$2p_\mu^2$	$8p_\mu^2 p_\nu^2$	$2p_\mu^2$	$8p_\mu^2 p_\nu^2$
$g^1 g^2 g^3 g^4$	$z_1 = z_2 = z_3 = z_4 =$	0	0	0	4	2	0	2	8	4	4	4	4	4	4	4	4	4	4
	$z_1 = z_2 = z_3 = z_4 =$	0	0	0	4	2	0	2	8	4	4	4	4	4	4	4	4	4	4
	$z_1 = z_2 = z_3 = z_4 =$	0	0	0	4	2	0	2	8	4	4	4	4	4	4	4	4	4	4
$g^1 g^2 g^3 g^4$	$z_1 = z_2 = z_3 = z_4 =$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	$z_1 = z_2 = z_3 = z_4 =$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	$z_1 = z_2 = z_3 = z_4 =$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$g^1 g^2 g^3 g^4$	$z_1 = z_2 = z_3 = z_4 =$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	$z_1 = z_2 = z_3 = z_4 =$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	$z_1 = z_2 = z_3 = z_4 =$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Für die Reihenfolge: $g^1 g^2 g^3 g^4$ wird $g^1 g^2 g^3 g^4$

$$Z_{\mu\nu} = -S_1 + 2p_\nu^2 \left[\frac{\lambda_\mu^2}{p_0^2 p_0^2} + \frac{\lambda_\nu^2}{p_{0\mu} p_{0\nu}} + \frac{\lambda_\mu^2 \lambda_\nu^2}{p_{0\nu}^2 p_{0\mu}^2} - \frac{1}{p_0^2 p_0^2} \right] + \frac{2\lambda_\mu^2}{p_{0\nu}^2 p_{0\mu}^2} - \frac{2\lambda_\nu^2}{p_{0\mu}^2 p_{0\nu}^2}$$

Darin bedeuten S_1, S_2 die Abkürzung:


$$S_1 = \lambda_\mu^2 \lambda_\nu^2 - \lambda_\mu^2 \frac{(p^1 p^2)^2 + (m c)^2}{p_0^2 p_0^2} + \lambda_\nu^2 \frac{(p^1 p^2)^2 + (m c)^2}{p_0^2 p_0^2} - \lambda_\mu^2 \frac{(p^1 p^2)^2 + (m c)^2}{p_0^2 p_0^2} + \lambda_\nu^2 \frac{(p^1 p^2)^2 + (m c)^2}{p_0^2 p_0^2}$$

Diese Ausdrücke werden bei weiterer Spezialisierung auf parallele Impulse, wie man durch Einsetzen von (8,8, 9) und Entwicklung nach g/m bis zur 4. Ordnung findet:

Vgl. Gl. (8,10).

Um die Zähler (5,15) durch die Nenner (5,14) zu dividieren, über die 6 Übergangsweg μ und die 24 Reihenfolgen (8,8, 9) zu summieren, muß man die Spalten der Tab. (8,10) mit den entsprechenden Spalten der Tab. (8,9) und dem Faktor 4 multiplizieren. Man erhält dann im Glied 4. Ordnung in g/m nach Mittelung über die Positronenrichtung p : [Vgl. S. 9, 7]

Annalen der Physik. 5. Folge. 24.

au lieu de calculer la diffusion photon-photon, ils calculèrent la diffusion d'un photon par un champ coulombien () (le processus de Delbrück). Dans ce cas il n'y a plus de divergence et donc pas de soustraction à opérer.

Une autre méthode fut élaborée par EULER et HEISENBERG⁽²⁴⁾ (déc. 1935). Elle consiste à calculer l'énergie des paires virtuelles créées par le champ appliqué classique A_μ en déterminant R soustraite et l'énergie correspondante $\int d^3x \text{Tr}_{\text{spin}} \langle \alpha | R | \alpha \rangle (\vec{\alpha}(\vec{p}-e\vec{A}) + \beta m + eV)$ considérée comme terme additif au hamiltonien classique du rayonnement. La forme de cette correction comme dans l'article d'Euler est déterminable a priori, il suffit donc d'étudier le cas de deux champs A_μ particuliers pour déterminer la dépendance de lagrangien effectif dans les invariants $\vec{E} \cdot \vec{B}$ et $E^2 - B^2$. Le calcul peut être effectué à tous les ordres en e et Heisenberg et Euler aboutissent à une belle expression intégrale explicite :

$$\delta \mathcal{L} = \frac{1}{8\pi^2} \int_0^{+\infty} \frac{d\lambda}{\lambda} e^{-\lambda m^2} \left[e^{\lambda ab} \frac{\text{sh } e a \lambda \text{ ch } e b \lambda}{\text{sh } e a \lambda \text{ sh } e b \lambda} - \frac{1}{\lambda^2} + \frac{e^2 \lambda^2}{3} (a^2 - b^2) \right] \text{ (comp. Réf. I p.196)}$$

avec $a^2 - b^2 = E^2 - B^2$ et $ab = \vec{E} \cdot \vec{B}$ dont le 4° ordre donne les valeurs de A et A' de Euler et Kockel.

* Euler et Heisenberg prennent \vec{E} et \vec{B} // et uniformes. Le champ \vec{E} uniforme présente des difficultés particulières (surmontables) en raison du paradoxe de Klein : même faible il provoque l'apparition de paires réelles. C'est pourquoi Weisskopf préféra utiliser un champ sinusoïdal.

Fig. Les graphes d'Euler (37) (voir p.192)

Licht-quanten Elek-tronen	Absorption von g^1	Absorption von g^2	Emission von $-g^3$	Emission von $-g^4$
$\mu = 1$	Paarerzeugung	Elektronensprung	Elektronensprung	Paarvernichtung
$\mu = 2$	Paarerzeugung	Elektronensprung	Positronensprung	Paarvernichtung
$\mu = 3$	Paarerzeugung	Positronensprung	Elektronensprung	Paarvernichtung
$\mu = 4$	Paarerzeugung	Positronensprung	Positronensprung	Paarvernichtung
$\mu = 5$	Paarerzeugung	Paarerzeugung	Zerstrahlung des Elektrons vom 1. mit dem Positron vom 2. Paar	Zerstrahlung des Positrons vom 1. mit dem Elektron vom 2. Paar
$\mu = 6$	Paarerzeugung	Paarerzeugung	Zerstrahlung des Positrons vom 1. mit dem Elektron vom 2. Paar	Zerstrahlung des Elektrons vom 1. mit dem Positron vom 2. Paar

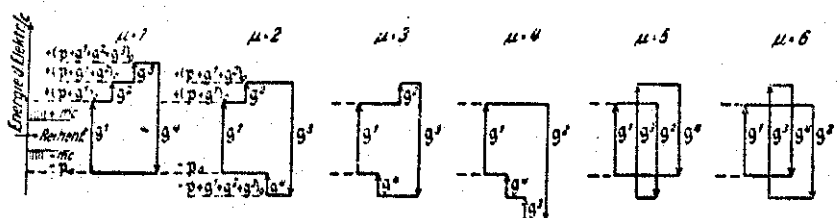


Fig. 1. Die 6 Übergangsmöglichkeiten

Le calcul de S (soustraction à R) est particulièrement difficile, mais comme les auteurs le font remarquer la détermination formelle de S repose sur l'hypothèse qu'il n'y a pas de polarisation proportionnelle au champ et donc l'effet de S est d'éliminer les termes correspondants de R. Cette remarque, généralisée, est à la base du point de vue de Weisskopf.

5.1.e. Le principe de renormalisation de charge

L'année suivante WEISSKOPF⁽³⁹⁾ (il avait collaboré avec Euler et Kockel) reprit le calcul du lagrangien effectif d'Euler et de Heisenberg (champs lentement variables) et le calcul du courant de polarisation induit par un champ faible*. Sur ces deux exemples il montra que les quantités physiques recherchées sont déterminées sans aucune ambiguïté si l'on fait l'hypothèse que les trois types suivants de contributions n'ont pas de signification physique :

- H₁ - L'énergie des électrons du vide dans l'espace libre
- H₂ - La charge et le courant des électrons du vide dans l'espace libre
- H₃ - Une polarisabilité du vide constante et uniforme et indépendante des champs

De plus les résultats obtenus sont identiques à ceux donnés par les prescriptions de soustraction de Heisenberg. La méthode est seulement plus simple car on a juste à reconnaître dans les expressions calculées les termes non physiques (divergents) et à les annuler (par exemple il faut annuler les termes de $j_{\mu}^{(pol)}$ indépendants du champ et ceux proportionnels au champ) et surtout les critiques de Pauli (et même de Heisenberg) sur l'ambiguïté de la détermination de l'opérateur S à soustraire à l'opérateur densité ne sont plus pertinentes.

L'hypothèse H₃ revient à effectuer une renormalisation infinie de la charge et des champs, donnant lieu à des résultats finis pour les phénomènes physiques observables de polarisation du vide.

* Les deux calculs correspondent à des approximations différentes. On ne peut donc pas déduire l'un de l'autre bien qu'en général $j_{\mu}^{(pol)}$ soit relié à $\int \mathcal{L}$.

Un autre point important de cet article est le calcul de la polarisation du vide dans le cas du champ scalaire chargé de Pauli et Weisskopf (voir chap. 5 § 3). Cette théorie rend possible la création de paires sans que la mer de Dirac ait un quelconque équivalent dans ce cas où toutes les particules sont des bosons. Et malgré l'absence d'une mer infinie, la polarisation du vide diverge tout autant que dans la théorie de Dirac. Le traitement de Weisskopf s'applique sans difficulté particulière à ce cas et donne des corrections à l'électrodynamique classique de forme analogue. La vraie cause de la divergence tient dans la possibilité de créer des paires à grand moment et non pas dans la présence d'une mer infinie. Et celle-ci pouvait être éliminée d'emblée en utilisant comme Fock⁽⁸⁶⁾ la forme symétrique (par conjugaison de charge) de la quantification de l'onde de Dirac. Nous montrerons en Appendice le parallélisme complet entre les calculs effectués avec le champ scalaire chargé, et le champ de Dirac - Fock. A Weisskopf, attaché aux images physiques, l'idée de la mer ne devait cependant pas trop déplaire et il n'expliqua pas un tel parallélisme. Mais il est vraisemblable que Pauli en était conscient et en 1941⁽¹¹⁰⁾ il se prononça explicitement en faveur de la méthode symétrique de quantification du champ de Dirac qu'il attribuait (à tort) à Heisenberg^{(22)*}.

Nous n'avons jusqu'à présent traité que de la polarisation du vide induite par un champ extérieur classique. Le problème couplé d'électrodynamique quantique est autrement difficile. Heisenberg et Euler dans les articles cités⁽²²⁾⁽²⁴⁾ examinèrent les conséquences du procédé de soustraction sur le problème couplé. La soustraction a pour effet de modifier le terme de couplage ($j \cdot A$) dans le hamiltonien de l'électrodynamique quantique en rajoutant des termes d'ordre 2, 3 et 4 dans le champ A_μ et ses dérivées. Mais si ces nouveaux termes ont pour effet de supprimer les divergences dans le cas particulier où A_μ est imposé de l'extérieur, il n'en est rien dans le cas général : il reste, montre Heisenberg, une divergence logarithmique pour l'énergie propre du photon, ce qui est particulièrement catastrophique pour une particule sans masse ! et la divergence habituelle de l'énergie propre de l'électron n'est pas affectée :

* Sans doute Pauli pensait-il à cette méthode quand en 1943 il désapprouva "l'affirmation de Dirac que dans la théorie des trous, où l'on part avec les états d'énergie négative occupés, les équations sont plus compliquées que dans la forme plus ancienne de la théorie utilisant un nombre constant de particules chargées qui font des transitions vers des états d'énergie négative".

"de telles énergies propres interviennent toujours quand on calcule les contributions à l'énergie d'ordre supérieur au deuxième, qui résultent d'une transition virtuelle dans un autre état avec retour dans l'état de sortie"⁽²⁴⁾. Autrement dit, quand il y a des boucles ça diverge. Le problème des divergences de l'électrodynamique modifiée par les soustractions de Heisenberg fut repris par SERBER⁽⁸⁰⁾ (fév. 1956). Serber critiqua la manière dont Heisenberg se servait de ses soustractions : les résultats étaient sensibles à la manière d'introduire "l'écart à la diagonale" ξ dans la matrice densité et dans les produits de champ (par exemple $\psi(x) \psi^\dagger(x) \rightarrow \psi(x - \xi/c) \psi^\dagger(x + \xi/2)$) et à la manière de brancher la perturbation (adiabatique ou brutale) ; en dehors de ces points techniques, il n'y avait pas de raisons que l'efficacité des soustractions s'étende à un domaine plus vaste que celui qui les déterminait (où le champ électromagnétique est une source extérieure classique).

Ainsi la difficulté fondamentale de l'électrodynamique quantique restait sans solution. Cependant l'équation de Dirac était devenue tout à fait fiable, puisque la difficulté des états d'énergie négative, puis la difficulté qui s'y était substituée dans la théorie des trous c'est-à-dire la divergence de la polarisation du vide, étaient en 1936 surmontées : on savait alors extraire la partie physique et finie de la polarisation du vide par une opération qu'on appelle aujourd'hui renormalisation de charge. Les prédictions expérimentales finies qui en résultèrent ne purent être vérifiées que bien plus tard, par exemple, quand on sut qu'une partie du déplacement mesuré par Lamb était due à la polarisation du vide.

5.2 LA CATASTROPHE INFRAROUGE

L'électrodynamique quantique présente deux types fondamentaux de divergences. Leurs origines et leurs méthodes de traitement sont très distinctes. Les unes, les divergences ultraviolettes sont déjà présentes dans la théorie classique et elles sont liées à une représentation ponctuelle de l'électron, les autres les divergences infrarouges, n'apparaissent que dans la théorie quantique et sont liées à l'impossibilité de distinguer expérimentalement des processus ne différant que par l'émission de photons nous indétectables.

5.2.a L'apparition de la catastrophe

Cette nouvelle catastrophe apparaissait clairement dès 1931 dans les formules calculées par MOTT⁽¹⁾ et SOMMERFELD⁽²⁾ pour la "Bremsstrahlung". La probabilité d'émettre un photon à la fréquence ω pour un électron diffusé par un centre coulombien se comporte en $d\omega/\omega$. Ces deux auteurs* calculaient en utilisant une méthode semi-quantique et s'intéressaient directement à la perte d'énergie en $\omega \cdot \frac{d\omega}{\omega}$ qui, elle, ne diverge pas**. De plus, comme nous allons le voir le traitement correct de la divergence infrarouge montre que la perte ainsi calculée est exacte. En 1934, BETHE et HEITLER⁽³⁾, étudiant les différentes causes de ralentissement des électrons de la matière, effectuèrent un calcul entièrement quantique de la Bremsstrahlung. Alors l'interprétation du résultat en terme de photons s'imposait : "cette probabilité [d'émettre un quantum de fréquence comprise entre ω et $\omega + d\omega$]

* Le calcul de Sommerfeld, utilisant les fonctions d'ondes exactes dans un potentiel coulombien, présente l'avantage d'être valable même lorsque les vitesses initiale et finale de l'électron sont faibles, c'est-à-dire lorsque le traitement du potentiel par l'approximation de Born n'est plus légitime.

** Plus précisément, si l'on néglige l'effet d'écran des électrons entourant le noyau diffuseur l'intensité spectrale I_ω du rayonnement émis est en $\omega \cdot \frac{\ln \omega}{\omega}$: elle diverge, mais son intégrale $\int I_\omega d\omega \sim \int \ln \omega d\omega$ converge, comme il se doit.

est en gros proportionnelle à $\frac{1}{\omega}$, et par conséquent devient très grande pour l'émission de quanta de basse énergie". Sans plus d'émoi les deux auteurs tirent la conclusion pratique "Nous avons donc tracé sur la figure 2 l'intensité du rayonnement émis, plutôt que la probabilité elle-même en fonction de la fréquence".

La possibilité d'émettre des photons infrarouges dans d'autres processus fut bientôt reconnue : la désintégration β du neutron peut aussi s'effectuer suivant $n \rightarrow p + e^- + \bar{\nu} + \text{photon I.R.}$

En 1936 KNIPP et UHLENBECK⁽⁴⁾ calculèrent l'énergie emportée par le rayonnement infrarouge. Félix BLOCH⁽⁵⁾, indépendamment et dans la même année obtint le même résultat ; de plus il donna pour la première fois une explication de la divergence infrarouge : "en intégrant sur la fréquence de zéro à sa valeur maximum (permise par la conservation de l'énergie) on obtiendrait un résultat infini pour la probabilité de la désintégration β radiative [...]. la situation a son origine dans le traitement du champ électromagnétique comme une petite perturbation. C'est à cause de ce traitement qu'il apparaît le dénominateur d'énergie $E' - E - \hbar\omega$ (E et E' énergies de l'électron avant et après l'émission du quantum) qui pour des valeurs suffisamment petites de ω devient arbitrairement petit et finalement conduit à un résultat divergent pour la probabilité du processus radiatif".

5.2.b La clé du problème : BLOCH et NORDSIECK⁽⁶⁾ (mai 1937)

"Note sur le champ de radiation de l'électron"

Après avoir ainsi condamné la méthode de perturbation, Bloch trouva quelques mois plus tard avec Nordsieck quoi lui substituer. Cet article de Phys. Rev. contient l'essentiel du traitement moderne du problème.

"Cette "catastrophe infrarouge" est évidemment indépendante des difficultés ultraviolettes de l'électrodynamique quantique ...". Car elle n'a pas de pendant classique. Toutefois la théorie classique, poursuivent les auteurs, indique la cause de la difficulté : l'intensité spectrale I_ω de Bremsstrahlung classique tend vers une limite non nulle quand $\omega \rightarrow 0$ et donc $\frac{I}{\omega}$, interprété comme le nombre de photons émis, diverge. Au niveau quantique il faut s'attendre à ce que la probabilité d'émettre un nombre

fini quelconque de photons soit nulle, et que seuls les processus accompagnés de l'émission d'une infinité de photons mous soient possibles. De surcroît la théorie classique donne l'image qui transposée au niveau quantique devient la méthode de Bloch et Nordsieck : le rayonnement émis classiquement lors de l'accélération brutale d'un électron peut être, dans la limite des basses fréquences, calculé comme la différence entre le champ propre (transformé de Lorentz du champ statique coulombien) de l'électron initial et celui de l'électron final. Quantiquement (dans la terminologie moderne) Bloch et Nordsieck vont habiller les électrons de leurs photons mous virtuels.

Disons tout de suite que le point de départ, c'est-à-dire l'impossibilité de traiter le problème dans un cadre perturbatif était erroné. Nous verrons comment il fut établi dans les travaux ultérieurs que les divergences peuvent se compenser ordre par ordre, à condition de faire bon usage de la théorie des perturbations, c'est-à-dire d'ajouter toutes les contributions à un ordre donné. Néanmoins cette erreur eut une conséquence très fructueuse, elle permit de montrer très tôt que les photons mous pouvaient être traités non perturbativement, un fait remarquable de l'électrodynamique quantique dont la découverte aurait pu être longtemps différée si Bloch et Nordsieck n'avaient cru à l'impossibilité du traitement perturbatif.

Le calcul de Bloch et Nordsieck se développe en deux étapes :

1 - Il faut d'abord chercher les états stationnaires du hamiltonien des électrons et du champ de radiation couplés. Le fait remarquable est que dans une approximation non relativiste ce problème est soluble de manière non perturbative*. Les nouveaux états stationnaires $|\vec{v}, \{n_\alpha\}\rangle$ décrivent un électron de vitesse \vec{v} accompagné de n_α photons réels du type α et d'une infinité de photons virtuels. Pauli et Fierz ultérieurement soulignèrent que les n_α diffèrent des N_α associés au champ quantique libre de rayonnement

* L'approximation utilisée par Bloch et Nordsieck est semi-relativiste : le hamiltonien explicitement diagonalisable est obtenu en substituant dans le hamiltonien de Dirac $H_D = -i\vec{\alpha}\cdot\vec{V} + \beta m - e\vec{\alpha}\cdot\vec{A}$ $\frac{c}{v} \vec{\alpha} \cdot \vec{x}$ et $\sqrt{1 - \frac{v^2}{c^2}} \vec{\alpha} \cdot \beta$ et en ne gardant que les composantes de Fourier du champ électromagnétique de fréquence $\omega < \omega_1$ afin que l'occupation des états d'énergie négative soit négligeable ($\omega_1 \sim mc^2$).

et que ce sont eux qui correspondent aux photons détectés expérimentalement puisque les états associés sont stationnaires. Pour que l'approximation soit bonne la fréquence des photons virtuels doit être coupée à une valeur ω_1 suffisamment basse.

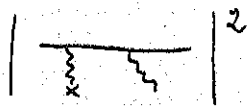
2 - Dans l'approximation de Born la probabilité de diffusion par le potentiel extérieur V est proportionnelle à $|\langle \vec{v}', \{n_\alpha\} | V | \vec{v}, 0 \rangle|^2 = P$ qui prend la forme $P = |\text{Amplitude de Mott}|^2 \times \prod_\alpha \left(e^{-c_\alpha} \frac{c_\alpha^{n_\alpha}}{n_\alpha!} \right)$ où c_α est une fonction de α et des polarisations et vitesses des électrons incident et sortant*. On trouve dans cette formule le facteur exponentiel typique e^{-c_α} interprété aujourd'hui comme la correction apportée à la diffusion de Mott par l'infinité des photons virtuels mous. Bloch et Nordsieck constatent comme souhaité que la probabilité d'émettre zéro ou un nombre fini de photons est nulle car $\sum_\alpha c_\alpha = +\infty$. La probabilité d'émettre un nombre quelconque de photons dans chaque mode redonne l'amplitude de Mott puisque $\sum_{n_\alpha} \frac{c_\alpha^{n_\alpha}}{n_\alpha!} = e^{c_\alpha}$. Et la perte d'énergie par Bremsstrahlung est celle que prévoit la théorie classique. A ce stade le problème pouvait sembler entièrement résolu. Pourtant le calcul précédent dissimulait une difficulté sérieuse, comme le virent Fierz et Pauli.

5.2.c La compensation des divergences au premier ordre de corrections radiatives

Quel est le rapport entre le résultat de Bloch et Nordsieck et le calcul perturbatif donnant la formule de Bethe et Heitler ? Une première réponse à cette question fut apportée en 1938 par BRAUNBECK et WEINMANN⁽⁷⁾. Ceux-ci

*On reconnaît (comme Fierz et Pauli) la forme d'une distribution des photons émis dans chaque mode, exprimant l'indépendance statistique de l'émission successive de photons : c_α représente la probabilité de l'émission d'un photon et $e^{-c_\alpha} \frac{c_\alpha^{n_\alpha}}{n_\alpha!}$ la probabilité d'émission indépendante simultanée de n_α photons. $\left(\sum_{\alpha/\omega \leq \omega_1} c_\alpha = c \frac{d\omega}{\omega} \text{ avec } c = \frac{2}{3\pi} \alpha_f \frac{(\Delta E)^2}{c^2} \right)$ ($\Delta \vec{v}$ variation de la vitesse de l'électron au cours du processus, α_f constante de structure fine). Rappelons qu'on appelle amplitude de Mott, l'amplitude relativiste de diffusion coulombienne dans l'approximation de Born et sans émission de photons secondaires.

calculèrent à l'ordre e^2 le produit $A = \pi_\alpha$ précédent d'abord pour zéro photon émis ($A = 1 + A_0$) puis pour un photon émis ($A = A_1$). A_1 s'identifie à la formule de Bethe et Heitler et présente la même divergence pour $\omega \rightarrow 0$; mais la somme $A = 1 + A_0 + A_1$ correspondant à l'amplitude de diffusion avec émission de zéro ou d'un photon ne contient plus de divergence. Toutefois elle contient $\ln \frac{\omega_1}{\omega_\ell}$ où ω_ℓ est la fréquence maximale du photon émis permise par la conservation de l'énergie, et ω_1 la coupure ultraviolette introduite par Bloch et Nordsieck. La valeur estimée de ω_1 donne une valeur négligeable aux corrections radiatives. Il n'en reste pas moins que leur détermination est entachée d'un certain arbitraire. Il est clair* aujourd'hui que A_0 et A_1 correspondent, dans la théorie des perturbations l'une aux corrections radiatives à la diffusion élastique et l'autre au processus inélastique à un photon sortant : respectivement $\left(\text{diagramme} + \text{diagramme} \right)^* \times \text{diagramme}$ et



. Les commentaires de Braunbeck et Weinmann sont plus

confus : ils semblent qu'ils n'avaient pas conscience de l'identité du coefficient exponentiel $e^{-c\alpha}$ à la série des perturbations de la diffusion élastique.

5.2.d L'effet de la coupure ultraviolette :

Dans le résultat de Braunbeck et Weinmann interviennent explicitement cette fréquence de coupure ω_1 et la fréquence limite ω_ℓ des photons réels émissibles. Il semble bizarre que ces deux fréquences n'interviennent pas dans le résultat non perturbatif de Bloch et Nordsieck. PAULI et FIERZ⁽⁸⁾, en 1938 expliquèrent ce paradoxe en montrant que la relation de conservation de l'énergie n'avait pas été tenue en compte. Les n_α ne peuvent différer de zéro que pour $\omega_\alpha < \omega_\ell$ alors que le produit π_α porte sur tous les modes tels que $\omega_\alpha < \omega_1$. Si $\omega_1 \gg \omega_\ell$ la partie exponentielle l'emportera toujours dans l'expression de A et la probabilité d'émettre un nombre quelconque de photons même infini devient

* En effet au même ordre d'approximation en terme de graphes de Feynman :

$$P = \left| \text{diagramme} + \text{diagramme} + \text{diagramme} \right|^2 + \left| \text{diagramme} \right|^2$$

$$= \left| \text{diagramme} \right|^2 \left[1 + \frac{1}{\left| \text{diagramme} \right|^2} \left(\left(\text{diagramme} + \text{diagramme} \right)^* \times \text{diagramme} + \text{c.c.} \right) + \left| \text{diagramme} \right|^2 \right]$$

est la probabilité au 2° ordre d'émettre zéro ou un photon

nulle : un électron ne serait jamais diffusé !

Outre cette constatation dramatique Fierz et Pauli donnent une méthode de calcul voisine de celle de leurs prédécesseurs, mais plus simple car entièrement non relativiste (le hamiltonien non relativiste est explicitement diagonalisable avec une coupure UV^{*}) et avec une introduction plus claire de la coupure U.V. sous forme d'une extension $\omega \sim \frac{c}{\omega_1}$ de la charge de l'électron.

La méthode utilisée par Pauli et Fierz pour diagonaliser le hamiltonien

$$\frac{P^2}{2m} - \frac{e}{m} \vec{P} \cdot \vec{A} + \frac{1}{2} \int (E^2 + B^2) d^3x$$

bien qu'elle ne soit dans leur article rien d'autre qu'une astuce de calcul, a pris par la suite une importance historique considérable. Les états propres correspondant aux électrons habillés étaient obtenus par une transformation canonique à partir des états nus. Réexprimé dans les nouvelles variables canoniques le hamiltonien présentait explicitement un terme de masse électromagnétique. Nous exposerons au chapitre 8 § 6 comment le lien de cette transformation avec la renormalisation de masse fut compris par Kramers et Opechowsky, puis par Tomonaga et Schwinger dans sa généralisation relativiste. La conclusion est très pessimiste : "la dépendance du résultat en ω_1 interdit une application à des électrons réels. Certes on pourrait penser [...] à remplacer ω_1 par une grandeur de l'ordre mc^2/\hbar , mais un tel procédé serait purement spéculatif".

En 1939, DANCOFF⁽⁹⁾ (voir aussi chap. 3 § 3) montra que le calcul entièrement relativiste des corrections radiatives à l'ordre e^2 ne sauvait pas la situation. Pour les grandes valeurs de ω , la contribution du processus inélastique s'annule ($\omega > \omega_1$). En revanche la correction radiative au processus élastique diverge logarithmiquement**.

* suffisamment basse pour que l'approximation dipolaire soit valable et pour que la contribution électromagnétique à la masse de l'électron des modes de fréquence $< \omega_1$ puisse être négligée. Soulignons que Pauli et Fierz sont conscients de l'ambiguïté de la définition de la masse.

** Tomonaga, lorsqu'il prononça sa conférence Nobel⁽⁸⁵⁾, attribua à Dancoff l'idée d'absorber la divergence de la probabilité de diffusion dans une modification de la masse ; et il prétendit même que seule l'erreur de calcul de Dancoff (le coefficient de la divergence logarithmique était inexact car Dancoff avait oublié les modes longitudinaux du champ de radiation) avait pu différer la naissance de la théorie de la renormalisation. Dans l'article de Dancoff nous n'avons trouvé aucune trace d'une telle entreprise. Le mot de renormalisation y est employé dans un sens bien différent : il s'agit seulement de la normalisation des fonctions d'onde initiale et finale de l'électron, changée parce qu'en langage figuré un électron n'a plus en présence du champ de rayonnement, la probabilité d'être lui-même, puisqu'il peut engendrer des photons et même des paires virtuelles. Certes Kramers avait déjà parlé à cette époque de redéfinition de la masse, mais Dancoff n'exploita ni le mot ni l'idée.

Alors qu'initialement Bloch et Nordsieck avaient affirmé l'indépendance des deux types de catastrophe, la situation devenait plus confuse : il semblait que l'émission infrarouge d'une charge accélérée devait dépendre dramatiquement de la forme de l'électrodynamique à très haute fréquence. Ce paradoxe ne devait être dissipé que par l'application du principe de renormalisation.

La compensation des divergences infrarouges fut prise comme critère par BETHE et OPPENHEIMER⁽¹⁰⁾ pour tester en 1946 la validité de la méthode de soustraction de Heitler (voir chap. 7 § 4) et en 1941 par Pauli et Jauch⁽¹²⁾ pour toute la théorie de Dirac de 1942 (voir chap. 7 § 3). Dans les deux cas le test fut négatif. L'article de Bethe et Oppenheimer présente l'avantage d'offrir, juste avant la théorie de la renormalisation, un résumé clair de la situation du problème infrarouge et l'on y trouve clairement exprimé le mécanisme de compensation ordre par ordre : il faut à un ordre donné tenir compte simultanément des processus réels et virtuels puisque l'expérience ne permet pas de distinguer entre deux processus qui ne diffèrent que par l'émission de photons mous indétectables ; la somme des deux types de processus ne présente plus de divergence infrarouge (la démonstration n'existait alors qu'au premier ordre, mais le résultat global de Bloch et Nordsieck laissait pressentir la généralisation). Dans la méthode de Bloch et Nordsieck la probabilité d'émettre un nombre fini de photons (réels) est nulle car cette théorie prend en compte une infinité de processus virtuels divergents à tous les ordres se résumant sous forme d'une $\exp(-\infty)$ pour donner 0.

Que l'on prenne le problème ordre par ordre ou globalement, il subsiste après élimination de la divergence infrarouge une dépendance dans une coupure ultraviolette.

Avant de donner un bref aperçu de l'évolution ultérieure du problème, mentionnons encore l'application de la méthode de Bloch et Nordsieck à l'effet Compton double (deux photons dans l'état final) par JOST⁽¹³⁾ en 1947.

L'effet Compton multiple était déjà exposé dans le cours de Heitler⁽¹⁴⁾ (2ème éd. 1944) mais sans mention spéciale du cas particulier où l'un des deux photons émergents appartient au spectre infrarouge. Le comportement divergent correspondant fut mis en évidence par Eliezer⁽¹¹⁾ en 1946 avec la théorie de Dirac de 1942. En l'appliquant à ce problème Jost donna de la méthode de Bloch et Nordsieck une formulation plus mathématique, avec transformation canonique (unitaire) et hamiltonien effectif, (généralisant ceux de Pauli et Fierz).

En 1947 l'origine physique de la catastrophe infrarouge et la bonne façon de poser le problème étaient connues. Les progrès ultérieurs après la théorie de la renormalisation montrèrent que les critiques de Fierz et Pauli étaient exagérément pessimistes : la coupure ultraviolette joue un rôle inessentiel dans le problème, elle disparaît complètement après renormalisation dans le traitement perturbatif et dans ce nouveau cadre la validité de l'approximation que constitue la méthode de Bloch et Nordsieck peut être précisée. Citons pour une description plus précise les excellents articles de Jauch et Rohrlich⁽¹⁵⁾ (1954) et de Morpurgo⁽¹⁶⁾ (1950).

Dans un autre contexte, nous reparlerons au chapitre 8 de la solution de Bloch et Nordsieck : la notion de champ propre de l'électron et d'états "renormalisés" (suivant l'expression de Bethe et Oppenheimer) par des photons virtuels jouèrent un rôle important dans la théorie de la renormalisation. N'était-ce pas au même congrès de Bologne d'octobre 1937 que Pauli et Fierz exposèrent leur traitement du problème infrarouge et Kramers insista sur la nécessité d'isoler la masse expérimentale dans les calculs d'électrodynamique.

5.3 LA RELATION ENTRE SPIN ET STATISTIQUE

Bien que le rôle fondamental des propriétés de symétrie en mécanique quantique soit apparu dès le début du développement de son appareil mathématique, les arguments issus de la théorie des groupes ne furent que progressivement acceptés et exploités. Dans la préface de la traduction anglaise de son célèbre ouvrage⁽⁴⁰⁾ Wigner remarque *"Quand la version originale allemande fut publiée pour la première fois en 1931, il y eut beaucoup de réticence de la part des physiciens pour accepter des arguments de théorie des groupes et le point de vue de la théorie des groupes. Il me plaît de constater que cette réticence s'est graduellement éteinte entre-temps et qu'en fait la nouvelle génération ne comprend pas les causes et la base de cette réticence"*. Prenant à l'extrême le point de vue "groupiste" on pourrait dire que la théorie des champs d'aujourd'hui n'est rien d'autre qu'une dynamique quantique des représentations de groupe de Lorentz-Poincaré. L'étude du lien entre spin et statistique sera pour nous le fil conducteur qui nous permettra d'explorer nombre d'applications des propriétés de symétrie à la jeune théorie quantique des champs. Dans la discussion historique sont intervenues l'invariance relativiste, les symétries discrètes d'inversion d'espace et de temps et la conjugaison de charge. Certains aspects de l'invariance relativiste liés à l'existence d'un tenseur d'énergie-impulsion conservé furent, nous l'avons vu, éclaircis par Heisenberg et Pauli dès la première étude systématique de la quantification des champs. Nous ne reviendrons pas non plus sur l'invariance de jauge et ses implications puisqu'elle n'intervient pas ici. De toute façon nous ne dresserons qu'un panorama incomplet des symétries étudiées avant 1947 (nous ne traiterons pas de l'échange du passé et du futur, ni de l'invariance conforme ...).

Comme Wolfgang PAULI fut l'inventeur du principe d'exclusion⁽⁴¹⁾, on ne s'étonnera pas de retrouver son nom tout au cours de la discussion du lien spin-statistique. A la fin de l'article original⁽⁴¹⁾ de 1925 Pauli conclut : *"le problème des fondements plus précis de la règle générale ici postulée concernant la présence d'électrons "équivalents" dans l'atome pourrait bien n'être soluble qu'après un nouvel approfondissement des principes fondamentaux de la mécanique quantique"*. Un espoir de Pauli qui

se trouva en bonne partie réalisé. Il est même probable que le nombre d'articles qu'il publia sur ce sujet ne donne qu'une faible idée de l'importance qu'il lui accorda. C'est pour la découverte du principe d'exclusion que le prix Nobel lui fut attribué en 1945. La cérémonie n'eut lieu qu'en 1947 et dans son discours de réception Pauli ne manqua pas d'expliquer comment la théorie quantique des champs, malgré toutes ses imperfections, avait pu justifier ce fait expérimental évident une fois connues les notions de spin et de statistique quantique : les particules de spin demi-entier vérifient la statistique de Fermi-Dirac et celles de spin entier vérifient la statistique de Bose-Einstein*. Quand fut connue la seconde quantification (voir chap. 2 § 1) la règle s'énonçait ainsi : les champs de spin entier (resp. demi-entier) doivent être quantifiés à l'aide de commutateurs (resp. d'anticommutateurs). A ce stade l'indiscernabilité des particules quantiques était expliquée : il fallait substituer la notion de nombre d'occupation d'un état à celle de particule, les deux statistiques d'Einstein-Bose et de Fermi-Dirac semblaient naturellement privilégiées par la quantification canonique, mais a priori les deux choix étaient formellement possibles pour chaque valeur du spin.

Afin de lever cette dernière ambiguïté quatre types d'arguments furent avancés (jusqu'en 1947) :

- la positivité de l'énergie permet de trancher pour le spin demi-entier.
- la localité des observables physiques (leur commutateur en deux points séparés par un intervalle du genre espace doit s'annuler) permet de régler le cas des spins entiers. Ces deux premiers arguments furent avancés par Pauli.
- l'impossibilité formelle de la quantification canonique avec des anticommutateurs pour des spins entiers (Iwanenko et Sokolov, de Wet)
- l'invariance par conjugaison de charge (Belinfante).

* Bien sûr les seuls exemples disponibles étaient alors l'électron, le proton et le photon pour ce qui est des particules élémentaires. Il ne fut même pas tout de suite clair que la statistique suivie par une particule complexe ne dépendait que de celle de constituants élémentaires.

Disons tout de suite que la preuve la plus générale (Pauli) s'obtient à l'aide des deux premiers postulats. Bien sûr elle n'est établie que pour les champs libres avant 1947 puisque le problème des champs en interaction souffrait des divergences ultraviolettes.

5.3.a Le cas des particules élémentaires connues (dans les années trente)

Quand en 1929 Heisenberg et Pauli dans leur premier article sur la mécanique quantique des champs d'onde⁽⁴²⁾ quantifièrent l'onde ψ de Dirac ils remarquèrent que les observables physiques (dont H) étaient quadratiques en ψ et que par conséquent des relations de commutation pour les observables pouvaient* se traduire alternativement par des relations de commutation ou d'anticommutation pour les ψ tout en préservant l'équation du mouvement $i \frac{\partial F}{\partial t} = [H, F]$ pour un champ F : c'est une conséquence de l'identité remarquable : $[BC, A] = B \{A, C\} - \{B, A\}C$. Ils ne purent préférer l'une des deux statistiques : "On voit que même du point de vue de la quantification des ondes et du traitement invariant relativiste du problème à plusieurs corps les deux types de solution, c'est-à-dire la statistique de Bose-Einstein d'une part et le principe d'exclusion ("Äquivalenzverbot") d'autre part, continuent d'apparaître formellement tout aussi justifiés l'un que l'autre, et nous ne pouvons donner une explication satisfaisante du choix de la seconde possibilité par la nature". Il leur manquait l'interprétation de Dirac des états d'énergie négative, c'est-à-dire la théorie des trous. (décembre 1929).

Jordan, dans son étude historique de l'hypothèse des quanta⁽⁴³⁾ avait eu connaissance avant leur publication des travaux de Heisenberg et Pauli mais il pensa au contraire : "Il n'est mathématiquement pas possible d'effectuer la quantification relativiste suivant Bose ou Pauli indifféremment ; seule la seconde possibilité est ouverte (pour les électrons). Le fait que les électrons obéissent au principe de Pauli et pas à la statistique de Bose est donc étroitement lié à la théorie de

* Jordan et Wigner n'avaient pas utilisé cet argument pour introduire des anticommutateurs (voir chap. 2 §1)

Dirac de l'électron tournant. Enfin il n'y a pas d'énergie de zéro pour les ondes de matière" (Jordan croyait vraisemblablement que la statistique de Fermi devait être préférée parce qu'elle éliminait cette énergie de zéro). Dans cet article s'adressant à un public large, Jordan ne pouvait donner des explications plus précises. Quoi qu'il en soit Heisenberg et Pauli condamnèrent dans une note (p. 30) de l'article cité le résultat de Jordan. "du reste dans aucun des deux cas (statistique de Bose ou de Fermi) n'interviendrait une énergie de zéro des ondes de matière" (puisque à cette époque $H = \psi^\dagger (\vec{\alpha}(\vec{p} - e\vec{A}) + \beta m) \psi$ et ψ ne contient que des annihilateurs). Jordan pensait aussi (dans le même article) que la mauvaise statistique était exclue dans le cas du champ électromagnétique parce qu'alors on aurait à la place de la fonction Δ (de Jordan-Pauli) une autre singularité n'admettant pas de détermination mathématique précise. Cette justification n'est là aussi pas la bonne (on aurait Δ_1 à la place de Δ). En 1929 la théorie quantique des champs n'était pas suffisamment mûre pour conclure*.

La situation avait changé en 1933. Le débat sur la mesurabilité des champs quantiques avait abouti à une conclusion positive de Bohr et Rosenfeld (voir chap. 4 § 3). Il était établi que dans la mesure où les charges -épreuves pouvaient être traitées classiquement, le champ électromagnétique quantique était accessible à la mesure. Si tel était le cas, conclua Pauli dans son article de "Handbuch" 1933⁽⁴⁴⁾, il fallait prendre des commutateurs pour la quantification, respectant ainsi la règle de commutation d'observables séparées par un intervalle du genre espace.

Dans la même année Dirac présenta sa théorie des trous au congrès Solvay. Lors de la discussion⁽⁴⁵⁾ Pauli s'exprima ainsi : "La théorie des lacunes m'a toujours paru très intéressante, à cause du rôle essentiel qu'y joue le principe d'exclusion. Tandis que ce principe n'était autrefois qu'une règle isolée dont la validité était indépendante de celle des autres bases de la théorie du quanta, la théorie des lacunes, introduite par Dirac pour échapper à la difficulté des masses négatives, aurait été impossible

* Pauli publia un commentaire⁽⁸⁹⁾ de l'article de Jordan où il signala comme une erreur le passage concernant le spin et la statistique.

si l'on n'avait pas voulu exclure toute fonction d'onde qui n'est pas antisymétrique". Ce raisonnement effectué dans la théorie intuitive des lacunes de Dirac, se transforme aisément dans la théorie moderne où l'on quantifie le champ de Dirac : l'énergie $H = \sum_{\alpha} E_{\alpha} (a_{\alpha}^{\dagger} a_{\alpha} - b_{\alpha} b_{\alpha}^{\dagger})$ ($E_{\alpha} > 0$, a_{α} et b_{α} annihilateurs d'électrons et de positrons resp.) n'est positive définie (pour sa partie opératorielle) que si b et b^{\dagger} anticommute (à un c-nombre près) $\rightarrow H = \sum_{\alpha} E_{\alpha} (a_{\alpha}^{\dagger} a_{\alpha} + b_{\alpha}^{\dagger} b_{\alpha}) + \text{c-const}$.

Ainsi en 1933 le problème "spin-statistique" était classé pour les spins $\frac{1}{2}$ et 1 et l'on pouvait s'estimer satisfait puisque toutes les particules connues à cette époque (électron, proton, photon, positron et neutrino) rentraient dans cette catégorie.

5.3.b Le développement de l'intérêt pour de nouveaux champs :

Il fallut attendre longtemps avant qu'une particule de spin nouveau fût découverte (le pion de spin 0 en 1947) et la découverte de gravitons de spin $\frac{3}{2}$ n'est pas pour demain : l'étude de nouveaux champs quantiques commencée par Pauli et Weisskopf en 1934 pouvait paraître d'un intérêt académique : il s'agissait d'une simple extension formelle. La possibilité de décrire ainsi les particules complexes, dans la mesure où l'on peut négliger leur structure interne, ne pouvait à elle seule justifier la recherche de nouvelles équations a priori complexes pour des spins élevés. Le stimulus vint surtout de la physique nucléaire (voir chap. 6 § 3). Dans ce domaine les données expérimentales s'accumulaient et la théorie des champs avait déjà remporté ses premiers succès avec l'interaction de Fermi⁽⁴⁶⁾ pour interpréter l'interaction β (1933). La théorie de YUKAWA de 1935⁽⁴⁷⁾ utilisait une particule scalaire pour décrire l'interaction nucléaire. Elle eut peu de succès lors de sa publication, mais deux ans plus tard on crut avoir découvert le méson de Yukawa : ce n'était que le μ et la confusion dura encore dix ans. Proca⁽⁴⁸⁾ élabora la théorie classique des champs vectoriels massifs, qui après quantification par Stueckelberg⁽⁴⁹⁾ et Kemmer⁽⁵⁰⁾ put être utilisée pour donner une version vectorielle massive de l'interaction nucléaire (Yukawa-Sakata-Taketani⁽⁵¹⁾, Stueckelberg⁽⁴⁹⁾, Kemmer-Fröhlich-Heitler⁽⁵³⁾, Bhabha⁽⁵⁴⁾). Enfin la théorie pseudoscalaire de Kemmer⁽⁵⁰⁾ date de la même année (1938) et c'est elle qui devait être la

bonne pour décrire le méson π . Les travaux cités suffisent à justifier l'intérêt d'un formalisme général s'appliquant à tout spin. Fierz dans son étude systématique des équations décrivant des particules de tout spin fit mention des travaux de Proca Stueckelberg et Kemmer, Yukawa ... et releva même des erreurs dans l'expression du courant (Sakata Yukawa⁽⁵⁴⁾).

Le côté dynamique et physique des spins élevés (ou du moins différents de 1/2 et de 1), on le voit, commença à susciter l'intérêt des théoriciens seulement à partir disons de 1934. L'aspect purement mathématique, c'est-à-dire l'étude des représentations du groupe de Lorentz fut développé bien plus tôt.

Les représentations du groupe de Lorentz (propre)* (L_+^\uparrow)

Dans "les équations d'onde de l'électron", un article qui suivait de près celui de Dirac, DARWIN fut un des premiers à voir ce que la covariance de l'équation de Dirac avait de singulier : "La théorie de la relativité ne repose sur rien d'autre que sur l'idée d'invariance, et on tire la notion de tenseur comme une conséquence nécessaire ; et il est plutôt déconcertant de trouver qu'apparemment quelque chose s'est échappé du filet, si bien qu'il existe des quantités physiques qu'il serait très incorrect, convenons-en, d'appeler tenseur". La suggestion qui suit, de lever la difficulté dans une théorie ultérieure donnant une origine électromagnétique ou /et gravitationnelle à la masse de l'électron est bien sûr erronée. L'expression de la loi de transformation du ψ de Dirac en terme de tenseurs ne saurait être comme le remarque Darwin que très artificielle. Landau, lui, exprimait cette bizarrerie en appelant le ψ de Dirac un demi-vecteur : à une rotation d'un angle θ dans l'espace physique correspond une "rotation" de $\frac{\theta}{2}$ des composantes du ψ : par exemple si la probabilité de mesurer $\frac{\hbar}{2}$ suivant l'axe Oz est 1, celle de mesurer $\frac{\hbar}{2}$ suivant un axe faisant un angle θ avec Oz est $\cos^2 \frac{\theta}{2}$ et non pas $\cos^2 \theta$. La bizarrerie est donc présente quand il s'agit du seul groupe des rotations spatiales. WIGNER et VON NEUMANN⁽⁵⁶⁾ montrèrent en décembre 1927

* On appelle ainsi la composante connexe de l'unité du groupe d'invariance de la forme $t^2 - \vec{x}'^2 = \alpha^2$, c'est-à-dire le groupe des transformations de Lorentz de déterminant +1 et conservant le signe du temps : ainsi les opérations de parité P ($\vec{x} \rightarrow -\vec{x}$) et d'inversion du temps ($t \rightarrow -t$) sont exclues.

que la théorie de Pauli de l'électron datant du mois de mai de la même année, était invariante par rotation spatiale et que les spineurs de Pauli correspondaient à l'utilisation de la représentation "bivaluée" de SO_3 par SU_2 (en notations modernes)*

Pour l'invariance relativiste la situation est tout à fait analogue : les représentations de L_+^\uparrow s'obtiennent à partir du groupe SL_2 des matrices 2×2 complexes de déterminant unité. La théorie mathématique correspondante était en gros contenue depuis longtemps dans la théorie des invariants des formes bilinéaires, dont celle de la relativité restreinte est un cas particulier, et dans la théorie des représentations de Frobenius et de von Burmide datant de la fin du 19ème (voir van der Waerden⁽⁵⁷⁾). Dans la première édition de 1928 du livre de WEYL "Gruppentheorie und Quantenmechanik"⁽⁵⁸⁾ on trouvait déjà décrites toutes les représentations de dimension finie de SL_2 :

$$L_+ \text{ s'obtient à partir de } SL_2 \text{ suivant } \begin{array}{ccc} SL_2 & \longrightarrow & L_+^\uparrow \\ A & \longrightarrow & \Lambda(A) \end{array}$$

$\Lambda(A)$ associe x' à x

suitant $\tilde{x}' = A^\dagger \tilde{x} A$, $\left\{ \begin{array}{l} \tilde{x} = \alpha_0 - \vec{\alpha} \cdot \vec{\sigma} \\ \tilde{x}' = \alpha'_0 - \vec{\alpha}' \cdot \vec{\sigma} \end{array} \right.$ $\vec{\sigma}$ désignant les trois matrices de Pauli

x' est bien réel car $\tilde{x}'^* = (\tilde{x}')^\dagger$ et $\tilde{x}'^\dagger = \tilde{x}'$ si $\tilde{x}^\dagger = \tilde{x}$
 et $\alpha'^2 = \det \tilde{\alpha}' = \det \tilde{\alpha} = \alpha^2$ puisque $\det A = 1$

* Le lecteur physicien est en général plutôt familier avec l'aspect infinitésimal du problème : dans les cours élémentaires de mécanique quantique on recherche les différentes représentations matricielles possibles des relations de commutation des composantes du moment cinétique $[L_x, L_y] = i L_z$ etc. , qui ne sont autres que les générateurs infinitésimaux d'une représentation du groupe des rotations. Ce point de vue était celui d'Elie Cartan⁽¹⁰⁹⁾ quand il découvrit en 1913 les représentations spinorielles du groupe des rotations. Pour le lecteur matheux voici le fait fondamental expliquant l'existence des spins 1/2 : le groupe SO_3 n'est pas simplement connexe ; en conséquence pour obtenir toutes ses représentations (au sens large) finies il faut partir de son groupe de revêtement SU_2 . De même L_+^\uparrow a pour revêtement SL_2 .

$\Lambda(A) = \Lambda(-A)$, d'où le terme de représentation "bivaluée" de L_+^{\uparrow} pour SL_2 . *

Puis H. Weyl donnait la manière d'obtenir les autres représentations irréductibles de dimension finie de SL_2 en faisant agir $SL_2 \times SL_2^*$ dans l'espace des polynômes symétriques de degré fixé à quatre indéterminées.

La description la plus utilisée aujourd'hui de ces représentations est due à B.L. VAN DER WAERDEN⁽⁵⁹⁾ (1929). Ehrenfest appela "spineurs" les quantités introduites par Pauli ou par Dirac et, rapporte van der Waerden⁽⁵⁷⁾, posa la question suivante lors de son passage à Göttingen pendant l'été 1929 : "Existe-t-il une analyse spinorielle, que tout physicien puisse apprendre comme l'analyse tensorielle et à l'aide de laquelle on puisse exprimer toute espèce possible de spineurs et toute équation invariante reliant des spineurs ?" Pour une explication plus détaillée du lien entre la théorie des invariants des formes quadratiques et le calcul spinoriel, le lecteur peut se référer à (57). Une représentation irréductible de dimension finie du groupe de Lorentz notée $\mathcal{D}(\frac{1}{2}, \frac{k}{2})$ s'obtient tout simplement en spécifiant l'action de SL_2 sur des "spineurs" $\xi_{\alpha_1, \alpha_2; \beta_1, \dots, \beta_k}$ symétriques par permutations des indices non pointés d'une part et des pointés d'autre part $\xi_{\alpha, \beta} \rightarrow \xi'_{\alpha', \beta'} = A_{\alpha\alpha'} A^{\beta\beta'} \xi_{\alpha\beta}$ (pour deux indices). Le cas du vecteur correspond à $\mathcal{D}(\frac{1}{2}, \frac{1}{2})$ car $\tilde{\xi}_{\beta\alpha} = \xi_{\alpha\beta}$ se transforme comme ci-dessus.

Les spineurs de Dirac se transforment suivant $\mathcal{D}(\frac{1}{2}, 0) \oplus \mathcal{D}(0, \frac{1}{2})$ van der Waerden exprima l'équation de Dirac en terme de spineurs :

$$\begin{cases} (i\partial_{\alpha\beta} - e A_{\alpha\beta})\varphi_{\beta} - m\chi_{\alpha} = 0 \\ (i\partial_{\alpha\beta} - e A_{\alpha\beta})\chi_{\alpha} + m\varphi_{\beta} = 0 \end{cases}$$

les deux spineurs χ_{α} et φ_{β} donnent les composantes de ψ

*Pressentant un rôle important de la théorie des groupes dans l'équation de Dirac, Pauli écrivit à Weyl le 29 janvier 1928 : "Récemment Dirac a réussi à mettre les équations de l'électron magnétique sous une forme invariante relativiste ... (son travail sortira dans le numéro de février des Proc. Roy. Soc.). Voilà encore pour vous un bel exemple pour la théorie des groupes !" (lettre [181]⁽⁸⁸⁾)

LAPORTE et UHLENBECK⁽⁶⁰⁾ donnèrent en 1931 l'expression des équations de Maxwell en notation spinorielle et étudièrent les invariants associés à ces différentes équations.

Dans l'équation de Dirac on gagnait ainsi une représentation manifestement covariante, mais bien sûr la covariance avait déjà été prouvée dès son invention par Dirac. La démonstration de Dirac⁽⁶¹⁾ utilisait simplement le fait que deux choix de γ_μ , disons γ_μ et γ'_μ obtenus à partir de γ_μ en transformant γ_μ comme un 4-vecteur, ne diffèrent que par une transformation unitaire S : $\gamma'_\mu = S \gamma_\mu S^{-1}$. Darwin⁽⁵⁵⁾ avait explicité la transformation du ψ dans le cas d'une transformation bidimensionnelle.

Tout comme dans l'étude de l'invariance par rotation, une autre démarche consiste à étudier les représentations des relations de commutation des opérations infinitésimales du groupe (algèbre de Lie). Ce procédé fut utilisé par MAJORANA⁽⁶²⁾ en 1932 pour aboutir à une représentation de dimension infinie (ainsi que "l'équation de Dirac" correspondante) contenant implicitement toutes les représentations irréductibles de dimension finie. Bien qu'il donnât l'espoir de trouver une dynamique invariante relativiste pour des particules de spin quelconque, cet article n'était guère utilisable puisque tous les spins se trouvaient mêlés.

En 1936 DIRAC⁽⁶³⁾ donna une solution très astucieuse de ce problème, faisant usage d'une écriture spinorielle des relations de commutation. Il trouva les équations invariantes pour des spins quelconques d'abord sous une forme analogue à celle utilisée pour l'électron puis, in fine sous forme spinorielle. Nous verrons bientôt les critiques que Fierz apporta à cette nouvelle théorie de Dirac.

Enfin une étude systématique et rigoureuse des représentations du groupe de Lorentz fut publiée par WIGNER⁽⁶⁴⁾ en 1939. Il apparut que l'essentiel des résultats connus sur les représentations de dimension finie était exact. Dans une étude ultérieure⁽⁶⁵⁾ Wigner* étudia les différentes équations d'onde possibles mais par une méthode originale délivrée de toute hypothèse a priori sur la forme de ces équations : il s'agit de trouver

* Cet article contient une bibliographie détaillée de la question.

les variétés invariantes relativistes correspondant à l'espace des solutions et de remonter ensuite à l'équation. Les résultats de Wigner s'accordaient avec ceux de Dirac et de Fierz, mis à part le cas particulier de masse nulle pour lequel on retrouve les solutions de Dirac et Fierz où chaque particule n'a que deux états de spin possible (hélicités + et -) mais aussi de nouvelles solutions avec une variable de spin continue, qui semblent dépourvues de signification physique.

L'utilisation de particules nouvelles en physique nucléaire (les mésons) l'étude mathématique des dynamiques classiques permises par la relativité, et enfin le problème de la quantification des champs correspondants progressèrent simultanément et avec d'importantes influences réciproques dans les années trente. Il nous reste à aborder le dernier problème et cette restriction sur la manière de quantifier les champs qu'apporte le "théorème" spin-statistique.

5.3.c Le champ scalaire de PAULI et WEISSKOPF : (1934)

Dans le cas du champ scalaire chargé, interagissant avec un champ électromagnétique, on ne peut utiliser l'argument valable pour le champ électromagnétique $F^{\mu\nu}$: ce dernier est directement mesurable et donc doit être quantifié avec des commutateurs. Inversement le champ scalaire φ ne peut être pris comme observable physique en raison de l'invariance de jauge qui rend la "phase" de φ inobservable : la théorie est invariante par la substitution $\varphi \rightarrow e^{i\epsilon\xi} \varphi$ et $A_\mu \rightarrow A_\mu - \partial_\mu \xi$. Il fallait donc dans ce cas trouver la contradiction apportée par une quantification à l'aide d'anticommutateurs.

Avant d'aborder ce point il convient de préciser comment Pauli et Weisskopf furent amenés à quantifier le champ scalaire et quelles conséquences remarquables ils en tirèrent. Il était clair en 1934 après le développement de la théorie des trous, la découverte du positron et le calcul de la polarisation du vide, que l'équation de Dirac ne pouvait avoir de signification cohérente dans le cadre du problème à un corps. Il fallait réinterpréter cette équation dans une théorie à plusieurs corps soit par la méthode intuitive de la mer de Dirac, ou mieux (tout au moins suivant l'Ecole russe) par la méthode formelle de la deuxième quantification de Wigner et Jordan. Ainsi tombait l'argument de Dirac qui l'avait conduit

à son équation et selon lequel il fallait trouver pour l'électron une équation différentielle du premier ordre pour que la théorie des transformations habituelles utilisée dans le problème à un corps puisse s'appliquer. (Voir chap. 2 § 2). Et l'équation de Klein-Gordon ne pouvait plus être rejetée suivant cet argument. L'absence de densité positive conservée de particules, qui avait conduit Dirac à rejeter la théorie scalaire, attira au contraire l'attention de Weisskopf : il soupçonna⁽⁸⁴⁾ (p. 11) que la croissance de la densité $|\varphi|^2$ au cours du temps avec conservation de la densité de charge $i(\psi^* \dot{\varphi} - \dot{\varphi}^* \psi)$ devait correspondre à une création de paires. Comme il n'avait pas l'habitude des opérateurs d'annihilation et de création il sollicita l'aide de Pauli. Même une citation bien choisie des "Maîtres Chanteurs" ne suffit pas à dissiper la mauvaise humeur momentanée de Pauli (" Ach, Meister, warum so viel Eifer warum so wenig Ruh, mich dünkt euer Urteil wäre reifer, hörtet ihr besser zu^{*} "). Pauli n'aimait pas Wagner, et Weisskopf ne fut mieux reçu que deux jours plus tard⁽⁸⁴⁾. Il résulta de cette fructueuse collaboration que la création de paires était possible pour des particules sans spin et sans équation de Dirac.

La quantification canonique appliquée au champ libre complexe de Klein-Gordon permit d'introduire naturellement des particules conjuguées de charge et ce sans faire appel à une mer infinie comme dans la théorie de Dirac. Pauli et Weisskopf insistèrent sur cette simplification, sans montrer toutefois qu'en réalité le même type de décomposition $\psi = \sum_{k, k_0 > 0} a_k e^{-ikx} + b_k^\dagger e^{ikx}$ (a_k annihilateur de quantum de charge négative, b_k créateur de quantum de charge positive) pouvait être utilisé pour le champ de Dirac et qu'à cet égard la théorie du champ de Dirac libre n'était pas plus complexe que celle du champ scalaire libre. IWANENKO et SOKOLOV⁽⁶⁷⁾, en 1937 insistèrent sur la similitude formelle des deux théories. Pauli et Weisskopf ne manquèrent pas de voir ce que donnait leur théorie pour les problèmes à la mode comme la polarisation du vide (voir § précédent) et la self-énergie de l'électron (dans un autre article de Weisskopf, voir chap. 3 § 3) ; malgré l'absence

* Ah, Maître, à quoi bon tant d'emportement et si peu de calme, il me semble que votre jugement serait plus mûr, si vous m'écoutez mieux.

du concept de mer infinie ils aboutirent à des résultats analogues, sans surprise. Pauli suivant son mépris de la "physique des soustractions" appela son papier écrit avec Weisskopf le papier "anti-Dirac", bien que les difficultés de fond soient les mêmes que dans la théorie des trous de Dirac. Les deux auteurs, à une époque où les particules scalaires n'étaient que fiction, publièrent dans les "Helvetica physica Acta", "revue vénérable mais pas couramment lue"⁽⁸⁴⁾.

Loin d'être le centre de l'article, la question du choix des commutateurs pour la quantification n'occupe qu'un passage très bref du texte : "Comme déjà mentionné dans l'introduction, un développement non contradictoire de la théorie ondulatoire scalaire relativiste n'est pas possible pour des particules satisfaisant au principe d'exclusion, car dans le cas de la statistique de Fermi, comme le montre une détermination plus précise du hamiltonien dans les variables a et b (annihilateurs "d'électrons" et de "positrons") on ne peut obtenir l'invariance relativiste du quadri-courant. Cela vient aussi du fait que des équations $\varphi(x) \varphi^\dagger(x') + \varphi^\dagger(x') \varphi(x) = 0$ et $\varphi(x) \varphi(x') + \varphi(x') \varphi(x) = 0$ il résulterait $\varphi(x) = 0$ et $\varphi^\dagger(x) = 0$ "^{(66) p.722}. Le premier argument nous est resté obscur (s'agit-il de la nullité du hamiltonien $H = \sum_k \frac{1}{2} E_k \{ \frac{1}{2} a_k, a_k^\dagger \} + \frac{1}{2} b_k, b_k^\dagger \}$, ou d'une violation de localité $[\rho(x), \rho(y)] \neq 0$ avec $(x-y)^2 < 0$?). En revanche le second exprime probablement le fait que lors de la quantification canonique d'un champ scalaire complexe par des anticommutateurs $\{ \varphi(x), \varphi^\dagger(x') \} = i \Delta(x-x')$ En particulier pour $x = x'$ $\varphi(x) \varphi^\dagger(x) + \varphi^\dagger(x) \varphi(x) = 0$ ce qui implique $\varphi = 0$.*

La question fut reprise de manière beaucoup plus détaillée par Pauli lors d'une conférence à l'Institut Poincaré⁽⁶⁸⁾ (1936). Pauli ne parla pas de l'argument précédent. Mais il montra que si l'on choisit pour les annihilateurs-créateurs les relations $\{ a_k, a_k^\dagger \} = 1$ etc. on obtient pour la densité de charge $[\rho(x), \rho(y)] \neq 0$ avec des intervalles $x - y$ du genre espace $(x - y)^2 < 0$. L'exigence de la causalité relativiste, suffit à

* Soit A un opérateur. Si $AA^\dagger + A^\dagger A = 0$ alors

$$\langle \psi | AA^\dagger + A^\dagger A | \psi \rangle = \|A|\psi\rangle\|^2 + \|A^\dagger|\psi\rangle\|^2 = 0 \text{ pour tout } |\psi\rangle \text{ donc } A = 0.$$

rejeter les anticommutateurs, mais il reste à savoir (c'est la critique de Wet ; voir plus loin) si les relations $\{a_k, a_k^+\} = 1$ sont effectivement réalisables et l'argument de l'article précédent semblait prouver que non. Pauli ne s'exprima pas sur ce point à ce moment.

L'année suivante IWANENKO et SOKOLOV⁽⁶⁷⁾ montrèrent que pour que l'équation d'évolution d'un champ F : $i \frac{\partial F}{\partial t} = [H, F]$ soit satisfaite dans la théorie scalaire avec anticommutateurs il fallait que $\{b_k, b_k^+\} = -\delta_{kk}$ ce qui n'est pas possible* ; l'argument est en fait exactement le même que celui de Pauli et Weisskopf, mais vu en transformée de Fourier : la quantification canonique avec des anticommutateurs n'est mathématiquement pas possible avec le lagrangien du champ scalaire.

En 1939 un jeune étudiant de Wigner, J. S. DE WET prépara un Ph. D. sur "le lien entre le spin et la statistique des particules élémentaires". De Wet⁽⁶⁹⁾ jugea "incorrecte" la preuve de Pauli exposée à l'Institut Poincaré et généralisa l'argument d'Iwanenko et Sokolov à tous les spins entiers. La démonstration comporte deux étapes :

1 - un lagrangien quadratique dans les dérivées du champ ne permet pas la quantification canonique avec des anticommutateurs : en effet pour que les équations d'évolution $i \frac{\partial F}{\partial t} = [H, F]$ soient satisfaites dans ce cas il faudrait que $\{\psi_i(x), \psi_i^+(x)\} = 0$ pour toute composante du champ ce qui le rendrait trivial. Le cas des lagrangiens linéaires dans les dérivées du champ (comme pour les spins 1/2 entiers) est singulier à cet égard car les champs conjugués ψ_i^+ sont aussi conjugués canoniques des ψ : l'anticommutateur $\{\psi_i(x), \psi_j^+(x)\}$ n'est plus nul.

2 - Le lagrangien pour un champ de spin entier est nécessairement quadratique dans les dérivées du champ : un spin entier est décrit par un tenseur et une expression scalaire quadratique par rapport à ce tenseur est nécessairement de degré pair dans les dérivées du champ.

La démonstration de de Wet, aussi séduisante qu'elle fût, présente deux faiblesses : elle ne s'applique que si la quantification canonique est effectivement possible, or justement la quantification canonique des champs

* puisque $b_k b_k^+ + b_k^+ b_k$ est positif.

de spin supérieur à un présente des difficultés en raison des contraintes de traces nulles et de divergences nulles qui permettent de n'avoir qu'un seul spin. Dans l'article de 1940, Pauli rejeta la démonstration de de Wet pour cette raison. Enfin de Wet suppose que le lagrangien de départ ne fait intervenir qu'un seul champ. Mais la distinction de ce cas et de celui à plusieurs champs est arbitraire : par exemple un champ scalaire nul peut être considéré comme un champ scalaire couplé à sa dérivée (de même qu'une équation différentielle du deuxième ordre peut toujours être considérée comme un système de deux équations différentielles du premier ordre), Pauli et Belinfante (voir plus loin) montrèrent même en utilisant cet artifice qu'il existait des théories de spin 0, satisfaisant la statistique de Fermi-Dirac (mais alors la localité est violée) et obtenues par quantification canonique. Le lagrangien utilisé est linéaire dans les dérivées des champs comme il se doit.

5.3.d La démonstration de Pauli en 1940 :

Les idées essentielles de cette démonstration furent exposées pour la première fois par FIERZ en 1938, qui reconnaît lui-même avoir été guidé par Pauli. Fierz⁽⁷⁰⁾ reprit le projet de Dirac de construction d'équations relativistes de spin quelconque. Il reprochait à Dirac d'avoir (dans l'article de 1936) identifié à tort le hamiltonien à $i\hbar \frac{\partial}{\partial t}$ ce qui ne saurait être le cas que pour une théorie à une particule, et une équation différentielle du premier ordre interprétable directement comme équation de Schrödinger. Sinon le vrai hamiltonien est celui dont les équations canoniques donnent l'équation étudiée. Enfin pour le problème couplé la prescription de Dirac du couplage minimal $i\partial_\mu \rightarrow i\partial_\mu - eA_\mu$ ne pouvait s'appliquer en général car elle violait la compatibilité des équations. Fierz se proposa de donner, d'abord dans le cas du champ libre, les expressions des tenseurs d'énergie-impulsion et du courant. Il fit pour cela un usage systématique des spineurs de van der Waerden. Au-delà du spin 1 les expressions de $T_{\mu\nu}$ et de j_μ souffrent d'une certaine ambiguïté ; seules les grandeurs intégrées sur tout l'espace : énergie, charge et impulsion totale sont univoquement déterminées. Qui plus est, on n'a plus la localité du courant ($[j_\mu(x), j_\nu(y)] = 0$ si $(x - y)^2 < 0$) pour des spins supérieurs à un. Il faut donc renoncer au

courant comme observable locale (ce qui ne signifie pas qu'il n'existe pas d'autres observables locales). Pour le couplage avec le champ électromagnétique des champs de spin quelconque la méthode suggérée par Dirac du couplage minimal ne pouvait être utilisée. Trouver un couplage qui préservât la compatibilité* des équations d'onde et des conditions supplémentaires libres était un problème très délicat qui fut résolu par FIERZ et PAULI⁽⁷¹⁾ dans la même année en introduisant un artifice : on utilise un nouveau lagrangien contenant le champ étudié et des champs auxiliaires de spin inférieur tel que les nouvelles équations du mouvement et la condition d'annulation des champs auxiliaires soient équivalentes à l'ancien système de l'équation d'onde et des conditions supplémentaires ; la prescription du couplage minimal peut être introduite au niveau du nouveau lagrangien.

Revenons à l'article de Fierz et à la démonstration que l'on y trouve du lien spin-statistique.

- pour les spins 1/2 entiers, l'énergie totale du champ classique n'a pas de signe défini, il est donc nécessaire de quantifier avec des anticommutateurs pour que l'énergie soit définie positive, comme dans le cas de l'équation de Dirac.

- les relations de commutation des composantes du champ doivent être de la forme $[F(x), F^*(y)]_{\pm} = \text{produit de dérivées} \times \Delta(x-y)$ (Δ de Jordan-Pauli) car les champs F et leurs commutateurs satisfont l'équation de Klein-Gordon et les commutateurs doivent s'annuler pour des intervalles du genre espace pour qu'il en soit de même pour les observables construites à partir de F . Il se trouve que ces deux considérations et la covariance des champs F suffisent à établir la forme de commutateur**. Notons qu'aucun usage n'est fait de la quantification canonique. Si le spin est maintenant entier et si l'on prend des anticommutateurs, le développement de Fourier de F et la

* Par exemple pour les spins entiers cette incompatibilité vient de ce que $[\pi^{\mu}, \pi^{\nu}] \neq 0$ avec $\pi_{\mu} = i\partial_{\mu} - e A_{\mu}$; or l'équation d'onde libre contient $\square = \pi^2$ et les conditions supplémentaires font intervenir ∂_{μ} et donc π_{μ} .

** Fierz n'est pas très précis à cet égard.

forme de la fonction Δ montrent que le signe de $\{a_k, a_k^+\}$ dépend du signe de k . Ce qui est absurde.

Cette démonstration était déjà très générale puisqu'elle ne faisait pas appel à la quantification canonique contrairement à celle de de Wet, qu'elle n'imposait pas que les représentations utilisées soient irréductibles (souligné par Fierz) contrairement à celle de Belinfante (voir plus loin).

Néanmoins le célèbre article de PAULI ⁽⁷²⁾ de 1940 "Le lien entre le spin et la statistique" donne une preuve plus générale et plus élégante, ne faisant nullement intervenir la forme explicite de $T_{\mu\nu}$ et j_μ , ni celle des équations.

"Ce papier est une partie d'un rapport préparé par l'auteur pour le congrès de Solvay 1939 et auquel de légères modifications ont été apportées. En raison de la période défavorable, le congrès n'a pas eu lieu et la publication des rapports a été reportée à une date indéfinie ..." ⁽⁷²⁾

Le premier point déjà discuté par Fierz consiste à préciser le lien entre la représentation irréductible $\mathcal{D}(\frac{j}{2}, \frac{k}{2})$ du groupe de Lorentz utilisée et la valeur du spin des particules de la théorie. Pour cela on cherche les ondes planes monochromatiques solutions de l'équation d'onde, dans le référentiel de repos ($\vec{k} = 0$ et $k_0 = m$) pour que le moment orbital soit nul. Les rotations spatiales (SU_2) permettent de classer les composantes de ces ondes planes en sous-représentations irréductibles. A chaque sous-ensemble irréductible de composantes ainsi déterminé correspond un type de particule et à chaque composante un état de polarisation. Une définition expérimentale du spin peut être le nombre de degrés de liberté d'une particule dans le référentiel au repos, c'est-à-dire le nombre de ces composantes. Avec cette définition on constate qu'à $\mathcal{D}(\frac{j}{2}, \frac{k}{2})$ correspondent les spins $\frac{j+k}{2}, \frac{j+k}{2} - 1, \dots, \frac{j-k}{2}$ * et seulement le spin $\frac{j+k}{2}$ si des

* Ce qui correspond au fait qu'une représentation $\mathcal{D}(\frac{j}{2}, \frac{k}{2})$ de SL_2 contient une représentation réductible $\mathcal{D}(\frac{j}{2}) \otimes \mathcal{D}(\frac{k}{2})$ de SU_2 qui (en appelant $\mathcal{D}(\frac{j}{2})$ la représentation irréductible de spin $\frac{j}{2}$ de SU_2) se décompose d'après la célèbre "loi de composition des moments cinétiques" en

$$\mathcal{D}(\frac{j+k}{2}) \oplus \dots \oplus \mathcal{D}(\frac{j-k}{2})$$

conditions supplémentaires convenables sont rajoutées à l'équation d'onde. Le cas de masse nulle nécessite un traitement spécial puisqu'il n'y a pas alors de référentiel de repos ; on peut alors définir le spin comme la valeur minimale du moment cinétique total d'une particule.

Pauli postule que les équations d'onde doivent être linéaires et homogènes. On peut préciser leur forme en introduisant la distinction entre les représentations du type

$$\begin{aligned} \mathcal{D}(p, q) & \text{ de caractère } + + & p \text{ et } q \text{ entiers} \\ \mathcal{D}(p + \frac{1}{2}, q) & \text{ de caractère } - + \text{ etc.} \end{aligned}$$

La dérivation est une forme linéaire elle a donc le caractère - - Les champs fondamentaux sont notés $U^{\pm\pm}$ suivant leur caractère.

Dans le cas de spin $\frac{1}{2}$ -entier les seuls champs intervenant sont du type U^{+-} ou U^{-+} et les équations peuvent se mettre sous la forme (on peut toujours se ramener au premier ordre)

$$\left\{ \begin{aligned} \sum \partial^{\mu\nu} U^{\mu\nu} &= \sum U^{\mu\nu} \\ \sum \partial^{\mu\nu} U^{\nu\mu} &= \sum U^{\nu\mu} \end{aligned} \right.$$

Un tenseur de rang pair bilinéaire dans les champs fondamentaux comme le tenseur d'énergie impulsion doit être de la forme

$$T^{\mu\nu} = \sum U^{\mu\alpha} U^{\alpha\nu} + \sum U^{\nu\alpha} U^{\alpha\mu} + \sum U^{\mu\alpha} \partial^{\beta\gamma} U^{\beta\gamma} + \sum U^{\nu\alpha} \partial^{\beta\gamma} U^{\beta\gamma}$$

Le système d'équations d'ondes est invariant par la substitution :

$$\left\{ \begin{aligned} \partial^{\mu\nu} &\rightarrow -\partial^{\mu\nu} \quad (\text{c'est-à-dire } x^{\mu} \rightarrow -x^{\mu} \equiv PT) \\ U^{\mu\nu} &\rightarrow i U^{\mu\nu} \\ U^{\nu\mu} &\rightarrow -i U^{\nu\mu} \end{aligned} \right. \quad \begin{aligned} & \text{Les } i \text{ sont introduits pour que cette substitution soit} \\ & \text{compatible avec la conjugaison complexe : } (U^{\mu\nu})^* = U^{\nu\mu} \end{aligned}$$

Pour cette substitution $T^{\mu\nu}$ change de signe $T^{\mu\nu} \rightarrow -T^{\mu\nu}$.

L'énergie n'a donc pas de signe défini pour un champ classique de spin $\frac{1}{2}$ -entier.

Il faut quantifier avec des anticommutateurs pour rendre l'énergie positive.

Dans le cas des spins entiers il faut utiliser l'hypothèse de localité des observables physiques et corrélativement des champs fondamentaux à partir desquels elles sont construites.

Le (anti)commutateur $[U(x), U'(y)]_{\epsilon} = K$ vérifie l'équation de Klein-Gordon, il se transforme suivant $\mathcal{D}(\frac{1}{2}, \frac{k}{2}) \otimes \mathcal{D}(\frac{1}{2}, \frac{k}{2})$ et il doit s'annuler pour $(x - y)^2 < 0$; il est donc nécessairement de la forme $K = \text{cte} \times \text{produit de}$

(j + k) dérivées $\times \Delta(x - y)$, car $\Delta(x)$ est la seule solution invariante de Lorentz nulle pour $x^2 < 0$ de l'équation de Klein-Gordon (à une constante multiplicative près).

La fonction Δ est impaire; en conséquence si j + k est pair (spin entier) K(x) est impaire et donc $K(0) = 0 = U(x)U^\dagger(x) + \epsilon U^\dagger(x)U(x)$ ce qui impose $\epsilon = -1$: il faut quantifier avec des commutateurs.

Notons que dans les deux cas la transformation PT : $x \rightarrow -x$ est utilisée.

Pauli conclue "selon notre point de vue, le lien entre le spin et la statistique est une des applications les plus importantes de la théorie de la relativité".

Nous abordons seulement maintenant le rôle de la conjugaison de charge dans ce problème, car contrairement à l'ordre chronologique des publications, les travaux de Belinfante et de Pauli et Belinfante sont postérieurs à l'établissement de la preuve que nous venons d'exposer.

5.3.e La conjugaison de charge :

On attribue souvent les premières études de cette symétrie à Pauli (1936) et à Kramers (1937). Pourtant, la symétrie que Weyl avait utilisée en 1930⁽⁷³⁾ pour prouver l'égalité des masses des électrons et des trous de Dirac n'était autre que l'invariance par conjugaison de charge. (voir chap. 2 § 2).

WEYL montrait d'abord qu'à partir de ψ tel que $(i\partial - eA - m)\psi = 0$ on peut construire ψ_c telle que $(i\partial + eA - m)\psi_c = 0$ en prenant $\psi_c = C\psi^*$ ($C = i\gamma^2$ en représentation de Dirac). Les équations de Dirac et Maxwell couplées sont invariantes par la substitution $\begin{cases} \psi \rightarrow \psi_c \\ e \rightarrow -e \end{cases}$

à condition que les spineurs anticommulent afin que $\int_{\mu} = e\bar{\psi}\gamma_{\mu}\psi$ soit laissé invariant. Non seulement Weyl avait donné, sans la nommer* l'expression de la conjugaison de charge, mais encore il montrait que la

* Il s'intéressait en premier lieu à l'échange du passé et du futur, mais sans la notion indispensable de transformation antiunitaire, développée par Wigner en 1932.

théorie n'était symétrique par cette opération qu'après quantification par des anticommutateurs.

La même opération fut réintroduite par PAULI en 1936⁽⁷⁴⁾ comme correspondance entre les états d'énergie positive et ceux d'énergie négative.

Mais l'étude de cette opération interprétée comme conjugaison de charge revient à KRAMERS (novembre 1937). La présentation est très proche de la présentation moderne (réf. I p. 85). L'effet de l'opération sur la théorie en deuxième quantification est étudié dans la version où le champ ψ est directement interprété en terme d'électrons et de positrons. Ainsi Kramers pouvait clarifier la théorie de MAJORANA⁽⁷⁶⁾ ; ce dernier, grâce à l'introduction de la représentation de Majorana des matrices γ_μ , avait donné une formulation symétrique quant au signe de la charge, de l'électrodynamique quantique. Grâce à Kramers il devenait clair que l'avantage de cette représentation était que $C = 1$ donc $\psi_c = \psi^*$, et la méthode de Majorana consistait à symétriser le courant $j^\mu \rightarrow \frac{1}{2} (j^\mu + j_c^\mu)$ (ce qui revient à utiliser le produit de Wick moderne $j^\mu \rightarrow : j^\mu :$). Majorana insista fortement sur la nécessité de ne plus parler "d'état d'énergie négative" et d'introduire d'emblée les antiparticules.

Les "undors" de BELINFANTE :

A l'université de Leyde, influencé par Kramers, Belinfante développa la théorie des "undors" dans une série d'articles de "physica"^(77, 78). Plutôt que d'utiliser les spineurs de van der Waerden, Belinfante préféra les "undors", se transformant comme un produit tensoriel de ψ de Dirac ; "undor" fut tiré du latin "unda". L'avantage des undors, bien qu'ils ne soient plus associés à des représentations irréductibles du groupe de Lorentz, est qu'ils permettent d'exprimer simplement l'effet de la conjugaison de charge.

L'équation de l'électron pour être symétrique par conjugaison de charge nécessite déjà la conjugaison complexe du ψ et donc la représentation $\mathcal{D}(\frac{1}{2}, 0) \oplus \mathcal{D}(0, \frac{1}{2})$ (le fait que l'opérateur de dérivation se transforme suivant $\mathcal{D}(\frac{1}{2}, \frac{1}{2})$ justifie à lui seul pour $m \neq 0$ l'introduction de deux types de spineurs (φ_β et χ_α). La construction de l'équation de Dirac satisfaite par un undor, du courant associé et de l'opération de conjugaison de charge généralisée s'effectue aisément sur le modèle du spin 1/2, avec des γ_μ et β

(tel que $j_\mu = e \psi^\dagger \beta \gamma_\mu \psi$) tels que $\gamma_\mu = \sum_{i=1}^3 \gamma_\mu^i$ et $\beta = \sum_{i=1}^3 \gamma_0^i$

pour un undor ψ de rang s .

La conjugaison de charge fait intervenir C tel que $(\gamma_\mu^i)^* C^{-1} = -\gamma_\mu^i$, $CC^+ = 1$ et elle s'exprime suivant
$$\begin{cases} \psi \rightarrow C \psi^* \\ e \rightarrow -e \end{cases}$$

Alors $j_\mu = e \psi^\dagger \beta \gamma_\mu \psi \rightarrow -e \psi^\dagger C^\dagger \beta \gamma_\mu C \psi^* = j_\mu^c$

or $C^\dagger \gamma_\mu C = -\gamma_\mu^*$ et $C^\dagger \beta C = (-1)^s \beta$ puisque β est un produit de s γ_μ^i

Donc $j_\mu^c = (-1)^s e \psi^\dagger \beta^* \gamma_\mu^* \psi^* = (-1)^s e \psi^\dagger (\beta \gamma_\mu)^\dagger \psi^*$ (car le courant est réel $\Rightarrow (\beta \gamma_\mu)^* = (\beta \gamma_\mu)^\dagger$)

$j_\mu^c = j_\mu$ si et seulement si les ψ commutent pour s pair (spin entier) et anticommulent pour s impair (spin 1/2 entier).

Belinfante précise qu'en fait j_μ et j_μ^c diffèrent d'un C -nombre ω^* ("la contribution de la mer pleine") ce qui nuit un peu à la rigueur de la démonstration ; mais on ne s'intéresse qu'à la partie opérationnelle du courant. Pour Belinfante, postuler l'invariance par conjugaison de charge impliquait la relation spin-statistique.

PAULI ne fut pas complètement satisfait par cette belle démonstration : Belinfante supposait qu'une particule de spin donné était décrite par un seul "undor". Pauli collabora avec Belinfante⁽⁷⁹⁾ pour montrer que si l'on supprimait cette hypothèse l'invariance postulée par conjugaison de charge ne suffisait pas à exclure les mauvaises statistiques. Deux contre-exemples sont proposés. Appelant I la positivité de l'énergie

II la localité

III l'invariance par conjugaison de charge

il existe une théorie scalaire de fermions avec I et III vérifiées (et pas II). Il existe une théorie de spin 1/2 de bosons avec II et III vérifiées (et pas I). En revanche, comme l'avait déjà montré Pauli I(II) suffit pour conclure dans le cas des spins 1/2 entiers (resp. entiers).

Pour trouver de nouveaux développements sur le thème "spin-statistique" il fallut attendre le développement de la théorie de la renormalisation et la nouvelle approche de Feynman de l'électrodynamique quantique. Dans les travaux de Schwinger de 1947-48 apparut le lien entre le choix de la

* puisque $[\psi(x), \psi^\dagger(x)]_\pm$ est un C -nombre infini.

statistique et la symétrie PCT. La formulation générale du théorème PCT et sa démonstration furent encore données par PAULI⁽⁸¹⁾ en 1955 et dédiées à Niels Bohr pour son 70ème anniversaire. Dans l'approche intuitive de FEYNMAN sans deuxième quantification le lien entre spin et statistique apparut sous un angle nouveau : les deux statistiques étaient formellement possibles, mais pour la mauvaise la probabilité que le vide reste inchangé était supérieure à un. Pauli⁽⁸⁰⁾ montra à quoi cela correspondait en théorie quantique des champs : la théorie de Feynman avec la mauvaise statistique est équivalente à une théorie de métrique indéfinie (à la Dirac, voir chap. 7 § 3) et dans ce type de théorie, il est possible de trouver par exemple un champ de Bose de spin 1/2 et d'énergie positive. Enfin avant 1947 seule était possible une démonstration pour les champs libres, en raison du problème ultraviolet. Plus tard la méthode de renormalisation, respectant les symétries fondamentales permit de généraliser le théorème spin-statistique. La preuve la plus générale et la plus rigoureuse en fut donnée dans le cadre de la théorie axiomatique des champs. Pour cette preuve et les références historiques le lecteur intéressé peut consulter le livre de Streater et Wightman "PCT, spin and statistic and all that". Le principe de la démonstration peut ainsi se résumer : l'invariance postulée de la théorie par le groupe de Lorentz propre peut être étendue à l'inversion PT grâce à un prolongement analytique des amplitudes et à une complexification du groupe de Lorentz (la composante connexe de l'unité du complexifié contient PT) et la symétrie PT n'est compatible avec la localité que si la bonne statistique est satisfaite ; c'est la positivité de l'énergie qui permet d'effectuer le prolongement analytique. Nous avons déjà remarqué que dans la démonstration de Pauli de 1940 la transformation PT joue un rôle privilégié.

Pour la période qui nous intéresse la démonstration du lien entre spin et statistique constitue un nouvel exemple de la puissance prédictive de la théorie quantique des champs avant même qu'elle pût avoir un statut mathématique satisfaisant. La fusion des hypothèses de la relativité restreinte et de celle de la mécanique quantique, en une théorie quantique des champs présente de grandes difficultés. Cette voie malheureusement difficile, est aussi heureusement étroite : la classe de théories satisfaisant aux deux types d'exigences relativistes et quantiques semble très réduite

(surtout avec l'exigence moderne de renormalisabilité) et de ce fait elle vérifie des propriétés très restrictives, étranges parfois comme le théorème spin-statistique où une propriété de la matière au repos se trouve reliée à l'invariance relativiste.

5.4. L'INVARIANCE RELATIVISTE MANIFESTE :

En 1929 Heisenberg et Pauli avaient pu donner non sans mal une preuve d'invariance relativiste de toute théorie quantique des champs. La preuve était difficile car la formulation canonique de la théorie faisait jouer un rôle apparemment privilégié au temps. Heisenberg, Pauli et Weisskopf ne se sont guère souciés d'obtenir une formulation plus intrinsèque. Après tout, la nécessité de respecter une covariance explicite à tous les stades du calcul se fit surtout sentir quand il fallut lever les ambiguïtés des soustractions par renormalisation. Avant guerre, seuls des critères esthétiques ou méthodologiques pouvaient justifier la recherche d'une extension manifestement covariante de la théorie de Heisenberg-Pauli. La formulation la plus belle devait être selon Tomonaga la plus propre à éclairer les difficultés fondamentales de la théorie des champs. On ne s'étonnera pas de trouver Dirac en tête de ces recherches, puisque l'harmonisation des idées quantiques et relativistes avait été très tôt le leitmotiv de ses travaux. Tomonaga fut seulement plus diraquien que Dirac en généralisant la théorie "multi-temps" de Dirac en théorie "super-multi-temps".

5.4.a. Deux articles essentiels de DIRAC : (91,93)

"Une mécanique quantique relativiste"⁽⁹¹⁾

Nous rappelons brièvement (voir chap. 2 § 4) les travaux de Dirac de 1932 qui aboutirent à la représentation d'interaction et aux temps multiples. Pour obtenir "une mécanique quantique relativiste" il fallait selon Dirac traiter dissymétriquement le champ électromagnétique et les particules chargées. Ce présupposé conduisait à une électrodynamique quantique élégante, relativiste dans sa formulation même grâce à l'introduction de temps différents pour chacune des particules chargées. Rosenfeld⁽⁹⁴⁾ eut tôt fait de démontrer que cette nouvelle théorie de Dirac était équivalente à la théorie de Heisenberg-Pauli (dans la version qui traite les électrons dans l'espace des configurations) : elle n'en différait que par le passage à la représentation d'interaction pour le champ électromagnétique (dans la terminologie moderne) et par l'introduction de temps multiples. Dirac, Fock et Podolsky⁽⁹²⁾ profitèrent de cette remarque pour parfaire la formulation de Dirac en utilisant la condition de jauge de Fermi et

pour souligner l'importance du rôle des temps multiples. Bloch, en 1939 donna l'interprétation physique des fonctions d'onde généralisées à plusieurs temps de Dirac : elles donnaient la probabilité d'observer la particule i à l'instant t_i au point x_i à condition que les différents $x_i = (t_i, \vec{x}_i)$ soient séparés par des intervalles du genre espace. STUECKELBERG⁽⁹⁶⁾(1938) suggéra brièvement une généralisation des temps multiples au cas où les électrons étaient décrits en 2° quantification il fallait introduire un temps local $t_{\vec{x}}$ pour le champ électronique $\psi(\vec{x}, t)$

"Le lagrangien en mécanique quantique" :⁽⁹³⁾(1933)

Cet article* est devenu très célèbre, avec un certain retard, grâce au calcul des intégrales de chemin de Feynman (1948). Il serait faux de croire qu'il resta inconnu de 1933 à 1948. L'essentiel de son contenu fut publié dans les éditions n° 2, 3 et 4 des "Principles" et Dirac lui-même développa des considérations générales du même genre en 1945⁽⁹⁷⁾. Dès sa parution l'article retint l'attention des théoriciens japonais.

Le lagrangien, pensait Dirac devait être plus fondamental que le hamiltonien pour deux raisons : il est associé à un principe variationnel il est un scalaire relativiste. Pour une formulation covariante de la mécanique quantique relativiste, il semblait plus habile d'établir une "correspondance" directement entre le formalisme lagrangien de la mécanique classique, et la mécanique quantique. Pour ce faire Dirac établit en premier lieu une correspondance entre les transformations canoniques classiques et leurs analogues quantiques. A toute transformation canonique classique $Q_1, P_1 \rightarrow Q_2, P_2$ correspond une fonction génératrice $F(Q_1, Q_2)$ telle que $P_1 = \frac{\partial F}{\partial Q_1}, P_2 = \frac{\partial F}{\partial Q_2}$ **. En mécanique quantique on peut trouver l'analogue de F : soit $U(q_1, q_2)$ défini par $\langle q_2 | q_1 \rangle = U(q_1, q_2)$ (les majuscules désignent les opérateurs et les minuscules, leurs valeurs propres) alors on a les relations entre opérateurs

* republié dans la sélection d'articles de Schwinger

** voir par exemple Landau et Lifchitz ; cours de mécanique

$$\begin{cases} p_1 = \frac{\partial U}{\partial q_1} \Big|_{\text{ord}} (Q_1, Q_2) \\ p_2 = -\frac{\partial U}{\partial q_2} \Big|_{\text{ord}} (Q_1, Q_2) \end{cases} \quad \text{le suffixe ord. signifiant}$$

que l'on a ordonné tous les Q_1 à droite des Q_2 dans le développement de Taylor des dérivées. L'évolution temporelle de t' à t'' est un cas particulier de transformation canonique $Q_{t'}, P_{t'} \rightarrow Q_{t''}, P_{t''}$, engendré par l'action $S(Q_{t'}, Q_{t''})$. Dirac écrit :

$$\langle q_{t''} | q_{t'} \rangle \quad \text{correspond à} \quad \exp i \int_{t'}^{t''} \frac{L dt}{\hbar}$$

$$\langle q_{t+dt} | q_t \rangle \quad \text{correspond à} \quad \exp i \frac{L dt}{\hbar}$$

Dirac note la proportionnalité pour $\hbar \rightarrow 0$ et utilise la décomposition $\langle q_{t''} | q_{t'} \rangle = \int \prod_{i=1}^m dq_i \langle q_{t''} | q_{t_m} \rangle \langle q_{t_m} | q_{t_{m-1}} \rangle \dots \langle q_{t_1} | q_{t'} \rangle$ avec $t_{i+1} - t_i \rightarrow 0$ pour préciser la correspondance, mais il ne dit nulle part que $\langle q_{t+dt} | q_t \rangle$ est proportionnel à $\exp i \frac{L dt}{\hbar}$.

Cette proportionnalité fut établie par Feynman : "L'analogie intime entre $\langle q_{t+dt} | q_t \rangle$ et la quantité $\exp \frac{iS}{\hbar} (q_{t+dt}, q_t)$ a été soulignée par Dirac à plusieurs occasions. En fait nous voyons maintenant que les deux quantités peuvent être proportionnelles à un degré suffisant d'approximation. Les remarques de Dirac furent le point de départ de ce développement" (Feynman⁽⁹⁵⁾ 1948). Dans un dernier paragraphe Dirac envisagea l'application de cette nouvelle "correspondance" à la théorie des champs. A défaut de donner des règles précises de quantification, elle donnait un nouveau cadre conceptuel appelé g.f.t. ("generalized transformation function") par Dirac. La g.f.t. devait être l'analogue quantique de $\exp i \int_{\sigma_1}^{\sigma_2} \dots$ étant l'intégrale d'action prise dans le 4-volume délimité par les hypersurfaces σ_1 et σ_2 , la valeur des champs sur σ_1 et σ_2 étant donnée. Pour des surfaces σ_1 et σ_2 du genre espace, les champs devaient être mesurables de manière compatible sur chacune des deux surfaces et la g.f.t. devait donner l'amplitude de probabilité de mesurer sur σ_2 telles valeurs du champ, ayant mesurées telles autres valeurs sur σ_1 . Par des chemins assez différents, Dirac arrivait à une notion d' hypersurface du genre espace analogue

aux temps multiples discrets introduits un an plus tôt.

5.4.b. Leur écho au Japon :

Lors de sa conférence Nobel (1966) Tomonaga a retracé le cheminement jusqu'à l'électrodynamique quantique moderne. L'article de Dirac⁽⁹¹⁾ sur la mécanique quantique relativiste attira tout de suite l'attention de Yukawa qui demanda à son jeune élève Tomonaga de recalculer la formule de Klein-Nishima dans cette nouvelle théorie. Avant même d'arriver au bout des calculs Tomonaga se rendit compte que le résultat devait être le même, et bientôt arrivèrent d'Occident les preuves d'équivalence de l'ancienne et de la nouvelle théorie.

En 1942, YUKAWA⁽¹⁰⁰⁾ (articles en japonais) sensible aux difficultés croissantes de la théorie des champs, critiqua la théorie de Heisenberg-Pauli et proposa de préciser et d'utiliser la g.f.t. de Dirac. Le concept de fonction d'onde, donnant la probabilité d'un état du champ dans tout l'espace à un instant t fixé, n'avait pas de signification relativiste puisque la simultanéité n'est pas conservée par transformation de Lorentz ; il fallait lui substituer une notion plus générale comme celle de g.f.t. Le projet de Yukawa, pensa Tomonaga, était trop ambitieux car il cherchait à définir la g.f.t. et son interprétation physique pour des hypersurfaces d'un genre quelconque, alors qu'on pouvait se limiter à la sous-classe invariante relativiste des hypersurfaces du genre espace.

TOMONAGA prit pour point de départ la théorie multi-temps de Dirac, et approfondit l'idée de Stueckelberg du temps local. Le premier article de Tomonaga traitant de ce problème fut publié en 1943⁽¹⁰¹⁾ et sa traduction anglaise parut seulement après la guerre en 1946⁽¹⁰²⁾. La covariance manifeste était obtenue en deux étapes : passage à la représentation d'interaction, puis introduction du temps local.

1° étape : Soit \bar{H}_1 et \bar{H}_2 les hamiltoniens de deux champs fondamentaux et \bar{H}_I le hamiltonien d'interaction. Ces hamiltoniens sont obtenus par intégration sur l'espace d'une densité $H(\vec{x})$: $\bar{H} = \int H(\vec{x}) d^3x$

L'équation de Schrödinger s'écrit $(\bar{H}_1 + \bar{H}_2 + \bar{H}_I) \psi = i\hbar \frac{\partial \psi}{\partial t}$

La représentation d'interaction est obtenue en posant $\psi = U \psi_I$

avec $U = e^{-i(\bar{H}_1 + \bar{H}_2)t}$ ψ_I vérifie $U^{-1} \bar{H}_I U \psi_I = i\hbar \frac{\partial \psi_I}{\partial t}$

Dans cette représentation les champs fondamentaux φ_1 et φ_2 suivent

l'évolution libre : donc $U^{-1} \bar{H}_I U$ n'est rien d'autre que l'intégrale spatiale de $H_I(\alpha)$ ($\alpha = (t, \vec{x})$) obtenue en substituant les champs quantiques libres $\varphi_1(\alpha)$ et $\varphi_2(\alpha)$ à $\varphi_1(\vec{x})$ et $\varphi_2(\vec{x})$. Dans un cas simple comme celui de l'électrodynamique ($H_I = e \bar{\psi} \not{A} \psi$), $H_I(\alpha)$ est un scalaire relativiste. Les commutateurs des champs libres peuvent être explicités : par exemple $[\varphi_1(\alpha), \varphi_1(\beta)] = i\Delta(\alpha-\beta)$ si φ_1 est un champ scalaire et ils font intervenir des fonctions Δ invariantes. Il reste seulement à mettre l'équation de Schrödinger sous une forme manifestement covariante.

2° étape : l'équation $\int d^3x H_I(\vec{x}; t) \psi_I(t) = i\hbar \frac{\partial \psi_I}{\partial t}$ est analogue à l'équation $\sum_i h_i(t) \psi_i(t) = i\hbar \frac{\partial \psi}{\partial t}$

($h_i(t)$ = hamiltonien de Dirac de la particule i dans le champ libre) à laquelle Dirac, Fock et Podolsky substituaient le système

$h_i(t_i) \hat{\psi}(t_1, \dots, t_m) = i\hbar \frac{\partial}{\partial t_i} \hat{\psi}(t_1, \dots, t_m)$ redonnant $\psi(t)$ en égalisant les temps : $\psi(t) = \hat{\psi}(t, t, \dots, t)$

Dans le cas présent le nouveau système s'écrit :

$$H_I(\vec{x}; t_{\vec{x}}) \hat{\psi}_I(\{t_{\vec{x}}\}) = i\hbar \frac{\partial}{\partial t_{\vec{x}}} \hat{\psi}_I(\{t_{\vec{x}}\})$$

ou en notations fonctionnelles plus correctes

$$H_I(\alpha) \hat{\psi}_I(\sigma) = i\hbar \frac{\delta}{\delta \sigma(\alpha)} \hat{\psi}_I(\sigma)$$

$\hat{\psi}_I$ est une fonctionnelle de l'hypersurface σ . L'équation précédente ne fait intervenir que des notions intrinsèques (indépendantes du référentiel). La conditions de compatibilité $[H_I(\alpha), H_I(\beta)] = 0$ est vérifiée si $(\alpha-\beta)^2 < 0$, autrement dit σ doit être du genre-espace. Tomonaga acheva son article en montrant que son équation donnait une restriction de la g.f.t. de Dirac : l'évolution entre σ_1 et σ_2 s'obtenait en intégrant l'équation de Tomonaga* :

$\hat{\psi}_I(\sigma_2) = \prod_{\sigma_1}^{\sigma_2} (1 - i H_I(\alpha) d^4\alpha) \hat{\psi}_I(\sigma_1)$ par découpage du 4-volume délimité par σ_1 et σ_2 en éléments d^4x eux-mêmes délimités par

* Heisenberg avait déjà écrit une formule analogue dans le cas particulier des surfaces $t = cte$ dans son étude (1938) des limitations de la mécanique quantique (voir chap. 4§).

des frontières du genre espace, et les éléments de matrices de l'opérateur d'évolution entre des états propres du champ sur σ_1 et les états propres du champ sur σ_2 avaient la même interprétation que la restriction de la g.f.t. à des surfaces du genre espace.

Les articles suivants de Tomonaga et de ses collaborateurs précisèrent l'application dans quelques cas particuliers du formalisme de Tomonaga. Koba Tati et Tomonaga⁽¹⁰³⁾ étudièrent le cas de l'électrodynamique quantique et la manière d'introduire la condition de Fermi $(\partial_\mu A) \psi = 0$ et ils démontrèrent l'équivalence⁽¹⁰⁴⁾ de leur formalisme et de l'ancien. Kanesawa et Tomonaga passèrent aux champs mésiques⁽¹⁰⁵⁾ puis au mélange champs mésiques et champ électromagnétique⁽¹⁰⁶⁾. Signalons enfin les travaux de Watanabe⁽¹⁰⁷⁾ et de Tanikawa⁽¹⁰⁸⁾, généralisant suivant l'espoir de Yukawa le g.f.t. de Tomonaga à des hypersurfaces en tout genre. Watanabe, pour préciser la structure de la g.f.t. eut l'idée d'identifier la g.f.t. infinitésimale à l'exponentielle de l'action mais il ne démontra pas que cette identification conduisait à la théorie usuelle. (Contrairement à Feynman).

L'équation de Tomonaga joua un rôle important après l'avènement de la renormalisation dans les travaux de Tomonaga et de Schwinger sur l'électrodynamique quantique.

APPENDICE Polarisation du vide sans la mer de Dirac (voir § 5.1.b)

Fock donna en 1933 une formulation de la théorie du positron donnant des rôles entièrement symétriques aux électrons et aux positrons.

Puisque dans cette version ainsi que dans la théorie de Pauli et Weisskopf il n'y a pas de "mer" perturbé par le champ appliqué, le lecteur peut se demander d'où sort la polarisation du vide :

champ de Dirac	champ scalaire chargé
$\psi(x) = \sum_{\alpha} a_{\alpha} u_{\alpha}(x) + b_{\alpha}^{\dagger} v_{\alpha}(x)$	$\varphi(x) = \sum_k a_k e^{-ikx} + b_k^{\dagger} e^{ikx}$

pour les champs libres.

Si l'on branche adiabatiquement le champ extérieur A_{μ} , en représentation de Heisenberg les champs deviennent :

$\psi(x) = \sum_{\alpha} a_{\alpha} u_{\alpha}^A(x) + b_{\alpha}^{\dagger} v_{\alpha}^A(x)$	$\varphi(x) = \sum_k a_k \varphi_k^A(x) + b_k^{\dagger} \varphi_k^{A*}(x)$
---	--

où les fonctions d'ondes introduites sont les fonctions propres dans le champ A_{μ} (pour l'équation de Dirac et l'équation de Klein-Gordon respectivement).

L'énergie de la matière est donnée par

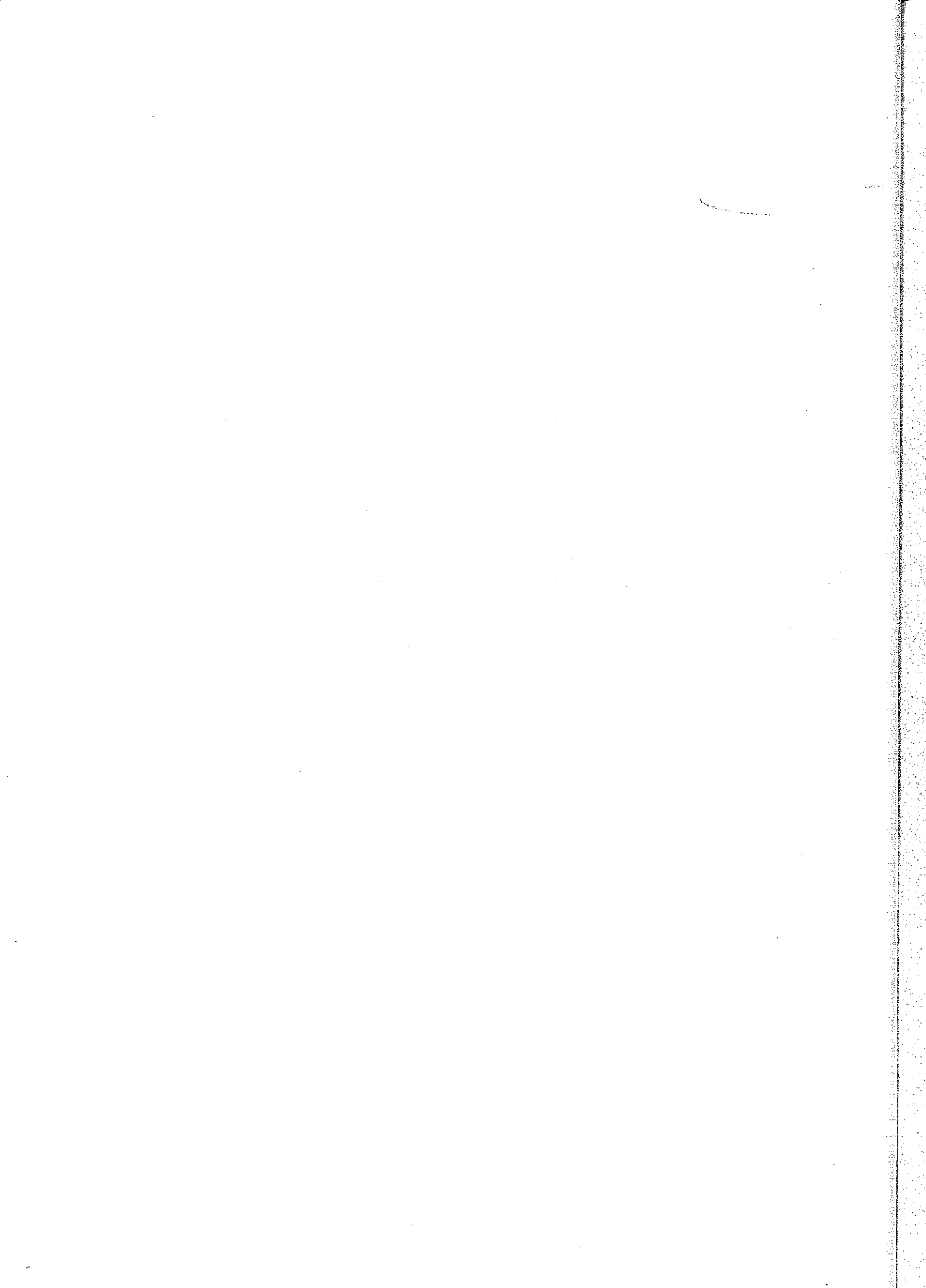
$H_{mat} = \sum_{\alpha} E_{\alpha} (a_{\alpha}^{\dagger} a_{\alpha} - b_{\alpha} b_{\alpha}^{\dagger})$ ($E_{\alpha} < 0$)	$H_{mat} = \sum_k \frac{1}{2} E_k [(a_k^{\dagger} a_k + a_k a_k^{\dagger}) + (b_k^{\dagger} b_k + b_k b_k^{\dagger})]$
--	--

en utilisant l'ordre naïf des créateurs et des annihilateurs. Ou encore

$H_{mat} = \sum_{\alpha} E_{\alpha} (N_{\alpha}^{+} + N_{\alpha}^{-}) - \sum_{\alpha} E_{\alpha}$	$H_{mat} = \sum_k (N_k^{+} + N_k^{-}) E_k + \sum_k E_k$
---	---

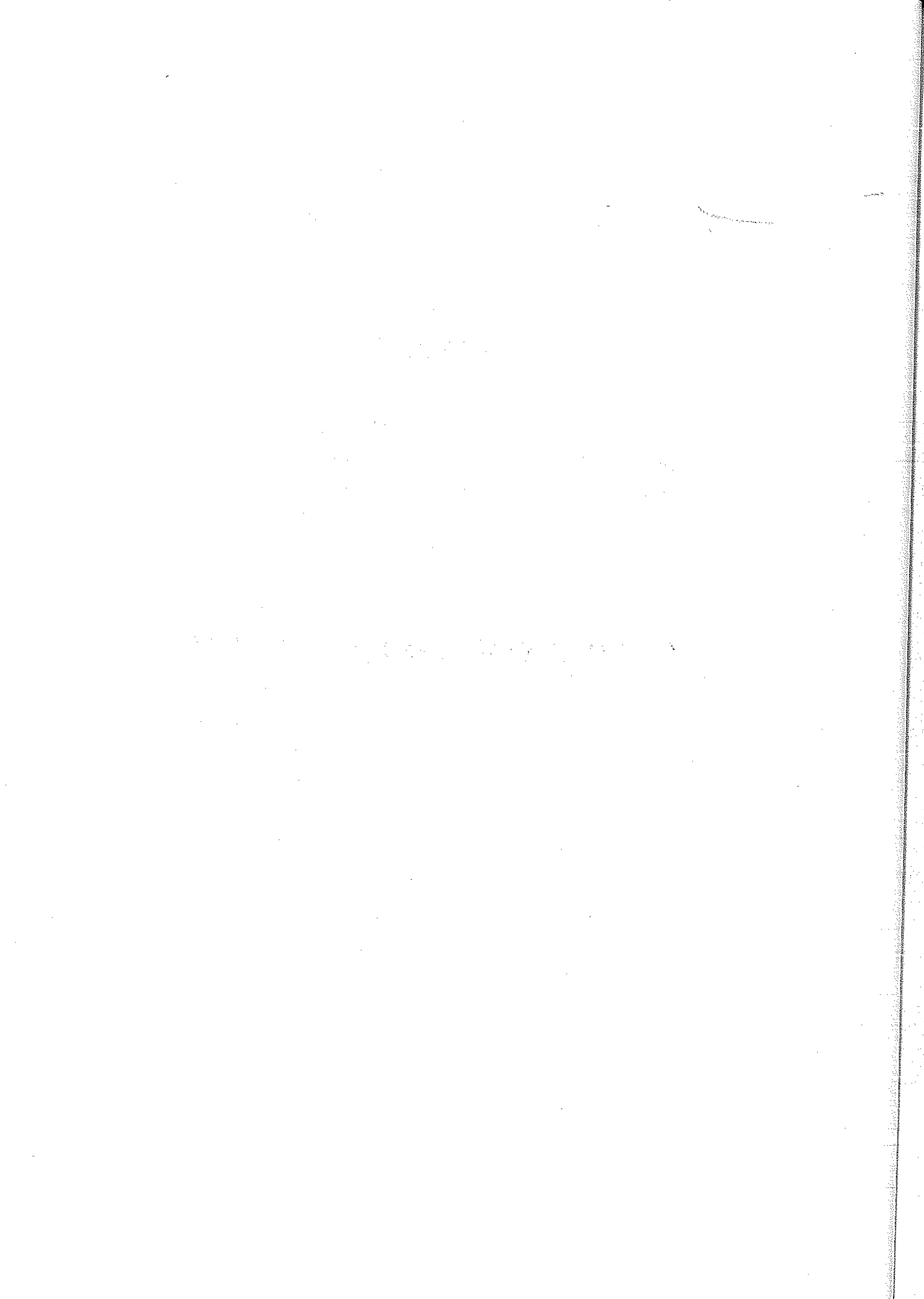
En représentation de Schrödinger les E_{α} (E_k) seraient les énergies propres non perturbées et les valeurs moyennes des nombres d'occupation N_{α}^{\pm} (N_k^{\pm}) évolueraient dans le temps (ainsi qu'en représentation d'interaction qui est celle utilisée aujourd'hui) et l'on pourrait annuler

les constantes infinies $\sum_{\alpha} E_{\alpha}$ et $\sum_k E_k$ (effet du produit de Wick) et inversement en représentation de Heisenberg les $\langle N \rangle$ restent nulles si l'espace était vide avant de brancher le champ (faible) et les E_{α} (E_k) sont les énergies propres perturbées en présence de A_{μ} . Dans ce dernier cas, il est essentiel de prendre en compte les termes infinis $-\sum_{\alpha} E_{\alpha}$ analogue à l'énergie de la mer de Dirac perturbée, et $\sum_k E_k$ énergie de zéro des "oscillateurs" perturbés qui contiennent à eux seuls l'effet du champ A_{μ} .



CHAPITRE 6

THEORIE QUANTIQUE DES CHAMPS ET HAUTES ENERGIES



Des années trente à l'immédiat après-guerre l'évolution de la physique nucléaire théorique et expérimentale fut si rapide et si riche qu'il serait vain de tenter ici de cerner tous les aspects de son histoire. Nous avons délibérément insisté sur l'avènement des méthodes de théorie quantique des champs en physique nucléaire, beaucoup plus que sur leur évolution ultérieure complexe. Bien qu'il n'existe sans doute pas d'étude historique couvrant globalement tout le sujet, il existe des travaux parfois excellents centrés sur des aspects particuliers de cette période de physique nucléaire. Nous les citerons à la fin de ce chapitre.

Le premier problème de théorie quantique des champs fut celui de l'électrodynamique quantique. Le formalisme développé par HEISENBERG et PAULI en 1929 était bien plus général et ouvrait la voie à l'utilisation de nouveaux champs. Cette ère nouvelle commença avec la théorie de FERMI (1933) de la désintégration β et avec la théorie mésique de YUKAWA (1935). Les concepts et méthodes nouveaux contenus dans la théorie de Dirac, Heisenberg et Pauli prenaient corps :

- les particules matérielles (même les fermions) n'étaient plus conservées : elles devenaient susceptibles de création ou d'annihilation lors d'un processus élémentaire, et c'est bien ce que virent les expérimentateurs dans leurs chambres de Wilson,

- l'interaction entre particules, au lieu d'être l'effet d'un champ de force devait être comprise comme résultant de l'échange virtuel de nouvelles particules,

- des champs nouveaux (scalaires, vectoriels massifs ...) devaient être introduits et le rôle des symétries pour établir la forme de leur interaction ou leur statistique fut mieux compris.

Les liens réciproques entre ces nouvelles théories et les expériences de physique nucléaire, malgré leur complexité, conduisirent d'abord à une confirmation des hypothèses de la théorie quantique des champs : les processus non divergents (dans l'approximation de Born) calculés en électrodynamique quantique et dans les nouvelles théories des champs finirent (non sans mal) par s'accorder très bien avec l'expérience, même jusqu'à des énergies très élevées (pour l'électrodynamique de Dirac appliquée aux rayons cosmiques).

Malheureusement cette base de confiance était sérieusement entamée par le problème des divergences de processus du deuxième ordre : avant même la théorie de la renormalisation il était clair que le problème des divergences était plus grave en physique nucléaire qu'en électrodynamique pure ; le degré de divergence était plus élevé, la constante de couplage méson-nucléon beaucoup trop grande pour qu'un développement perturbatif fût pris au sérieux, et dans certains cas comme la théorie de Fermi, la dimension de la constante de couplage semblait introduire une longueur critique (Bohr, Heisenberg) en dessous de laquelle la théorie perdait tout sens. Cette gravité des divergences était d'autant plus gênante que les effets du deuxième ordre attendus, au vu des données expérimentales (par exemple le moment magnétique anormal des nucléons), étaient importants. La solution de ces difficultés devenait plus urgente.

La physique nucléaire de cette période est marquée par une série d'hésitations, de tâtonnements, de confusions, cependant très riches en découvertes. Le degré de confiance attribué aux formules de l'électrodynamique de Dirac variait d'un théoricien à l'autre, les conséquences à tirer des résultats expérimentaux violant ces formules étaient mal déterminées ; DUHEM a bien vu en d'autres temps qu'une expérience négative ne permettait pas de localiser l'hypothèse fautive dans la théorie physique ; ainsi l'absorption anormale des rayons cosmiques très énergétiques pouvait venir d'un échec de la théorie de Dirac ou d'une analyse incorrecte du fonctionnement de l'appareil de mesure, ou d'une confusion de nature des particules observées.

Dans la diversité des points de vue adoptés par les théoriciens purent être découvertes de nouvelles particules (le positron, le neutron, le méson^{*}) et purent naître de nouvelles théories (théories mésiques, explosions de Heisenberg, théorie de couplage fort ...). Des idées nouvelles, plus ou moins importées de l'électrodynamique pure (méthode de Bloch et Nordsieck), comme celle de nuage virtuel de mésons, d'énergie propre mesurable (isobares), de

* (le méson μ , identifié à tort au méson de Yukawa).

moment magnétique anormal, préparèrent le terrain au principe de renormalisation. Les techniques de calcul s'améliorèrent et sans doute Tomonaga et Schwinger, rompus à la physique nucléaire, en ont-ils profité. Cependant le caractère complexe et parfois confus de la physique nucléaire d'avant-guerre cachait une des origines profondes des difficultés rencontrées ; on comprend que le principe de renormalisation soit apparu sur le terrain de l'électrodynamique pure.

6.1 DES SPECTRES β A LA THEORIE DE FERMI :

6.1.a Le neutrino

Les particules β émises par certains noyaux radioactifs présentent un spectre d'énergie continu, contrairement aux radiations α et γ . Ce fait expérimental était bien connu dans les années vingt (Chadwick⁽¹⁾ 1914). La largeur du spectre fut d'abord attribuée à des effets secondaires de ralentissement dans la matière, ou à l'émission de photons secondaires indétectés, car elle était beaucoup trop grande par rapport à la largeur naturelle associée à la durée de vie de l'élément instable par la quatrième relation d'incertitude. En 1925 deux expérimentateurs du groupe de Rutherford à Cambridge, ELLIS et WOOSTER⁽²⁾ estimèrent les effets secondaires hypothétiques et les trouvèrent trop faibles pour être tenus responsables de l'anomalie. Ils donnèrent deux ans plus tard⁽³⁾ une preuve expérimentale plus évidente de leur point de vue : d'une part ils obtinrent la courbe de densité spectrale des rayons β émis par Ra E (cette source présentait l'avantage de n'émettre que des rayons β) et d'autre part ils mesurèrent par une méthode microcalorimétrique l'énergie totale dissipée lors de l'émission dans une enceinte très absorbante : la comparaison de l'intégrale de la courbe spectrale et de cette énergie totale montrait qu'aucun autre rayonnement primaire ou secondaire n'avait été émis. Dès lors la continuité des spectres β devenait plus énigmatique. Ellis et Wooster s'en tiraient par une hypothèse ad hoc bien fragile : ils supposaient qu'il existait des électrons non quantifiés dans le noyau. Dans le groupe de Tübingen en Allemagne Lise MEITNER⁽⁴⁾ (1922) avait défendu l'hypothèse des photons γ secondaires. Au vu des nouveaux résultats d'Ellis et Wooster elle reprit des expériences analogues avec ORTHMANN⁽⁵⁾ (1929) et se rallia à leur opinion.

Niels BOHR, convaincu par ces expériences, alla jusqu'à remettre en cause le principe de conservation de l'énergie⁽⁶⁾. PAULI, plus méfiant, critiqua violemment un point de vue aussi extrême : il écrivit à Klein en février 1929 [216]⁽⁷⁾ : "Quant à moi je suis assez sûr que les rayons γ sont la cause de la continuité des spectres β et que Bohr est complètement dans l'erreur quand il suppose à ce propos que la conservation de l'énergie est violée ! Je crois aussi que les expériences calorimétriques pèchent quelque part et que les

rayons γ n'ont échappé aux expérimentateurs qu'en raison de leur maladresse. Mais je comprends trop peu la physique expérimentale pour pouvoir établir ce point de vue ...". Heisenberg et Pauli venaient en effet d'achever leur théorie quantique des champs d'onde (voir chap. 2 § 3) et avaient traité l'émission β comme application. Leur conclusion, surestimant la probabilité d'émission de photons secondaires, était bien sûr erronée.

Du 1er au 4 juillet 1929 eut lieu à Zürich un colloque de physique expérimentale où Lise MEITNER eut l'occasion de convaincre Pauli : "à Zürich Madame L. Meitner nous a fait un beau compte-rendu du côté expérimental de la question et elle m'a presque persuadé que la continuité des spectres β ne peut être expliquée par des processus secondaires (émission γ , etc.). Nous ne savons donc vraiment pas ce qui se passe !" (lettre [231] de Pauli à Bohr datée du 17 juillet 1929). Le 4 décembre 1930 Pauli, dans une lettre ouverte à Geiger et Meitner fit l'hypothèse d'une particule nouvelle de spin $\frac{1}{2}$ très pénétrante qui traverserait les parois des enceintes utilisées par Meitner. Cette hypothèse jugée très hardie par Pauli lui-même ne fit l'objet d'aucune publication avant le congrès Solvay de 1933^{**}. Pauli reformula oralement son hypothèse à Pasadena au printemps 1931 et à Ann Arbor au cours de l'été. Pauli appelait alors sa nouvelle particule le "neutron" ; l'usage de ce mot pour décrire l'hypothétique particule lourde découverte en 1932 par Chadwick, était alors confiné à l'école de Rutherford. En octobre 1931 se tint le colloque de Rome, Fermi fut aussitôt séduit par la proposition de Pauli, et il proposa d'appeler la nouvelle particule "neutrino", diminutif italien de "neutrone" : l'énergie moyenne des rayons β émis exigeait une masse faible (devant celle du proton) pour le neutrino.

6.1.b La découverte du neutron :

RUTHERFORD prononça le 3 juin 1920 sa "Bakerian lecture"⁽⁸⁾. Ce fut l'occasion de résumer ses idées sur le noyau atomique, alors imaginé comme

* "Vas-tu encore maltraiter le malheureux principe de conservation de l'énergie" écrivit Pauli à Bohr [277]⁽⁷⁾ (mars 1929).

** Toutefois dès 1931 Carlson et Oppenheimer (voir plus loin) prirent au sérieux l'hypothèse de Pauli pour expliquer la pénétration anormale des rayons cosmiques.

un agrégat d'électrons, de protons et de particules α . Une généralisation naturelle de la notion d'isotope le conduisit à prévoir l'existence d'un noyau de masse atomique 1 et de charge 0, qui serait l'édifice nucléaire le plus simple (après le proton), constitué d'un électron et d'un proton intimement (c'est-à-dire bien plus que dans l'atome d'hydrogène) liés : *"Dans certaines conditions cependant, il doit être possible qu'un électron se combine beaucoup plus étroitement avec le noyau d'hydrogène pour former une espèce de doublet neutre"*⁽⁸⁾. Par la suite de nombreuses expériences furent entreprises au Cavendish Laboratory pour chercher le "neutron" (ainsi baptisé par Rutherford). Citons seulement la première effectuée par Glasson⁽⁹⁾ en 1921 à l'aide d'une décharge électrique.

James CHADWICK travaillait dans le laboratoire de Rutherford et il était sans doute imprégné de cette longue tradition de recherche du neutron. Aussi il donna des expériences de Joliot et Curie de 1931 une interprétation très différente de celle de leurs auteurs. BOTHE⁽¹⁷⁾ et Mme CURIE⁽¹⁸⁾ avaient indépendamment constaté que la radiation émise par le béryllium et par le bore bombardés par des particules α issues du Polonium était anormalement pénétrante. Irène CURIE et Frédérik JOLIOT⁽¹⁰⁾ étudièrent à la fin de la même année l'interaction de cette radiation avec diverses cibles. L'effet le plus important était obtenu en prenant la paraffine comme cible et alors on observait l'éjection de protons détectés dans une chambre à ionisation. La radiation du béryllium en raison de sa neutralité électrique était assimilée à des rayons γ très durs. Le défaut de conservation de l'énergie pouvait être attribué aux erreurs expérimentales considérables. Curie et Joliot concluaient : *"Il paraît donc établi par ces expériences qu'un rayonnement électromagnétique de haute fréquence est capable de libérer, dans les corps hydrogénéés, des protons animés d'une grande vitesse"*. L'analyse de Chadwick fut toute différente. Dans une lettre⁽¹¹⁾ à "Nature" du 17 février 1932 intitulée "l'existence possible du neutron" il réinterpréta les résultats de Curie et Joliot^(11 et 12), tout en donnant une confirmation expérimentale de son hypothèse. L'éjection de protons par la radiation du béryllium n'était pas conforme à la formule de Klein-Nishima, et son énergie calculée connaissant les défauts de masse de Be^9 et C^{13} différait de la valeur expérimentale dans l'hypothèse d'une réaction $\text{Be}^9 + \alpha \rightarrow \text{C}^{13} + \gamma$. Chadwick proposa la réaction $\text{Be}^9 + \alpha \rightarrow \text{C}^{12} + n$ et raffermir cette hypothèse du neutron par l'étude expérimentale de la

diffusion par d'autres cibles que la paraffine et par la mesure de l'énergie des atomes émis.

6.1.c Les théories de Heisenberg et Majorana du noyau atomique :

La découverte du neutron par Chadwick fut providentielle pour résoudre l'une des difficultés majeures de la physique nucléaire de cette époque. Une première difficulté, la continuité des spectres β avait amené Pauli à formuler l'hypothèse du neutrino. Inversement, la découverte du neutron permit d'élaborer un modèle du noyau bien plus satisfaisant que l'ancienne image d'un état lié de protons et d'électrons (et éventuellement de particules α). En prenant le proton et l'électron comme seuls constituants on tombait sur de graves difficultés : on avait du mal à faire tenir tant d'électrons et de protons en un si petit volume, on n'expliquait pas le maximum de stabilité à $Z \approx A/2$, et les noyaux de Li^6 et N^{14} obéissaient à la statistique de Bose, contrairement à toute attente.

Le 21 avril 1932 IWANENKO⁽¹³⁾ proposa dans une lettre à "Nature" "d'admettre que les neutrons jouent aussi un rôle important dans la constitution du noyau". On évitait ainsi d'évoquer (comme Bohr) une perte d'individualité des électrons dans le noyau (plus de spin, plus de moment magnétique). Une théorie du noyau à partir du neutron du proton (et éventuellement de particules α) pourrait donner du poids à l'hypothèse du neutron.

Une telle théorie fut d'abord proposée par Heisenberg⁽¹⁴⁾ en juin 32. Le noyau était constitué de neutrons et de protons soumis à l'interaction p-p coulombienne, et à des interactions "d'échange" (Platzwechsel) par analogie avec H_2^+ : dans la molécule H_2^+ , en raison de la symétrie d'échange des deux protons les états propres de l'énergie sont symétrique et antisymétrique et leur différence d'énergie donnée en première approximation par une intégrale de recouvrement d'orbitales atomiques centrées sur les deux noyaux ; si l'on imagine un instant que n est constitué de $p+e^-$, l'édifice np est analogue à H_2^+ et mn analogue à H_2 . Chez Heisenberg cette analogie avait une valeur purement heuristique contrairement à l'idée initiale de Rutherford du neutron. Avec cette représentation des interactions nucléaires le hamiltonien devait comporter des termes transformant n en p et réciproquement : et pour cette raison Heisenberg introduisit la notion d'isospin : n et p étaient formellement

considérés comme deux états d'isospin différents d'une même particule. Le succès du modèle de Heisenberg tenait à ses prédictions correctes du spin et de la stabilité ($\frac{Z}{A}$) du noyau. L'année suivante Majorana⁽¹⁵⁾, alors à Leipzig, essaya de se débarrasser de la notion d'échange, qu'il jugeait trop artificielle, en partant des exigences empiriques concernant le spin et la densité des noyaux. Il aboutit à un hamiltonien un peu différent de celui de Heisenberg (signe des termes d'échange n-p). L'origine des forces nucléaires restait obscure.

6.1.d Le 7ème Congrès Solvay⁽¹⁶⁾ (octobre 1933)

"Structure et propriétés des noyaux atomiques"

Les principaux acteurs cités dans les paragraphes précédents : Bohr, Chadwick, Fermi, Heisenberg, Irène Joliot-Curie, Joliot, Meitner, Pauli et Rutherford étaient présents. Ce congrès marque une étape importante dans l'histoire de la physique nucléaire. Les expérimentateurs, grâce aux progrès dans l'étude des rayons cosmiques et de la radioactivité venaient de découvrir deux nouvelles particules : le positron (voir chap. 2 § 2) et le neutron, et la théorie nucléaire, avec les travaux récents de Gamow, de Heisenberg et Majorana, prenait corps. D'un côté Dirac triomphant envisageait de nouvelles conséquences de sa théorie du positron (Voir chap. 5 § 1) de l'autre un schéma cohérent de description des phénomènes nucléaires commençait à se dessiner. L'hypothèse de neutrino s'intégrait parfaitement dans ce schéma : elle permettait conjointement à celle du neutron de préserver les grandes lois de conservation. L'ajoutant à une liste de particules connues (n, p, γ , e^+ , e^-) déjà plus longue, elle devenait moins hardie. C'est dans la discussion du rapport de Heisenberg sur la constitution du noyau qu'on en trouve le premier exposé publié (P. 324). Pauli s'exprima ainsi : "les lois de conservation restent valables, l'expulsion des particules β étant accompagnée d'une radiation très pénétrante de particules neutres, qui n'a pas été observée jusqu'ici ... Quant aux propriétés de ces particules neutres, les poids atomiques des éléments radioactifs nous apprennent tout d'abord que leur masse ne peut pas dépasser beaucoup celle de l'électron ... Il est possible que la masse propre des neutrinos soit zéro ... Il me paraît admissible que les neutrinos possèdent un spin $\frac{1}{2}$ et qu'ils obéissent à la statistique de Fermi ... l'hypothèse qu'ils possèdent un moment magnétique ne me paraît

pas du tout fondée". Chadwick, commentant les résultats négatifs des expériences de recherche du neutrino conclua : "il est certain que le neutrino, s'il existe, sera excessivement difficile à déceler". La position de Bohr avait évolué, mais il ne renonçait pas à considérer une violation du principe de conservation de l'énergie comme possible : "Etant données les difficultés théoriques que soulèvent la solution de ce problème, c'est peut-être une question de goût que de savoir quel point de vue l'on préfère. Aussi longtemps que nous n'aurons pas de nouvelles données expérimentales, il est sage de ne pas abandonner les lois de conservation, mais, d'autre part personne ne sait quelles surprises nous attendent encore".

6.1.e La théorie de Fermi : (décembre 1933)

Deux mois après le congrès de Solvay, Fermi publia un bref article de quatre pages dans la revue italienne "Ricerca Scientifica"⁽¹⁹⁾ certainement très peu lue à l'étranger. Les rédacteurs de "Nature" avaient refusé l'article jugé trop spéculatif ! Dans l'ancienne théorie du noyau, l'électron préexistait à son éjection lors de la désintégration β . Après la nouvelle théorie de Heisenberg le problème de la complexité ou de l'éléментарité du neutron restait ouvert et il fut en effet l'objet d'une discussion au congrès de Solvay. Heisenberg avait "l'impression" que le neutron devait être élémentaire sinon on l'aurait déjà vu se désintégrer. Bohr avec le soutien de Perrin rejetait la question : "A mon avis, le sens qu'on doit attacher à la distinction entre particules élémentaires et particules complexes ne peut pas être indiqué sans ambiguïté". Fermi, lui, fonda sa théorie sur "l'hypothèse que les électrons émis par les noyaux n'existent pas pendant la désintégration mais sont créés ainsi que les neutrinos de manière analogue à la formation d'un quantum lumineux qui accompagne un saut quantique d'un atome" (résumé de "Tentativo di una teoria dell' emissione dei raggi "beta")⁽¹⁹⁾.

Dans son électrodynamique quantique (voir chap. 13) Fermi, tout comme Dirac, avait résisté à l'emploi de la deuxième quantification pour les fermions, et, selon Rasetti, ce dernier formalisme lui était peu familier. Mais il devenait indispensable pour décrire la création de l'électron et du neutrino au cours d'un processus élémentaire : on ne pouvait l'éviter à la manière de Dirac dans la théorie du positron. La conservation du

nombre de particules matérielles (électrons, neutrinos) devait être abandonnée. Et le formalisme de Klein et Jordan des ondes quantifiées par des anticommutateurs était prêt pour cela.

Dans l'article ultérieur plus détaillé de Nuovo Cimento⁽²⁰⁾ et dans sa traduction allemande⁽²¹⁾ Fermi développait sa théorie à partir de trois hypothèses :

- a) la non-conservation du nombre d'électrons ou de neutrinos,
- b) l'interprétation du neutron et du proton comme deux états quantiques d'une même particule (isospin de Heisenberg),
- c) le choix d'un hamiltonien d'interaction tel que toute transition de neutron à proton s'accompagne de l'émission d'un électron et d'un neutrino.

Plus précisément Fermi choisissait pour cette interaction la forme $H_I = g [\bar{\tau}_- \bar{\Psi}_e \delta_0 \Psi_\nu + \text{herm. conj.}]$, g étant la constante de couplage à tirer de l'expérience, τ_- la matrice de Pauli transformant un état neutronique en état protonique, Ψ_ν (Ψ_e) le champ quantifié de neutrinos (resp. d'électrons). Nous avons utilisé les notations modernes* pour que le lecteur reconnaisse le choix de l'interaction vectorielle : si l'on ne fait pas l'approximation non relativiste de Fermi pour le neutron et le proton l'interaction est en $(\bar{\Psi}_p \delta^\mu \Psi_n \bar{\Psi}_e \delta_\mu \Psi_\nu + h.c.)$. Fermi, dira-t-on, aurait pu choisir l'interaction scalaire plus simple. Son choix fut guidé par l'analogie électromagnétique : interaction $j_\mu A^\mu$ (ou $j_0 A^0$ dans l'approximation non relativiste).

Bien qu'elle contînt deux arbitraires, la constante de couplage et les fonctions d'onde des nucléons dans le noyau, la théorie de Fermi permettait de démontrer la constance approchée du produit de la durée de vie et d'une

* Fermi utilisait une forme bilinéaire en Ψ_ν et Ψ_e et non pas sesquilineaire comme aujourd'hui $(H_I = g [\bar{\tau}_- \bar{\Psi}_e \delta^\mu \Psi_\nu + h.c.])$ δ étant une matrice 4-4 choisie pour que $\bar{\Psi}_e \delta^\mu \Psi_\nu$ soit la première composante d'un 4-vecteur). Un neutrino (et non un antineutrino comme aujourd'hui) était créé lors de $n \rightarrow p + e^- + \nu$.

certaine fonction de l'impulsion maximale des électrons émis, ainsi que la forme de la distribution d'énergie de ces électrons. Ces deux prédictions étaient bien vérifiées par l'expérience et la seconde avait à condition de prendre la masse du neutrino négligeable devant celle de l'électron.

Il est peu de théories en physique nucléaire qui aient duré aussi longtemps que la théorie de Fermi de l'émission β . Utilisée dans l'approximation de Born (les ordres supérieurs divergent sans rémission puisque la théorie de Fermi n'est pas renormalisable) elle reste même aujourd'hui une bonne approximation des théories à bosons vecteurs (Weinberg-Salam etc.). L'histoire* de son évolution et de ses modifications commence avec le travail de GAMOW et TELLER⁽²²⁾ (1936). L'invariance relativiste laisse le choix entre cinq formes possibles de l'interaction de Fermi. Gamow et Teller utilisèrent cette liberté pour améliorer l'accord avec l'expérience. Plus tard furent introduits des termes violant la parité (YANG et LEE 1956), et la confirmation expérimentale de leur existence surprit bien des théoriciens. Il fallut attendre 1959⁽²³⁾ pour que fut détecté le neutrino par la réaction $\bar{\nu} + p \rightarrow n + e^+$. De tels processus inverses de la désintégration β avaient été prévus dès 1934 par BETHE et PEIERLS⁽²⁴⁾ mais, en raison de leur section efficace très faible (10^{-44} cm^2) il fallut attendre les flux de neutrinos intenses des réacteurs nucléaires pour les observer. En 1934 Bethe et Peierls devaient conclure : "il n'y a pas de possibilité pratique d'observer le neutrino".

* Pour une histoire plus détaillée du neutrino à partir de la théorie de Fermi nous renvoyons le lecteur à la réf. 69 (WU).

6.2. LES RAYONS COSMIQUES :

Nous rappellerons d'abord grossièrement la description actuelle des rayons cosmiques. Les rayons cosmiques "primaires" venant du cosmos (du soleil ou d'ailleurs) et entrant dans l'atmosphère terrestre sont en majeure partie constitués de protons d'énergie très variable, pouvant atteindre des valeurs énormes (jusqu'au joule !). De tels protons interagissent immédiatement avec l'air pour donner les rayons cosmiques secondaires. Nous donnerons une version simplifiée de la suite des réactions qui peuvent se produire. Dans une première collision le proton incident crée des pions π_0 , π^+ , π^- qui se désintègrent rapidement suivant $\pi_0 \rightarrow \gamma\gamma$, $\pi^+ \rightarrow \mu^+ + \nu_\mu$ et $\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$. Les électrons (et positrons) créés suivant $\gamma \rightarrow e^+e^-$ et par la désintégration des muons lents $\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$ et $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$ constituent la composante molle (aisément absorbable par le plomb) des rayons cosmiques secondaires. Ceux produits en haute atmosphère n'atteignent en général pas le niveau du sol. Les muons rapides, grâce à la dilatation du temps peuvent descendre en basse altitude : ils constituent la composante dure, peu absorbée par le plomb.

Il fallut attendre les années cinquante pour qu'un tel schéma soit envisageable ; au début de l'histoire des rayons cosmiques l'ignorance de l'existence puis du nombre et de la nature des mésons posa bien des énigmes. Les seules particules simples connues étaient l'électron, le photon et le proton et dans l'hypothèse alors retenue que ces seules particules pouvaient intervenir dans les rayons cosmiques, seule l'électrodynamique entrait en jeu, et pas encore une théorie spécifique des forces nucléaires. Cependant les spécialistes des rayons cosmiques utilisaient des méthodes de détection (compteurs Geiger, chambre de Wilson) développées pour l'étude de la radioactivité des noyaux et s'intéressaient au même type de propriétés : courbure dans un champ magnétique, absorption et ralentissement dans divers matériaux ... Pour les expérimentateurs de la physique nucléaire naissante, les rayons cosmiques étaient vraiment une pluie d'or tombant du ciel : ils disposaient d'une source naturelle de particules énergétiques, permettant de tester les théories pour des longueurs très petites, alors que les expériences usuelles utilisant des sources radioactives ou les accélérateurs de l'époque donnaient des énergies bien plus faibles.

Ainsi on pouvait explorer l'univers microscopique avec une résolution meilleure que le Fermi (10^{-15} m), dans ce domaine des très courtes distances où justement la théorie des champs semblait échouer en raison des divergences ultra-violettes. Comme l'a bien vu D. Cassidy⁽⁶⁹⁾ les théoriciens adoptèrent face à ce problème des divergences deux types d'attitude : les uns cherchèrent à prouver que l'électrodynamique de Dirac était encore applicable pour le genre de phénomène mis en jeu par les rayons cosmiques, les autres jugèrent indispensable de modifier la théorie pour que change le statut des divergences, pour que naissent des alternatives révolutionnaires à la théorie quantique des champs usuelle.

6.2,a L'origine de l'étude des rayons cosmiques*

Tout au début de ce siècle, il apparut que l'air contenu dans une enceinte fermée présentait une certaine conductivité malgré toutes les précautions pour éliminer les radiations ionisantes. Il était difficile de savoir si cette conductivité résiduelle venait de la radioactivité terrestre, ou d'une autre origine. En 1910 le Suisse GOCKEL entreprit un voyage en ballon à 4 500 m d'altitude, muni d'un électroscope et d'une chambre d'ionisation. La radiation observée était plus intense en haute altitude ! La variation de l'intensité avec l'altitude fut étudiée plus précisément en 1911-12 par l'Allemand HESS. On pensa d'abord que les rayons cosmiques étaient constitués de rayons γ très durs, jusqu'à ce que J. CLAY en 1927 prouvât l'effet de latitude, c'est-à-dire la déflexion des rayons cosmiques par le champ magnétique terrestre. Les techniques de détection ne cessèrent de s'améliorer (en particulier avec l'utilisation de méthodes de coïncidences) et les expériences de détection et d'absorption en basse et haute altitude dans des paysages et des conditions météorologiques plus ou moins favorables se multiplièrent, donnant une matière très riche aux théoriciens.

* Nous passons très vite sur cette période de découverte des rayons cosmiques. Louis Leprince-Ringuet en a donné un récit passionnant⁽²⁴⁾.

6.2.b Le premier échec de l'électrodynamique de Dirac à prédire les ralentissements

En 1931 la théorie du positron était toute jeune et nul ne songeait encore à l'utiliser dans l'étude des phénomènes d'absorption des électrons cosmiques. Il n'y avait même pas de calcul relativiste complet de ce phénomène.

CARLSON et OPPENHEIMER⁽²⁵⁾ utilisèrent une généralisation de la méthode* de Møller utilisant les potentiels retardés (voir chap. I § 4) et l'équation de Dirac pour calculer la perte d'énergie d'un électron par ionisation des atomes. La Bremsstrahlung, plus importante à haute énergie, n'était pas envisagée. L'ionisation observée expérimentalement était beaucoup trop faible. Selon Carlson et Oppenheimer, confiants dans leur calcul, il fallait en déduire qu'une partie des particules cosmiques n'étaient pas des électrons, mais plutôt des particules neutres magnétiques faiblement ionisantes, comme celles que Pauli avaient proposé à la conférence de Ann Arbor de l'été 1931. D'où le titre de la lettre⁽²⁵⁾ et de l'article⁽²⁶⁾ (1932) de Carlson et Oppenheimer : "sur le parcours des électrons rapides et des neutrons". Il s'agit bien sûr du neutrino, encore appelé par Pauli neutron, et encore objet de spéculations concernant sa masse et son moment magnétique.

Pendant ce temps en Allemagne HEISENBERG⁽²⁸⁾ calcula le freinage d'électrons rapides dans la matière et trouva par d'autres méthodes que Carlson et Oppenheimer des résultats peu différents. La valeur théorique du freinage était trop faible. *"Comme raison physique de ce désaccord on pourrait peut-être considérer le rayonnement nécessairement émis pendant les processus de collision, qui a été négligé dans les théories antérieures"*. Mais Heisenberg n'avait pas confiance en le résultat que donnerait la théorie de Dirac pour ce processus du deuxième ordre.

Le calcul de HEITLER⁽²⁵⁾ de juin 1933 montra que la Bremsstrahlung dominait effectivement à haute énergie. Le ralentissement calculé semblait en accord avec les faits. Mais entre-temps le positron avait été découvert et de nouvelles difficultés surgirent.

* Bethe⁽²⁷⁾ fit indépendamment à Rome un calcul analogue avec la même méthode ; l'article ne comporte pas de confrontation aux données expérimentales.

6.2.c La théorie de Dirac du positron échoue elle aussi :

A la fin de 1932 le positron fut découvert par ANDERSON (voir chap. 2) et presque en même temps BLACKETT et OCCHIALINI⁽³⁰⁾ découvrirent la production de gerbes par les rayons cosmiques dans la matière : ces gerbes consistaient en un ensemble de traces d'électrons et de positrons qui semblaient partir d'une même origine située à l'intérieur du matériau absorbant*. Les beaux clichés de ces auteurs étaient le fruit d'une astuce expérimentale : ils opéraient la détente de la chambre de Wilson en coïncidence** avec la particule incidente, au lieu de procéder aléatoirement.

En juin 1933 OPPENHEIMER et PLESSET⁽³¹⁾ calculèrent l'émission de paires par les rayons γ dans la matière. Le résultat s'accordait bien avec les observations pour les γ issus d'une source radioactive (ThC") mais pour la borne supérieure de pénétration des rayons γ cosmiques très énergétiques ils trouvèrent "un désaccord net" avec l'expérience. Ce calcul grossier fut repris par HEITLER et SAUTER⁽³²⁾, puis par BETHE et HEITLER⁽³³⁾ dans une étude systématique des différentes causes d'absorption des radiations par la matière. Pour la création de paire et l'absorption de rayons γ par la matière les prédictions théoriques s'avérèrent excellentes ("meilleures que prévu") pour les rayons γ d'énergie bien connue émis par des sources radioactives. Néanmoins un désaccord subsistait avec l'expérience, cette fois pour le ralentissement des électrons par Bremsstrahlung. Les électrons d'Anderson⁽³²⁾ de 300 MeV étaient beaucoup plus pénétrants que prévu. Bethe et Heitler conclurent "la perte d'énergie des électrons rapides... donne le premier exemple d'échec de la mécanique quantique pour un phénomène extra-nucléaire. Nous croyons que le rayonnement des électrons libres sera un des tests les plus directs de toute électrodynamique quantique future". Mais ils se trompaient en accusant l'électrodynamique

* Blackett et Occhialini pensaient que les paires résultaient de la désintégration de noyaux légers produits par les particules incidentes.

** La technique des coïncidences avait été importée d'Italie (groupe de ROSSI) par Occhialini.

de Dirac. Nous avons là une belle illustration de la thèse de Duhem suivant laquelle une expérience négative ne nous dit pas quelle est l'hypothèse fautive dans la théorie physique : on ne peut pas tester séparément les différentes hypothèses. Dans le cas de l'absorption anormale des rayons cosmiques très énergétiques plusieurs hypothèses pouvaient être accusées : celles de l'électrodynamique de Dirac, celles concernant la nature des particules observées, celles concernant l'analyse du dispositif expérimental. Il se trouve que les premières étaient valides dans ce domaine, mais pas les deux dernières. Des expérimentateurs, comme Blackett et Anderson savaient combien peu nombreuses et peu sûres étaient leurs mesures d'absorption dans ce domaine des très hautes énergies et se gardèrent de condamner la théorie de Dirac.

6.2.d Nouvelles théories pour interpréter l'absorption anormale :

Néanmoins dans les mois qui suivirent la confiance en la théorie de Dirac diminua encore en raison de nouvelles mesures contradictoires, et surtout de considérations de WILLIAMS⁽³³⁾, WEIZSÄCKER⁽³⁴⁾ et LANDAU⁽³⁵⁾ prouvant que la formule d'absorption de Bethe et Heitler était utilisée dans son domaine de validité même pour les très hautes énergies. Au lieu d'utiliser la méthode conventionnelle de calcul de la Bremsstrahlung dans le référentiel du proton cible, ces auteurs interprétaient le phénomène dans le référentiel de l'électron incident comme la diffusion de l'onde électromagnétique transformée de Lorentz du champ statique du proton. Il suffisait alors d'appliquer la formule de Klein-Nishina pour les différentes composantes de Fourier de cette onde dans le domaine de validité reconnu de cette formule. Le résultat était identique à celui de Bethe et Heitler, mais la nouvelle démonstration assurait sa validité même pour les électrons cosmiques les plus énergétiques.

Face à cette situation certains théoriciens comme OPPENHEIMER⁽³⁶⁾ et NORDHEIM⁽³⁷⁾ introduisirent des hypothèses ad hoc de coupure ou de champ critique. HEISENBERG⁽⁷⁷⁾, lui, adopta une solution plus radicale en développant en 1936 la théorie des explosions : les gerbes de Blackett étaient interprétées comme des explosions qui devaient être prédites par la théorie de Fermi : Heisenberg remarquait ingénieusement que la constante

de couplage g de Fermi était homogène au carré d'une longueur $g = \lambda_0^{2*}$ et que par conséquent les différents termes de la série des perturbations devaient se comporter en $(\frac{\lambda_0}{\lambda})^m$, λ étant la dimension caractéristique du problème étudié (par exemple la longueur d'onde d'un électron cosmique) pour $\lambda > \lambda_0$ la suite des termes décroissait, mais pour $\lambda < \lambda_0$ elle croissait : les processus d'émission multiples (à l'origine des gerbes) devenaient plus probables qu'une diffusion simple : d'où l'explosion. Un argument semi-classique "grossier" justifiait encore le comportement explosif : pour une équation d'onde non linéaire avec terme d'interaction en φ^4 , la collision de paquets d'ondes d'énergie supérieure à un certain seuil conduisait à un étalement explosif du paquet (en dessous du seuil le terme d'interaction est négligeable devant le terme cinétique). L'idée des explosions fut bien accueillie par les spécialistes des rayons cosmiques jusqu'aux nouvelles expériences de Anderson.

6.2.e La victoire de la théorie de Dirac : les cascades

Au congrès de Londres d'octobre 1934⁽³⁸⁾ ROSSI ainsi que AUGER et LEPRINCE-RINGUET établirent la distinction entre deux composantes - dure et molle - des rayons cosmiques. Rossi, le premier, avait constaté qu'une partie des rayons cosmiques était absorbée par 8 cm de plomb et que le reste traversait allègrement 10 cm de plomb. Auger et Leprince-Ringuet firent des observations identiques en altitude et éventuellement dans une grotte creusée dans le glacier de Jungfrauoch pour étudier l'effet d'absorption par la glace. Ces expérimentateurs ne firent pas d'hypothèse de nature sur cette composante dure, mais ils conclurent que la composante molle, responsable des gerbes, était créée dans l'atmosphère par la composante dure : sinon elle aurait été totalement absorbée avant d'arriver en basse altitude.

En 1936 ANDERSON et NEDDERMEYER⁽³⁹⁾ firent de nouvelles mesures de ralentissement par le plomb des rayons cosmiques, mais cette fois en distinguant soigneusement la composante molle de la composante dure et en corrigeant un vice de leur dispositif expérimental antérieur qui laissait

* Contrairement à la constante de couplage $\frac{e}{\sqrt{4\pi}}c$ de l'électrodynamique qui est sans dimension.

les particules fortement ralenties (et donc fortement incurvées par le champ magnétique) indétectées. La formule de Bethe et Heitler était maintenant très bien vérifiée pour la composante molle, même à haute énergie (au moins jusqu'à 300 MeV) mais pas pour la composante dure.

Les théoriciens OPPENHEIMER⁽⁴⁰⁾, BHABHA et HEITLER^(41,42) puis CARLSON et OPPENHEIMER⁽⁴³⁾ montrèrent que les formules de Bethe et Heitler relatives à la Bremsstrahlung et à la création de paire rendaient compte de l'existence de gerbes, sans explosion de Heisenberg. Les différentes particules d'une gerbe ne devaient pas être considérées comme issues d'un même point, mais comme produites en cascade par une série d'émissions de photons et de créations de paires du genre



Heisenberg concéda qu'une grande partie des gerbes pouvait être interprétée sans explosion. Mais il restait le problème de la composante dure. Était-ce vraiment des protons ? (Williams⁽³³⁾, Bhabha⁽⁴²⁾).

6.2.f La découverte du méson :

En octobre 1934, à la conférence de Londres déjà citée⁽³⁸⁾, ANDERSON et NEDDERMEYER avaient constaté un ralentissement anormal de certains rayons cosmiques ; il est vrai que cette anomalie, nous l'avons vu, ne reposait qu'en partie sur la confusion des deux composantes, mais les auteurs n'en étaient pas encore conscients. Ils rejetèrent l'assimilation de ces rayons anormaux à des protons proposée par Williams. En 1936 la situation s'était de beaucoup clarifiée. Les frères MONTGOMERY⁽⁵⁰⁾ avaient à l'issue de leurs expériences rejeté l'hypothèse d'une composante dure constituée de protons. Anderson et Neddermeyer⁽³⁹⁾, voyant que la théorie de Dirac-Bethe et Heitler s'appliquait bien à la composante molle du rayonnement, étaient peu tentés de la considérer comme fautive pour l'autre composante. Il devenait clair que l'absorption anormale était plus une question de nature des particules qu'une question d'énergie. Anderson et Neddermeyer prirent alors au sérieux l'hypothèse que les particules pénétrantes n'étaient pas des électrons : "*il est important d'identifier les particules pénétrantes*". Leurs expériences de Pasadena à la recherche de protons primaires avaient donné des résultats négatifs ;

il semblait exclu d'assimiler une grande partie de la composante dure à des protons. De nouvelles expériences d'absorption des mêmes auteurs⁽⁴⁵⁾, de CRUSSART et LEPRINCE-RINGUET⁽⁴⁴⁾ montrèrent encore plus clairement la séparation en deux groupes des particules cosmiques. Anderson et Neddermeyer résumèrent ainsi la situation : "L'interprétation du groupe pénétrant rencontre de graves difficultés, mais elle semble à présent limitée aux hypothèses suivantes : a) qu'un électron possède quelque propriété autre que sa charge et sa masse rendant compte de pénétrations souvent anormalement grandes dans les éléments lourds ou b) qu'il existe des particules de charge unité mais de masse ... plus grande que celle d'un électron et beaucoup plus petite que celle d'un proton : cette hypothèse rendrait compte aussi bien de l'absorption anormale que de l'ionisation observée. Pour autant que la charge et la masse sont les seuls paramètres caractérisant l'électron l'hypothèse b) semble la meilleure hypothèse de travail"⁽⁴⁵⁾ (30 mars 1937). L'ionisation observée sur certaines traces reproduites dans l'article précédent⁽³⁹⁾, ainsi que la longueur et la courbure des traces donnaient de nouveaux arguments en faveur de l'hypothèse b). Les expériences les plus décisives (au dire même d'Anderson et de Neddermeyer) furent conduites par STREET et STEVENSON, à l'aide d'une série de plaques de plomb parallèles disposées dans une chambre de Wilson plongée dans différents champs magnétiques. Ils annoncèrent leurs résultats au congrès de Washington de mai 1937 : "... il est évident que les particules pénétrantes ne peuvent être décrites comme des électrons obéissant à la théorie de Heitler, et que les protons ne peuvent en être une fraction appréciable".

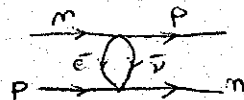
Le 1er juin 1937 OPPENHEIMER et SERBER⁽⁴⁷⁾ envoyèrent une lettre à la Phys. Rev. où les expériences d'Anderson Neddermeyer et Street et Stevenson étaient interprétées comme une "indication claire" de l'existence d'une nouvelle particule de masse comprise en gros entre 50 et 100 MeV et qui pouvait être identifiée à la particule introduite par YUKAWA en 1935. Pauli et Weisskopf avaient développé en 1934 (voir chap. 5 § 3) la théorie quantique d'un champ scalaire et Yukawa avait (indépendamment) introduit des particules scalaires dont l'échange serait à l'origine des interactions nucléaires. Dans la même lettre Oppenheimer et Serber expliquèrent pourquoi ils avaient jusque-là tenu à l'écart la théorie de Yukawa : la divergence des processus du second ordre rendait caduque cette théorie pour l'interaction

p-p et pour le moment magnétique des nucléons. Disons tout de suite que l'identification de la nouvelle particule découverte dans les rayons cosmiques au méson de Yukawa était une "plaisanterie qui dura dix ans" (Oppenheimer)⁽⁴⁹⁾. Les expérimentateurs avaient découvert le méson μ et non le méson π . Mais cette erreur avait déjà un premier rôle positif : elle donnait crédit à la théorie de Yukawa. Nous verrons plus loin comment cette erreur put encore avoir une fonction positive dans le développement de la théorie mésique.

STUECKELBERG⁽⁴⁸⁾ écrivit de Genève le 6 juin à Phys. Rev. pour exprimer une conclusion identique à celle des Californiens : "Il semble hautement probable que Street et Stevenson, Neddermeyer et Anderson ont réellement découvert une nouvelle particule élémentaire qui a été prédite par la théorie". Yukawa dans une note publiée un peu plus tard⁽⁵¹⁾ donna lui-même son analyse des expériences américaines et européennes récentes et il crut reconnaître son méson.

6.3 LA THEORIE MESIQUE :

En juin 1934 TAMM et IWANENKO écrivirent chacun une lettre à Nature, l'un de Moscou, l'autre de Leningrad pour signaler l'une des conséquences de la théorie de Fermi. Les neutrons et les protons devaient interagir par échange d'une paire virtuelle électron-neutrino



même que deux électrons peuvent interagir par échange d'un photon*. Cette généralisation de la notion d'interaction par échange d'une particule est à la base de la description moderne des différents types d'interaction connus. Elle est une description métaphorique de certains termes de la théorie des perturbations, maintenant universellement répandue dans le langage de la théorie des champs. Bien que l'idée d'échange fût la bonne, Tamm et Iwanenko furent bien obligés de constater que l'échange de paires électron-neutrino donnait un ordre de grandeur bien trop faible pour l'interaction neutron-proton.

6.3.a La théorie de Yukawa : (février 1935)

Pour pallier à cet inconvénient, Yukawa décida d'introduire une nouvelle particule dont l'échange fût responsable des interactions nucléaires. Le premier article⁽⁵²⁾, datant de février 1935 fut publié (en anglais) dans une revue japonaise assez peu diffusée. Toutefois l'école japonaise de physique théorique était réputée et l'article de Yukawa fut lu aux Etats-Unis**. Nous suivrons maintenant le plan de l'article de Yukawa.

Pour Hideki Yukawa l'idée de Tamm et d'Iwanenko constituait un progrès important par rapport au "Platzwechsel" (échange) purement heuristique de Heisenberg (voir § 1). L'analogie avec l'électrodynamique quantique serait encore plus parfaite si l'interaction nucléaire se faisait par échange d'une seule particule au lieu d'une paire électron-neutrino. Et le

* La même hypothèse fut avancée par Heisenberg en 1934 au cours d'une leçon non publiée au Cavendish laboratory (réf. 23).

** D'après Oppenheimer⁽⁴⁹⁾ l'idée de Yukawa ne fut connue en Europe qu'après les nouvelles expériences sur les rayons cosmiques.

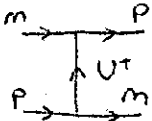
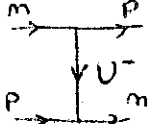
choix de la nature de cette particule permettrait de satisfaire aux exigences empiriques.

Yukawa lorsqu'il publia ce premier article ne connaissait probablement pas la théorie de Pauli et Weisskopf de la quantification d'un champ scalaire chargé. Il dut se contenter d'une présentation semi-quantique et semi-quantitative du champ scalaire responsable de l'interaction. Il n'avait pas non plus à sa disposition le champ vectoriel massif inventé par Proca en 1936 et qui aurait lui aussi satisfait à la statistique. Il choisit donc ce qu'il y avait de plus simple : un champ scalaire massif chargé.

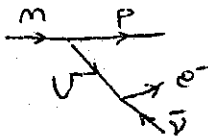
Dans la théorie classique, ce champ, donnant le potentiel d'interaction entre nucléon et proton, de même que le champ électromagnétique donne le potentiel électrostatique, devait satisfaire dans le vide l'équation $(\square + \lambda^2)U = 0$ pour que le comportement exponentiel en $-g^2 \frac{e^{-\lambda r}}{r}$ introduit par Heisenberg et Majorana pour l'interaction nucléon-nucléon se retrouve ; il devait être créé par les transitions $n \rightarrow p$ d'où l'équation en présence du champ $\psi = \begin{pmatrix} \psi_n \\ \psi_p \end{pmatrix}$ associé aux nucléons : $(\square + \lambda^2)U = g \bar{\psi} \tau_{\pm} \psi$ g désignant la constante de couplage et τ_{\pm} la matrice de Pauli faisant passer d'un état neutronique à un état protonique. Le champ ψ dans l'approximation non relativiste devait satisfaire à une équation de Schrödinger-Pauli avec deuxième membre correspondant à l'interaction avec le champ U . L'élimination du champ U dans le cas d'une paire neutron-proton redonnait un hamiltonien d'interaction n-p analogue à celui de Heisenberg.

Après quantification le champ U devait s'analyser en terme de bosons chargés de masse $m_U = \lambda \frac{\hbar}{c}$ (en utilisant l'interprétation de de Broglie des ondes planes) estimée à 200 masses électroniques à partir de la valeur empirique de λ introduite dans la théorie de Heisenberg. Ce lien tout à fait général entre la portée de l'interaction et la masse de la particule échangée, est devenu bien familier aux physiciens d'aujourd'hui. Pourquoi n'avait-on encore jamais vu de particule U ? Yukawa répondait : l'énergie utilisée dans les expériences usuelles de diffusion par des noyaux n'étaient pas suffisante pour créer ces particules massives.

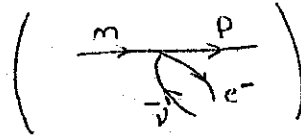
L'interaction neutron-proton, dans la théorie quantique devait être représentée comme un échange virtuel de U_+ , U_- suivant (avec les notations

modernes)  ou  et l'interaction n-n et p-p

comme des processus du deuxième ordre (en g^2).

La désintégration β fut elle aussi réinterprétée par Yukawa comme résultant d'un échange de U suivant  . La masse élevée

de U, montrait-il, assurait que cette nouvelle théorie donne le même résultat que l'interaction de Fermi



En conclusion Yukawa émettait l'idée que sa nouvelle particule pourrait bien se trouver dans les rayons cosmiques.

Yukawa, exploitant l'idée issue de l'électrodynamique quantique de l'interaction par échange d'un boson intermédiaire, venait de prédire l'existence d'un tel boson pour les interactions fortes et aussi pour les interactions faibles. Dans le premier cas, on crut l'avoir découvert en 1937 et le découvrit réellement en 1947, dans le deuxième cas les bosons intermédiaires des interactions faibles, aujourd'hui supposés différents de ceux des interactions fortes, restent encore à découvrir.

Les nouvelles expériences de 1936-1937 sur les rayons cosmiques, nous l'avons vu, confirmèrent l'hypothèse de Yukawa. Dans la note de 1937⁽⁵¹⁾ Yukawa annonçait des calculs plus poussés à l'aide de la théorie scalaire chargée de Pauli et Weisskopf dont il avait pris connaissance.

En décembre 1937 ces calculs étaient achevés⁽⁵³⁾ ; le hamiltonien nucléaire effectif qui en résultait, présentait un signe opposé à celui donné par l'expérience pour les termes d'échange n-p. Yukawa se proposait d'explorer plus tard le cas d'autres types de champs.

6.3.b Autres théories mésiques :

En 1938 plusieurs nouveaux types de théorie mésique furent proposés. La dépendance en spin des interactions nucléaires semblait exiger un méson vectoriel plutôt que scalaire. La théorie de PROCA⁽⁵⁵⁾ du champ vectoriel

massif fut quantifiée et utilisée par KEMMER⁽⁵⁶⁾ et BHABHA⁽⁵⁷⁾, STUECKELBERG⁽⁵³⁾, YUKAWA SAKATA TAKETANI⁽⁵⁴⁾, FROHLICH, HEITLER et KEMMER⁽⁵⁹⁾ : la théorie de Yukawa jouissait d'un succès explosif. La théorie pseudoscalaire, celle qui a survécu, fut inventée par KEMMER⁽⁶⁰⁾ dans la même année. Elle fut reprise en 1940 par MOLLER et ROSENFELD⁽⁷⁶⁾ dans un mélange pseudoscalaire-vectorel, quand la théorie vectorielle fut détrônée par la mesure expérimentale du moment quadrupolaire du deutéron dans le laboratoire de Rabi.

Une autre étape importante fut l'introduction de mésons neutres par KEMMER⁽⁶¹⁾ en avril 1938. L'universalité des interactions nucléaires fut reconnue en 1936 par BREIT et FEENBERG⁽⁶²⁾ et par CASSEN et CONDON⁽⁶³⁾ : les nouvelles expériences de diffusion nucléon-nucléon, entre autres, montraient que l'interaction p-p était attractive et du même ordre de grandeur que n-p. L'hypothèse de l'indépendance de charge fut formalisée par Cassen et Condon⁽⁶³⁾ en introduisant un hamiltonien nucléaire invariant par rotation du spin isotopique. Quand se développèrent les premières théories mésiques les idées originales de Heisenberg concernant l'échange n-p dominaient encore, ce qui explique que l'on n'ait utilisé que des mésons chargés et que seul Kemmer ait songé à introduire des mésons neutres en 1938.

6.3.c Effets du deuxième ordre et divergences :

Le problème des divergences en électrodynamique pure pouvait être momentanément repoussé pour un théoricien pragmatique. Même les expériences sur les rayons cosmiques à très hautes énergies semblaient conformes à ce que l'on pouvait calculer de fini dans cette théorie.

En revanche dans la théorie mésique des effets physiques importants et mesurables comme l'interaction p-p ou le moment magnétique anormal de l'électron et du proton semblaient imprévisibles puisqu'ils étaient du deuxième ordre et par conséquent divergents et OPPENHEIMER⁽⁴⁷⁾ s'en était plaint peu après la publication de l'article de Yukawa.

Avec la théorie de Kemmer des mésons neutres le problème disparaissait en partie pour l'interaction n-n et p-p maintenant calculables dans

l'approximation de Born suivant

$$\frac{m \quad m}{m \quad \left\{ \begin{array}{c} U^0 \\ \end{array} \right\} \quad m}$$

. Mais pour le moment

magnétique des nucléons la difficulté resta longtemps entière.

Dès 1935 WICK eut l'idée de considérer que la plus grande partie du moment magnétique d'un nucléon était due à un nuage virtuel de paires électron-neutrino entourant le nucléon, engendré par l'interaction de Fermi. YUKAWA, SAKATA et TAKETANI estimèrent plus tard, dans la théorie mésique vectorielle avec coupure U. V. les valeurs des moments magnétiques du proton et du neutron mais tombèrent sur un résultat trop faible. D'autres tentatives furent plus heureuses (voir chap. 3 § 4 à propos du g-2 de l'électron) mais la méthode utilisée était pour le moins suspecte.

6.3.d Les théories de couplage fort :

La valeur élevée de la constante de couplage nucléaire (ainsi que sa dimension) rendaient absurde la méthode de perturbations usuelle. Les tentatives pour appliquer la méthode non perturbative de Bloch et Nordsieck (voir chap. 5 § 1) aux théories mésiques échouèrent en raison de la complexité algébrique de l'interaction. Alors HEISENBERG⁽⁶⁵⁾ (1939) et WENTZEL⁽⁶⁶⁾ (1940) proposèrent un développement en puissances décroissantes de la constante de couplage, effectué dans la théorie classique. En raison de la force du couplage les mésons virtuels entourant un nucléon nu devaient avoir des nombres quantiques très élevés, et un calcul classique devait se justifier. Une des premières conséquences de cette théorie était une valeur faible indépendante de la constante de couplage, et dépendant seulement de la taille du nucléon (supposée finie au départ du calcul* pour assurer la finitude des résultats) pour la section efficace de diffusion méson-nucléon. Les données empiriques pour cette diffusion donnaient une valeur très faible devant celles données par le calcul perturbatif usuel en raison de l'identification erronée du pion au muon, et cette erreur donna la justification paradoxalement la plus ferme des théories de couplage fort. Dans ce cadre l'un des résultats les plus importants fut d'abord décrit par Wentzel : l'existence d'états excités "isobares" du nucléon dont la vérification expérimentale par

* Ce qui explique que les théories de couplage fort aient été souvent désignées comme théories de "source étendue".

BRÜCKNER⁽⁶⁷⁾ pour la résonance $p(3/2, 3/2)$ attendit 1952.

Nous ne détaillerons pas ici la longue histoire des théories de couplage fort, bien qu'elle constitue une tentative en partie réussie (avec la prédiction des isobares) de théorie des champs non perturbative, mais hélas aussi non quantique. Soulignons cependant que ces théories ignoraient (avant 1947) le principe de renormalisation, bien qu'un effet comme l'existence d'isobares fût un effet d'énergie propre.

Autres tentatives pour juguler le problème des divergences :

Heisenberg engagea pendant la guerre le programme de matrice S , dans une perspective "observationiste" analogue à celle de la mécanique des matrices. (Voir chap. 4 § 7). La théorie de Heitler du "radiation damping" ainsi que les théories de Wentzel-Dirac et de Bhabha (partant d'une élimination des divergences au niveau classique) donnèrent, elles, un moyen effectif de calculer des valeurs finies pour les processus courants de physique nucléaire. Nous exposerons l'histoire de ces théories, ainsi que d'autres alternatives à la théorie quantique des champs usuelle au chapitre 7

. Il est évident que toute approche nouvelle de l'électrodynamique quantique ou de la théorie des champs en général pouvait avoir des conséquences dans le domaine de la physique nucléaire. De fait le principe de renormalisation et le critère de renormalisabilité changèrent beaucoup la face des problèmes de physique nucléaire.

Bibliographie :

Neutrino, théorie de Fermi :

Wu⁽⁶⁷⁾ : "The neutrino".

Rasetti⁽⁶⁸⁾ : Introduction aux articles de Fermi sur l'émission β dans les œuvres complètes de Fermi.

Bethe et Bacher⁽⁷³⁾, Bethe⁽⁷⁴⁾, Livingston et Bethe⁽⁷⁵⁾ : "Nuclear Physics" A, B et C.

Rayons cosmiques :

David Cassidy⁽⁶⁹⁾ : "Cosmic ray showers, high energy physics and quantum field theories".

Louis Leprince-Ringuet⁽²¹⁾ : "Les rayons cosmiques".

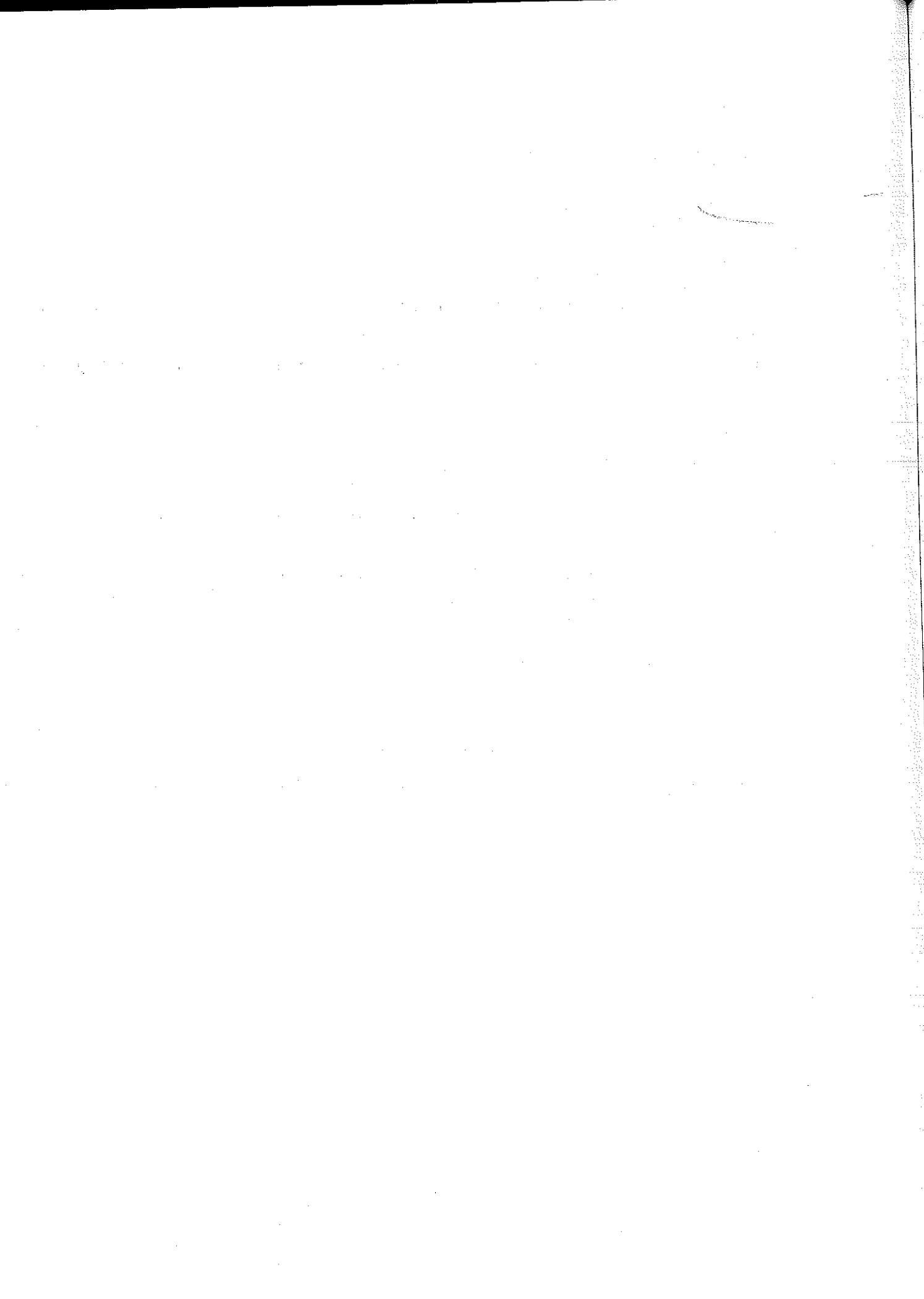
Oppenheimer⁽⁴⁹⁾ : "Thirty years of mesons".

Théories mésiques :

Wentzel⁽⁷⁰⁾ : "Quantum theory of fields".

Wentzel⁽⁷¹⁾ : "Recent research in meson theory".

Bethe, de Hoffmann⁽⁷²⁾ : Meson and fields II (old strong coupling theory).



CHAPITRE 7

ALTERNATIVES THEORIQUES

Nous avons choisi dans ce chapitre quelques exemples importants de théories ou de bases théoriques proposées comme alternatives à la théorie quantique des champs divergente. Il n'est pas possible d'épuiser ce sujet en quelques pages. Parmi nos nombreuses omissions, citons la théorie neutrinique de la lumière, les théories de coupure relativiste (Wataghin, Scherzer, Feynman...), une théorie de l'électron étendu de Peierls et Mac Manus, l'essai de quantification de l'espace par Snyder, l'électrodynamique de Wheeler et Feynman avec action à distance... Cependant les exemples choisis suffisent à montrer quelques fonctions importantes de la recherche d'alternatives : une meilleure compréhension de l'origine des difficultés de la théorie orthodoxe (c'est-à-dire des divergences)

- La refonte ou le développement de théories plus anciennes inachevées.

- L'introduction de concepts nouveaux souvent réintégrés dans des théories futures (matrice S , probabilités négatives....)

On comprendra ainsi mieux pourquoi dans un éventail aussi large de recherches, la solution moderne du problème des divergences ne fut pas entrevue dans les théories les plus ambitieuses, (mathématiquement ou conceptuellement), mais dans une modeste réinterprétation (due à Kramers), de la théorie classique de l'électron.

Avant 1947, les inventeurs de ces nouvelles théories, durent en général se contenter de quelques succès formels ou esthétiques assez peu en rapport avec les efforts d'imagination et de calcul. Plusieurs grands théoriciens, dans l'attente d'une nouvelle idée ou d'une révélation expérimentale, commentèrent les tentatives inachevées ou non corroborées de leurs collègues (voir la bibliographie donnée à la fin de ce chapitre).

La conférence de Cambridge, du 22 ou 27 juillet 1946, permit une confrontation des différentes analyses de la situation. Voici quelques extraits de la préface de Nicholas Kemmer datée de décembre 1947 :

"Ce fut la plus grande conférence sur les particules fondamentales tenue de ce côté-ci de l'Atlantique après la fin de la guerre, et l'évènement fut vraiment important... Echanges de points de vue sous forme de discussions libres, et renouement d'amitiés longtemps interrompues, toujours de grande valeur dans une conférence, n'avaient jamais été plus nécessaires.... Lors de la session d'ouverture, purement théorique, le mot

"difficultés" ressortait justement des titres des exposés [Pauli : "Difficultés des théories des champs et de la quantification des champs"- Dirac : "Difficultés de l'électrodynamique quantique" Bohr : "Problèmes de la physique des particules élémentaires"]. L'écoulement des années de guerre n'a rien changé au fait fondamental : le sujet discuté transcende les limites de la théorie établie et aucun point de départ pleinement satisfaisant n'est connu : pourtant bien des papiers montrent à l'évidence un développement et une classification des idées. Les papiers des deux dernières sessions [Rosenfeld, Wentzel, Heitler, Møller, Stueckelberg....] en témoignent" .

7.1 THEORIES MONISTES ET THEORIES MIXTES :

Les théories "monistes", ainsi nommées par Taketani et Sakata (Voir chap 8 § 6), traitent les différents phénomènes physiques comme des manifestations d'une seule entité physique, en l'occurrence le champ électromagnétique. Il faut les distinguer de théories unitaires comme celles d'Einstein, de Weyl et d'Eddington qui unissent les phénomènes électromagnétiques et gravitationnels en une théorie géométrique non Riemannienne, avec parfois la prétention de prédire des phénomènes quantiques sans l'hypothèse supplémentaire de quantification. Les théories monistes ici-envisagées devaient toutes être soumises à la quantification dans l'opinion de leurs auteurs. Le lecteur s'étonnera peut être que nous réunissions dans un même paragraphe deux types apparemment très dissemblables de théorie. En vérité les théories mixtes sont formellement étroitement apparentées aux théories monistes ; il faut distinguer le principe heuristique et le contenu mathématique des théories qui en découlent.

Le point de vue moniste en électromagnétisme est aussi vieux que la théorie de Thomson-Abraham-Lorentz qui voyait une origine purement électromagnétique à l'inertie de l'électron. Mais, suivant Poincaré, un tel modèle donnait des électrons explosifs, dont les parties se repoussaient mutuellement ; il fallait ajouter des forces de cohésion de nature nouvelle et la belle idée moniste s'effondrait.

7.1a La théorie de Mie de 1912

En 1912 aucune théorie cohérente ne pouvait expliquer des phénomènes quantiques nouveaux comme les spectres atomiques. Gustav MIE concluait : " *Il me semble indispensable pour les progrès ultérieurs de nos connaissances de donner à la théorie de la matière de nouvelles bases* " Comme nouveau point de départ Mie proposait, à titre d' "essai" "de décrire toutes les manifestations du monde matériel à partir des seuls états reconnus de l'éther, c'est-à-dire les champs électriques et magnétiques et le courant de charge " (84). Le Lagrangien de la théorie de Mie (84,85,86) avait la forme $L(A) = -\frac{1}{4} f^2 + \delta L(A, f)$ avec $f_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Le courant et le champ électromagnétique $F^{\mu\nu}$ étaient définis par $j = \frac{\partial L}{\partial A}$ et $F = -2 \frac{\partial L}{\partial f}$ afin que l'équation obtenue en variant

par rapport à A l'action $S = \int L d^4x$ soit identique à l'équation de Maxwell $\partial_\mu F^{\mu\nu} = j^\nu$. La théorie de Mie, on s'en doute, ne fut en mesure de décrire aucun des phénomènes spécifiquement quantiques connus des expérimentateurs. Dans les années trente, comme l'importance de l'invariance de jauge de l'électrodynamique était connue, le rôle absolu donné par Mie au potentiel vecteur A_μ via la définition du courant $j = \frac{\partial L}{\partial A}$ pouvait sembler suspect.

7.1b La théorie de Born et Infeld :

Max BORN expliquait (91), en 1933, la proéminence de la vision dualiste opposant matière à rayonnement d'abord par l'échec des conceptions monistes ("unitarian stand-point") d'Abraham-Lorentz et de Mie et aussi par les succès importants de la mécanique quantique de l'atome faisant usage de la conception dualiste. Le problème des divergences ultra-violettes amenait à reconsidérer la question. Born était *"convaincu de la grande supériorité philosophique de l'idée unitaire"* Son ambition était de donner une nouvelle théorie électromagnétique et aussi une nouvelle méthode de quantification mieux adaptée à cette théorie (et ce dès les premiers articles (87,89) de 1933) écrits juste après l'émigration forcée de Born de Göttingen à Cambridge). La théorie de Born était un cas particulier de la théorie de Mie de 1912, mais avec un lagrangien invariant de jauge, c'est-à-dire sans dépendance en A_μ : $L = \frac{1}{a^2} (\sqrt{1+2a^2 L_0} - 1)$ avec $L_0 = -\frac{1}{4} f^2$ (donc $L \sim L_0$ habituel pour a petit). Ce lagrangien non quadratique conduisait à des équations du mouvement non linéaires, admettant des solutions à singularités ponctuelles interprétables comme des charges d'énergie propre finie. Le courant de Mie $\frac{\partial L}{\partial A}$ étant nul, les singularités remplissaient le rôle de charges. La solution statique à symétrie sphérique donnait la relation $r_e = \sqrt{ae}$ entre le "rayon" de l'électron et sa charge e . Et l'étude du mouvement des singularités donnait $mc^2 = (1.2361\dots) \frac{e^2}{4\pi r_e}$ pour leur masse. Ces deux relations permettaient d'exprimer le paramètre a en fonction des seules grandeurs expérimentales m et e .

Ultérieurement (en 1934) Born et son étudiant polonais Leopold INFELD donnèrent deux justifications a priori de la forme du lagrangien (88,91). La première justification reposait sur une analogie avec la dynamique d'une particule en relativité restreinte. Le lagrangien

d'une telle particule s'écrit $L = -mc^2 \left(\sqrt{1 - \frac{v^2}{c^2}} - 1 \right)$. L'introduction du radical est liée à l'existence de la vitesse limite c . Born et Infeld pensaient que de même l'hypothèse d'un "champ absolu" limite supérieure des champs électriques et magnétiques devait conduire à l'expression analogue du lagrangien de Born, a désignant la valeur du champ absolu. La deuxième justification était plus ingénieuse. L'action la plus simple que l'on puisse former à l'aide d'un seul tenseur A_{μ} et qui soit invariante par un changement quelconque de coordonnées (ainsi que la relativité générale l'exige) est de la forme $S = \int \sqrt{-\det A} d^4x$. Identifiant la partie symétrique g de A au tenseur métrique et la partie antisymétrique $a^2 f$ au champ de Mie, il vient en coordonnées galiléennes $\det A = \det (g + a^2 f) = -(1 - \frac{1}{2} a^2 f^2 - a^4 \det f)$ et en négligeant le terme en a^4 , $S = \int \sqrt{1 + 2a^2 L_0} d^4x$ conformément à l'"Ansatz" original de Born. Le mouvement des singularités calculé par Born et Infeld dans cette théorie (avec les termes en a^4) (en négligeant les accélérations) donnait une image apparemment correcte de l'électron classique. Born put même donner une explication (90,91) de l'anomalie observée par Regener (95) : les rayons cosmiques traversaient sans peine toute la profondeur du lac de Constance (voir chap. 6 § 2 pour les absorptions anormales analogues observées par Anderson en 1933), car selon Born et Infeld ces rayons γ ne voyaient qu'une partie des charges étendues sur des distances r_q .

SCHRÖDINGER (97) (1935) s'intéressa à cette nouvelle théorie classique et la reformula en terme des vecteurs complexes $\vec{E} + i\vec{H}$ et $\vec{B} - i\vec{D}$ (\vec{E} et \vec{B} composantes de $f_{\mu\nu}$ et \vec{D} et \vec{H} composantes de $F_{\mu\nu}$) pour mieux faire apparaître certaines propriétés de symétrie : les champs électrique et magnétique jouaient des rôles complètement symétriques et par ailleurs à chaque solution singulière des équations on pouvait associer une solution d'énergie négative (pour la charge), comme à propos de l'équation de Dirac. Par ces traits contraires à l'expérience, la théorie classique de Born et Infeld se révélait défectueuse. En 1938 Schrödinger collabora avec Born (98) pour donner une nouvelle estimation de la masse des singularités en tenant compte du spin.

Pendant que Born et Infeld travaillaient à une théorie électromagnétique non linéaire, EULER et KOCKEL (96) obtinrent avec l'électrodynamique quantique normale, et son vide polarisable, et par

des voies totalement différentes un lagrangien effectif du même type que celui de Born et Infeld. (voir Chap. 5 § 1) . Plus précisément

$$\delta L = + \frac{1}{90\pi} \frac{1}{\alpha} \frac{1}{E_0^2} [(E^2 - B^2)^2 + 7\vec{E} \cdot \vec{B}] \text{ chez Euler et Kockel}$$

et

$$\delta L = + \frac{(1.2361..)^4}{8} \frac{1}{E_0^2} [(E^2 - B^2)^2 + 4\vec{E} \cdot \vec{B}] \text{ chez Born et Infeld.}$$

$$(E_0 = e/4\pi r_0^2, r_0 = \frac{\alpha}{m})$$

L'identification des coefficients des crochets donnaient même un ordre de grandeur correct pour la constante de structure fine. La coïncidence pouvait frapper les esprits, mais Born ne perdait pas de vue que sa théorie devait atteindre le stade de la quantification avant d'être comparée à ses rivales.

Une première méthode de quantification avait été exposée par Born dès les premiers articles ⁽⁸⁵⁾ de 1933. Elle consistait en une généralisation de l'équation du Hamilton-Jacobi, inspirée par le théorème d'indépendance de Hilbert. Malheureusement la généralisation par Born de ce théorème à des systèmes à plusieurs degrés de liberté (en l'occurrence une infinité) était fautive, comme le montra WEYL ⁽¹⁰⁵⁾ l'année suivante. De toute manière l'avantage d'une formulation symétrique pour les coordonnées d'espace et de temps était détruit par l'obscurité mathématique de la généralisation quantique. Born et Infeld durent revenir aux bonnes vieilles méthodes de quantification de Heisenberg et Pauli pour donner une preuve formelle ⁽⁹²⁾ d'invariance relativiste de la version quantique de leur théorie, puis pour définir les opérateurs \vec{p}, \vec{q} et \vec{M} (impulsion, position et moment cinétique) attachés à une singularité. * Les définitions étaient plus ou moins artificielles et la séparation entre moment intrinsèque et moment orbital ne pouvait être assimilée à la séparation analogue pour l'électron de Dirac. De plus la menace de divergences bien pires (non renormalisables) que celles de l'électrodynamique quantique normale planait sur cette théorie. En commentaire à sa correspondance avec Einstein ⁽¹⁰⁶⁾, Born écrivit : " Nous renonçons à cette

* Pryce ⁽⁹⁴⁾ étudia le même problème.

théorie parce que nous ne réussîmes pas à la mettre en accord avec les principes de la théorie quantique du champ. Du moins cette tentative fut-elle la première à lever à l'aide d'une théorie non linéaire les difficultés apparues en microphysique ". Au bout de l'impasse Born se tourna vers le principe de réciprocité dont nous ne parlerons pas ici.

7.1c. La théorie de Bopp de 1940

L'histoire des théories monistes continue avec " une théorie linéaire de l'électron" (99) de Fritz BOPP. L'idée moniste séduisait Bopp sans que les tentatives de Mie avec son A_μ absolu ou de Born et Infeld avec leurs non linéarités compliquées, le satisfassent pleinement. Bopp ne se faisait pas trop d'illusion sur l'unité profonde des théories avec singularités. Pour connaître complètement le mouvement des singularités il fallait chez Born et Infeld comme chez Bopp ajouter aux équations du champ de Mie $f_{\mu\nu}$ une hypothèse "d'unité énergétique" mettant toute l'énergie dans le champ et aucune aux points singuliers. Cette hypothèse assurait aussi la "cohésion" des particules : la force totale exercée par les champs au point singulier (force propre) devait s'annuler.

Le lagrangien de Bopp était le lagrangien quadratique le plus simple qui différait légèrement du lagrangien normal $-\frac{1}{4} f^2$:

$$L = -\frac{1}{4} \left[f^2 + \frac{1}{\chi^2} (\delta_\rho f_{\mu\nu})^2 \right]$$

Ce lagrangien était aussi privilégié par son interprétation en terme de champ de Maxwell + champ de Proca assurant la cohésion. En effet $F = -2 \frac{\partial L}{\partial f}$ peut se décomposer en somme $F_M^{\mu\nu} + F_P^{\mu\nu}$ d'un champ de Maxwell et d'un champ de Proca, et le tenseur d'énergie impulsion en différence $T_{\mu\nu} = T_{\mu\nu}^M - T_{\mu\nu}^P$. Le signe - devant $T_{\mu\nu}^P$ contredit la positivité de l'énergie et interdit une interprétation physique de la décomposition suivant une remarque de Stueckelberg. Bopp, c'est bien naturel, vit plutôt les avantages de sa théorie : tout en restant linéaire elle donnait une valeur finie aux énergies propres et elle permettait une étude détaillée des réactions de rayonnement en raison de sa relative simplicité.

En 1943 Bopp ⁽¹⁰⁰⁾ donna une formulation plus générale des théories de force propre nulle (analogue à celle de Stueckelberg : voir § 5). L'équation de Lorentz généralisée des charges contenait bien sûr des dérivées d'ordre supérieur et des solutions parasites que Bopp chercha à éliminer systématiquement. Il effleura le problème de la quantification. Celle-ci présentait une nouvelle difficulté : la présence de dérivées temporelles d'ordre supérieur à deux dans le lagrangien.

Ce problème fut résolu indépendamment en 1941 par PODOLSKY et KIKUCHI ⁽¹⁰²⁾ qui s'intéressaient à une théorie de Podolsky ⁽¹⁰¹⁾ identique à celle de Bopp de 1940, visant elle aussi à rendre les énergies propres finies. Grâce à une astuce due à Ostrogradsky (prendre les dérivées \dot{q} comme de nouvelles variables indépendantes) la théorie pouvait être mise sous forme canonique.

7.1 d Théories mixtes et champs f de Pais

Une fois extraits le champ de Maxwell et le champ de Proca du champ de Bopp de 1940, et une fois déduites du principe d' "unité énergétique" l'équation de Lorentz généralisée donnant le mouvement des singularités, la théorie de Bopp devenait une théorie dualiste mixte. PAULI et VILLARS ⁽⁸³⁾ distinguèrent deux sortes de théories mixtes : les "formalistes" où les différents champs couplés aux charges contribuent avec des signes différents au tenseur d'énergie impulsion, et les "réalistes" où ce tenseur est une somme positive des différentes contributions. Les théories mixtes de Stueckelberg (voir § 5) sont de ce dernier type.

A la fin de la guerre, PAIS ^(103,104) examina le problème de la quantification des théories mixtes des deux types. Les "champs f" de Pais assurant la cohésion étaient couplés ainsi que le champ de Maxwell au champ de Dirac ψ . La quantification du tout (compte tenu de la théorie du positron) donnait les résultats suivants :

• Pour un seul champ f de Proca soustrait (à la Bopp), l'énergie propre de l'électron était finie à tous les ordres de la série des perturbations. En langage moderne la régularisation de Pauli-Villars de l'énergie propre de l'électron est possible avec une seule masse auxiliaire (il faut au moins deux masses pour régulariser les diagrammes de polarisation du vide).

• Pour un seul champ f scalaire neutre ajouté (à la manière "réaliste") on peut rendre finie l'énergie propre d'un électron au 1er ordre (en e^2 et en f^2 si f désigne le contenu de couplage électron-scalaire). Une telle compensation ne pourrait se produire dans la théorie classique puisque les énergies propres positives infinies du champ de Maxwell et du champ f s'y ajoutent ; en revanche dans la théorie quantique la contribution des champs f à l'énergie propre est négative. La compensation des divergences ne se serait bien sûr pas maintenue.

Dans cette deuxième version du champ f , Pais alla jusqu'à calculer l'effet pasternack * . La valeur obtenue était finie, mais elle dépendait de la masse du champ f , à défaut d'une renormalisation de masse. L'absence des quanta f dans la nature s'expliquait par leur grande instabilité (suivant $f \rightarrow e^+ e^-$).

Nous verrons au chap.8 comment les théoriciens japonais menèrent des calculs analogues à ceux de Pais. L'introduction de masses auxiliaires pour rendre convergente l'énergie propre de l'électron (c'est-à-dire "régulariser") est aussi due à Stueckelberg (voir § 5) (sous une forme identique à la forme définitive de Pauli et Villars ⁽⁸³⁾ et indépendamment à Feynmann (voir chap § § 6). Ils'avéra que tout l'intérêt de cette méthode de "régularisation" résidait dans le respect des invariances fondamentales (relativiste et de jauge).

En quelque sorte les théories mixtes avaient atteint leur but : éliminer les divergences. Mais la théorie finie résultante, en raison de la non positivité de l'énergie (et de l'instabilité correspondante du vide) n'admettait pas d'interprétation physique immédiate. Il fallait encore renormaliser et faire tendre les masses auxiliaires vers l'infini pour obtenir l'électrodynamique quantique moderne.

* Aujourd'hui appelé déplacement de Lamb ; voir Chap § § 3

7.2 A PARTIR DE THEORIES CLASSIQUES DE L'ELECTRON PONCTUEL

7.2.a La limite de Wentzel de 1933 :

Ce procédé inventé par WENTZEL fut utilisée à nouveau par Dirac en 1939 sous une forme légèrement différente sous le nom de λ -limite et nous la nommerons ainsi dès maintenant.

Nous avons vu au Chap. 23 comment DIRAC, FOCK et PODOLSKY (17) (D.F.P.) utilisèrent une théorie à temps multiples pour donner une formulation explicitement covariante de l'électrodynamique quantique (les électrons étant des électrons de Dirac en 1° quantification). Bien que cette théorie fût directement quantique, Wentzel eut l'idée ingénieuse d'exploiter la notion de temps multiples pour améliorer d'abord la théorie classique. Pour échapper aux divergences de l'électron ponctuel, il était d'usage dans la théorie classique d'utiliser comme Lorentz, Abraham, Poincaré..., un modèle de l'électron de rayon fini. "Inutile d'expliquer", écrivit Wentzel, "qu'il serait plus satisfaisant de considérer une "particule élémentaire" d'emblée comme sans extension".

La difficulté bien connue de l'électrodynamique de Maxwell-Lorentz vient de ce que le champ électromagnétique total à distance finie d'une électron diverge lorsque cette distance tend vers zéro. Wentzel utilisa la possibilité offerte par la théorie "multi-temps" de considérer des "distances du genre temps" au lieu du genre espace : la théorie de D.F.P.

donnait une extension du champ électromagnétique $F_{\mu\nu}^W(x, t_1, \dots, t_n)$ pour des valeurs du temps x^0 du champ et des temps t_i des particules distinctes et il se trouve que la limite $\lim_{\substack{\lambda \rightarrow 0 \\ \lambda^2 > 0 \\ \lambda_0 > 0}} \frac{1}{2} (F_{\mu\nu}^W(z_i + \lambda) + F_{\mu\nu}^W(z_i - \lambda)) = f_{\mu\nu}^i$

($z_i = (t_i, \vec{z}_i)$ 4-vecteur de la particule i ; t_1, t_2, \dots, t_n fixés) est finie et non ambiguë. Pour l'obtenir il suffit de résoudre celles des équations de D.F.P. qui déterminent A_μ en fonction des sources

dans l'approximation non quantique ($\hbar \rightarrow 0$)^{*}. La force propre ainsi obtenue, $e \frac{dz_i^\mu}{ds_i} \cdot f_\mu^{iv}$, s'identifie à la force de freinage de rayonnement de Lorentz-Abraham^{**}.

Il n'y a plus de terme d'inertie infini ni de termes complémentaires dépendant de la structure de l'électron.

Dans le premier article⁽¹⁵⁾ Wentzel aborda la généralisation quantique de sa théorie d'une manière peu commode : il essaya de donner un sens au développement en puissances de $\frac{d\vec{z}_i}{dt_i}$ (maintenant opérateur) du champ propre. (Les calculs précédents de la limite de Wentzel ne valaient que pour des \vec{z}_i et $\frac{d\vec{z}_i}{dt}$ c- nombres). Wentzel avait deux possibilités présentant chacune leurs difficultés - dans une théorie où $[\vec{z}_i, \frac{d\vec{z}_i}{dt}] \neq 0$ le problème de l'ordre de ces opérateurs - dans la théorie de l'électron de Dirac où $[\vec{z}_i, \vec{\alpha}_i] = 0$ le problème de la convergence de la série en puissance de \vec{z}_i donnant le champ propre ($|\frac{d\vec{z}_i}{dt}| = |\vec{\alpha}_i| = 1$ est plus grand que le rayon de convergence). "Ainsi il ne semble pas exclu que la raison de nos difficultés se trouve du côté de la mécanique ondulatoire relativiste plutôt que dans la théorie du champ électromagnétique"⁽¹⁵⁾. En 1933 la théorie des positrons était toute récente, et le problème de la polarisation du vide ou encore d'une théorie relativiste cohérente de l'électron ne devait être en bonne partie résolue que l'année suivante. Rien d'étonnant alors si Wentzel accusait la théorie de l'électron plutôt que l'électrodynamique quantique dans son ensemble.

Le second article de Wentzel sur la λ -limite donna une généralisation quantique plus effective du procédé et montra quelles divergences il permettait d'éliminer. Plutôt que de chercher à expliciter

* Plus précisément Wentzel utilise l'équation $\frac{\partial A_\mu}{\partial s_i} = -e_i \Delta(x-z_i) \frac{dz_\mu}{ds_i}$ obtenue en écrivant $\frac{\partial A_\mu}{\partial t_i} = i [H_i, A_\mu]$

avec $H_i = H_i^c + e_i (\varphi - \frac{d\vec{z}_i}{dt_i} \cdot \vec{A}) = H_i^c + e_i \frac{ds_i}{dt_i} A_\nu \frac{dz_\nu}{ds_i}$ et $[A_\mu(x), A_\nu(y)] = i g_{\mu\nu} \Delta(x-y)$ (Les commutateurs, dans la version non quantique, sont à interpréter comme des crochets de Poisson).

** $e \frac{dz_i^\mu}{ds_i} f_{\mu\nu}^i$ propre = $\frac{e^2}{6\pi} (\dot{z}_\nu - \dot{z}^2 z_\mu)$ Wentzel fait d'abord les calculs dans le référentiel tangent à l'électron et donne seulement à la fin l'expression covariante.

f_{uv} propre, il valait mieux n'exprimer la limite qu'au fur et à mesure des calculs de processus physiques. Pour ce faire Wentzel modifiait les équations habituelles de l'électrodynamiques de D.F.P. en modifiant le terme du couplage :

$$e_i \frac{d\vec{z}_i^\mu}{ds_i} A^\mu \longrightarrow e_i \frac{d\vec{z}_i^\mu}{ds_i} \langle A^\mu \rangle$$

$\langle A^\mu \rangle$ désignant une forme explicitement invariante relativiste de la limite de Wentzel obtenue grosso modo en faisant la moyenne pour tous les λ à l'intérieur du cône positif de la λ -limite.

Lors de l'élimination des degrés de liberté longitudinaux du champ, les termes d'énergie propre coulombienne en $\frac{e^2}{|\vec{r}_i - \vec{r}_j|}$ disparaissent grâce à la λ -limite. Mais la contribution des modes transverses à l'énergie propre divergeait encore. Waller (18) (voir chap. 3 § 2) avait obtenu une intégrale divergente du type $\int_0^\infty d\nu (A + B\nu)$. Grâce à la λ -limite le terme en $\int_0^\infty A d\nu$, disparaît. Ce terme, comme le remarqua Wentzel est en fait non quantique, et correspond au champ de coulomb divergent d'une charge ponctuelle classique entraînée dans le mouvement. L'autre terme contient \hbar et s'annule quand $\hbar \rightarrow 0$, il est spécifiquement quantique et lié aux fluctuations du vide.

Même après cette tentative ingénieuse le principal écueil subsistait. Weisskopf, lorsqu'il calcula en 1934 l'énergie propre de l'électron dans la théorie du positron eut l'espoir que la λ -limite et l'effet de polarisation du vide puissent conjointement donner une valeur finie de l'énergie propre de l'électron. Dans son premier calcul faux (voir chap 3 § 3) Weisskopf trouva le même degré de divergence que Waller et cet espoir s'évanouit. Mais il resurgit lorsque Weisskopf, suivant la remarque de Furry, trouva la divergence logarithmique. Plus tard PAULI (20) put montrer que la λ -limite avait un effet plutôt nuisible dans la théorie du positron*. L'explication tenait à l'incompatibilité de la λ -limite annulant toute interaction sur des temps inférieurs à $\sqrt{\lambda^2}$.

* la divergence quadratique de Waller réapparaissait.

et de la polarisation du vide donnant naissance à une infinité de paires virtuelles dans le voisinage spatio-temporel d'un électron.

Il faut plutôt chercher l'intérêt de la λ -limite au niveau de la théorie classique . Signalons déjà que la limite de Wentzel conduit aux mêmes équations de l'électrodynamique classique que la théorie de Dirac de 1938. Mais les fondements sont moins généraux que dans cette théorie de Dirac . Wentzel avait utilisé une démarche assez curieuse : partant d'une théorie quantique (Dirac Fock Podolsky) il réussit à améliorer la théorie classique correspondante et il quantifia à nouveau . Par là-même, il servit de relais entre Dirac et Dirac puisque Dirac incorpora la limite de Wentzel, désormais baptisée λ -limite, dans sa théorie de 1938. Pour atteindre la généralité de celle de Dirac, il manquait à la théorie de Wentzel un bilan de l'énergie-impulsion du champ et de la matière chargée, et une démonstration explicite de la structure hamiltonienne des équations qu'il utilisait. Mais les résultats étaient équivalents et surtout l'indépendance de structure, bientôt chère à Kamers, était proclamée .

7.2b La théorie de Dirac de 1938 (1)

Pour remédier aux difficultés persistantes de l'électrodynamique quantique il fallait d'abord donner une construction plus satisfaisante de la théorie classique : tel fut le point de vue de Wentzel, de Kramers, et de Dirac en 1938. Après sa dernière contribution pourtant importante à la théorie de la polarisation du vide en 1934, DIRAC trouva intolérable le degré de complexité auquel était parvenu l'électrodynamique quantique. Ses succès lui paraissaient maigres en comparaison aux insurmontables problèmes de divergences. Une expérience de SHANKLAND (2) suffit en 1936 pour qu'il reniât toute cette théorie qu'il avait pourtant créée et développée en bonne partie. Shankland croyait avoir observé une violation de la conservation de l'énergie dans des processus atomiques, et Dirac comme Bohr, Kramers et Slater en 1924, ou comme Bohr, à propos de l'émission β *, n'hésita pas à remettre en question ce robuste principe et les théories quantiques relativistes qui en faisaient usage : "*Puisque le seul propos de l'électrodynamique quantique- mis à part celui d'unifier les hypothèses de la théorie du rayonnement- est de rendre compte du genre de coïncidences que précisément les expériences de Shankland viennent d'exclure, nous pouvons l'abandonner sans regret- en fait, en raison de son extrême complexité, beaucoup de physiciens seront très contents d'en voir la fin*" (3). Alors Dirac se tourna vers des problèmes purement formels (équations d'onde relativistes (4), variables complexes en mécanique quantique (5) ou vers la cosmologie (6) (7) et même l'épistémologie (8). En 1938 Dirac vint à penser (1) : "*une nouvelle idée physique est maintenant requise, une idée intelligible à la fois dans la théorie classique et dans la théorie quantique, et notre voie d'accès la plus facile est de nous limiter à la théorie classique*".

* Alors Dirac s'était fermement opposé à Bohr sur ce point : "Ma propre opinion sur cette question est que je préférerais garder à tout prix la conservation rigoureuse de l'énergie et que j'abandonnerais même l'idée d'une matière constituée d'atomes séparés et d'électrons plutôt que la conservation de l'énergie" Lettre de Dirac à Bohr du 26 novembre 1929.

Le retour à l'électrodynamique classique déjà effectué par Born donnait lieu à trop de "complexité" et d'arbitraire selon Dirac. Au lieu, comme Born, de modifier les équations de Maxwell, Dirac pensa qu'il valait mieux les garder quitte à introduire des termes de "soustraction" comme dans la théorie du positron. Dirac donnait à son choix une justification d'ordre épistémologique : plutôt que d'utiliser comme Lorentz ou Born "un modèle fondé sur des idées physiques ordinaires" il vallait mieux établir à priori "un schéma mathématique bien défini et cohérent et en accord avec les principes fondamentaux". Ainsi Dirac put obtenir pour la première fois une électrodynamique parfaitement invariante relativiste, délivrée du problème des divergences et déduite d'un nombre restreint d'hypothèses. Et cette nouvelle théorie de Dirac devait devenir la pierre angulaire de toute électrodynamique classique rigoureuse. En quoi était-elle nouvelle ? Dirac répond ⁽¹⁾ : "Nous travaillerons tout le temps avec les équations de Maxwell habituelles, et ainsi nous n'aboutirons à aucune équation essentiellement nouvelle. Nous chercherons plutôt à obtenir une interprétation physique satisfaisante pour des équations déjà bien connues. On peut se demander pourquoi ce problème n'a pas été résolu bien avant. Beaucoup de travaux passés ont exploré les implications générales de la théorie de Maxwell, mais ce bien avant la découverte de la mécanique quantique en 1925 ; alors les gens cherchaient surtout à savoir comment un électron pouvait rester sur une orbite atomique sans rayonner - question dont la réponse se situe hors de la théorie classique, comme nous le savons maintenant - et ainsi ils ne se soucièrent pas de trouver l'interprétation la plus naturelle de leurs équations".

Jusqu'en 1938, il n'y avait guère d'interprétation satisfaisante des équations de Maxwell-Lorentz, en dehors du modèle de Lorentz-Poincaré-Abraham (voir chap 3 § 1 et pour plus de détails l'article de Pais ⁽⁹⁾) d'un électron de rayon fini. Dans ce modèle l'utilisation d'une distribution de charge continue permettait d'éviter le problème des divergences associées à un électron ponctuel, et de séparer la force propre exercée par le champ de l'électron sur l'électron de celle exercée par le champ extérieur. Dirac ⁽¹⁾ critiqua ce modèle parce qu'il ne pouvait être valable qu'approximativement, seulement pour les champs et accélérations suffisamment faibles pour lesquels la structure de l'électron n'intervient pas, et parce que l'hypothèse d'une masse purement électromagnétique pour les particules fondamentales (associée à l'énergie

propre coulombienne de l'électron étendu) ne pouvait tenir face à la découverte récente du neutron et des masses négatives de l'électron de Dirac. Dirac renvoyait aussi aux arguments de FRENKEL (1925) en faveur d'un électron ponctuel. Frenkel reprochait à l'électron de Lorentz sa complexité sa rigidité incompatible avec la relativité restreinte et surtout son instabilité : *"je tiens cette énigme (et les questions liées) pour un problème purement scholastique. Il résulte d'une extension aveugle aux particules élémentaires de la matière (les électrons) d'un principe de division qui une fois appliqué à des systèmes composés (atomes etc...), a justement conduit à ces dernières particules élémentaires. Les électrons sont indivisibles physiquement et géométriquement. Ils n'ont pas d'étendue dans l'espace. Il n'y a pas de forces intérieures entre les éléments d'un électron, car de tels éléments n'existent pas. L'explication électromagnétique de la masse tombe aussi, ainsi que toutes les difficultés qui sont liées à l'établissement d'équations du mouvement d'un électron plus exactes à partir de la théorie de Lorentz"* (1). Frenkel insista sur la nécessité de distinguer champ extérieur et champ total, mais ne donna pas d'expression générale du champ extérieur, laissant ainsi son exigence d'un électron ponctuel au stade du dogme.

Signalons, puisque Dirac s'y réfère, le travail important de FOKKER (12) (1929), donnant une dynamique rigoureuse invariante relativiste pour un système isolé d'électrons ponctuels. Cherchant un principe variationnel, les variables du champ une fois éliminées, pour décrire ce système, Fokker se rendit compte qu'il fallait utiliser non pas les potentiels retardés, mais la semi-somme des potentiels retardés et des avancés. La théorie de Fokker constitue un cas particulier de la théorie de Dirac ($A_{\mu}^{\text{in}} + A_{\mu}^{\text{out}} = 0$) *; mais le point de départ de Dirac en 1938 fut très différent.

* Elle inspira aussi plus tard Wheeler et Feynmann (107) dans leur théorie de l'électrodynamique avec action à distance.

Bref résumé de la théorie de Dirac (1) :

(Cas d'un seul électron)

Le champ électromagnétique total $F^{\mu\nu}$ se compose du champ incident $F_{in}^{\mu\nu}$ et du champ retardé créé par les sources $F_{ret}^{\mu\nu}$.
 $F^{\mu\nu} = F_{in}^{\mu\nu} + F_{ret}^{\mu\nu}$. Dirac définit $F_{out}^{\mu\nu}$ par $F^{\mu\nu} = F_{out}^{\mu\nu} + F_{adv}^{\mu\nu}$,
 $F_{adv}^{\mu\nu}$ désignant le champ avancé créé par les sources.

Si les sources fonctionnent pendant un intervalle de temps limité pour $t = -\infty$ $F^{\mu\nu} = F_{in}^{\mu\nu}$ } d'où l'interprétation physique de
 $t = +\infty$ $F^{\mu\nu} = F_{out}^{\mu\nu}$ } $F_{in}^{\mu\nu}$ et $F_{out}^{\mu\nu}$
 comme champs incidents et émergents.

Leur différence $F_{rad}^{\mu\nu} = F_{out}^{\mu\nu} - F_{in}^{\mu\nu} = F_{ret}^{\mu\nu} - F_{adv}^{\mu\nu}$ représente le "champ de radiation" créé par les sources. Ce champ a la propriété remarquable d'avoir une limite finie au lieu de l'électron donné par :

$$F_{\mu\nu}^{rad} = \frac{e}{3\pi} \left(\ddot{z}_\mu \dot{z}_\nu - \dot{z}_\nu \ddot{z}_\mu \right)$$

$z_\mu(s)$ étant la ligne d'univers de l'électron paramétrée par le temps propre s , et le point désignant $\frac{d}{ds}$.

Les équations de Maxwell s'écrivent :

$$\partial_\mu F^{\mu\nu} = j^\nu \quad j^\nu(x) = e \int_{-\infty}^{+\infty} \frac{dz^\nu}{ds} \delta^4(x-z) ds$$

Il reste à déterminer l'équation du mouvement de l'électron.

Pour cela Dirac considère le flux d'énergie et d'impulsion à travers la surface d'une sphère (dans le référentiel propre de l'électron) centrée sur l'électron entre les instants s_1 et s_2 . Ce flux, par conservation de l'énergie-impulsion, doit être égal à la différence entre l'énergie-impulsion $B_\mu(s_2)$ contenu dans la sphère à l'instant s_2 et la même quantité à l'instant s_1 : $B_\mu(s_1)$.

Ou encore :

$$\int \partial_{\mu\nu} ds^\nu = B_\mu(s_2) - B_\mu(s_1)$$

$$\partial T_E(s_1, s_2)$$

$T_E(s_1, s_2)$ désignant l'hypersurface (à 3 dim) définie par les conditions précédentes et $\theta_{\mu\nu}$ le tenseur d'énergie-impulsion du champ électromagnétique $F^{\mu\nu}$.

L'évaluation de l'intégrale de surface dans la limite

$$\epsilon \rightarrow 0 \text{ donne : } \int_{S_1}^{S_2} \left(\frac{1}{2} \epsilon^{-1} \frac{d^2 z_\mu}{dS^2} - e \frac{dz_\nu}{dS} f_\mu^\nu \right) dS = B_\mu(S_2) - B_\mu(S_1)$$

ou encore : $\frac{1}{2} \epsilon^{-1} \ddot{z}_\mu - e \dot{z}_\nu f_\mu^\nu = \dot{B}_\mu$ avec $f_{\mu\nu} = F_{\mu\nu} - \frac{1}{2} (F_{\mu\nu}^{ret} + F_{\mu\nu}^{adv})$

Pour B_μ Dirac prend le choix "le plus simple" $B_\mu = k \dot{z}_\mu^*$

L'équation du mouvement de l'électron devient

$$m \ddot{z}_\mu = e \dot{z}_\nu f_\mu^\nu \quad \text{à condition de poser } m = k + \frac{1}{2} e^2 \epsilon^{-1}.$$

Comme $f_{\mu\nu} = F_{\mu\nu}^{in} + \frac{1}{2} F_{\mu\nu}^{rad}$ l'équation obtenue

ne présente pas de divergence et s'écrit encore en utilisant l'expression

de $F_{\mu\nu}^{rad}$ au lieu de l'électron et les identités $\dot{z}^\nu = 1$, $\dot{z}_\nu \dot{z}^\nu = 0$, $\dot{z}_\nu \ddot{z}^\nu + \ddot{z}_\nu \dot{z}^\nu = 0$

$$m \ddot{z}_\mu - \frac{e^2}{6\pi} \ddot{z}_\mu - \frac{e^2}{6\pi} \ddot{z}^\nu \dot{z}_\nu = e \dot{z}_\nu F_{\mu\nu}^{in}$$

C'est l'équation aujourd'hui nommée de "Dirac-Lorentz". Elle est identique à l'équation de Lorentz-Abraham dans la limite de rayon nul. Mais dans le modèle de Lorentz-Abraham le rayon de l'électron devait rester fini (pour que la masse de l'électron soit purement électromagnétique) et l'équation de "Dirac-Lorentz" était considérée comme approximative, alors que dans la théorie de Dirac "il y a de bonnes raisons de la croire exacte, dans les limites de la théorie classique" (1)

Cette équation, aussi belle soit elle, présente un certain nombre de bizareries qui rendent son interprétation physique un peu délicate. Dirac ne se découragea pas pour autant. Il fut au contraire séduit par la nouveauté du rapport introduit entre cette théorie mathématique et l'expérience ("le plus beau trait de la théorie").

L'équation est du 3ème ordre et non pas du 2ème, il faut donc connaître z, \dot{z} et \ddot{z} à l'instant origine pour calculer le mouvement ultérieur d'un électron. En champ incident nul, l'équation est analytiquement soluble et l'on constate l'existence de solutions exponentiellement et indéfiniment accélérées ("run-away solutions")**. Pour éliminer ces solutions non physiques Dirac ajouta une hypothèse supplémentaire à sa

* Ce choix peut être en fait justifié. Voir Rohrlich (13).

** Ainsi nommées par Peierls.

théorie : l'accélération dans la limite $t \rightarrow +\infty$ devait s'annuler.

Mais même après l'élimination de ces solutions parasites, la théorie présente quelques aspects étranges. Une impulsion électromagnétique envoyée vers un électron excite cet électron dès qu'elle s'en approche à des distances de l'ordre de $r_0 = \frac{e^2}{4\pi mc^2}$ et cet électron rayonne aussitôt, si bien que le signal transmis aura une avance de l'ordre de $\frac{r_0}{c}$ sur un signal se propageant à la vitesse c . Selon Dirac il ne fallait pas voir là une violation de la causalité mais plutôt une manifestation de l'extension finie de l'électron, bien que la théorie de Dirac parte d'objets formellement ponctuels et sans structure. Il restait la propagation de signaux physiques plus rapides que la lumière : *"c'est une modification fondamentale des idées ordinaires de relativité et il faut l'interpréter ainsi : il est possible pour un signal d'être transmis plus vite que la lumière à l'intérieur de l'électron. La taille finie de l'électron réapparaît maintenant en un sens nouveau, l'intérieur de l'électron étant une région de violation non pas des équations du champ électromagnétique, mais de quelque propriété de l'espace-temps."* Il est très étonnant de voir que contrairement aux idées reçues, une théorie peut être parfaitement invariante de Lorentz et admettre cependant la propagation de signaux à une vitesse supérieure à c . Aujourd'hui cette propriété de la théorie de Dirac est plutôt vue comme une non-localité dans le temps^{*}; pour Rohrlich⁽¹³⁾ il vaut mieux considérer que les signaux lumineux se propagent toujours à c , mais qu'ils peuvent se propager vers le passé, en raison du rôle que jouent les potentiels avancés dans la théorie de Dirac. Or la somme de deux vecteurs du genre lumière peut donner un vecteur du genre espace^{**}, on voit que même avec cette interprétation formellement différente il

* qui se manifeste clairement quand on écrit l'équation de Dirac-Lorentz sous forme intégral-différentielle.

** Des considérations analogues ont été utilisées récemment par certains pour justifier la non-localité de la mécanique quantique. Cependant les genres de non-localité contenus dans la théorie quantique et dans la théorie de Dirac sont fondamentalement différents puisque le second permet de propager de l'information à une vitesse supérieure à c et pas le premier.

existe la possibilité dans la théorie de Dirac de transmettre de l'information à une vitesse supérieure à c . Mais de tels effets ne seraient observables que sur une échelle de temps $(\frac{r_0}{c})$ où les phénomènes quantiques prédominent et il vaut mieux abandonner l'espoir de les observer.

En 1943 ELIEZER (14), élève de Dirac, montra que dans la théorie de Dirac il n'existait pas de solution du problème de Kepler où l'électron tombât sur le noyau, comme la perte d'énergie de l'électron par rayonnement semblait l'exiger. La difficulté, considérée comme sérieuse par Dirac et Eliezer était en réalité plutôt académique, car l'électron tombe effectivement pour des distances au noyau supérieures à r_0 ; les anomalies n'interviennent que pour des distances infra-nucléaires.

La généralisation à plusieurs électrons ne posait aucun problème et Dirac put même généraliser le principe variationnel de Fokker (12) pour qu'il donne exactement les équations de sa théorie. Dirac pensait ainsi avoir un bon point de départ pour la quantification de sa théorie; il subsistait seulement la difficulté de mettre les équations sous forme hamiltonienne (déjà constatée par Fokker).

Comme Dirac avait utilisé pour établir l'équation du mouvement de l'électron des considérations énergétiques, on devait pouvoir définir pour le système électrons- rayonnement un tenseur d'énergie impulsion fini et conservé. La démonstration fut donnée dès 1938 par PRYCE (59) et elle fut utile plus tard à Bhabha et à Stueckelberg dans d'autres théories classiques de particules ponctuelles.

La théorie mésique classique de Bhabha (1939)

On pourrait croire que la théorie de Dirac de 1938 n'était que du domaine des mathématiques pures avant d'atteindre le stade de la quantification. Bien au contraire elle fut utilisée et généralisée par BHABHA dès 1939 pour calculer des sections efficaces de diffusion méson- nucléon et pour argumenter contre la conception de HEISENBERG d'une "longueur universelle" (voir chap. 4 § 6). Pour Heisenberg l'universalité des divergences de la théorie quantique des champs devait avoir un pendant empirique: elle devait traduire l'existence d'une longueur universelle au delà de laquelle les conceptions ordinaires d'espace et de temps perdraient leur sens, et qui serait le seuil

d'apparition de phénomènes nouveaux comme les "explosions". A Cambridge les divergences n'étaient qu'un artefact d'une méthode de calcul ou d'une théorie impropre. Dans une lettre à "Nature" (60) Bhabha put donner des prédictions claires, sans divergence dans un domaine d'énergie où Heisenberg (et même depuis peu Heitler) ne croyaient guère plus à la théorie des champs mésiques .

Bhabha (42) sut habilement généraliser les théories de Dirac au cas de nucléons ponctuels et sans spin (fictifs) couplés à un champ vectoriel massif neutre * . L'équation de Dirac -Lorentz devenait un peu plus complexe mais rendait possible les calculs de diffusion classique sans aucune divergence. L'approximation non quantique devait être valable si les propriétés d'impulsion des quanta individuels pouvaient être négligées, c'est-à-dire pour une impulsion des mésons diffusés très inférieure à la masse du nucléon. Cette restriction comprenait la plupart des données expérimentales, même pour les rayons cosmiques. Pour des mésons ultra-relativistes la section efficace de diffusion se comportait maintenant en $\frac{g^4}{\mu^2}$ au lieu du $\frac{g^4 p^2}{\mu^4}$ de la théorie habituelle (g constante de couplage, p impulsion du méson, μ sa masse), conformément à la valeur faible observée expérimentalement. De plus aucun phénomène d'explosion n'apparaissait. Si vraiment il existait une longueur universelle à la Heisenberg, elle ne pouvait être associée, pensait Bhabha, à la seule masse de méson et elle devrait plutôt trouver son origine dans les effets de spin ou de charge ici négligés.

Signalons une propriété intéressante de cette théorie de Bhabha. La masse du nucléon définie comme le coefficient de la dérivée temporelle du second ordre dans l'équation du mouvement de la particule soumise à un champ incident sinusoïdal de fréquence ω , dépend de cette fréquence. Car le terme analogue au terme de freinage de l'équation de

* La théorie de l'interaction des mésons chargés ne pouvaient avoir d'analogue classique, puisqu'un méson chargé, en interagissant avec un nucléon, changeait son identité (proton \leftrightarrow neutron).

Dirac-Lorentz contient une contribution dépendant de ω à l'inertie de la particule. La relation entre la masse à basse énergie M' et la masse M à haute énergie est $M' = M - \frac{1}{2} g^2 \mu$.

Une suite importante de la théorie de Bhabha fut la réponse de Heisenberg, qui souhaita "corriger" la théorie de Bhabha en étudiant toujours une théorie classique, mais avec source étendue contrairement à Bhabha, et commença ainsi les développements en "couplage fort" (voir chap. 6 § 3). Nous ne détaillerons pas ici ces théories.

Parallèlement au travail de Bhabha, BOOTH et WILSON ⁽⁶²⁾ (1940) calculèrent les mêmes sections efficaces (pour une théorie mésique vectorielle neutre) dans la version quantique perturbative de la théorie. Là encore la diffusion à haute énergie était faible, et cette valeur faible n'était obtenue que grâce à l'absence de charge et de spin. Suivant une nouvelle idée de Bhabha, HEITLER et MA ⁽⁶³⁾ (1940) introduisirent de manière ad hoc dans le calcul perturbatif de la diffusion des mésons chargés des états intermédiaires des nucléons de charge $-e$, $2e$ etc... pour obtenir, grâce à une compensation des contributions des nucléons virtuels de charge positive et négative, un comportement à haute énergie des sections efficaces analogue à celui de la théorie neutre. Mais cette compensation ne se produisait que pour la partie de l'interaction indépendante du spin des nucléons ("g₁-diffusion"). L'autre partie ("g₂-diffusion") pouvait être atténuée dans la théorie classique comme le montra Bhabha ^(64,65) en 1940, (alors à l'institut des sciences de Bangalore (Inde)) en utilisant une nouvelle généralisation de la théorie de Dirac de 1938 établie par lui-même et par CORBEN, donnant les équations du mouvement exactes d'un dipôle ponctuel décrit par un 6-vecteur $S_{\mu\nu}$. Selon Bhabha l'amortissement provoqué par les termes de freinage était responsable de cet affaiblissement de la g₂-diffusion. Nous verrons plus loin comment Heitler retint cette idée au niveau quantique pour donner simultanément aux deux types de diffusion (g₁ et g₂) une valeur raisonnable.

7.2c Théorie de Dirac et λ -limite

La difficulté technique de mettre les équations de la théorie de Dirac données par un principe variationnel à la Fokker, sous forme hamiltonienne était liée à la singularité en $\delta((z_i - z_j)^2)$ de l'interaction entre 2 électrons i et j contenue dans l'action de Fokker. Dirac (14) (1939) (conférence à l'Institut Poincaré) introduit un 4-vecteur λ de genre temps $\lambda^2 > 0$ et tel que $\lambda_0 > 0$ pour faire la substitution $\delta(x^2) \rightarrow \theta(x_0) \delta((x+\lambda)^2) + \theta(-x_0) \delta((x-\lambda)^2)$ signifiant que deux particules ne devaient interagir que lorsque l'une était légèrement (à λ -près) écartée vers l'extérieur du cône de lumière de l'autre. L'idée était tout à fait analogue à celle de Wentzel.

Ainsi Dirac aboutit aux équations canoniques suivantes :

$$\left\{ \begin{array}{l} \frac{dz_\mu^i}{ds_i} = \left\{ z_\mu^i, F_i \right\} \quad (1) \\ \frac{dp_\mu^i}{ds_i} = \left\{ p_\mu^i, F_i \right\} \quad (2) \\ \frac{dA_\mu}{ds_i} = \left\{ A_\mu, F_i \right\} \quad (3) \end{array} \right. \quad \text{avec} \quad \left\{ \begin{array}{l} \left\{ z_\mu^i, p_\nu^i \right\} = -g_{\mu\nu} \quad \text{etc.} \\ \left\{ A_\mu(x'), A_\nu(x'') \right\} = \frac{1}{2} g_{\mu\nu} \left(\Delta(x' - x'' + \lambda) + \Delta(x' - x'' - \lambda) \right) \\ \text{pour les crochets de Poisson des} \\ \text{variables dynamiques} \end{array} \right.$$

et la condition supplémentaire $\partial A = -\frac{1}{2} \sum_i e_i (\Delta(x - z_i - \lambda) + \Delta(x - z_i + \lambda))$

Ce système d'équations, à la limite $\lambda \rightarrow 0$ est équivalent au système équations de Maxwell + équation de Dirac-Lorentz, en prenant :

$F_i = -\frac{1}{2m_i} ((p_i - eA_i)^2 - m_i^2)$ * En effet les équations (1) et (2)

donnent l'équation de Dirac-Lorentz avec $f_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ et

l'équation (3) (assortie des relations fondamentales de commutation) donne

$A_\mu = \frac{1}{2} (A_\mu^w(x+\lambda) + A_\mu^w(x-\lambda))$, ainsi que Dirac le remarqua. Les propriétés

du champ de Wentzel assuraient qu'à la limite $\lambda \rightarrow 0$ $f_{\mu\nu}$ soit identique au $f_{\mu\nu}$ défini dans le § précédent.

* ou $F_i = -p_i^2 + m_i$ pour des électrons de Dirac.

L'utilisation de la λ -limite par Dirac différait un peu de celle de Wentzel : elle était introduite au niveau des commutateurs du champ au lieu d'apparaître dans le terme de couplage des équations du mouvement des électrons. La manière de Dirac avait l'avantage de donner des calculs plus simples après quantification, mais donnait des équations du mouvement rigoureusement équivalentes à celles de Wentzel. *

Lors de la même conférence à l'Institut Poincaré, Dirac aborda la généralisation quantique. Les équations quantiques étaient exactement celles de Dirac-Fock-Podolsky (voir chap 2 § 4) à l'introduction-près de la λ -limite dans la fonction Δ . Dirac calcula la fonction d'onde d'un seul électron et du champ électromagnétique au 2ème ordre en e . Il constata comme Wentzel que l'énergie propre coulombienne disparaissait grâce à la λ -limite; l'énergie propre transverse divergeait encore, mais grâce à l'utilisation d'un formalisme explicitement covariant relativiste Dirac put voir "que cet infini était tellement simple qu'on pouvait l'éliminer en ajoutant une constante à la fonction de Hamilton " :

$$-p + m \longrightarrow -p + m + \frac{e^2 \hbar}{4\pi \epsilon_0 m \lambda^2}$$

En termes modernes la λ -limite constituait une méthode de régularisation covariante ** donnant une expression de la masse renormalisée en fonction de la masse nue (les électrons étant traités en 1° quantification) Waller en 1930, puis Weisskopf en 1934 n'avaient pu donner d'expression covariante de l'énergie propre d'un électron car la coupure utilisée dans l'espace des k n'étaient pas covariante ***. Dirac conclut

* La théorie de Dirac a été critiquée en 1945 par T. Lewis (28), mais cette critique visant la justesse mathématique du formalisme canonique de Dirac reposait sur une confusion de notation (29).

** Weisskopf avait conscience que l'énergie propre devait être covariante (Note dans (23))

** A condition de soumettre λ à la même transformation de Lorentz que x .

ainsi sa conférence : "Il est donc prouvé que la théorie permet d'éliminer tous les infinis dans le cas d'un seul électron, jusqu'à l'ordre e^2 ". On pouvait croire que Dirac crut le problème des infinis définitivement réglé. En réalité Dirac, nous allons le voir, chercha à annuler les divergences qui subsistaient malgré la λ -limite, sans doute parce qu'il n'avait pas conscience qu'une redéfinition de la masse (expression qu'il ne prononce pas) était le seul effet divergent, et que tout processus physique calculable devenait fini après cette redéfinition.

Pourtant Dirac, dans ces travaux, de 1939, avait anticipé les travaux d'après guerre de Tomonaga Schwinger et Pauli-Villars sur deux points fondamentaux : utilisation d'une formulation explicitement covariante de l'électrodynamique, méthode de régularisation covariante*.

Rapport entre la théorie de DIRAC de 1938-1939 et la théorie de KRAMERS de 1937.

Il nous semble important d'examiner les similitudes et les différences de ces deux théories, car la théorie de Kramers est considérée (à juste titre) comme l'origine la plus importante du principe de renormalisation alors que celle de Dirac était à peine mentionnée quand furent trouvés en 1947 des effets radiatifs finis. De plus nous n'avons pas trouvé de rapprochements explicites dans la " littérature".

Nous exposerons la théorie de Kramers et ses développements ultérieurs au chap 3. Schématiquement les équations de l'électrodynamique classique de Dirac et de Kramers s'écrivent en notations relativistes** et en omettant les indices tensoriels :

$$\text{Dirac} \left\{ \begin{array}{l} m\ddot{\mathbf{z}} = e v \cdot \mathbf{f} \\ \partial \cdot \mathbf{F} = \mathbf{j} \\ \mathbf{f} = \mathbf{F} - \frac{1}{2} (\mathbf{F}_{\omega} + \mathbf{F}_{\text{ret}}) \end{array} \right. \quad \text{Kramers} \left\{ \begin{array}{l} m\ddot{\mathbf{z}} = e v \cdot \mathbf{F}_{\text{ext}} \\ \partial \cdot \mathbf{F} = \mathbf{j} \\ \mathbf{F}_{\text{ext}} = \mathbf{F} - \mathbf{F}_{\text{coul}} \end{array} \right.$$

* Le lecteur se demande peut-être pourquoi après la naissance de la théorie de la renormalisation on n'a pas fait usage de la λ -limite : nous avons déjà donné la réponse, la λ -limite n'est pas compatible avec une description en 1° quantification des particules matière.

** que Kramers n'utilisait pas.

F_{coul} désignant le champ identique au champ coulombien de l'électron dans son référentiel tangent.

Ces deux systèmes d'équations diffèrent car $F_{\text{coul}} \neq \frac{1}{2} * (F_{\text{av}} + F_{\text{ret}})$, mais on peut les confondre dans l'approximation non relativiste où se place Kramers. $\frac{1}{2} (F_{\text{av}} + F_{\text{ret}})$ est une généralisation relativiste de F_{coul} . *

Une des conséquences les plus spectaculaires de l'identité approximative des deux théories et de leurs généralisations quantiques est l'identité des prédictions pour la partie non relativiste du déplacement de Lamb **

Une fois reconnue leur analogie formelle, les théories de Kramers et de Dirac dépendaient aussi de certains présupposés : il fallait pour ces deux théoriciens comprendre d'abord mieux la théorie classique avant de penser à sa quantification et il fallait donner des équations indépendantes de la structure de l'électron.

Sur ce point important de l'indépendance de structure on trouve une première divergence entre les exposés de Kramers et ceux de Dirac : pour ce dernier l'indépendance de structure devait être obtenue en ne considérant que des électrons rigoureusement ponctuels ; pour Kramers elle devait simplement résulter d'une séparation convenable entre champ extérieur et champ propre et de l'introduction au fondement même de la théorie des seules quantités observables \int_K masse physique et champ extérieur . Kramers ne faisait pas d'hypothèse sur la constitution

* Par définition de F_{coul} , $F_{\text{coul}} = F_{\text{ret}} = F_{\text{av}}$ si l'électron est animé d'un mouvement uniforme et F_{coul} et $1/2 (F_{\text{ret}} + F_{\text{adv}})$ sont symétriques par $t \rightarrow -t$. Pour plus de détails, voir l'article de Kramers au Congrès Solvay 1947.

** Kramers a montré en 1948 que sa théorie donnait le même résultat que le calcul de Bethe . Il reste à établir que la λ -limite donne aussi le même résultat. Avec les notations introduites au chap 8 pour le calcul de Bethe, le déplacement est une somme d'intégrales : $\int_0^{\infty} k dk \frac{1}{E_m - E_n + k}$

La λ -limite transforme cette intégrale en $\lim_{\lambda \rightarrow 0} \int_0^{\infty} \cos(\lambda_0 k) dk \frac{k}{E_m - E_n + k}$
décomposant $\frac{k}{E_m - E_n + k}$ suivant $\frac{1}{E_m - E_n + k} - \frac{E_m - E_n}{E_m - E_n + k}$, le 1er terme, correspondant au contre-terme de masse de Bethe, donne 0, et le 2ème donne une divergence logarithmique en $\ln \lambda_0$; la coupure $\lambda_0 \sim \frac{1}{mc}$ redonne le résultat de Bethe.

intime de l'électron, il admettait seulement la possibilité d'écrire des équations indépendantes de sa structure reliant les quantités physiques observables, et la possibilité d'absorber ainsi toutes les divergences : ce point de vue constitue l'essence du principe de renormalisation. Au contraire la préoccupation principale de Dirac était plus formelle : il s'agissait pour lui d'établir un système d'équation sans divergence pour des charges ponctuelles.

L'impact historique des théories de Kramers et de Dirac se mesure bien plus aux présupposés et aux idées qui guidèrent leurs auteurs qu'à leur contenu prédictif. En effet, ces idées étaient susceptibles d'une généralisation. Le principe de renormalisation de Kramers, transposé au niveau quantique, fournit en 1947 un nouveau point de départ pour la théorie quantique des champs.

7.3 LA METRIQUE INDEFINIE

La λ -limite, comme on peut aisément s'en convaincre, a pour effet d'introduire dans les intégrales portant sur l'impulsion k des photons virtuels un facteur $\cos \lambda_0 k$. Les divergences du type $\int^{\infty} k^n dk$ deviennent $\lim_{\lambda_0 \rightarrow 0} \int^{\infty} k^n \cos(k\lambda_0) dk$, les intégrales étant entendues au sens des distributions. Si n pair la procédure rend l'intégrale convergente. Ainsi les divergences d'origines classiques liées au caractère ponctuel de l'électron linéaires en k ($n=0$) disparaissent ; mais les divergences quadratiques ($n=1$) de Waller subsistent. Pour éliminer les divergences résiduelles de rang impair, DIRAC assortit en 1941 la λ -limite d'une nouvelle méthode de quantification faisant usage d'une métrique "indéfinie" (c'est-à-dire non positive) dans l'espace de Hilbert des états.

Le premier exposé de cette nouvelle méthode se trouve dans la "Bakerian Lecture" de Dirac soutenue le 19 juin 1941 ⁽²⁴⁾. Cet exposé et le suivant ⁽²⁵⁾ datant de 1943, sont de lecture difficile, pas seulement en raison de l'interprétation physique délicate, mais aussi en raison de la conception très symboliste des mathématiques qu'a Dirac : la métrique indéfinie est introduite de manière très formelle, et la nécessité de probabilités négatives pour représenter l'algèbre formelle utilisée par Dirac * n'apparaît qu'en fin des calculs. PAULI

* L'utilisation de la représentation de Fock (voir chapitre 1) masque les propriétés métriques de l'espace des états.

clarifia le contenu mathématique de cette nouvelle théorie de Dirac dans un article de 1943 ⁽²⁶⁾; pour une meilleure compréhension nous utiliserons sa présentation.

7.3.a Les présupposés philosophiques de Dirac :

Le succès de la démarche de Heisenberg inventeur de la mécanique des matrices servit plusieurs fois d'exemple historique de résolution d'une crise scientifique. Nous verrons plus tard (§ 5 et chap.8 § 2) comment Heisenberg lui-même, et sans doute aussi Kramers, tentèrent d'imiter cette démarche. On trouve la même référence à Heisenberg en tête de la "Bakerian lecture" de Dirac " *Heisenberg, en découvrant la mécanique quantique, a introduit un nouveau point de vue sur la nature de la théorie physique. Auparavant on considérait toujours comme essentiel qu'il existât une description détaillée ("detailed description") de ce qui se passe dans les phénomènes naturels, et on utilisait cette description pour calculer des résultats comparables à l'expérience. Heisenberg fit valoir qu'il était suffisant d'avoir un schéma mathématique par lequel calculer d'une manière consistante les résultats de toute expérience. Une description détaillée dans le sens traditionnel est inutile et peut même être impossible à établir".* Privé des intuitions données par une description détaillée, il fallait à Dirac une nouvelle heuristique. Comme dans les précédents travaux de Dirac l'interprétation physique continuait d'être postérieure au "jeu mathématique". Ainsi Dirac s'opposa à Eddington lors de sa première leçon de Dublin ⁽²⁵⁾: " *ces leçons ainsi que celles d'Eddington ont pour but d'unifier la relativité et la théorie quantique, mais elles absorbent la question d'un point de vue différent. La méthode d'Eddington est de clarifier d'abord les idées physiques et ensuite de construire progressivement un schéma mathématique. Notre méthode est juste le contraire - construire d'abord un schéma mathématique et essayer ensuite de trouver son interprétation physique. Cette méthode devrait être la plus facile car elle commence par la tâche la plus facile [...]* Le schéma, pour être acceptable, doit être simple et beau ("neat and beautiful") et le nombre de tels schémas que peuvent donner les mathématiques pures est très limité".

Ces présupposés de Dirac, d'ordre épistémologique, permettent de comprendre comment il fut amené à cette théorie étrange de 1942 où apparaissent des photons d'énergie négative avec une probabilité négative, et que la plupart des théoriciens - avec l'exception notoire de Pauli - refusèrent même d'examiner de près. *"Les énergies négatives et les probabilités négatives ne devraient pas être considérées comme un non-sens. Elles sont des concepts mathématiquement bien définis, comme une somme d'argent négative, puisque les équations exprimant les propriétés importantes des énergies et des probabilités peuvent être utilisées mêmes lorsque celles-ci sont négatives. Ainsi les énergies et les probabilités négatives devraient être considérées seulement comme des choses qui n'interviennent pas dans les résultats expérimentaux. L'interprétation physique de la mécanique quantique relativiste obtenue par une extension naturelle de la théorie non relativiste introduit ces choses et elle est donc en contradiction avec l'expérience. Nous devons par conséquent trouver des manières de modifier ou de compléter cette interprétation"*. (Dirac⁽²⁴⁾).

La théorie de la mer de Dirac était un artifice mathématique destiné à éliminer les énergies négatives de la théorie de l'électron. Au fur et à mesure de son développement Dirac trouva repoussante la complexité de cette mer infinie*. Et dans sa *"Bakerian lecture"* il proposa une nouvelle interprétation de l'équation de Dirac sans mer, en introduisant un *"monde hypothétique"* presque entièrement saturé de positrons, c'est-à-dire avec seulement un petit nombre d'état d'énergie négative occupés, et en faisant l'hypothèse que les probabilités de transition étaient les mêmes dans ce monde hypothétique que dans le monde réel. Une telle théorie donnait évidemment des résultats identiques pour

* Pauli, à juste titre nous semble-t-il, critiqua cette opinion de Dirac : "je désapprouve Dirac quand il dit que les équations sont plus compliquées quand on part avec les états d'énergie négative occupés, que dans l'ancienne forme de la théorie utilisant un nombre constant de particules chargées faisant des transitions vers les états d'énergie négative" (20).

les processus élémentaires usuels (diffusion $e^- - e^+$, création de paires, annihilation....) ; l'absence de polarisation du vide dans cette mer presque vide n'était pas bien gênante en l'absence de toute manifestation expérimentale de ce genre de phénomène.

Pour les bosons se posait le problème des probabilités négatives. Pauli et Weisskopf, en 1934 (voir chap.5§3) avaient montré comment la 2ème quantification permettait de résoudre le problème pour un champ scalaire. Il est probable que la quantification du champ scalaire chargé guida Dirac dans sa nouvelle quantification du champ électromagnétique ; dans les deux cas la condition d'hermiticité des champs quantiques n'est pas vérifiée (comme nous allons le voir). Pour éliminer les divergences de l'électrodynamique, il ne fallait pas hésiter, selon Dirac, à introduire des probabilités négatives, quitte à changer in fine l'interprétation physique de la théorie.

7.3.b La notion de métrique indéfinie et son application à la quantification du champ électromagnétique (d'après Pauli (20)).

Le métrique indéfinie de Dirac-Pauli n'est autre qu'une forme quadratique non dégénérée η dans l'espace de Hilbert des états. L'introduction de η permet de définir une conjugaison par rapport à η à la place de la conjugaison hermitique usuelle par rapport à la norme dans l'espace de Hilbert *.

Alors il devient possible de quantifier un oscillateur harmonique de deux manières différentes : pour un oscillateur de hamiltonien $H = \frac{1}{2} (p^2 + q^2)$ avec les relations canoniques $[p, q] = -i$ les relations de "réalité" $p = \check{p}$ et $q = \check{q}$ sont interprétables de deux manières :

* Le conjugué d'un ket $|\psi\rangle$ est alors $\langle\psi|\eta$ et la valeur moyenne d'une observable O est $\langle\psi|\eta O |\psi\rangle$ dans l'état $|\psi\rangle$.

1°) $\check{p} = p^+$ et $\check{q} = q^+$ alors on introduit les annihilateurs et créateurs $a = \frac{1}{\sqrt{2}} (p - iq)$ $a^+ = \frac{1}{\sqrt{2}} (p + iq)$, qui ont pour action sur les états propres $|n\rangle$ de valeur propre $n + \frac{1}{2}$ du hamiltonien $a |n\rangle = \sqrt{n} |n-1\rangle$ et $a^+ |n\rangle = \sqrt{n+1} |n+1\rangle$

2°) $\check{p} = -p^+$ et $\check{q} = -q^+$: la conjugaison utilisée est antihermitique. Alors H s'écrit $H = -(N + \frac{1}{2})$ N ayant pour spectre les entiers naturels, et les opérateurs $b = \frac{1}{\sqrt{2}} (p+iq)$ et $\check{b} = \frac{1}{\sqrt{2}} (p-iq)$ jouent le rôle d'annihilateurs et de créateurs :

$$b|n\rangle = \sqrt{n} |n-1\rangle \text{ et } \check{b}|n\rangle = -\sqrt{n+1} |n+1\rangle$$

Une définition de η assurant que la conjugaison $\check{}$ s'identifie à la conjugaison antihermitique pour p et q est $\eta |n\rangle = (-1)^n |n\rangle$. La probabilité d'avoir n quanta dans un état $|\psi\rangle$: $\langle \psi | \eta | \psi \rangle$ est négative et vaut $(-1)^n |\psi_n|^2$ si $|\psi\rangle = \sum_n \psi_n |n\rangle$.

Pour simplifier l'exposé nous considérerons seulement la nouvelle quantification d'un champ scalaire, mais le schéma s'adapte aisément au cas du champ électromagnétique. Un champ scalaire neutre libre A(x) tel que $(\square + \mu^2)A = 0$, $A = A^*$ ayant pour décomposition de Fourier $A = \int_{\vec{k}} \alpha_k (a_k e^{-ikx} + \check{a}_k e^{+ikx})$ ($k_0 = |\vec{k}|$) et α_k désigne le coefficient de normalisation) se quantifie usuellement suivant

$$[a_k, \check{a}_k] = 1 \text{ et } \check{a} = a^+ \text{ et alors le hamiltonien s'écrit } H = \sum_{\vec{k}} k_0 (N_k + \frac{1}{2})$$

Dirac utilise simultanément les deux possibilités de quantification d'un oscillateur harmonique et il écrit

$$A = \frac{1}{\sqrt{2}} \left[\sum_{\vec{k}} \alpha_k (a_k e^{-ikx} + \check{a}_k e^{ikx}) + \sum_{\vec{k}} \alpha_k (b_k e^{ikx} + \check{b}_k e^{-ikx}) \right]$$

Le premier morceau étant quantifié suivant la méthode 1°) et le second suivant 2°). H s'écrit maintenant $H = \sum_{\vec{k}} k_0 (N_k^+ - N_k^-)$, N_k^+ et N_k^- correspondant respectivement aux photons d'énergie positive et négative.

Avec cette nouvelle méthode de quantification, une intégrale divergente à une boucle de photon dans la théorie couplée s'écrit

$$I = \int_0^{+\infty} f(k_0) dk_0 + \int_0^{+\infty} f(-k_0) (-1) dk_0$$

le premier terme correspond à la probabilité positive d'émettre et de

réabsorber un photon virtuel d'énergie positive, et le deuxième terme celle négative (d'où le (-1)) d'émettre un photon d'énergie négative. On a encore $I = \int_{-\infty}^{+\infty} f(k_0) dk_0$ au lieu de $I = \int_{-\infty}^{+\infty} f(k_c) dk_c$ dans l'ancienne théorie et les divergences du type $\int_{-\infty}^{+\infty} k_0^n dk_0$ avec n impair disparaissent.

7.3.c Interprétation de Dirac de sa nouvelle théorie

Ainsi conjuguant la λ -limite et la métrique indéfinie Dirac aboutissait à une théorie finie (ordre par ordre). Restait le problème de l'interprétation, donnée dans les articles (24) et (25).

L'introduction des photons d'énergie négative se fait par l'adjonction de "variables redondantes" à la théorie usuelle. Pour s'en convaincre il suffit de remarquer que $B = \frac{1}{\sqrt{2}} \left[\sum_k \alpha_k (a_k e^{-ikx} - \frac{v}{a_k} e^{ikx}) + \sum_k (b_k e^{ikx} - \frac{v}{b_k} e^{-ikx}) \right]$ commute avec A et donc avec toutes les variables dynamiques intervenant dans le hamiltonien total. Au lieu de se placer dans une représentation diagonalisant N_k^+ et N_k^- simultanément, on peut se placer dans une représentation diagonalisant les opérateurs N_k^a et N_k^b correspondant aux quanta des champs A et B. Dans cette nouvelle représentation on retrouve l'électrodynamique quantique usuelle en se restreignant aux états vides de quanta b. Il faut donc s'attendre à ce que les divergences usuelles de l'électrodynamique soient maintenant reportées dans la transformation canonique faisant passer de l'ancienne à la nouvelle représentation. Dirac donna une interprétation physique dans le "monde hypothétique" de la première représentation : l'état initial, dans cette interprétation était le vide $N_k^+ = N_k^- = 0$ et l'état final était interprété en comptant les photons d'énergie positive comme créés dans le processus de diffusion et ceux d'énergie négative comme absorbés. Il restait encore une difficulté : la probabilité qu'un nombre impair de photons soit absorbé était négative. Cette difficulté fut levée par Dirac d'une manière assez tortueuse, il exprima l'état $N_k^+ = N_k^- = 0$ dans la représentation N_k^a, N_k^b et utilisa les lois d'Einstein sur l'absorption et l'émission du rayonnement pour établir une relation entre les amplitudes de transition calculées dans le monde hypothétique et les coefficients d'Einstein.

* On trouve pour la probabilité d'avoir N_k^a photons dans cet état (-2) N_k^a !

7.3.d La critique de Pauli :

PAULI fut très intéressé par la nouvelle méthode de Dirac, mais il discuta dans un échange de lettres le bien fondé de l'interprétation physique de Dirac. Celle-ci faisant usage des relations d'Einstein introduisait un élément extérieur à la théorie, une sorte de nouvelle recette de correspondance comme on en avait dans l'ancienne théorie quantique. Dirac répondit en montrant que la considération d'états initiaux plus généraux que $N_k^+ = N_k^- = 0$ donnait une justification des relations d'Einstein. L'article de Pauli " *La nouvelle méthode de Dirac de quantification du champ*" de 1943 (20) montre bien son intérêt pour cette nouvelle théorie, mais la conclusion restait critique : "Le côté arbitraire des règles de traduction des résultats concernant le monde hypothétique en résultats concernant le monde réel et le manque d'unicité de règles semblent indiquer que de nouvelles idées et des changements plus radicaux du formalisme actuel seront nécessaires pour obtenir une théorie quantique vraiment satisfaisante du champ électromagnétique". La critique de Pauli se durcit au congrès de Cambridge de 1946 (30) "l'interprétation des résultats dans le monde réel semble essentiellement incorrecte... De surcroît la théorie avec les "probabilités négatives" donne dans le monde réel des matrices S qui en général ne sont pas unitaires ; ce doit être considéré comme une grave objection de principe ". En effet il nous semble que la violation probable de l'unitarité constitue le principal écueil de la théorie de Dirac.

7.3e Une nouvelle version de la théorie de Dirac (1946)

Nous avons vu que la nouvelle méthode de quantification de Dirac conduisait à des calculs finis ordre par ordre dans la représentation N_+, N_- des nombres de photons dans le "monde hypothétique" et qu'en conséquence les divergences étaient réintroduites par la transformation unitaire faisant passer de cette représentation à la représentation usuelle N_a, N_b des nombres de photons réels N_a et redondants N_b . Dirac pensa que cette réapparition des divergences était liée au traitement ordre par ordre et qu'une solution non perturbative finie dans une représentation devait l'être aussi dans une autre. La représentation N_+, N_- ne permettait pas de trouver aisément une solution non perturbative aux équations de l'électrodynamique quantique. Alors Dirac eut l'idée d'utiliser une

troisième représentation diagonalisant la partie annihilatrice M de A ($A=M+\check{M}$) ; comme $[M(x'), M(x'')] = 0$, $M(x)$ peut être effectivement diagonalisé simultanément en tout point. Dans cette représentation les équations de l'électrodynamique se simplifient beaucoup et l'existence de solutions (non perturbatives) devient évidente. Il reste à repasser dans la représentation N_a, N_b . Dirac échoua pour cette dernière étape : les calculs étaient trop complexes. La conclusion de Dirac dans la leçon de Dublin de 1946 ⁽³⁴⁾ fut plutôt pessimiste : la première méthode d'interprétation du "monde hypothétique souffrait de trop d'arbitraire, et la dernière méthode se heurtait à trop de difficultés mathématiques. A la conférence de Cambridge de 1946, Dirac ⁽³⁰⁾ soutint que les difficultés de l'électrodynamique, aussi grandes soient-elles, étaient purement mathématiques, que les solutions (non perturbatives) devaient exister, qu'elles étaient seulement trop difficiles à calculer "

7.3f. Applications de la λ -limite et de la métrique indéfinie

En électrodynamique pure les conséquences de la nouvelle théorie de Dirac furent explorées par l'élève indien de Dirac, Jayaratnam ELIEZER. Ses travaux portèrent tant sur la théorie classique de Dirac (instabilité d'un système hydrogénoïde classique, modification de l'équation de Dirac-Lorentz en vue de la suppression des solutions parasites), que sur la nouvelle méthode de quantification de 1942. Dans ce cadre-ci, il vérifia la finitude des ordres supérieurs à 2 de la série des perturbations pour un électron libre, et aussi pour un électron lié, en présence du champ de rayonnement.

Il trouva un déplacement nul des niveaux atomiques. Ce calcul (1946) ⁽⁷²⁾ précédait de peu l'expérience de Lamb ; un peu plus tard il aurait constitué une difficulté sérieuse de la nouvelle théorie

* L'argument utilisé au § précédent pour montrer l'élimination des divergences d'ordre pair (ou impair) n'est rigoureuse que pour des intégrands rationnels ; il était insuffisant pour l'électron relativiste.

de Dirac. Ces travaux d'Eliezer sont condensés dans un article de 1947 de Rev. Mod. Phys. (33).

La λ -limite, en tant que méthode covariante d'élimination de certaines divergences devait pouvoir s'appliquer aux théories mésiques. Nous avons déjà vu quel profit Bhabha tira d'une généralisation de la théorie de Dirac de 1938 au champ vectoriel massif. L'intérêt de Pauli pour cette méthode de Dirac et de Wentzel fut sans doute en grande partie suscité par la gravité du problème des divergences en physique nucléaire (voir chap. 6 § 3). En 1943 Pauli demanda à JAUCH de calculer avec la λ -limite le moment magnétique anormal des nucléons. Les photons d'énergie négative n'intervenaient pas explicitement dans les calculs, mais ils servaient à justifier la finitude des ordres supérieurs. Le calcul de Jauch (34) effectué dans les théories mésiques pseudoscalaires et vectorielles donnait une valeur fixée pour les moments magnétiques du proton et du neutron : la λ -limite suffisait à éliminer la divergence du type $\int_0^{\infty} \frac{k^4 dk}{(k^2 + \mu^2)^2}$ en raison de la parité de l'intégrand*. Malheureusement la valeur trouvée pour les neutrons était positive contrairement à la mesure d'Alvarez-Bloch.

Dans la même année PAULI (35) montra le lien entre la λ -limite et les sources étendues utilisées dans les théories de couplage fort (voir chap. 6 § 3) : Une redéfinition du champ mésique permettait de faire passer la fonction d'extension de la source, du terme d'interaction spin-champ mésique au commutateur du champ mésique**. La λ -limite correspondait à un cas particulier de théorie de source étendue d'extension infinie (alors que la théorie usuelle divergente correspond au cas où l'extension de la source tend vers zéro). Dans la théorie classique, la λ -limite devait donner, comme toute théorie de source étendue en couplage fort, la prédiction d'isobares. Mais la nouvelle condition d'existence d'isobares (sur la masse des mésons et la constante de couplage)

* Une forme rationnelle aussi simple de l'intégrand était obtenue grâce à l'hypothèse d'une inertie infiniment grande des nucléons (les seuls degrés de liberté résiduels étant ceux de spin), par conséquent non relativistes.

** Plus précisément une interaction d'un champ scalaire ϕ avec un spin \mathcal{F} d'extension $U(x)$ suivant $i \int U(\vec{x}) \vec{\sigma} \cdot \vec{\nabla} \phi d^3x$ avec $\int U(\vec{x}) d^3x = 1$ est équivalente à une modification $[a_{\vec{k}}, a_{\vec{k}'}] = G(\vec{k}) \delta(\vec{k} - \vec{k}')$ des relations de commutation du champ mésique aux $G(\vec{k}) = v(\vec{k}) v(-\vec{k})$, $v(\vec{k}) = \int U(\vec{x}) e^{-i\vec{k} \cdot \vec{x}} d^3x$.

n'était pas vérifiée. Dix ans avant la découverte expérimentale d'isobares, ce n'était pas très grave. La λ -limite, par rapport aux théories habituelles de source étendue présentait deux avantages : elle était covariante relativiste, elle n'introduisait pas de paramètre arbitraire associé à la taille du nucléon. Appliquée à la théorie métrique mixte pseudo scalaire-vectorielle de Møller-Rosenfeld (voir chap 6 § 3), elle donnait des résultats finis au 2ème ordre et la finitude des ordres supérieurs devait être assurée par la métrique indéfinie. Il restait la grave difficulté du moment magnétique du neutron calculé par Jauch.

Un autre obstacle à la nouvelle théorie de Dirac vint avec l'étude des divergences infra-rouges dans cette théorie par JAUCH et PAULI (37,39) en 1944 puis par ELIEZER (40) en 1947 dans d'autres processus que la "Bremsstrahlung". Avec l'interprétation physique de Dirac de sa théorie où la notion de photons liés à l'électron envisagés par Bloch et Nordsieck (voir chap. 5 § 2) prenait difficilement place, Pauli montra que la probabilité d'émettre pendant la diffusion un nombre fini de photons était finie alors que celle d'émettre un nombre arbitraire de photons infra-rouges, correspondant au cas physique, divergeait.

La direction de recherche prise par Dirac en 1938 avait pu au début donner un nouvel espoir de chasser les divergences de la théorie des champs, mais de nouvelles difficultés d'interprétation ou d'exploitations'accumulèrent jusqu'à ce qu'en 1947 le principe de renormalisation donne un nouveau point de départ. De la métrique indéfinie inventée par Dirac reste aujourd'hui la quantification de Bleuler et Gupta du champ électromagnétique , donnant la première formulation rigoureuse de la méthode de Fermi (voir chap. 2 § 3).

7.4. L'AMORTISSEMENT PAR RAYONNEMENT DE HEITLER :

HEITLER fut le plus actif à promouvoir une théorie du "Radiation damping" dépourvue de divergences, à tous les ordres de la série des perturbations, et il lui a laissé son nom, bien que les contributions du physicien polonais GÓRA et de A.H. WILSON à l'élaboration de cette théorie soient tout aussi importantes. Outre Manche Heitler et Wilson avaient été impressionnés par les travaux de Bhabha (64) (voir § 3) de 1939 montrant que dans une théorie non quantique des mésons neutres généralisant la théorie électromagnétique de Dirac de 1938, l'effet de l'amortissement par rayonnement (analogue du terme de freinage de l'équation de Dirac-Lorentz) diminuait de beaucoup l'interaction méson-nucléon à haute énergie.

L'absorption de la composante dure des rayons cosmiques se révéla anormalement faible dans l'hypothèse alors retenue que les mésons constituaient cette composante. (voir chap. 6 § 3). Les effets d'amortissement, pensèrent WILSON et HEITLER en 1941 (indépendamment l'un de l'autre et simultanément) * pouvaient expliquer cette anomalie. Ainsi se développa un procédé systématique pour retenir dans la série des perturbations des termes correspondant à des effets d'amortissement et annuler les termes divergents. Heitler était bien conscient de l'arbitraire de ce procédé : *"La méthode ci-dessous décrite sera fondée sur l'omission des termes divergents. Jusqu'à maintenant aucune raison profonde ne peut être donnée pour justifier la petitesse de ces termes. Par conséquent, cette théorie doit être considérée comme heuristique, mais en raison de l'accord excellent avec l'expérience il est probable que la future électrodynamique quantique se réduira essentiellement à notre théorie"* (44) Ce dernier commentaire était trop optimiste. Il est aujourd'hui certain que les termes divergents annulés par Heitler contribuent après renormalisation souvent autant que les autres termes d'amortissement. L'accord avec l'expérience était certes bon, mais les expériences de diffusion méson-nucléon étaient bien trop rares et imprécises pour suffire à justifier le procédé de Heitler,

* Sokolov (53) développa aussi en URSS le même genre d'arguments.

Tout comme Dirac, Heitler parvenait à une théorie quantique sans divergence. Le procédé d'élimination des divergences était plus artificiel chez Heitler, mais chez Dirac l'interprétation de la théorie via "le monde hypothétique" était artificielle. Heitler doutait que la future théorie pût prendre la théorie classique comme base, "car l'électrodynamique quantique est essentiellement un problème de trois constantes universelles La petite valeur $e^2/4\pi\hbar c = 1/137$ suggère qu'aucune signification physique ne peut être attribuée à la limite $\hbar \rightarrow 0$ (c, e finis). Le point de départ doit être [...] $\epsilon \rightarrow \infty$ ou $e \rightarrow 0$ " (44). A défaut d'une théorie de base plus solide, Heitler n'était pas trop mécontent de sa théorie "heuristique".

7.4.a Antécédents :

Dans leurs articles sur la largeur des niveaux atomiques, WEISSKOPF et WIGNER (43) avaient pu extraire un résultat fini pour un problème d'amortissement grâce à l'approximation suivante : ils ne considéraient que des états intermédiaires virtuels d'énergie approximativement égale à celle de l'état initial *, ce qui permettait un "Ansatz" exponentiel (en $e^{-\frac{\Gamma}{\hbar} t}$) avec Γ réel, pour la population du niveau excité ; la partie imaginaire de Γ , correspondant au déplacement normalement infini d'après Oppenheimer, devenait zéro. WALLER (45) en 1934 fut le premier à généraliser la méthode de Wigner et Weisskopf à d'autres processus comme les diffusions photon-électron ou électron-électron. Là encore l'"Ansatz" exponentiel et la sélection des états intermédiaires permettait le calcul d'effets d'amortissement finis (de petites corrections dans le cas présent). Mais le calcul de Waller manquait de généralité et l'effet des conditions initiales du problème de diffusion n'était pas traduit

* Ainsi le déplacement au 2ème ordre $\delta E_n = P.P. \langle \psi_n | V \frac{1}{E_n - H_0} V | \psi_n \rangle$ du niveau n du hamiltonien non perturbé H_0 , par la perturbation V, normalement divergent, devenait zéro car la dite approximation revenait à ne considérer que les dénominateurs d'énergie voisins de zéro, lesquels ne contribuent pas à l'intégrale sur les états intermédiaires à cause de la partie principale.

correctement. Le physicien polonais GÓRA^(46,47) écarta cette difficulté grâce à un branchement adiabatique de la perturbation, mais il aboutit à un résultat faux (par rapport à la théorie de Heitler) pour le problème de la diffusion électron-photon à défaut d'avoir tenu compte des états intermédiaires contenant un photon virtuel (et pas seulement les photons réels incidents ou émergents).

7.4.b Le procédé de Heitler et Wilson

Seul Wilson avait eu connaissance de l'article de Waller, mais Heitler et Wilson connaissaient tous deux celui de Góra et pensèrent qu'il résolvait "en principe" le problème quantique général de l'amortissement par rayonnement. Les articles de Wilson⁽⁴⁸⁾ et de Heitler⁽⁴⁹⁾, publiés simultanément après confrontation en 1941, se ressemblent beaucoup. Les deux portaient de l'équation de Schrödinger projetée sur les états propres du hamiltonien libre en représentation d'interaction pour aboutir à une équation intégrale (appelée équation intégrale de Heitler) portant sur les coefficients des exponentielles dans un "Ansatz" à la Wigner-Weisskopf. Les divergences étaient éliminées en annulant l'analogue de la partie imaginaire de Γ . Heitler sélectionnait la solution satisfaisant aux conditions initiales du problème de diffusion en utilisant un branchement adiabatique de l'interaction comme Góra, et Wilson préférait considérer pour cela des états électroniques faiblement liés à un centre coulombien, faisant tendre in fine cette liaison vers zéro. Il fut rassurant de constater que les corrections ainsi apportées à la diffusion de Born étaient suffisamment faibles pour que sa vérification dans les expériences sur les électrons et les γ cosmiques ne fut pas remise en question. L'intérêt de la nouvelle théorie de Heitler et Wilson résidait plutôt dans l'application à la théorie mésique. Wilson montra que les effets d'amortissement électromagnétique dans la diffusion méson-nucléon étaient importants comme on l'espérait, et Heitler put aussi montrer un affaiblissement notable de la probabilité de cette diffusion sous l'effet de l'amortissement dû aux mésons virtuels.

En 1942 HEITLER et PENG⁽⁵⁰⁾ donnèrent une prescription plus précise et plus générale pour annuler les divergences de la théorie quantique des champs ainsi qu'une preuve de l'invariance relativiste de leur méthode de soustraction. Il fallait retenir seulement le premier terme

non nul (non divergent) du noyau de l'équation intégrale de Heitler .
 Pour préciser , nous utiliserons une formulation équivalente plus
 moderne due à GÓRA en termes d'opérateurs de diffusion ⁽⁵¹⁾ que Pauli ⁽⁵²⁾
 retint dans les exposés qu'il fit de la théorie de Heitler * . L'origine de
 la théorie des états stationnaires de diffusion, se trouve, rappelons-le ,
 dans un article de DIRAC de 1927 et elle fut reprise pour le problème
 de l'absorption des neutrons lents par BREIT et WIGNER en 1936. L'équa-
 tion intégrale de Heitler concerne la matrice T définie par $S = 1 - 2\pi iT$,
 et elle s'écrit :

$$T = K(1 - i\pi T)$$

avec : $\langle b | K | a \rangle = \delta(E_b - E_a) \langle b | (1 - VP_0)^{-1} V | a \rangle$

et
$$P_0 = P.P. \frac{1}{E_a - H_0}$$

Cette équation résulte formellement de l'équation de Schrödinger
 (voir démonstration en appendice). La prescription de Heitler consiste
 à ne retenir dans le développement de $\frac{1}{1 - VP_0} V = V + VP_0 + VP_0VP_0V + \dots$
 que les premiers termes non divergents. Le développement peut être considéré
 comme un développement en puissances du scalaire α ; il en résulte que
 tous ses termes se transforment de la même manière sous l'effet d'une
 transformation de Lorentz et que la sélection de certains termes ne
 nuit pas à la covariance. L'expression de S (ou de T) en fonction de
 $K : S = \frac{1 - i\pi K}{1 + i\pi K}$ (à comprendre comme développement en puissances de K) ne
 contient pas de divergences si K tronqué n'en contient pas. Toutefois
 le S tronqué contient des processus virtuels à tous les ordres représen-
 tant un phénomène d'amortissement puisque le 2ème terme de $T = K - i\pi K^2 + \dots$
 viole la règle du bilan détaillé $\langle b | T | a \rangle = \langle a | T | b \rangle$. Remarquons de plus

* Indépendamment MA et HSÜEH (Pékin) ⁽⁵⁷⁾ utilisèrent les états stationnaires
 de diffusion pour le même problème en 1944.

(cette considération est absente des premiers articles de Heitler) que la matrice S tronquée est encore unitaire . En effet S unitaire \iff
 $K = K^+$ or tous les termes du développement de K sont hermitiques.

7.4.c Critiques de la théorie de Heitler

Les critiques les plus virulentes ne furent publiées qu'après guerre. En 1946 BETHE ET OPPENHEIMER ⁽⁵⁶⁾ dans leur étude de "la réaction du rayonnement sur la diffusion des électrons dans la théorie de Heitler". frappèrent un coup fatal sur la théorie de Heitler. Nous avons déjà cité cet article pour la clarification qu'il apporta du problème infra-rouge (voir chap. 5 § 2) . Mais son but principal était de montrer que la théorie de Heitler, contrairement à l'électrodynamique quantique usuelle, donnait lieu à des difficultés infra-rouges insurmontables. En bref la compensation des divergences infra-rouges entre émission - absorption de photons mous virtuels et émission de photons mous réels ne pouvait plus s'opérer puisque la théorie de Heitler tronquait la contribution des photons virtuels mous sans toucher à celle des photons réels.

Pauli, au congrès de Cambridge de 1946 ⁽³⁰⁾ ne fut pas plus tendre pour la théorie de Heitler que pour celle de Dirac. Outre les difficultés infra-rouges, la théorie de Heitler donnait des résultats non conformes à l'expérience pour l'interaction de plusieurs électrons et pour le moment magnétique des nucléons. Après la théorie de la renormalisation elle devait être abandonnée par Heitler lui-même : La troisième édition (1953) de la théorie quantique du rayonnement ne mentionne même plus le procédé de soustraction de Heitler. Au congrès Solvay de 1948 ⁽⁵⁸⁾ Heitler reconnut que seuls les termes primitivement divergents annulés dans sa théorie pouvaient rendre compte d'effets comme le déplacement de Lamb et le moment magnétique des nucléons.

7.5 LES THEORIES DE MATRICE S

La matrice S , limite de l'opérateur unitaire d'évolution sur une durée infinie, est apparue en théorie quantique relativiste sous deux aspects : d'abord avec les travaux de Heisenberg et de Stueckelberg comme une méthode alternative à la théorie quantique des champs en crise, puis comme un objet privilégié d'étude calculable à partir de champs quantiques avec Dyson. Les deux conceptions rivales -objet primitif, ou objet déductible d'une théorie des champs- se prolongent jusqu'à nos jours.

Une matrice unitaire de diffusion fut introduite pour la première fois par WHEELER ⁽⁶⁶⁾ en 1937 dans un problème particulier de physique nucléaire. Wheeler cherchait par une méthode variationnelle ("resonating group structure method") les meilleures fonctions d'ondes nucléaires (d'états liés ou de diffusion) décomposables en produit de fonctions d'onde de groupes stables (comme les particules α etc.). A ce stade le rôle de la matrice de diffusion était purement mathématique, et il nous semble juste d'attribuer à Heisenberg la prise de conscience de l'importance d'une telle matrice en physique des particules.

7.5a La matrice S de Heisenberg

HEISENBERG partit de présupposés philosophiques concernant la mesurabilité et l'observabilité des grandeurs physiques. C'est pourquoi nous avons discuté les deux premiers articles de Heisenberg (67,68) sur la matrice S au chap. 4. Ces travaux de 1942 constituaient une proposition positive de Heisenberg pour échapper à l'impossibilité d'une description spatio-temporelle détaillée qu'interdisait la "longueur universelle". Les seules grandeurs accessibles à l'expérience devaient être les états asymptotiques d'un système dans une expérience de diffusion, ou encore les énergies propres d'un système lié. L'opérateur S concentrait l'étude sur le premier type d'observation.

Les "Ansätze" :

La difficulté essentielle de cette nouvelle approche était le nombre insuffisant de contraintes imposées à la matrice S . L'unitarité et l'invariance relativiste ne pouvaient suffire à spécifier cette matrice. Il fallait faire un "Ansatz" sur la forme de S et le comparer

à l'expérience. Les "Ansätze" de Heisenberg ⁽⁶⁸⁾ supposaient une forme simple pour la matrice hermitique η (matrice des déphasages dans le problème particulier de diffusion d'une particule par un potentiel) telle que $S = e^{i\eta}$. Un choix très simple de η serait par exemple en notation moderne $\eta = g : \int \varphi^\dagger d^4x : ,$ la permutation des créateurs et des annihilateurs éliminant toute divergence. Dans d'autres travaux datant de 1944-1945, mais publiés bien plus tard ⁽⁶⁹⁾ à cause de la guerre (toutefois des tirés à part circulèrent dans quelques grands centres étrangers) Heisenberg examina la possibilité de correspondance de la matrice S avec une théorie classique ondulatoire décrite par une équation intégrodifférentielle non linéaire du type $\int d\alpha' A(\alpha, \alpha') \varphi(\alpha') + \int d\alpha' d\alpha'' B(\alpha, \alpha', \alpha'') \varphi(\alpha') \varphi(\alpha'') + \dots = 0 ;$ la non linéarité était empruntée aux théories de Born et Infeld ou à l'électrodynamique du vide polarisable, et le caractère intégrodifférentiel aux théories de Stueckelberg et de Bopp. La structure hamiltonienne, conduisant aux divergences habituelles, était abandonnée, mais la quantification restait possible en postulant des relations de commutations entre l'impulsion-énergie P_μ définie à partir d'un tenseur d'énergie impulsion $T_{\mu\nu}$ conservé : $[P_\mu, \varphi(x)] = i\pi \partial_\mu \varphi^*$. Heisenberg ne put déterminer si ce nouveau formalisme pouvait être développé sans contradiction en dehors des cas dégénérés.

L' "Ansatz" correspondant à la théorie de Heitler de l'amortissement fut présenté comme tel par HEITLER lui-même à la conférence de Cambridge de 1946 ⁽³⁰⁾ mais il ne semble pas que Heisenberg lui ait accordé beaucoup d'attention. Les travaux de Stueckelberg, surtout, apportèrent des restrictions plus fortes aux choix de S , comme nous le verrons bientôt.

* Podolsky et Kikuchi ⁽⁷⁰⁾ développèrent une quantification analogue dans un problème faisant intervenir une fonction de Lagrange avec des dérivées temporelles secondes (voir § 1).

Etats liés et prolongement analytique

Le second type d'observables reconnu par Heisenberg : les énergies discrètes des états liés (et leurs durées de vie) semblaient exclus d'une théorie de la matrice S . La difficulté fut levée par Kramers, comme Heisenberg le nota ⁽⁶⁹⁾ : "H.A. Kramers a trouvé le lien entre les états stationnaires discrets et la matrice S et me l'a exposé dans une discussion amicale à Leyde en 1942". L'argument de Kramers généralisé plus tard par Heisenberg était le suivant dans le cas du problème de diffusion d'une onde S par un potentiel central : à de grandes distances r du centre de diffusion, l'onde incidente a la forme $\frac{e^{-ikr}}{r}$ et l'onde émergente $\frac{e^{ikr}}{r}$. Les valeurs propres $S(k)$ de S (déphasages) déterminent le coefficient de l'onde émergente. Dans les problèmes simples la fonction d'onde prolongée analytiquement en k est encore solution de l'équation de Schrödinger. A un k imaginaire pur $k = -i\alpha$ ($\alpha > 0$) correspond un état lié si $S(-i\alpha) = 0$ et les énergies discrètes correspondantes seront $-\frac{\alpha^2}{2m} = \frac{k^2}{2m}$ (particule de masse m non relativiste).

Cette remarque de Kramers montra pour la première fois l'importance de considérations d'analyticité dans les problèmes de matrice S . L'argument dut être raffiné par la suite car la condition $S(-i\alpha) = 0$ n'était pas suffisante comme le démontra MA en 1946 ⁽⁷¹⁾. TER HAAR ⁽⁷²⁾ montra que les zéros redondants servaient à compenser les pôles (en k) des coefficients du développement asymptotique en puissances de $\frac{1}{r}$ de la fonction d'onde, et MØLLER ⁽⁷²⁾ donna (toujours en 1946) un moyen d'écarter les zéros redondants. Ce travail de Møller ne fut pas sa seule contribution à la théorie de la matrice S ; Møller simplifia considérablement la preuve d'unitarité et donna une preuve de l'invariance relativiste à partir des propriétés de covariance des fonctions d'onde asymptotiques (voir (72) et conférence de Møller à Cambridge ⁽³⁰⁾ en 1946).

7.5b Stueckelberg et la matrice S :

Ernest C.G. STUECKELBERG de Briedenbach a apporté une importante contribution à la théorie quantique relativiste et il a plusieurs fois anticipé les développements modernes de la théorie quantique des champs. Rappelons son introduction ⁽⁷⁴⁾ en 1938, bien

avant Tomonaga d'un temps local et de la représentation d'interaction pour une formulation explicitement covariante de l'électrodynamique quantique. Nous allons voir la nouveauté de ses travaux effectués à partir de 1941. Les difficultés de la guerre et la publication en français des articles dans la revue "Helvetica Physica Acta" ont malheureusement limité la diffusion des idées de Stueckelberg aux frontières de la Suisse, encore que le groupe de Zürich comprenait des savants aussi prestigieux que Pauli, Wentzel, et que Heisenberg lui rendit visite.

Le chemin de Stueckelberg vers la matrice S fut très différent de celui de Heisenberg. Il commença par l'étude de la théorie classique de l'électrodynamique.

Un nouveau modèle de l'électron ponctuel : (1941)

La théorie de Lorentz-Abraham de l'électron n'apportait à la diffusion Thomson que des corrections négligeables devant les corrections quantiques ; et elle n'avait guère besoin d'être améliorée au niveau classique. Cependant Stueckelberg ⁽⁷⁵⁾ pensait, comme Dirac, qu'un approfondissement de la théorie classique permettrait de mieux comprendre les difficultés au niveau quantique et il proposa une théorie mixte où des charges ponctuelles se trouvaient simultanément couplées au champ électromagnétique et à un champ scalaire neutre ; à condition de prendre la même constante de couplage et une masse $M = \frac{2m}{\alpha}$ pour les scalaires, toute divergence disparaissait de la théorie classique. L'équation du mouvement des électrons était analogue à celle de Lorentz, le paramètre r_0 de développement des corrections à la force de Lorentz était remplacé par la longueur de Yukawa $\frac{h}{MC}$. Stueckelberg envisagea aussi des théories mixtes plus générales. Celle obtenue en ajoutant au champ électromagnétique un champ scalaire et un champ de Proca s'identifiait à la théorie de Dirac-Pryce de 1938 en faisant tendre simultanément les masses du scalaire et du vecteur vers zéro. *

* La théorie de Bopp de 1940, quant à elle, ne rentrait pas dans cette catégorie puisque Bopp soustrayait les tenseurs d'impulsion énergie d'un champ vectoriel non massif et d'un champ de Proca, au lieu de les ajouter comme Stueckelberg (pour que l'énergie soit toujours positive).

Plus tard, en 1944, Stueckelberg ⁽⁷⁶⁾ publia une généralisation de cette nouvelle théorie classique. Abandonnant l'idée de la mixité il déduisait par des arguments très généraux de conservation et de covariance une équation du mouvement de l'électron sans divergence. Il y avait une infinité de manières différentes de soustraire la partie $\frac{1}{2}(A_{ret} + A_{adv})$ divergente (et indépendante du champ incident) du champ appliqué à l'électron et la forme générale de l'équation du mouvement devait être $m \eta \left(\frac{d^2}{dt^2} \right) \vec{z} - m \lambda_0 \frac{d^3 \vec{z}}{dt^3} = \vec{f}^{in}(t)$ pour une charge de masse m soumise à la force de Lorentz du champ incident $\vec{f}^{in}(t) = e \vec{E}_{in} + \frac{d\vec{z}}{dt} \wedge \vec{B}_{in}$, et avec $\lambda_0 = \frac{e^2}{6\pi c^3}$, $\eta = 1$ redonne l'équation de Lorentz-Dirac. Selon Stueckelberg tout l'arbitraire de la théorie classique débarassée de ses divergences était contenu dans la fonction η ou encore dans la fonction g de "prémonition" obtenue en inversant l'opérateur différentiel $\eta \left(\frac{d^2}{dt^2} \right) - \lambda_0 \frac{d}{dt}$ et donnant une action de \vec{f}^{in} sur la charge avant que le champ incident ne l'atteigne : $m \frac{d^2}{dt^2} \vec{z} = \int g(t') \vec{f}^{in}(t-t') dt$.

Prémonition et matrice S :

Ce genre de prémonition avait déjà été remarqué par Dirac dans sa théorie de 1938 pour des temps de l'ordre de $\frac{r_0}{c}$. Au lieu de se ramener comme Dirac à une mécanique hamiltonienne grâce à la \hbar -limite, Stueckelberg ⁽⁷⁶⁾ utilisa l'effet de prémonition pour justifier l'introduction de la matrice S. Dans le problème classique comme dans le problème quantique les conditions initiales devaient être données à un temps $-T$ précédant l'arrivée du champ incident pour que le mouvement ultérieur puisse être calculé sans ambiguïté. La matrice S de Stueckelberg transformait le vecteur d'état $\psi(-T)$ du système à $-T$ en $\psi(t)$ à t . La correspondance entre les théories classique et quantique, à défaut d'un principe hamiltonien, était obtenue en identifiant l'évolution $S^{-1}AS$ d'une observable A à l'évolution résultant des équations classiques. Comme d'habitude, des divergences apparaissaient ; Stueckelberg (comme Heisenberg) ^{*}

* Et comme Jordan et Klein en 1927 (voir chap. 2 § 1) dans un contexte (nombre de particules conservé) où le procédé ne posait aucun problème.

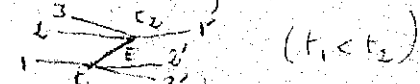
donna des prescriptions de permutation des créateurs et des annihilateurs pour pallier à ces divergences. A chaque fonction g de prémonition correspondait une prescription bien définie. Par exemple la théorie de Góra-Heitler s'obtenait en prenant $\eta = 1$ (équation de Dirac-Lorentz).

La mécanique asymptotique

Ce procédé contenait trop d'arbitraire. Dans un article ultérieur (77) Stueckelberg développa un "double principe de correspondance" donnant un rôle privilégié aux évolutions asymptotiques quasi périodiques. Dans les règles de correspondance devaient intervenir deux types de limite $h \rightarrow 0$ et aussi $\tau_0 \rightarrow 0$, τ_0 étant le temps caractéristique de prémonition. La limite $\tau_0 \rightarrow 0$ permettait de passer d'une mécanique "fonctionnelle" non locale dans le temps à la mécanique hamiltonienne traditionnelle. Wentzel (79) douta que ces règles suffisent à déterminer complètement les règles de soustraction.

Causalité et fonction de Green causale :

Dans une brève communication (78) (1946) Stueckelberg restreint le choix possible de matrices S par une nouvelle exigence de causalité. Cette exigence était satisfaite dans la dernière version de sa théorie, mais pas dans les "Ansätze" de Heisenberg du type $S = \exp - g : \int \phi^4 d^4x :$

Une telle matrice S autorise des processus du type  (t₁ < t₂)

la différence $t_1 - t_2$ étant macroscopique et les traces de 1,2',3' et de 1',2,3 permettant d'inférer au niveau macroscopique des lieux différents pour les deux vertex. La théorie interdit de décrire en détail ce qui se passe pendant l'interaction et donc entre les instants t_1 et t_2 .

Pourtant le principe de conservation de l'énergie appliqué au niveau macroscopique permettrait dans certains cas d'affirmer l'existence d'une particule d'énergie $E < 0$ dans l'état intermédiaire, ce qui est absurde.

Cette exigence de causalité (rejet des particules intermédiaires réelles d'énergie négative), contenait explicitement l'usage de la fonction de Green causale (appelée aujourd'hui propagateur de Feynman) dans la construction de la matrice S : une inhomogénéité $\delta^4(x)$ ne pouvait que créer une particule d'énergie toujours positive dans le futur ou en absorber une dans le passé.

La fonction D^C de Stueckelberg fut introduite dans une lettre à Phys. Rev. du 2 juin 1949 annonçant un calcul, avec Rivier, du moment magnétique du neutron. D_C est la fonction de Green invariante impaire de $\square+m^2$: $(\square+m^2)D_C = \delta^4$ et $D^C(-x) = D^C(x)$, dont la transformée de Fourier s'écrit : $\frac{1}{k^2-m^2+i\epsilon}$. RIVIER (81) (1949) explicita et commenta les calculs. La fonction D^C pouvait être introduite soit axiomatiquement comme étant la seule fonction de Green invariante et causale au sens de Stueckelberg, ou encore en appliquant la théorie des perturbations covariantes (dans un formalisme hamiltonien).

Régularisation.

La lettre du 2 juin 1948 contenait une autre nouveauté. Pour être rendue finie, la série des perturbations devait être modifiée en utilisant au lieu de la fonction D_C , la fonction obtenue en remplaçant dans l'expression de D_C en transformée de Fourier

$$\frac{1}{k^2-m^2+i\epsilon} \quad , \quad PP \frac{1}{k^2-m^2} \quad \text{par} \quad PP \frac{1}{k^2-m^2} + \sum_{i=1}^N c_i \frac{1}{k^2-m_i^2}$$

Dans le problème particulier du moment magnétique du neutron au 2ème ordre, la limite $m_i \rightarrow \infty$ donnait un résultat fini. Il est clair aujourd'hui que dans le cas général cette limite, sans renormalisation, conduit à des résultats divergents. L'introduction des masses auxiliaires pour rendre finis les termes de la série des perturbations sans affecter leur covariance relativiste est connue aujourd'hui sous le nom de régularisation de Pauli-Villars.

Les hommages de Dyson (82) et de Pauli et Villars (83) à Stueckelberg montrent bien l'influence de son oeuvre dans la naissance de l'électrodynamique quantique moderne. Le chemin suivi à partir d'une électrodynamique classique non divergente et non locale dans le temps peut paraître tortueux. Il conduit pourtant à une construction de la matrice S en terme de fonction de Green causale, anticipant sous une forme moins lisible et moins systématique les travaux de Feynmann et de Dyson.

APPENDICE

=====

Démonstration moderne de l'équation intégrale de Heitler.

Définissons R par $S = 1 - 2\pi iT$ et $\langle b|T|a \rangle = \delta(E_b - E_a) R_{ba}$. Il est bien connu que R satisfait à l'équation intégrale de la diffusion : $R = V + V G_0 R$

avec $G_0 = \frac{1}{E_a - H_0 + i\epsilon} = P_0 - i\pi \delta(E_a - H_0)$ $P_0 = P.P. \frac{1}{E_a - H_0}$

D'où $R = V + V P_0 R - i\pi V T \Rightarrow (1 - V P_0) R = V(1 - i\pi T)$

$\Rightarrow T = K(1 - i\pi T)$ avec $\langle b|K|a \rangle = \delta(E_b - E_a) \langle b|(1 - V P_0)^{-1} V|a \rangle$

Bibliographie

- Report of Cambridge Conference Vol. I "Fundamental particles"
Taylor and Francis (1947)
Exposés de Bohr, Pauli, Dirac, Heitler, Møller et Stueckelberg
- Rapports du 8° Conseil Solvay (1948) - Stoops 1950
Exposé de R. Peierls et de R. Oppenheimer
- Wentzel, Rev. Mod. Phys. 19, 1 (1947)
- Sakata, Prog. Th. Phys. 2, 145 (1947)
- Kramers, Ned T. Natuuk 11 (1944) 134, traduit dans Coll. Scient. Papers
(conf. d'Utrecht) N.H. Ed.
- Schweber, Bethe, de Hoffmann, Mesons and Fields I § 20b - Peterson 1955

CHAPITRE 8

NAISSANCE DE L'IDEE DE RENORMALISATION

8.1 VERS UN PRINCIPE DE RENORMALISATION :

La notion de renormalisation remonte à l'étude du mouvement des corps dans les fluides. Dans un fluide même parfait, l'inertie d'un corps en mouvement est modifiée par "l'entraînement" d'une partie du fluide environnant. Par exemple l'équation du mouvement d'une sphère (de volume V , de masse m_0 soumise à une force f dans un fluide parfait incompressible et irrotationnel de densité ρ s'écrit $\vec{f} = (m_0 + \frac{\rho V}{2}) \vec{a}$. Autrement dit des expériences de mécanique réalisées dans le fluide feront intervenir la masse renormalisée $m = m_0 + \frac{\rho V}{2}$ au lieu de la masse nue m_0 . Les considérations de STOKES (1) datant de 1842 guidèrent probablement J. J. THOMSON quand il introduisit en 1881 (voir chap. 3 § 1) la notion de masse électromagnétique d'une sphère chargée en mouvement lent : les analogies mécaniques fleurissaient alors en Angleterre. Plus tard LORENTZ attribua à l'électron une masse d'origine purement électromagnétique. La renormalisation devenait responsable de toute son inertie. Dès le XIX^e siècle l'électrodynamique dans les milieux diélectriques apporta la notion de renormalisation de charge. Une charge plongée dans un milieu diélectrique polarise le milieu dans son voisinage et ce phénomène diminue la charge effective mesurée à l'aide de charges d'épreuve situées en dehors de ce voisinage. Ce phénomène de renormalisation de charge fini et calculable ne posait aucun problème conceptuel, contrairement au cas de la masse électromagnétique sensible à la structure de l'objet chargé.

Quand on parle de principe de renormalisation en théorie quantique des champs on entend généralement deux idées différentes :

i) les valeurs mesurées de la masse et de la charge dans des expériences à basse énergie diffèrent des valeurs "nues" correspondant à un débranchement hypothétique de l'interaction électromagnétique.

ii) la substitution des valeurs observables de la masse et de la charge aux paramètres nus dans l'expression formelle des autres grandeurs mesurables absorbe toutes les divergences.

La première idée, nous venons de le voir, est assez ancienne, elle remonte à STOKES 1842, et il serait léger de croire que les grands théoriciens responsables des premiers développements de la théorie quantique des champs n'en étaient pas conscients : ils connaissaient trop bien l'électrodynamique/

classique. La seconde, malgré la nette anticipation de KRAMERS, ne s'imposait vraiment qu'en 1947. Les infinis de la théorie étaient vraiment considérés comme catastrophiques et on n'envisageait pas que la seule élimination de paramètres nus suffît à rendre la théorie finie.

Dans ce paragraphe nous chercherons à dégager des études précédentes (chap. 5 et 6) ce qui dans le premier combat contre certaines divergences de l'électrodynamique - la polarisation du vide et les divergences infra-rouges - se rapprocha du principe de renormalisation. Le principe se trouva énoncé sous ses deux aspects, dans les travaux de WEISSKOPF sur la polarisation du vide, mais il ne concernait alors que la renormalisation de charge. Seulement la première idée (i) contenue dans ce principe était présente dans les études du problème infra-rouge : la notion de charge habillée faisait son apparition et la masse électromagnétique sortait des calculs.

8.1.a Polarisation du vide et renormalisation de charge :

Au congrès de SOLVAY d'octobre 1933, DIRAC mit clairement en évidence le phénomène de polarisation du vide sous l'effet d'un champ extérieur et conclut que les charges usuellement observées devaient différer de celles introduites dans le hamiltonien de Dirac. L'effet d'écran prévu par Dirac n'était fini que grâce à une coupure convenable. Dirac était conscient de difficultés de calcul dans un cas plus général : on avait à soustraire de infinis correspondant à la mer perturbée et à la mer au repos. Dans une autre publication Dirac décrivit une manière générale d'opérer cette soustraction en identifiant les termes singuliers de la distribution perturbée. Cette identification, même une fois perfectionnée par HEISENBERG souffrait encore d'ambiguïtés que l'évocation des différentes invariances (relativiste, de jauge) ne suffisait pas à éliminer. L'annulation de la partie singulière divergente avait pour effet d'éliminer tout problème de polarisabilité uniforme. WEISSKOPF remarqua que c'était là son seul effet qu'il valait mieux inverser la situation : déclarer tout phénomène de polarisabilité uniforme inobservable et en déduire la partie singulière à annuler. Un tel phénomène réduisant toutes les charges observées d'un facteur constant rendait inobservable la charge nue introduite au départ des calculs. Les termes correspondant à ce phénomène étaient aisément reconnaissables

par exemple le lagrangien effectif décrivant les propriétés dispersives du vide avait la forme $\int \mathcal{L} = \int \mathcal{L}_{\text{ini}} + \infty \times (E_0^2 - B_0^2)$, le 2° terme divergent étant absorbable dans une renormalisation des champs \vec{E} et \vec{B} et corrélativement de la charge e^* . Le principe de renormalisation de charge, ainsi introduit par WEISSKOPF (1936) gagnait en simplicité et en vraisemblance par rapport aux prescriptions de Heisenberg-Dirac. Le problème des divergences de la polarisation du vide par un champ extérieur classique est relativement simple et les traités modernes n'y consacrent pas plus de quelques lignes, justement parce que la nature des divergences apparaît clairement dans le résultat final. La situation est moins simple quand on considère l'effet de la renormalisation de masse : la structure des termes divergents, par exemple dans les corrections radiatives à la diffusion, est plus complexe.

8.1.b Problème infra-rouge et champ propre :

Il résulta de la première étude de BLOCH et NORDSIECK que l'électron ne pouvait être complètement dissocié de son champ électromagnétique propre, contenu dans la définition des états stationnaires du système couplé électron + Rayonnement. Un électron diffusé par un potentiel extérieur émettait nécessairement une infinité de photons mous, sous l'effet de l'altération du champ propre. Les états stationnaires du système électron + rayonnement devaient, en raison même de leur stationnarité être pris comme les états physiques observables : on habillait ainsi les électrons d'un nuage de photons mous virtuels. Ce point de vue fut formalisé par PAULI et FIERZ par l'introduction d'une transformation canonique faisant passer des états nus aux états habillés. Une généralisation de cette transformation canonique, nous le verrons, joua un rôle important dans les théories de TOMONAGA et de SCHWINGER. Pauli et Fierz avaient déjà constaté qu'elle permettait d'isoler dans le hamiltonien transformé un terme de masse électromagnétique. Il manquait encore à Pauli et Fierz de deviner que ce terme de masse contenait à lui seul toutes les divergences et que le reste du

*Rappelons que le produit eA^μ est inchangé par la renormalisation.

hamiltonien donnait des résultats finis. Mais il était clair que la notion d'électron isolé, coupé de l'interaction électromagnétique, était dépourvue de sens physique. DANCOFF, puis BETHE et OPPENHEIMER parlèrent de "renormalisation" de la fonction d'onde de l'électron : un électron sous l'effet de l'interaction électromagnétique a une probabilité inférieure à un d'être lui-même et la probabilité complémentaire d'être accompagné de photons virtuels.

8.1.c Renormalisation en physique nucléaire :

Si les corrections "radiatives" en électrodynamique quantique ne sont que de petits effets (après renormalisation), il en va tout autrement en physique nucléaire. Dans la théorie mésique le couplage est trop fort pour qu'on puisse négliger les corrections apportées par l'émission-absorption de mésons virtuels et la nécessité de les évaluer était plus pressante. De nombreux procédés furent inventés (voir chap. 7) pour extraire des résultats finis comparables à l'expérience, mais rien qui ressemblât à un principe de renormalisation, qui d'ailleurs eût été caduc dans ce cas d'une théorie non renormalisable. Néanmoins les effets de renormalisation devenaient une réalité expérimentale : il fallait rendre compte par ex. du moment cinétique du proton et du neutron. Le problème de divergences, aussi académique qu'il fût dans l'électrodynamique d'avant-guerre, touchait des problèmes concrets en physique nucléaire. Divergences des théories mésiques et divergences de l'électrodynamique quantique posaient des problèmes parallèles ; il fallait, suivant Heisenberg ou les théoriciens de Nagoya, les résoudre simultanément ; à moins que la solution indépendante de l'un n'éclaire l'autre.

8.2 LES IDEES DE KRAMERS :

A la fin du chapitre sur les divergences ultra-violettes, nous avons cité plusieurs commentaires sur la nature de ces divergences. Aucun des physiciens cités ne fit aussi tôt que Kramers une analyse aussi juste et aussi proche de la solution future du problème. UHLENBECK⁽¹⁾ rapporte "*Je peux encore entendre Kramers se lamenter ... il ne pouvait vraiment*

pas comprendre pourquoi la théorie de Dirac du rayonnement était si bonne, bien qu'elle ne fût pas de distinction suffisante entre le champ propre de l'électron et le champ extérieur, bien qu'elle ne rendît pas compte de la masse électromagnétique infinie ; et bien que la correspondance avec la théorie classique de l'électron ne fût pas claire du tout. Il fut conscient plus tôt et plus profondément que la plupart de ses contemporains des imperfections de la théorie". Kramers dépasse le stade de la critique en élaborant sa propre théorie électromagnétique classique puis quantique ; la distinction entre champ propre et champ extérieur et l'idée de renormalisation de masse y jouent un rôle essentiel, dès 1937, dix ans avant la découverte du déplacement de Lamb. BETHE qui donna la première évaluation théorique du Lamb-Shift reconnut l'influence des idées de Kramers : "Le fait que l'énergie propre d'une particule libre ne peut conduire qu'à une modification de sa masse est la base de la solution moderne de la difficulté des divergences. La suggestion originale fut faite par KRAMERS en 1947* en relation avec l'électrodynamique classique. Il fit remarquer que nous ne pouvons jamais observer expérimentalement m_0 , la masse nue de l'électron, mais seulement $m_0 + \delta m$, la masse observée notée m . Retournant au point de vue original de Heisenberg** sur la mécanique quantique, il suggéra que seule la quantité observable m devrait jouer un rôle dans la théorie et non pas les quantités séparées m_0 et δm . Le principe est connu comme le "principe de renormalisation de masse" ... Stimulé par les idées de Kramers, BETHE (1947) interpréta le déplacement de Lamb comme un effet de l'interaction de l'électron avec le champ de radiation. En fait Bethe calcula simplement l'énergie propre d'un électron lié dans l'atome. Mais, suivant l'idée de Kramers, une grande partie de cette énergie propre est déjà prise en compte si l'on utilise la masse observée m de l'électron dans le calcul plutôt que la masse nue m_0 . Par conséquent, le vrai déplacement de niveau est la différence entre l'énergie

* Il serait plus exact, nous semble-t-il, de faire remonter cette suggestion à 1937. La date de 1947 réfère à l'application des idées de Kramers lui-même au calcul du déplacement de Lamb (déjà calculé par Bethe) par lui-même exposé au congrès Solvay de 1947 - (voir plus loin la lettre de Bethe).

** Cette paternité n'est pas exprimée par Kramers dans les articles où il introduit la renormalisation de masse.

propre d'un électron lié et celle d'un électron libre"⁽²⁾. Ce texte de Bethe (et Schweber et de Hoffmann) ainsi que la lettre de Bethe reproduite au § 5, suffisent à prouver l'importance historique du rôle de Kramers*.

Nous présentons ici la théorie de Kramers de l'électrodynamique telle qu'elle fut développée à partir de 1937 par Kramers et ses élèves, et nous montrerons en quoi elle diffère de l'approche moderne et pourquoi elle donne pour le déplacement de Lamb une valeur identique à celle donnée par Bethe en 1947.

8.2. a La théorie de Kramers de 1937 :

Dans la seconde partie d'une monographie consacrée à l'étude de la constitution de la matière, Kramers exposa la "théorie quantique de l'électron et du rayonnement"⁽³⁾. En électrodynamique classique il est bien connu que le mouvement d'une charge est déterminé par le champ extérieur appliqué à cette charge ; le champ propre de la charge (coulombien lorsqu'elle est au repos) diverge au lieu de la charge lorsque celle-ci est ponctuelle, et il ne doit pas intervenir dans la formule de Lorentz. Dans la théorie de DIRAC de 1927 toute considération de ce genre était absente. Kramers ne put se contenter d'exposer dans sa monographie la théorie de Dirac, sans lui trouver une meilleure justification. Il annonce dans l'introduction : "nous soulignons le fait que la théorie du rayonnement de Dirac ne peut être considérée comme une quantification de la théorie électromagnétique classique, mais - contrairement à cette dernière - elle n'est en mesure de donner que le calcul de l'interaction "séculaire" entre le rayonnement et les particules". La théorie de Kramers repose sur la distinction entre champ propre et champ extérieur. Pour disons un électron (la théorie se généralise sans peine à un nombre quelconque d'électrons) l'action d'un champ extérieur \vec{A}_{ext} (en jauge $\varphi_{\text{ext}} = 0$) imposé est décrite par le hamiltonien $H_{\text{mat}}(\vec{r}, \vec{p} - e\vec{A}_{\text{ext}})$. L'électron est non relativiste et il est décrit en 1°

* Le nom de Kramers n'était pas mentionné dans l'article de Bethe sur le déplacement de Lamb.

quantification. Le champ propre de Kramers est défini comme le champ qui se réduit dans le référentiel propre de la particule au champ statique coulombien, et le champ extérieur n'est autre que la différence du champ total et du champ propre : $\vec{A}_{tot} = \vec{A}_{prop} + \vec{A}_{ext}$. Dans le problème d'électrodynamique couplé c'est le champ total qui est donné par les équations de Maxwell en fonction du courant j_μ de l'électron. Les équations de ce problème s'écrivent :

$$\left\{ \begin{array}{l} \partial_\mu F_{ext}^{\mu\nu} = j^\nu - \partial_\mu F_{prop}^{\mu\nu} \\ m \ddot{\vec{r}} = e (\vec{E}_{ext} + \vec{v} \wedge \vec{B}_{ext}) - \vec{\nabla} U(\vec{r}), \text{ pour } H_{mat} = \frac{(\vec{p} - e\vec{A}_{ext})^2}{2m} + U(\vec{r}) \end{array} \right.$$

Ces équations doivent être mises sous forme canonique pour pouvoir passer à la théorie quantique. Malheureusement, constate Kramers le terme $\partial_\mu F_{prop}^{\mu\nu}$ contient $\ddot{\vec{r}}$, ce qui semble contredire le fait que le hamiltonien total ne doit contenir que des dérivées premières par rapport au temps. (La difficulté n'est pas insurmontable comme le montreront les travaux ultérieurs). Toutefois si l'on s'intéresse seulement à l'évolution séculaire du système, c'est-à-dire sur des temps très grands devant le temps caractéristique du mouvement électronique (par exemple la période atomique) alors la moyenne temporelle $\overline{\partial_\mu F_{prop}^{\mu\nu}}$ s'annule et le comportement séculaire du système sera donné correctement par les équations issues de l'hamiltonien de Dirac $H_D = \frac{\vec{p} - e\vec{A}_{ext}}{2m} + \frac{1}{2} \int (E_{ext}^2 + B_{ext}^2) d^3x + U(\vec{r})$ ($U(\vec{r})$ est un potentiel extérieur fixé, comme par exemple celui d'un centre coulombien) dans lequel on a substitué \vec{A}_{ext} à \vec{A} . La théorie du rayonnement de Dirac obtenue par quantification canonique de cette théorie hamiltonienne est donc valide dans l'approximation séculaire.

Lors du congrès de Bologne du 18 octobre 1937 (célébration du bicentenaire de la naissance de Luigi Galvani) Kramers exposa sa théorie (4) en insistant sur les traits originaux de sa nouvelle approche : "Dans ma précédente publication j'ai cherché à représenter la théorie de telle sorte que la question de la structure et de l'extension finie des particules n'intervienne pas explicitement et que la grandeur qui est introduite comme masse particulaire soit d'emblée la masse expérimentale". L'élimination de

la structure au profit de l'introduction des paramètres expérimentaux : voilà énoncées les idées de base de la théorie de la renormalisation. Le H_{mat} introduit dans la théorie a une forme déterminée par des expériences en régime quasi stationnaire ; ainsi la masse m qui intervient dans l'expression

$$H_{\text{mat}} = \left(\vec{p} - \frac{e \vec{A}_{\text{ext}}}{c} \right)^2 + U$$
 est la masse expérimentale. Les expressions formelles de l'hamiltonien total auxquelles aboutissent Kramers et Dirac sont identiques mais les interprétations sont différentes : chez Kramers le champ qui intervient est seulement le champ extérieur et la théorie n'est valable que dans l'approximation séculaire, alors que chez Dirac le champ A_{μ} désigne le champ total, et la théorie, s'il n'y avait la difficulté des divergences, prétendrait être valide sur toute échelle de temps. Pour Kramers la théorie de Dirac est fautive en ce que l'énergie du champ propre de l'électron y est comptée deux fois : une fois comme contribution à la masse expérimentale dans $\frac{p^2}{2m}$ (chez Dirac aussi, bien sûr m désigne la masse expérimentale) et une fois comme partie de $\frac{1}{2} \int (E^2 + B^2) d^3x$, énergie du champ total. Il faut donc réinterpréter la théorie de Dirac et alors on ne doit plus s'étonner qu'elle conduise à des divergences lors du calcul d'ordres supérieurs qui sont dépourvus de sens dans l'approximation séculaire.

Quelles sont les critiques que l'on peut faire à cette première version de la théorie de Kramers ? L'approximation séculaire est trop limitative : elle ne permet "d'avoir confiance qu'aux résultats des calculs perturbatifs de premier ordre, et seulement dans le cas où la probabilité d'occurrence d'un processus élémentaire croît très peu pendant la période caractéristique du système"⁽⁴⁾. Pour conclure Kramers signale que la méthode de BLOCH et NORDSIECK, et celle de PAULI et FIERZ exposée au même congrès de Bologne, donnent l'espoir de traiter correctement des processus multiples. L'idée de Bloch et Nordsieck de ne considérer que des états finaux réels, habillés d'une infinité de photons mous, est bien à rapprocher de l'idée de Kramers d'exprimer la théorie à l'aide de paramètres empiriques. De plus nous verrons à propos de l'article d'OPECHOWSKY l'analogie entre le formalisme de Fierz et Pauli et les développements ultérieurs de la théorie de Kramers. La théorie de Kramers est non relativiste, et l'on savait bien en 1937, grâce aux travaux de Weisskopf, que les divergences (logarithmiques) de la théorie

entièrement relativiste différaient profondément de celles de la théorie non relativiste. Il pouvait sembler préférable de traiter ces divergences directement au niveau quantique, et l'avenir révéla que c'était bien la démarche la plus fructueuse. Kramers, lui, pensait (comme Dirac Voir chap. 7 § 2) qu'il valait mieux guérir d'abord la théorie classique de ses divergences et quantifier après. L'idée de tenir compte de la réaction de rayonnement dans la théorie classique de l'électron et de séparer dans cette réaction la contribution à la masse électromagnétique de l'électron n'était bien sûr pas nouvelle, elle datait du siècle précédent avec les travaux de Lorentz (voir chap. 3 § 1). Ce que Kramers apporta de nouveau, ce fut d'une part l'idée d'éliminer la structure de l'électron, d'autre part d'utiliser la théorie résultant de cette élimination comme base de correspondance pour la théorie quantique.

8.2.b Les travaux des disciples de Kramers :

L'approximation séculaire, nous l'avons dit, était bien trop restrictive. Il s'avéra possible de mettre les équations de Kramers sous forme canonique dans un cas plus général.

Le premier article de SERPE (déc. 1939) ne traite pas encore de ce problème : le calcul présenté est entièrement classique. La théorie de WIGNER et WEISSKOPF⁽⁸⁾ de l'émission spontanée (1930 voir chap. 1 § 4) prévoyait un déplacement des niveaux atomiques, il est vrai négligeable devant l'écart de ces niveaux dans l'approximation à deux niveaux. Si comme OPPENHEIMER (voir chap 3 § 2) on tient compte des autres niveaux le déplacement diverge : autrement dit l'énergie propre d'un électron lié est infinie. Serpe montra que la théorie de Wigner et Weisskopf est l'analogue de correspondance d'une théorie classique erronée donnant la même valeur pour le déplacement, et ne donnant plus de déplacement du tout (pour un oscillateur harmonique* et dans l'approximation dipolaire) une fois corrigée à la manière de Kramers. Il fallait donc selon Serpe, rejeter l'évaluation de Wigner et Weisskopf du déplacement. Nous allons maintenant détailler un peu plus les raisonnements de Serpe.

*Le cas de l'oscillateur harmonique, moins réaliste que celui du potentiel coulombien, avait aussi été étudié par Weisskopf et Wigner dans leur 2° article⁽⁹⁾.

- La théorie classique de l'oscillateur harmonique interagissant avec le champ électromagnétique et avec le même hamiltonien que celui de Wigner et Weisskopf donne un déplacement analogue à celui calculé dans la théorie quantique à deux niveaux. Les équations, alternativement, peuvent être résolues en modifiant la masse, de telle sorte que le déplacement s'annule : rien d'étonnant si dans la théorie de Dirac la masse électromagnétique est comptée deux fois.

- Dans la théorie de Kramers, où le champ propre est soustrait du champ total, il n'y a plus de déplacement de niveau. La différence entre les deux théories peut être explicitée en utilisant le modèle de Lorentz de l'électron, décrit par une densité $\rho(r)$ d'extension a . Une fois éliminé le champ créé par la densité ρ le mouvement de l'électron dans le potentiel harmonique $\frac{1}{2} K r^2$ est décrit par :

$$m_0 \ddot{\vec{r}} + K \vec{r} = -\frac{4}{3c^2} W_{\text{self}} \ddot{\vec{r}} + \frac{2}{3} \frac{e^2}{4\pi c^3} \ddot{\vec{r}} + \text{autres sup.} \quad (W_{\text{self}} = \frac{1}{8\pi} \int \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r}-\vec{r}'|} d^3r d^3r')$$

. Dans la première théorie (Wigner-Weisskopf classique) m_0 est considéré comme la masse expérimentale. Pour des solutions du type $e^{i\omega t} e^{-r/\tau}$

$\omega \approx \sqrt{\frac{K}{m_0 + \frac{4}{3c^2} W_{\text{self}}}} \neq \omega_0 \left(\omega_0 = \sqrt{\frac{K}{m_0}} \right)$. Dans la théorie de Kramers le terme $-\frac{4}{3c^2} W_{\text{self}} \ddot{\vec{r}}$, correspondant à la réaction du champ propre de Kramers n'existe pas, il est contenu dans la définition de la masse expérimentale

$m = m_0 + \frac{4}{3c^2} W_{\text{self}}$. Le dernier terme en $\ddot{\vec{r}}$, réaction du champ extérieur, est responsable de l'amortissement.

Dans un article ultérieur⁽⁷⁾ (nov. 1940) SERPE examina ce que donnait la quantification de la théorie de Kramers pour le problème précédent. La difficulté pour mettre sous forme canonique le formalisme de Kramers (le champ extérieur dépend de $\ddot{\vec{r}}$) est levée en utilisant la substitution $\ddot{\vec{r}} \rightarrow -\omega^2 \vec{r}$, légitime si l'oscillateur est peu perturbé. Faisant de plus l'approximation dipolaire ($e^{i\vec{k}\cdot\vec{r}} \approx 1$) dans le développement de Fourier de A) Serpe put mettre les équations sous forme canonique. Le déplacement quantique est alors nul dans l'approximation des deux niveaux et la largeur naturelle est celle donnée par Wigner et Weisskopf. On peut regretter que Serpe ne se soit pas dispensé de cette dernière approximation (des deux niveaux), et qu'il ait étudié le seul cas de l'oscillateur harmonique : sinon il aurait pu comme Kramers en 1947 trouver la formule de Bethe du déplacement.

Un autre développement, cette fois plus formel, de la théorie de Kramers fut fourni par OPECHOWSKI⁽¹⁰⁾ en novembre 1940. La méthode d'Opechowski est plus générale que celle de Serpe car elle s'applique pour un potentiel extérieur $U(\vec{r})$ quelconque, mais le hamiltonien qu'il trouve n'est valable que dans l'approximation dipolaire, non relativiste, et au 1^o ordre seulement. D'après une note d'Opechowski, il semble que Kramers lui-même avait alors traité un cas plus général, mais il n'en reste pas de trace publiée. Le hamiltonien d'Opechowski s'écrit

$$H = \frac{p^2}{2m} + U(\vec{r} - \frac{e}{m} \vec{z}) + \frac{1}{2} \int [(\text{rot } \vec{A})^2 + \vec{E}_1^2] d^3x$$

Les couples de variables conjuguées étant (\vec{r}, \vec{p}) et (A, \vec{E}_1) ($\text{div } \vec{A} = 0$) \vec{z} est le vecteur de Hertz du champ \vec{E} défini par $\text{rot rot } \vec{z} = \vec{E}_1$ et $\text{div } \vec{z} = 0$.

Comme Opechowski le remarqua ce hamiltonien avait déjà été utilisé par PAULI et FIERZ⁽¹¹⁾ (voir chap. 5 § 2) dans un autre contexte, comme hamiltonien équivalent au hamiltonien de Dirac dans l'approximation non relativiste et dipolaire. Il résultait de cette remarque, sans plus de calcul, que la théorie d'Opechowski redonnait les résultats de Dirac sur l'émission et l'absorption du rayonnement par un atome. Dans la dite approximation le hamiltonien de Dirac s'écrit

$$H_D = \frac{p^2}{2m} - \sum_{\alpha} \hbar \omega_{\alpha} (\vec{E}_{\alpha} \cdot \vec{p}) P_{\alpha} + U(\vec{r}) + \frac{1}{2} \sum_{\alpha} (P_{\alpha}^2 + Q_{\alpha}^2) \hbar \omega_{\alpha}$$

les Q_{α} désignant les coordonnées des oscillateurs harmoniques fictifs associés au champ, P_{α} l'impulsion conjuguée et \vec{E}_{α} le vecteur polarisation du mode α convenablement normalisé. Intéressés par le problème infra-rouge Pauli et Fierz voulaient prendre comme états initial et final les états propres du hamiltonien H_D avec $U = 0$, pour habiller ainsi les électrons de photons non relativistes. Il était alors commode d'utiliser la transformation canonique*

$$\left\{ \begin{array}{l} \vec{r} = \vec{r}' - \sum_{\alpha} \vec{E}_{\alpha} Q_{\alpha} \\ Q_{\alpha} = Q'_{\alpha} \end{array} \right\} \left\{ \begin{array}{l} \vec{p} = \vec{p}' \\ P_{\alpha} = P'_{\alpha} + \vec{E}_{\alpha} \cdot \vec{p}' \end{array} \right.$$

qui transformait H_D en $H' = \frac{p'^2}{2m} + U(\vec{r}') - \sum_{\alpha} \vec{E}_{\alpha} Q'_{\alpha} + \frac{1}{2} \sum_{\alpha} (P'_{\alpha}{}^2 + Q'_{\alpha}{}^2) \hbar \omega_{\alpha} + \delta H$ avec $\delta H = -\frac{1}{2} \sum_{\alpha} (\vec{E}_{\alpha} \cdot \vec{p}')^2 \hbar \omega_{\alpha} = \frac{\mu}{m} \frac{p'^2}{2m}$ et $\mu = \frac{4}{3} \frac{e^2}{4\pi} \frac{\omega_1}{c^3}$ (ω_1 coupure U.V.) Fierz et Pauli choisissaient la coupure ω_1 de telle sorte que $\mu \ll m$

* engendrée par la transformation unitaire $U = e^{i\vec{p}' \cdot \sum_{\alpha} \vec{E}_{\alpha} Q_{\alpha}} = e^{i\frac{e}{m} \vec{p}' \cdot \vec{z}}$

et alors H' était équivalent à H_D . Le hamiltonien ainsi obtenu en négligeant δH est identique à celui d'Opechowski mais son interprétation est différente : il n'est valable selon Pauli et Fierz que dans l'hypothèse d'une masse électromagnétique de l'électron faible, c'est-à-dire d'un grand rayon, alors que celui d'Opechowski est indépendant de la structure de l'électron.

En avril 1944 KRAMERS⁽¹¹⁾ prononça à Utrecht (symposium sur les particules élémentaires) une conférence sur les difficultés fondamentales d'une théorie des particules. Seule une part modeste de l'exposé est consacrée à la théorie de Kramers, Serpe, et Opechowski, pour signaler que la divergence des termes spectraux est *"exactement parallèle à la divergence de la masse électromagnétique"* (Serpe) et pour montrer que cette nouvelle théorie est plus conforme au principe de correspondance puisqu'elle *"correspond"* à la théorie classique de l'électron de Lorentz. Elle n'avait été développée que dans une approximation non relativiste : *"une investigation ultérieure devrait déterminer si continuant dans cette direction on peut développer une théorie qui fournisse une formulation satisfaisante et invariante relativiste du problème électron-rayonnement"*. Kramers présente impartialement les autres approches (voir chapitre précédent), n'étant pas sûr que la mauvaise définition de paramètres observables soit la seule source des difficultés en théorie quantique des champs.

8.2.c Le calcul de Kramers du déplacement de Lamb :

Il fallut attendre le conseil de Solvay de 1948 pour que KRAMERS se décidât, une fois découvert le déplacement de Lamb, à étendre le formalisme d'Opechowski aux ordres suivants et calculât le déplacement des raies spectrales dans l'approximation dipolaire non relativiste.

La méthode suivie ici par Kramers est un peu différente de sa méthode initiale : les différents champs (extérieur, total et propre) sont définis dans un modèle d'électron étendu pour pouvoir localiser plus précisément les divergences et la dépendance dans la structure. Ainsi Kramers peut montrer que dans l'approximation dipolaire sa définition du champ extérieur conduit à des équations indépendantes de la structure. Elles peuvent être mises sous forme hamiltonienne :

$$H = \frac{p^2}{2m} + U(\vec{n} - \frac{e}{m} \vec{Z}) + \frac{1}{2} ((\text{rot } \vec{A})^2 + E^2) + \frac{e^2}{2m} (-A^2 + 2 \vec{E} \cdot \vec{Z})$$

(les notations sont les mêmes que pour le hamiltonien d'Opechowski).

Il est aisé de vérifier que ce hamiltonien conduit à la formule de Bethe pour le déplacement de Lamb, sans qu'aucune soustraction soit nécessaire.

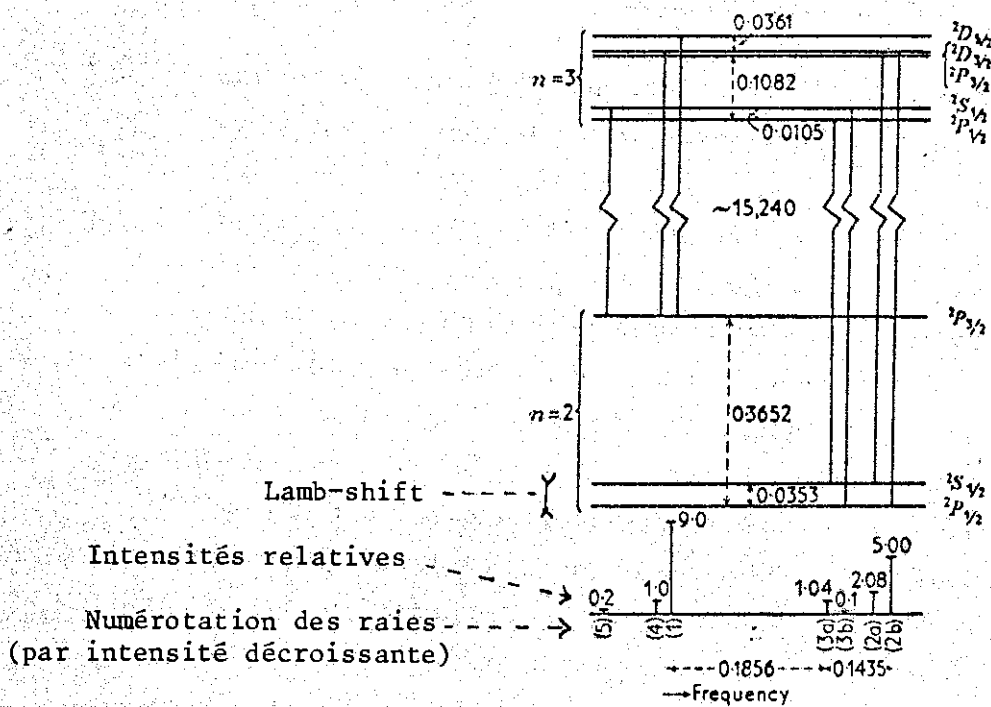
La première question qui vient à l'esprit est : quel est le lien formel entre ce hamiltonien et le hamiltonien habituel pour que le déplacement de Lamb calculé soit le même ? Cette question fut posée par DIRAC au cours de la discussion : *"Peut-on obtenir une relation entre votre nouveau hamiltonien et le hamiltonien habituel par une transformation canonique ?"* OPPENHEIMER donna la réponse précise : dans l'approximation utilisée (dipolaire non relativiste) une telle transformation existe (c'est la transformation de Fierz et Pauli que l'on caractériserait aujourd'hui par $U = e^{i \frac{e}{m} \vec{p} \cdot \vec{Z}}$) et la différence entre les deux hamiltoniens est un terme de masse électromagnétique comme nous l'avons vu à propos du hamiltonien d'Opechowski. Dès lors rien d'étonnant si le hamiltonien de Kramers redonne le résultat de Bethe : il ne fait qu'introduire naturellement le contre-terme de masse dans la théorie habituelle.

Bethe et Kramers dans leurs calculs de déplacement de Lamb partaient tous deux des mêmes idées de base dues à Kramers : éliminer la structure de l'électron, exprimer les grandeurs physiques recherchées en fonction des paramètres expérimentaux. Il se trouve que la méthode de Kramers, attachée à quantifier une théorie classique juste, se révéla moins générale et moins fructueuse que la méthode maintenant établie qui consiste à renormaliser au niveau quantique et qui conduisit Bethe à son résultat. Après tout, d'un point de vue plus formel qui semble s'imposer dans la théorie relativiste, la base de correspondance de l'électrodynamique quantique moderne - onde de Dirac couplée au champ électromagnétique - ne présente pas de divergence, puisque le concept de particule n'apparaît pas à ce niveau ; les divergences n'apparaissent qu'au niveau quantique, comme un effet des fluctuations du vide et la mise en garde de Kramers : ne pas quantifier une théorie classique fautive, n'intervient plus. Il reste que l'idée de renormalisation, dans son sens le plus général, est une idée de Kramers.

8.3. LE DEPLACEMENT DE LAMB-RETFERD

Dès sa naissance la théorie quantique des champs put embrasser un ensemble très vaste de phénomènes. Plus tard sous l'impulsion de FERMI et de YUKAWA ses applications s'étendirent à la physique nucléaire. La théorie relativiste de l'électron et la prédiction de l'existence du positron furent spectaculairement confirmées par l'expérience en 1933. La théorie de Fermi s'accordait bien avec les expériences d'émission β . On crut aussi avoir découvert le méson de Yukawa. Néanmoins jusqu'en 1947 l'expérience ne put montrer la voie à travers les difficultés de la théorie des champs : tant sa construction formelle que ses difficultés mathématiques se présentèrent surtout comme des problèmes internes à une extension naturelle de la mécanique quantique de Heisenberg, Schrödinger et Dirac, survenant au bout de l'exploration systématique des conséquences des principes fondamentaux de la mécanique quantique et de la relativité. Ce n'est qu'en 1947, avec la découverte du déplacement de Lamb que l'expérience joua un rôle actif et stimula les théoriciens déjà en possession d'une version naïve du formalisme de la théorie des champs et de l'idée de base de renormalisation : au conseil Solvay de 1948 Oppenheimer s'exprima ainsi : *Ces développements auraient pu être appliqués aux déplacements de niveaux n'importe quand durant les quinze dernières années, pourtant ils requièrent l'impulsion de l'expérience, ils devaient être stimulés et vérifiés*"⁽¹³⁾.

Une nouvelle fois, en 1947, l'atome d'hydrogène joue son rôle de paradigme de la théorie quantique : l'expérience de LAMB et Retherford mit en évidence une nouvelle propriété du spectre de l'atome d'hydrogène : les niveaux $2S_{1/2}$ et $2P_{1/2}$ n'étaient pas confondus, contrairement à la prédiction de l'équation de Dirac. Plus précisément Lamb et Retherford donnèrent la première preuve décisive du fait que le niveau $2S_{1/2}$ est au dessus de la valeur calculée par Dirac ; il est déplacé d'environ 1000 MHz (la valeur admise aujourd'hui est 1058 MHz). Nous donnons le schéma de niveaux de l'atome d'hydrogène, auquel le lecteur pourra se reporter pour comprendre les notations utilisées par la suite.



D'après Series (14)
(1957).

FIG. 7. Fine structure of H_{α} (D_{α}) showing calculated Lamb shifts.
(Compare Fig. 2, p. 33.) Intervals in cm^{-1} .

Les résultats de cette expérience remarquable furent discutés lors de la conférence sur les fondements de la mécanique quantique de l'île Shelter (près de New York) du 2 au 4 juin 1947 et donnèrent lieu à une première estimation théorique de Bethe, comme nous le verrons bientôt.

C'est le développement de la technique du radar pendant les années de guerre, auquel LAMB et KUSCH participèrent activement, (revoir conférences Nobel^(20,21)) qui fournit les sources d'onde centimétriques (klystrons et magnétrons) nécessaires à l'exploration détaillée de la structure fine de l'hydrogène. Mais l'expérience de Lamb est l'aboutissement de plusieurs techniques expérimentales raffinées : la méthode des jets atomiques, l'utilisation des radiofréquences, la méthode de résonance qui furent développées avant guerre. Lamb et Rethford eurent à résoudre, tant par des évaluations théoriques que par des astuces expérimentales, les nombreuses difficultés que comportait leur projet ambitieux. Avant de décrire l'existence d'un déplacement, dès les années 30. Le lecteur

désireux de s'informer plus précisément trouvera d'excellents articles de revue : un de LAMB⁽¹⁵⁾ (1951) sur l'histoire du déplacement, un de KELLOG et MILLMAN⁽¹⁶⁾ (1946) sur la méthode de résonance et un de GORDY⁽¹⁷⁾ (1948) sur la spectroscopie microonde ; le cours de D.E.A. de Jean BROSSEL à l'ENS Ulm* donne une idée très précise des difficultés expérimentales qui furent rencontrées dans la réalisation d'une expérience comme celle de Lamb et Retherford.

Jusqu'en 1947 le sujet du déplacement éventuel des raies spectrales était particulièrement mouvant. Ni la théorie, ni l'expérience n'étaient en mesure de trancher et les estimations théoriques fluctuaient suivant les résultats changeants des expériences. Ce climat d'incertitude, pimenté de controverses théoriques (comme celle à propos de Fröhlich Heitler et Kahn relatée plus loin) nous apparaît comme un des éléments importants de la crise de la théorie quantique des champs.

8.3a. Les premières prédictions théoriques d'un déplacement.

DIRAC, dans la théorie générale de la diffusion⁽¹⁸⁾ qu'il publia en 1927, démontra l'existence d'un déplacement des niveaux du système diffusant et en donna l'expression**. Dans une note de leur premier article sur la forme des raies d'émission WIGNER et WEISSKOPF⁽⁸⁾ (1930) réaffirmèrent l'existence d'un tel effet pour les niveaux d'un atome ; cet effet était fini dans l'approximation à deux niveaux. S'affranchissant de cette dernière approximation OPPENHEIMER⁽¹⁹⁾ (1930) montra que ce déplacement divergeait et qu'on ne pouvait donc prendre au sérieux la prédiction de l'électrodynamique quantique de l'époque. L'équation de Dirac déjà établie en 1928 donnait un spectre théorique de l'hydrogène atomique en parfait accord avec l'expérience. L'étape suivante, s'il y en avait une, serait donc pour les expérimentateurs de mettre en évidence des écarts par rapport à la théorie de Dirac. L'atome d'hydrogène, comme il est le plus simple et qu'il était calculable de manière entièrement analytique tout au moins jusqu'à la théorie de Dirac, fut pendant toute l'histoire de la spectroscopie un sujet d'étude très exploité. En plus, une mesure précise de son spectre permet de remonter à des constantes fondamentales comme le Rydberg.

$$** \int \tilde{E}_m = \langle m | V \left(\frac{1}{E-H} \right)_{P.P.} V | m \rangle$$

p.p. = partie principale
 v = perturbation

en notation moderne

* non publié

8.3.b. Le doublet de l'hydrogène et l'effet Pasternack :

La première évidence expérimentale d'une structure fine du spectre de l'hydrogène fut obtenue avec le doublet de la série de Balmer H_{α} . Ce doublet était constitué de deux paires de raies 1,4 et 2,3 (voir le schéma de niveau)*. En 1887, Michelson et Morley donnèrent une première estimation (0.253 cm^{-1}) de la séparation du doublet). Dans les années suivantes les mesures se multiplièrent, les techniques expérimentales (la spectrométrie) s'améliorèrent, par ailleurs la théorie de Dirac - DARWIN de l'atome d'hydrogène fut publiée en 1928 si bien qu'en 1932 HOUSTON et HSIEH⁽²³⁾ purent soupçonner que la séparation mesurée était trop petite par rapport à la prédiction de DIRAC^{**}. "Une explication possible de ce fait, comme BOHR et OPPENHEIMER nous l'ont fait remarquer[...] est que l'effet de l'interaction entre le champ de rayonnement et l'atome a été négligé dans le calcul des fréquences. On ne sait pas actuellement comment combiner l'atome et le champ en un seul système, mais il est clair qu'à l'issue d'une telle combinaison, on devrait s'attendre à un déplacement des niveaux et il devrait être de l'ordre de α fois la séparation de structure fine. C'est l'ordre de grandeur de l'effet observé⁽²³⁾". A posteriori on est frappé de la justesse de ces commentaires, mais le fait expérimental était loin d'être évident et d'autres mesures (que nous ne citons pas, voir réf. 15) semblaient être en accord avec la théorie de Dirac. L'obstacle majeur à une détermination précise de la séparation du doublet était l'effet Doppler : dans toutes les mesures avant guerre on utilisait un tube à décharge d'hydrogène sous pression réduite plongé dans l'air liquide ; l'analyse du spectre d'émission se faisait à l'aide d'un interféromètre de Fabry-Pérot (à l'exception de Drinkwater etc) et il fallait une théorie de la forme de raie (essentiellement l'élargissement Doppler) pour remonter à la séparation.

En 1937 HOUSTON utilisa des mesures du même type pour H_{α} et D_{α} (le deutérium, étant plus lourd, a l'avantage de présenter un effet Doppler un peu réduit) tout en perfectionnant l'analyse ; la valeur de la séparation demeurait en désaccord avec celle de DIRAC, quoique légèrement plus faible que la précédente mesurée. Un an plus tard R.C. WILLIAMS⁽²⁵⁾, put séparer la composante 2 de la composante 3 dans le cas

* On ne fut en mesure de séparer les composantes de chaque paire que beaucoup plus tard (2 et 3 par Williams en 1938)

** Les résultats publiés concernent aussi H_{β} , H_{γ} , H_{δ} et H_{ϵ} .

du deutérium. Cette fois l'écart à la théorie de Dirac était bien plus criant : 0.130 cm^{-1} au lieu de 0.108 cm^{-1} pour ν_{23} puisque le Lamb-shift (0.03 cm^{-1}) se trouvait alors comparé à 0.108 et non plus à la séparation du doublet 0.317 cm^{-1} .

PASTERNAK⁽²⁹⁾, dans une lettre à la Phys. Rev., se référant aux travaux précédents ainsi qu'à des résultats non publiés de HOUSTON et ROBINSON (sur D_α) montra que les différents écarts observés, compte tenu des intensités relatives de 2a 2b et de 3a 3b s'expliquaient en supposant que le niveau $2S_{1/2}$ est déplacé de 0.03 cm^{-1} vers le haut par rapport à la prédiction de Dirac (cette valeur correspond à l'ordre de grandeur vrai).

Les mesures précédentes furent effectuées dans des universités américaines. En Angleterre DRINKWATER, RICHARDSON et W.E. WILLIAMS⁽²⁶⁾ aboutirent en 1940 à des conclusions différentes. Signalons qu'ils essayèrent sans succès d'utiliser un jet atomique pour éliminer l'effet Doppler : l'excitation était inefficace et le signal trop faible.. Revenant à la méthode traditionnelle de la lampe à décharge ils obtinrent des résultats différents pour les intensités (composante 3) et en conséquence une interprétation différente du spectre. Il subsistait des déplacements anormaux ; les auteurs les attribuèrent à une contamination du deutérium par l'hydrogène et au spectre moléculaire et H_2 et D_2 . GIULOTTO⁽²⁷⁾ en mars 1947 se plaça dans des conditions expérimentales où l'intensité du spectre moléculaire pouvait être sûrement négligée, et souligna la persistance de l'écart à la théorie de Dirac.

L'alternance des résultats contradictoires, la manière détournée d'évaluer le déplacement suscitèrent plutôt la méfiance des théoriciens. Pourtant on peut dire rétrospectivement que dès 1937, avec les mesures de Houston et Williams le déplacement de Lamb était clairement visible : une analyse moderne des différentes causes d'erreurs conduirait certes à des marges plus larges que celles données par ces auteurs mais l'effet subsisterait (Lamb⁽²⁰⁾).

8.3.c. Essais d'interprétation théorique de l'effet Pasternack :

Nous ne traiterons pas ici des calculs d'énergie propre puisqu'en raison même de leur divergence, ils ne pouvaient fournir une interprétation satisfaisante du déplacement des raies.

Signalons d'abord brièvement deux effets parasites possibles : l'un, l'effet Stark dû au champ électrique du tube à décharge, fut mentionné par WILLIAMS⁽²⁵⁾, mais il était vraiment négligeable. DRINKWATER et WILLIAMS⁽²⁶⁾ accusèrent le spectre moléculaire de l'hydrogène, comme nous l'avons déjà dit.

En 1932 KEMBLE et PRESENT⁽²⁸⁾ interprétèrent les résultats de Houston et Hsieh en invoquant une valeur finie du rayon de l'électron. Dans ce cas la distribution de charge devait être modifiée au voisinage de l'électron, il en résultait un déplacement des niveaux S (car perturbativement la structure de l'électron n'intervient que lorsqu'il est près du noyau, et seules les orbitales S ont une densité non nulle au centre). En prenant un rayon de l'électron $a \approx 5 \cdot 10^{-14}$ m le niveau $2S_{1/2}$ est déplacé de 0.02 cm^{-1} en accord avec les mesures de Houston et Hsieh. La valeur $5 \cdot 10^{-14}$ m parut étrangement grande à Kemble et Present devant le rayon classique de l'électron $r_c = 2 \cdot 10^{-15}$ m, d'autant que la théorie de Born (voir chap. 7 § 1) prévoyait $a \approx r_0$.

UEHLING, en 1935 (voir chap. 5 § 1) calcula le déplacement des niveaux S dû à la polarisation du vide et constata que le signe était mauvais et l'ordre de grandeur trop faible par rapport à l'estimation de Kemble et Present. On sait aujourd'hui que la polarisation du vide contribue effectivement (-27 MHz sur 1058 MHz) au déplacement de Lamb mais qu'elle n'est pas l'effet principal.

A défaut de trouver dans l'électrodynamique pure une explication satisfaisante, plusieurs chercheurs proposèrent une origine nucléaire du phénomène. La structure hyperfine des raies prévue par FERMI⁽⁶⁴⁾ ne put être tenue pour responsable. BECHERT et MEIXNER, tenant compte des corrections relativistes et de l'effet d'entraînement du noyau, montrèrent en 1935 que cette structure était trop petite; par ailleurs ils évaluèrent l'effet d'entraînement du noyau dans la théorie relativiste et conclurent négativement. Houston et Hsieh⁽²³⁾ avaient déjà constaté que la structure hyperfine calculée par BETHE⁽³⁰⁾ pouvait être négligée; WILLIAMS en 1938 montra expérimentalement qu'il en était bien ainsi: les observations n'étaient pas modifiées en présence d'un champ magnétique, alors qu'elles

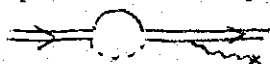
l'auraient été dans le cas contraire.

Avec l'article de 1939 de FRÖHLICH HEITLER et KAHN⁽³¹⁾ (FHK) commence une longue polémique sur l'effet de la dissociation du proton en paire neutron-"mésotron". En langage moderne ces auteurs calculèrent la contribution du graphe







au potentiel électrostatique,


où l'électron interagit virtuellement avec le pion virtuel π^+ . Naturellement un tel processus diverge puisqu'il comporte une boucle, mais FHK furent capables d'extraire une contribution finie en soustrayant un terme de la forme ∞/R obtenu en prenant formellement le coefficient de $1/R$ pour R grand dans l'expression de l'amplitude, c'est-à-dire en retirant seulement la partie non coulombienne de l'énergie d'interaction avec l'électron situé à la distance R^{**} . Le déplacement du niveau $2S_{1/2}$ qui en résulte est 0.03 cm^{-1} (miracle !) conformément à l'estimation de Pasternack. FHK étaient en partie conscients des limites de leur calcul : ils s'intéressaient à un effet répulsif pour des distances fractions du rayon classique de l'électron r_e , or justement la théorie mésique présentait des difficultés pour les distances inférieures à r_e .

(divergences). On pourrait encore formuler d'autres critiques : les auteurs n'utilisent pas une méthode de perturbation correcte, ils devraient prendre ne compte tous les graphes qui interviennent au même ordre ( etc.) dans la théorie mésique et aussi les processus analogues en électrody-

* Cette notation est un peu abusive car les auteurs utilisaient une théorie mésique vectorielle.

** On sait bien aujourd'hui que seule diverge la partie du graphe qui correspond aux valeurs limites des impulsions externes: $p_1 - p_2 \rightarrow 0$ ou $k \rightarrow 0$ pour le "photon" coulombien c.à.d.

$R \rightarrow \infty$. En électrodynamique quantique le type analogue de divergence de  est compensé par les divergences de  et  renormalisés de masse. Si FHK avaient eu l'idée d'étudier la modification du potentiel coulombien en électrodynamique pure due au graphe  et avaient appliqué la même méthode consistant à retenir la partie non coulombienne ils auraient trouvé la bonne théorie du déplacement de Lamb (sans toutefois apporter une justification correcte de l'élimination des infinis).

namique pure ( etc.) . De toutes façons, comme on le sait aujourd'hui, la théorie mésique ne s'applique effectivement pas à cette échelle, et pis, elle n'est pas renormalisable.

Sous la direction de Robert OPPENHEIMER, à l'université de Californie, le jeune Willis Eugene LAMB effectuait alors une thèse sur les propriétés électromagnétiques des systèmes nucléaires. Il publia une lettre critique dans Phys. Rev. ⁽³²⁾ (1939) fondée sur l'interprétation physique que FHK donnaient de leur théorie : le proton à l'intérieur du noyau a une probabilité p d'être dissocié sous forme n , π^+ et $1-p$ d'être lui-même, l'électron interagit donc directement avec le proton, ou avec π^+ ; mais, remarque Lamb, on ne peut trouver d'effet répulsif aux courtes distances que si la probabilité p est supérieure à un, ce qui est bien le cas chez FHK pour qui elle diverge, mais ce qui est absurde. FHK auraient pu rétorquer à Lamb qu'il ne fallait pas s'attarder à une telle imagerie puisqu'il s'agissait de processus virtuels*, mais ils critiquèrent ⁽³³⁾ plutôt l'utilisation de la notion de densité de probabilité pour des mésons et sur une échelle de longueur de l'ordre de r_0 : alors intervient le phénomène de création de paires $\pi^+\pi^-$; la seule quantité qui reste définie est la densité de charge et pour elle on n'a pas de limite supérieure. LAMB ⁽³⁴⁾ refusa la réplique car les mésostrons émis, étant lourds, devaient être considérés comme non relativistes et dans ce cas la notion de probabilité avait un sens ; enfin rajouta Lamb FHK n'avaient pas le droit d'utiliser un résultat (répulsion près du centre) conditionné par la non validité de la théorie des perturbations : la répulsion disparaît si l'on introduit une coupure U.V. telle que les probabilités restent inférieures à l'unité. KAHN ⁽³⁵⁾, en 1941, se rallia au point de vue de Lamb : il n'y avait pas de raison sérieuse de croire à un déplacement des raies spectrales lié à l'interaction nucléaire. Le champ répulsif trouvé par FHK était si grand qu'il dépassait le seuil de création de paires (e^+e^-) et une modification plus raisonnable du potentiel au voisinage du centre conduisait comme

* Il est typique à cette époque de chercher une interprétation concrète aux termes de la série des perturbations : les processus "virtuels" invoqués ne différaient en fait de processus réels que par la violation de la conservation de l'énergie (dans l'ancienne théorie des perturbations) qui est effectivement permise pendant des temps courts d'après la 4^o relation d'incertitude.

l'avait montré SOMMERFELD⁽³⁶⁾ à grand renfort de fonctions hypergéométriques dégénérées (1941) à un déplacement ridiculement petit. Signalons pour l'anecdote que BLATT reprit en 1945 le calcul de FHK avec la nouvelle théorie de Dirac (de 1942) et trouva un résultat convergent et faible.

Il est facile aujourd'hui de voir que toute cette discussion se déroulait bien loin de la vérité - les interactions nucléaires n'ont pas grand chose à voir avec le déplacement de Lamb - mais elle montre quelle importance certains théoriciens et expérimentateurs apportaient à cet effet hypothétique : "Vraisemblablement la question de l'écart à la loi de Coulomb et de la structure fine exacte de l'hydrogène atteint juste les limites de validité de la mécanique quantique actuelle. Cela rend désirable une investigation expérimentale détaillée de la question et une comparaison avec nos théories actuelles. L'information tirée de mesures exactes de la structure fine de l'hydrogène sera probablement importante pour le développement de la future électrodynamique quantique"⁽³¹⁾. La dernière phrase est prophétique (Fröhlich, Heitler et Kahn).

8.3.d Le développement des techniques expérimentales :

La technique des jets atomiques et moléculaires

Cette technique consiste à éjecter une vapeur à travers une série de diaphragmes pour obtenir un jet directif. Ce dispositif permet d'étudier d'éventuelles déflexions du jet et il élimine aussi l'effet Doppler pour le rayonnement émis perpendiculairement au jet. Les jets atomiques connus depuis DUNOYER furent utilisés pour la première fois en physique atomique par STERN⁽³⁸⁾ et GERLACH⁽³⁴⁾ (1921 à 1924) pour étudier la déflexion d'atomes d'argent par un gradient de champ magnétique. De 1931 à 1934, BREIT⁽⁴⁰⁾, RABI^(40,41,42), KELLOG^(41,42), RAMSEY⁽⁴¹⁾, ZACHARIAS^(41,42) et COHEN⁽⁴³⁾ généralisèrent la méthode à l'étude des spins nucléaires : on obtenait ainsi la valeur du spin et le facteur de Landé nucléaire.

La résonance magnétique

Cette méthode, comme elle permet l'exploration directe de la structure Zeeman, apporta un progrès spectaculaire en physique atomique. Le jet atomique polarisé par un gradient de champ magnétique traverse un

champ magnétique homogène B. Dans cette dernière zone, il est aussi soumis à une onde radio fréquence qui induit, lorsque sa fréquence est accordée, des transitions entre les sous-niveaux Zeeman dans le champ B et égalise ainsi la population des différents sous-niveaux. La détection est effectuée sur le jet après un gradient de champ de signe opposé au premier. L'expérience fut réalisée sur une idée de RABI par RABI, MILLMAN, KUSCH et ZACHARIAS^(44,45). Les premières expériences de ce type étaient des expériences de résonance magnétique nucléaire faisant appel à des radio-fréquences de l'ordre du Mégahertz. Elles ouvrirent la voie à toute une série d'études expérimentales fines de l'atome, dont l'expérience de Lamb exécutée dans le laboratoire de Rabi à l'université de Columbia est un exemple. Les variantes futures, tout en gardant le principe général polarisation - égalisation par radio fréquence - détection, utilisèrent d'autres domaines de fréquence (Lamb), d'autres méthodes de polarisation (métastabilité chez LAMB, pompage optique chez KASTLER et BROSSEL ...) d'autres types de détection (transfert ionisant chez LAMB, absorption de la radio-fréquence chez BLOCH et PURCELL ...).

Les ondes centimétriques

Leur utilisation en spectroscopie d'absorption fut suggérée pour la première fois dès 1928 par GROTRIAN dans son livre : "Représentation graphique des spectres des atomes"⁽⁴⁶⁾ (p. 37) : les règles de sélection n'interdisent pas de transition entre niveaux de même nombre quantique principal, les longueurs d'ondes émises ou absorbées lors de telles transitions devraient se situer dans le domaine centimétrique (structure fine). Malheureusement on ne disposait pas alors de générateurs d'ondes centimétriques de qualité suffisante. En 1932 BETZ⁽⁴⁷⁾ observa une absorption à 3 cm dans l'hydrogène excité correspondant probablement au doublet de H_{α} (2,74 cm). En 1935 Haase étudia par la même méthode l'absorption dans le niveau $n = 3$ mais obtint un résultat négatif. La difficulté, pensait-on*, venait de la trop faible intensité des générateurs à éclateur

* En réalité l'absorption ne pouvait être que très faible du fait que les populations des différents niveaux $n = 2$ dans une lampe à décharge sont approximativement égales (voir Lamb⁽²⁰⁾).

("Löschfunktensender"). Lamb (voir (20)) eut connaissance des travaux de ces pionniers. Une autre expérience d'absorption centimétrique fut réussie en 1934 par CLEETON et WILLIAMS mais cette fois sur une transition moléculaire : la célèbre inversion de l'ammoniac (1,25 cm). Ils utilisaient des tubes Westinghouse de 9 cm sous-dimensionnés (du type magnétron). D'après Gordy "pour autant que je sache un seul papier sur la spectroscopie microonde existe dans la littérature d'avant-guerre" ; il se trompait puisqu'il y avait aussi Betz et Haase, mais probablement pas de beaucoup. Les progrès les plus importants dans ce domaine furent réalisés pendant la guerre. La construction des radars nécessitait des générateurs microondes (klystrons et magnétrons) puissants et de fréquence bien définie. La plupart des physiciens américains ou immigrés aux Etats-Unis travaillèrent soit à la bombe atomique soit à l'amélioration de la technique du radar pendant la seconde guerre mondiale. Plusieurs expérimentateurs du groupe de RABI au laboratoire de Columbia, dont KUSCH et LAMB entre autres, améliorèrent les techniques microondes. SCHWINGER développa la théorie des guides d'onde.

8.3.e L'expérience de Lamb-Retherford :

Après la guerre Lamb eut l'occasion de lire un livre de HERZBERG où les expériences de Betz et Haase se trouvaient mentionnées. Utilisant les connaissances acquises au "Columbia Radiation Laboratory" * pendant la guerre, par exemple en étudiant l'absorption des ondes centimétriques dans la vapeur d'eau atmosphérique, il eut l'idée de reprendre ces expériences avec les nouvelles sources microondes, mais après un examen plus approfondi il estima que les niveaux $n = 2$ dans une lampe à décharge ne présentaient pas de différence de population suffisante pour que l'absorption de la radio fréquence soit mesurable. Une nouvelle idée, celle d'utiliser la métastabilité du niveau $2S^{1/2}$, lui permit au bout d'une année de réflexion d'imaginer le dispositif expérimental qu'il réalisa avec

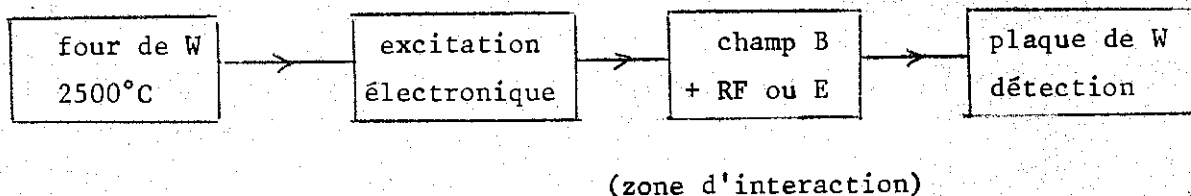
* C'est dans ce même laboratoire dirigé par Rabi que se déroulèrent les expériences de Lamb et Retherford ainsi que celles de Kusch sur le g-2.

son élève RETHERFORD. Beaucoup d'incertitudes, souligne Lamb⁽²⁰⁾, planaient sur les chances de succès de cette expérience, en particulier sur la production effective de métastables et sur la méthode de détection des métastables (nous y reviendrons), et il fallut sans doute une certaine audace pour se lancer dans cette grande entreprise.

Au printemps 1947 l'expérience donna des résultats positifs. Ceux-ci furent discutés du 2 au 4 juin 1947 sur l'île de Shelter où se trouvaient réunies quelques unes des grandes figures de la physique de l'époque, notamment Kramers et Bethe. Un premier compte-rendu rapide de l'expérience fut publié le 18 juin 1947 dans Phys. Rev..

Le premier article de Lamb et Retherford⁽⁵¹⁾ (18 juin 1947)

Le dispositif permet d'étudier les transitions radio-fréquence entre le niveau $2S_{1/2}$ et les niveaux $2P_{1/2}$ et $2P_{3/2}$, ou plus exactement entre les composantes Zeeman de ces différents niveaux. Les fréquences de résonance mesurées s'interprètent en supposant que le niveau $2S_{1/2}$ est environ 1000 MHz en dessus de la valeur donnée par Dirac, les niveaux $2P_{1/2}$ et $2P_{3/2}$ étant inchangés. Le dispositif expérimental peut être ainsi schématisé :



- Le four de tungstène à 2500°C permet une dissociation à 60 % de l'hydrogène moléculaire en hydrogène atomique : $H_2 \rightarrow H^* + H^*$; la sortie est équipée de diaphragmes produisant un jet étroit.

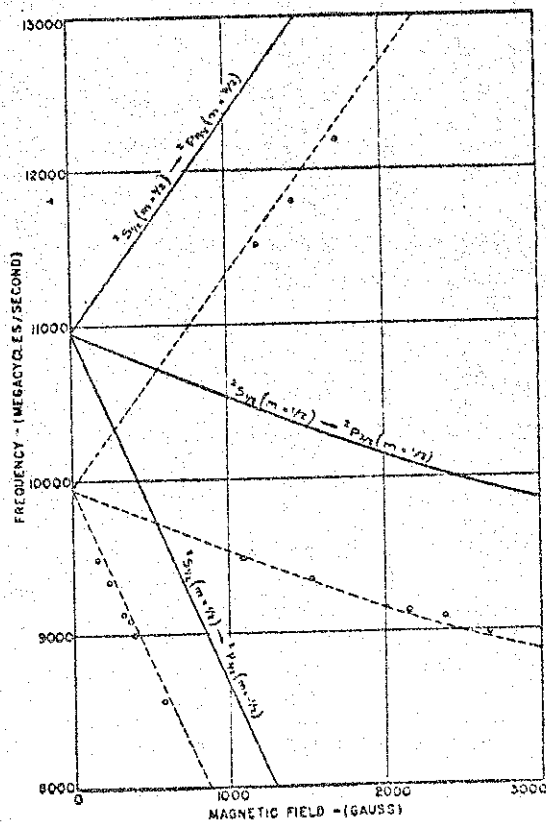
- Le jet est excité perpendiculairement par un faisceau d'électrons accélérés sous environ 11 V. On obtient en gros 1 % de métastables $2S_{1/2}$ qui n'ont pas le temps de se désintégrer en vol puisque leur durée de vie est de 1/7 s. En revanche le niveau $2P_{1/2}$ se désexcite en $1.6 \cdot 10^{-9}$ s.

- La détection des métastables est assurée par l'éjection d'électrons d'une plaque de tungstène qu'ils percutent. Les courants de l'ordre de 10^{-14} A sont mesurés à l'électromètre.

- Si dans la zone d'interaction on applique un champ électrique E les niveaux $2S_{1/2}$ et $2P_{1/2}$ se trouvent mélangés par la perturbation et la durée de vie des métastables se trouve brusquement réduite, le courant observé décroît. Cette première partie de l'expérience visait à établir que le signal mesuré par le détecteur était bien dû aux métastables et à vérifier la formule de Breit et Teller que nous donnons plus loin. Si maintenant, au lieu de E, une radio fréquence accordée sur la transition $2S_{1/2} \rightarrow 2P_{1/2}$ ou $2S_{1/2} \rightarrow 2P_{3/2}$ est appliquée, on observe encore une décroissance du courant. En réalité ces expériences sont effectuées en présence d'un champ magnétique pour différentes raisons que les auteurs expliqueront plus tard, et ce sont les composantes Zeeman dans ce champ qui interviennent.

Nous reproduisons ci-contre la figure publiée dans l'article original où l'on peut voir les droites Zeeman correspondant aux différentes transitions. Les fréquences de résonance mesurées (traits pointillés) sont environ 1000 MHz plus bas que la valeur théorique (trait plein).

A la fin de l'article les auteurs annoncent une publication plus détaillée de leurs résultats. Cette publication fut différée jusqu'en 1950. La précision de la première expérience était de 100 MHz et les auteurs prévoyaient déjà un nouvel appareil qui permettrait de l'amener à 10 MHz.



puis à 1 MHz : "le programme était imposant et il rencontra des difficultés inattendues et très longues à surmonter. En conséquence le papier promis deux ans plus tôt fut différé"⁽⁵³⁾. Avant le premier exposé détaillé de la première expérience parurent de très brèves indications des résultats de la seconde⁽⁵²⁾.

Les difficultés de la réalisation de l'expérience

Non contents de donner les détails techniques de leurs expériences, Lamb et Retherford dans une série impressionnante de trois articles publiés à un an d'intervalle de 1950 à 1952, firent eux-mêmes l'historique de l'étude du doublet de l'hydrogène et des techniques expérimentales qu'ils utilisèrent⁽⁵³⁾, et aussi une analyse théorique très poussée des différentes causes d'erreur possibles. Les deux derniers articles^(54,55) présentent surtout les raffinements introduits sur le second appareil, nous nous intéresserons plutôt au premier où l'on trouve la justification et l'analyse des différentes parties du dispositif grossièrement décrites en 1947. La figure qui suit représente le premier appareil.

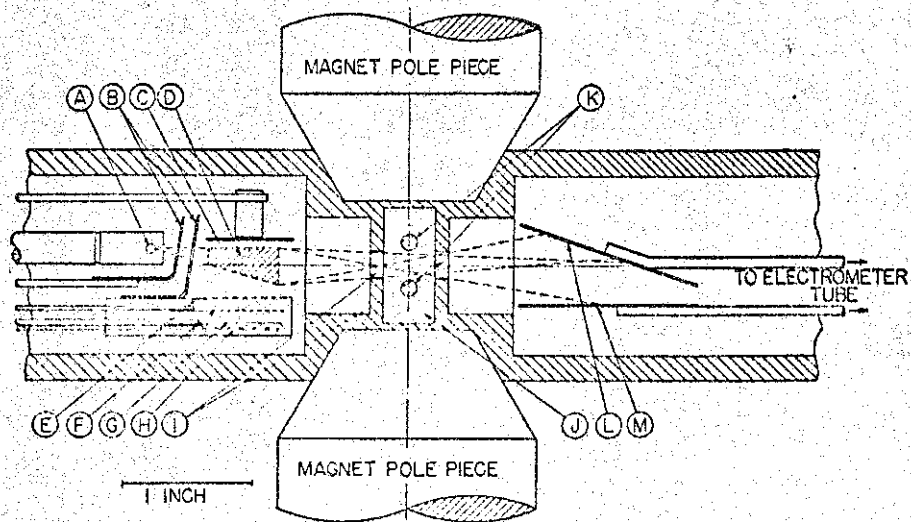


FIG. 19. Cross section of apparatus. A. Tungsten oven of hydrogen dissociator. B. Shields. C. Anode of electron bombarder. D. Bombardment region. E. Accelerator grid of electron bombarder. F. Control grid. G. Cathode of electron bombarder. H. Heater for cathode. I. Slits. J. Wave guide. K. Quenching wires and transmission lines. L. Metastable detector target. M. Electron collector.

- La première difficulté est liée à la production des métastables. La métastabilité du niveau $2S_{1/2}$ fut un sujet longuement controversé par théoriciens et expérimentateurs (voir Lamb⁽¹⁵⁾). La métastabilité prédite par les règles de sélections (pour des processus de désexcitation au premier ordre) pouvait être remise en question par une petite perturbation extérieure. Si les niveaux $2S_{1/2}$ et $2P_{1/2}$, comme le prévoyait la théorie de Dirac étaient rigoureusement dégénérés, un champ électrique même faible suffisait à mélanger les deux niveaux et à rabaisser considérablement la durée de vie de $2S_{1/2}$. En 1933 dans son article de "Handbuch"⁽³⁰⁾ BETHE donna un calcul de la durée de vie en présence d'une perturbation V , et en 1940 BREIT et TELLER⁽⁵⁶⁾ donnèrent une formule plus raffinée valable dans le cas où les deux niveaux $2S_{1/2}$ et $2P_{1/2}$ sont séparés de $\hbar\omega$:

$$\tau_S = \tau_P \left[\hbar^2 \left(\omega^2 + \frac{1}{4} \gamma_P^2 \right) / |V|^2 \right] \quad \gamma_P \text{ désignant la largeur}$$

naturelle du niveau p , et V l'élément de matrice de la perturbation entre les niveaux $2S_{1/2}$ et $2P_{1/2}$. Si $\hbar\omega$ était nul, les champs électriques parasites, comme le champ motionnel (champ B vu du repère de l'atome) suffiraient à détruire la métastabilité. Mais avec la valeur réelle $\hbar\omega \approx 1000$ MHz, τ_S restait suffisamment grand (500 fois plus). Cependant avant la mesure, Lamb et Retherford "ne prenaient pas la possibilité d'une levée naturelle de la dégénérescence très au sérieux". Ce fut une première raison pour utiliser un champ B qui écarte les composantes Zeeman et augmente ainsi artificiellement le $\hbar\omega$ de la formule de Breit et Teller. Le champ magnétique permettait par ailleurs d'écarter du jet toute particule chargée qui troublerait le détecteur. Mais son avantage essentiel, comme il apparut au cours de l'exécution de l'expérience était la possibilité de le faire varier continûment : la condition de résonance n'est pas exacte, en raison de la largeur naturelle des niveaux P et de la sous-structure hyperfine (au total ~ 200 MHz), une mesure précise de fréquences de résonance nécessitait donc le tracé de la courbe de résonance et le pointé de son centre ; pour obtenir une telle courbe un balayage continu en fréquence n'était pas possible avec les générateurs de l'époque, mais le champ B , lui pouvait être balayé sans problème.

- Une incertitude assez lourde pesait sur la détection: le processus d'émission d'un électron à la surface d'un métal est complexe et son calcul peu sûr. Il est vrai que WEBB en 1924 l'avait utilisé pour le mercure métastable, mais les données pour l'hydrogène étaient inexistantes. Le signal obtenu par Lamb et Retherford se révéla bien plus faible que prévu, mais il était largement mesurable.

- Il n'est pas question de reproduire ici, même partiellement toutes les études théoriques de Lamb-Retherford visant à estimer les différents effets parasites. Entre autres le rôle de la structure fine fut soigneusement examiné.

Les raffinements ultérieurs abaissant la marge d'erreur à 1 MHz constituèrent un test important de la nouvelle électrodynamique quantique. Mais les premiers résultats de Lamb-Retherford présentaient à eux seuls une évidence expérimentale telle que les théoriciens devaient s'efforcer de l'expliquer. Ils y parvinrent aussitôt.

8.4 LE MOMENT MAGNETIQUE ANORMAL DE L'ELECTRON

L'électron possède un spin et corrélativement un moment magnétique (responsable de "l'effet ZEEMAN anormal"). Les premières déterminations expérimentales de ce moment magnétique aboutirent à la valeur $g \left(\frac{e}{2m} \right) \frac{\hbar}{2}$ avec $g = 2$. Cette valeur inattendue du facteur de Landé g (il vaut 1 pour le magnétisme orbital) fut justifiée par DIRAC en 1928 par son équation ; elle apparaît alors comme une conséquence de l'invariance relativiste. Les premières expériences qui établirent l'existence d'un écart "g-2" sont contemporaines de celles de Lamb. Toutefois leur lien avec la renaissance soudaine de l'électrodynamique quantique est beaucoup plus lâche. *"Historiquement la stimulation de l'expérience vint entièrement de la mesure du déplacement de Lamb ; l'évidence d'anomalies magnétiques ne reçut son interprétation correcte qu'en conséquence de la prédiction théorique d'un moment magnétique de spin additionnel"* affirme Schwinger⁽⁵⁸⁾. La première preuve expérimentale indirecte d'un magnétisme anormal de l'électron date de mai 1947, à peu près en même temps que les résultats de la première expérience de Lamb-Retherford. Mais les conclusions étaient beaucoup moins tranchées, une multitude d'interprétations restaient possibles dans cette situation expérimentale complexe. L'importance de l'évidence du g-2 à cette époque tient surtout à l'assurance supplémentaire qu'elle donna aux nouveaux théoriciens de l'électrodynamique quantique qu'ils étaient sur la bonne voie. La valeur $\frac{g-2}{2} = \frac{\alpha}{2\pi}$ était une des prédictions les plus fondamentales des calculs de Schwinger, en ce qu'elle concernait l'électron libre, indépendamment de l'existence d'un centre coulombien, contrairement au déplacement de Lamb.

8.4.a Possibilité théorique

La théorie de DIRAC de l'électron avait emporté un tel succès qu'il fallut du temps pour qu'on commençât à douter d'une de ses prédictions les plus nettes et les plus spectaculaires : la valeur 2 du facteur de Landé. Pourtant dès 1933, dans son article de "Handbuch⁽⁵⁹⁾" PAULI montra que l'exigence de Dirac du couplage minimal $(i\partial_\mu \rightarrow i\partial_\mu - eA_\mu)$ n'était pas nécessaire. L'équation de Dirac pouvait être modifiée en

$(i\cancel{\partial} - e\cancel{A} - m + \mu_e \frac{1}{2} \sigma_{\mu\nu} F^{\mu\nu}) \psi = 0$ ($\sigma_{\mu\nu} = \frac{i}{2} [\gamma_\mu, \gamma_\nu]$) sans violer l'invariance relativiste ni l'invariance de jauge, le terme supplémentaire modifiant de μ_e le moment magnétique de l'électron*. Seul le critère de simplicité permettait de préférer l'équation de Dirac à cette nouvelle équation. (Il ne pouvait être question à cette époque de renormalisabilité).

L'expérience ne donnait pas non plus de raison de s'intéresser à un moment magnétique anormal de l'électron. Il en allait tout autrement pour le proton et le neutron dont les moments magnétiques diffèrent des valeurs données par l'équation de Dirac par des facteurs respectifs $\gamma_p \approx 2.8$ et $\gamma_n \approx -1.9$. Dès sa naissance la théorie mésique de YUKAWA donna l'espoir d'attribuer une partie du moment magnétique au nuage de mésons virtuels qui entouraient le nucléon. FRÖHLICH HEITLER et KEMMER firent plusieurs tentatives de ce genre. Ayant exploré plusieurs types de mésons (scalaires, pseudoscalaires, vectoriels) et de couplages ils conclurent en 1938⁽⁶⁰⁾ que la théorie mésique vectorielle donnait les meilleurs résultats. Les auteurs calculaient l'énergie du nucléon dans un champ magnétique homogène B au 2^o ordre de la série des perturbations. Le terme linéaire en B donnait le moment magnétique. Bien sûr le résultat divergeait, mais une coupure

* On sait bien aujourd'hui que cette forme doit être utilisée pour tenir compte du moment magnétique anormal. Ce dernier peut être important pour des particules complexes ou soumises à d'autres types d'interaction (nucléaires). (Le proton et le neutron par exemple). Weisskopf⁽⁵⁸⁾, à propos de la mesure par Stern (1933) du moment magnétique du proton raconte une anecdote amusante et instructive quant à "l'arrogance" des physiciens : "il y eut un séminaire du groupe théorique de Göttingen et Otto Stern vint à Göttingen de Hamburg pour parler des mesures du moment magnétique du proton qu'il était sur le point d'achever. Il décrivit son appareil, mais sans donner le résultat. Il prit un morceau de papier et demanda à chacun de nous : "quelle est votre prédiction pour le moment magnétique du proton ?". Tous les théoriciens de Max Born à Victor Weisskopf dirent "bien sûr la grande force de l'équation de Dirac est qu'elle prédit un moment magnétique d'un magnéton de Bohr pour une particule de spin 1/2". Alors il nous demanda d'écrire les prédictions ... deux mois plus tard il revint pour parler de l'expérience achevée, la valeur était 2.8. Il projeta la feuille de nos prédictions sur l'écran. Ce fut une expérience lamentable".

raisonnable donnait de bonnes estimations. Il ne fallait retenir de ce genre de calcul que l'idée d'attribuer aux processus virtuels une part de moment magnétique : elle expliquait que les corrections à la valeur de Dirac soient à peu près de signe opposé pour le proton et le neutron, par ailleurs elle aurait pu être appliquée à l'électrodynamique pure si la nécessité s'en était fait ressentir.

8.4.b Premiers signes expérimentaux d'un g-2 non nul

Comme pour le déplacement de Lamb, on peut étudier rétrospectivement les mesures de facteur de Landé d'avant-guerre et chercher dans le flou des erreurs expérimentales une tendance au g-2. Par ce jeu quelque peu futile on découvre⁽⁶¹⁾ que lors d'une série de mesures destinées à améliorer la connaissance de $\frac{e}{m}$ à partir des figures Zeeman d'états singlet, KINSLER⁽⁶²⁾ en 1934 trouva pour la somme des facteurs de Landé de $2p^5 3s^1 P_1$ (Ne) et $2p^5 3s^3 P_1$ (Ne) la valeur 2.5017 ± 0.0016 au lieu de la valeur théorique 2.5 qui résultait de $g = 2$ pour l'électron ; avec la valeur aujourd'hui connue de g-2 la prédiction théorique est de 2.5012.

Nous avons déjà parlé de la méthode de résonance magnétique inventée par RABI et mise au point par RABI, MILLMAN, KUSCH et ZACHARIAS^(44,45), et dit qu'elle constituait un moyen très précis d'explorer la structure Zeeman des atomes et des molécules, et d'en déduire constantes de structure hyperfine et facteur de Landé*. Dans ce même laboratoire de Rabi à l'université de Columbia, ces méthodes ne cessèrent de s'améliorer, et profitèrent après la guerre de l'amélioration des techniques de radiofréquence. Les premières mesures de facteurs de Landé s'effectuaient à partir de la mesure de la fréquence des transitions Zeeman et du calibrage du champ magnétique. Cette dernière opération, à l'époque des premiers travaux de Rabi, Millman, Kusch et Zacharias s'effectuait à l'aide d'une bobine et d'un galvanomètre balistique avec une précision d'environ 0.2 %. En 1941 MILLMAN et KUSCH éliminèrent cette cause d'erreur en faisant le rapport entre la fréquence Zeeman étudiée et la fréquence Zeeman d'une transition radiofréquence calculable d'un atome

*de manière générale on appelle facteur de Landé le rapport entre moment magnétique et moment cinétique à $e/2m$ près.

alcalin dans le même champ magnétique. Ainsi les facteurs de Landé inconnus (en général nucléaires) étaient exprimés en fonction du facteur de Landé de l'électron, supposé égal à 2. La valeur obtenue dans cette hypothèse (inexacte) pour le facteur de Landé nucléaire de ${}^7_3\text{Li}$ était en désaccord de 0.1 % avec la valeur obtenue directement par calibration du champ magnétique⁽⁴⁵⁾. Lors de sa conférence Nobel (1955) Polykarp KUSCH rappelle pourquoi il n'avait pu conclure en 1941 : *"Voici peut-être un bon commentaire sur les hasards de la physique expérimentale : aucun effet significatif ne nous avait échappé, mais l'erreur dans l'inductance mutuelle que nous avons utilisée pour la calibration du champ magnétique était 0.2 % au lieu de 0.1 %"*.

8.4.c L'expérience de Nafe Nelson et Rabi

Le 19 mai 1947, NAFE NELSON et RABI⁽⁶⁵⁾ envoyèrent une lettre à la Phys. Rev. donnant les premiers résultats de leur étude de la structure hyperfine du niveau $2S_{1/2}$ de l'hydrogène et du deutérium. Pour ces atomes, les plus simples qu'il soit, la structure hyperfine est explicitement calculable. La théorie de Fermi⁽⁶⁴⁾ relie la structure hyperfine (en champ $B = 0$) au moment magnétique nucléaire μ_N et au magnéton de Bohr μ_0 suivant :

$$h\nu = \frac{8\pi}{3} \left(\frac{2I+1}{I} \right) \mu_N \mu_0 \psi^2(0) \quad (I \text{ spin nucléaire, } \psi \text{ fonction d'onde électronique}).$$

La mesure de ν donnait donc le moment magnétique μ_N du proton ou du deutéron. Les valeurs obtenues différaient de celles de Millman et Kusch⁽⁶³⁾ d'environ 0.2 % et cet écart valait cinq fois l'erreur probable sur les mesures. *"Seules de nouvelles expériences pourront décider si le désaccord entre la théorie et l'expérience est dû à quelque facteur inconnu dans la théorie de l'atome d'hydrogène (c'était bien ça !) ou simplement à une erreur sur l'estimation de l'une des constantes naturelles comme α "*. (Nafe Nelson Rabi)⁽⁶⁵⁾. Les résultats furent confirmés en septembre 1947 par NAGLE, JULIAN et ZACHARIAS⁽⁶⁶⁾ qui virent *"un violent désaccord avec la formule de Fermi"* puis en décembre de la même année par Nafe et Nelson⁽⁶⁷⁾ après quelques améliorations apportées à leur expérience. Dans cette dernière publication, ils soulignaient la multiplicité d'interprétations de ce désaccord : on pouvait invoquer le calcul de Schwinger (publié entre-temps

mais aussi éventuellement une erreur dans la valeur de α ou du rydberg, un effet de polarisation du vide, un effet de déplacement de niveau, une conséquence possible de la théorie mésique, un effet relativiste d'entraî-
nement du noyau ... Enfin une confusion supplémentaire était introduite par le fait que le rapport $\frac{v_H}{v_D}$ ne concordait pas avec la valeur théorique, et là on ne pouvait accuser la valeur de α ou de R_∞ , ni celle de μ_V . A. BOHR⁽⁷²⁾ résolut plus tard la question en montrant que cet effet résultait du centrage de la fonction d'onde sur le proton à l'intérieur du deutéron.

BREIT fut le premier dans ce dédale d'interprétations à opter pour une modification du moment magnétique de l'électron* (lettre à Phys. Rev. du 29 septembre 1947)⁽⁶⁸⁾. Il reprenait le vieil argument de Pauli de la possibilité d'un terme supplémentaire $\mu_e \frac{1}{2} \sigma^{\mu\nu} F_{\mu\nu}$ pour l'opérateur de Dirac. Ce terme conduit à substituer $\mu_0 \mu_e$ à μ_0 ($\mu_0 > 0$ magnéton de Bohr) dans l'expression des moments magnétiques électroniques et selon Breit $\mu_0 + \frac{\mu_e}{2}$ à μ_0 dans l'expression de la structure hyperfine de l'hydrogène et du deutérium. Cette dernière déduction était fautive, comme le nota SCHWINGER,⁽⁶⁹⁾ en raison d'une banale erreur de calcul ; la valeur empirique de $\frac{\mu_e}{\mu_0}$ estimée par Breit était nécessairement fautive : - 0.0056 au lieu de - 0.0012. Pour comprendre, sachant que $\frac{\mu_e}{\mu_0} \approx 0.1\%$, l'écart de 0.2 % trouvé par Rabi Nafe et Nelson il faut voir que μ_0 intervient deux fois dans le calcul de $h\nu$ par la formule de Fermi : une fois explicitement, une fois à travers μ_N puisque Kusch et Millman avaient mesuré $\frac{\mu_N}{\mu_0}$ plutôt que μ_N . BREIT publia lui-même un erratum⁽⁷⁰⁾ après avoir pris connaissance de la critique de Schwinger, puis une étude⁽⁷¹⁾ détaillée des conséquences de l'existence d'un moment anormal. Dès sa première lettre sur ce sujet Breit avait vu que le terme $i\vec{\alpha} \cdot \vec{E}$ contenu dans $\frac{1}{2} \sigma^{\mu\nu} F_{\mu\nu}$ contribue un peu au déplacement de Lamb.

8.4.d L'expérience de Foley et Kusch

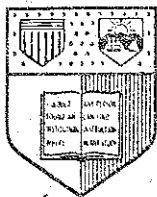
L'hypothèse de Breit fut prise au sérieux par RABI qui imagina de tester une conséquence plus directe de l'existence d'un moment magnétique anormal

* Breit appela cette modification moment magnétique "intrinsèque" de l'électron contrairement à la terminologie en usage aujourd'hui. Le moment magnétique normal de l'électron de Dirac était vu comme conséquence de la "Zitterbewegung"

de l'électron : elle devait affecter la valeur d'un rapport de facteurs de Landé bien choisi. Les expériences furent réalisées par FOLEY et KUSCH. La première expérience (nov 1947) ⁽⁷³⁾ - "les auteurs expriment leur reconnaissance aux professeurs Breit et Rabi pour leurs suggestions qui stimulèrent ces expériences" - portait sur le gallium (isotopes 69 et 71) et donnait pour le rapport $\frac{g_J(^2P_{3/2})}{g_J(^2P_{1/2})}$ une valeur différente de la valeur théorique 2 obtenue en prenant les valeurs conventionnelles des moments magnétiques orbital et de spin de l'électron, et en faisant l'hypothèse d'un couplage \vec{L}, \vec{S} pur (Russel-Saunders)*. Pour obtenir le rapport cherché avec la précision requise il faut tenir compte de la structure hyperfine, et heureusement Foley et Kusch disposaient des études de cette structure effectuées sur Ga ⁶⁹ et Ga ⁷¹ par BECKER et KUSCH ⁽⁷⁴⁾. Toutes ces expériences utilisaient la méthode de résonance sur jet atomique suivant le dispositif original : polarisation par gradient de champ magnétique, interaction avec la radio-fréquence dans une zone de champ magnétique homogène puis détection des transitions provoquées par la RF à l'aide d'un gradient de champ de signe opposé au premier. Foley et Kusch aboutirent à $g_s - 2 = 0.00229 \pm 0.00008$.

Il restait un point suspect : la validité de l'hypothèse d'un couplage \vec{L}, \vec{S} pur. Pour s'en assurer FOLEY et KUSCH ⁽⁷⁵⁾ entreprirent une autre série de mesure faisant intervenir un autre atome, le radium, pour lequel les interactions de configuration, si elles intervenaient devaient être différentes. La mesure de $g_J(Na\ ^2S_{1/2})/g_J(Ca\ ^2P_{1/2})$ s'écartait encore de la valeur théorique 3 et donnait $g_s - 2 = 0.00244 \pm 0.00006$, résultat compatible avec le précédent. La description détaillée de ces expériences fut publiée l'année suivante ⁽⁶¹⁾. La deuxième lettre ⁽⁷⁵⁾ date du 26 décembre 1947 et la lettre de Schwinger donnant $\frac{g-2}{2} = \frac{\alpha}{2\pi} \approx 0.00116$, quatre pages plus loin dans la Phys. Rev., date du 30 décembre 1947. Schwinger avait eu connaissance des dernières mesures de Kusch et Foley qui permettaient une vérification plus précise de sa prédiction que celles de Nafe Nelson et Rabi.

* Si en première approximation on néglige le moment magnétique nucléaire, le hamiltonien d'interaction paramagnétique s'écrit $H = g_L \mu_o L \cdot B + g_S \mu_o S \cdot B$. Dans une multiplicité de J donné ($\vec{J} = \vec{L} + \vec{S}$) $H_J = g_J \mu_o J_z B$ ($B // O_z$) avec $g_J = \alpha_L g_L + \alpha_S g_S$ et α_L et α_S donnés par le théorème de Wigner-Eckart (dans le cas du couplage \vec{L}, \vec{S} pur où l'on ne prend en compte que l'électron de valence seul) ne dépendent que de J, L et S. Si l'on prend $g_L = 1$ (ce que Schwinger justifia plus tard) la donnée d'un rapport de g_J (déduite d'un rapport de fréquences Zeeman) fixe la valeur de g_S .



Floyd R. Newman
Laboratory of Nuclear Studies
Cornell University

[Voir page 351]

Newman Lab. Cornell Univ.
Ithaca, New York 14853
607-256-4397

December 7, 1981

Mr. Olivier Darrigol
83 rue Broca
75013 Paris
France

Dear Mr. Darrigol:

I have your letter of October 28, 1981, and will try to answer your main questions.

Shelter Island is a small island at the east end of Long Island. The conference there took place in June 1947. Most of the members were about my age or slightly younger, some were about ten years older, especially I.I. Rabi and Kramers. At least three important members were about ten years younger than I, namely, Schwinger, Feynman and Marshak. The total number of physicists at the conference were about thirty, nearly all of them theorists. Kramers gave a talk at the conference on his ideas of Mass Renormalization. The infinite self-energy, he said, is unobservable because it is part of the mass of the particle. Only when the particle is accelerated will there be any observable effects, namely, the difference of the self-energy of the accelerated particle and one at rest. Already during the conference I decided I wanted to try and calculate the Lamb-shift using Kramer's idea. I do not know whether Kramers had published anything on the subject before 1947, or after.

I did not go to the Solvay conference in 1948 and did not submit any paper to that conference. This explains why there is no paper by me in the proceedings of the conference.

There were two repetitions of the Shelter Island conference, one at Pocono in 1948, and one at Old Stone-on-Hudson in 1949, both arranged by Oppenheimer. In 1950 these conferences were replaced by the Rochester conferences, arranged by Marshak. The Rochester conferences started with about 75 and then grew to more than 100 participants. After about 1960, they became international and still bigger.

If you want any further information, please write me again. I am not planning any trip to Europe until 1984.

Yours sincerely,

H. A. Bethe

Hans A. Bethe

HAB/ns

8.5 LA CONFERENCE DE L'ILE DE SHELTER

Elle eut lieu du 2 au 4 juin 1947. Elle marque une étape importante dans l'histoire de la théorie quantique des champs. Expérimentateurs du groupe de Rabi et théoriciens de l'électrodynamique quantique s'y trouvèrent confrontés. Nous n'avons d'autre document* sur le déroulement de cette conférence que la lettre que BETHE nous a aimablement adressée. Nous la reproduisons ici intégralement. Nous ferons seulement remarquer que les idées de Kramers semblent avoir eu peu d'impact avant cette conférence puisque ses publications concernant l'électrodynamique étaient ignorées de Bethe. Si l'on en croit la rumeur populaire le premier calcul de Bethe du déplacement de Lamb aurait été effectué dans le train qui le ramenait de New York.

* Le professeur SCHWEBER nous a récemment fait parvenir un tiré à part intitulé : "Some notes for a history of quantum field theory" donnant des informations plus précises sur les grandes conférences américaines de 1947 (Note rajoutée après la rédaction de cette thèse).

8.6 LES PREMIERS CALCULS FINIS DE CORRECTIONS RADIATIVES

En juin 1947 les résultats de l'expérience de Lamb-Retherford furent rendus publics. Ce fut le point de départ de toute une série de calculs théoriques puis d'une reformulation de l'électrodynamique quantique. Les dirigeants, particulièrement aux Etats-Unis, avaient pris conscience pendant la guerre de l'importance du pouvoir scientifique et les conditions matérielles de la recherche s'améliorèrent. La plupart des théoriciens américains et certains immigrés avaient abandonné pendant la guerre l'orientation initiale de leurs travaux, pour mettre au point l'arme nucléaire ou pour perfectionner la technique du radar. A l'issue de cette sombre période les physiciens connurent un enthousiasme nouveau pour réattaquer d'anciens problèmes fondamentaux, comme celui des divergences de l'électrodynamique. Les nouvelles investigations de ce problème eurent lieu pour la plupart aux Etats-Unis où s'était regroupée une grande partie de la communauté internationale des physiciens*, et ils furent presque tous publiés dans la "Physical Review" ou présentés aux congrès de la société américaine de physique. Toutefois les développements ingénieux réalisés indépendamment dans ce domaine par les Japonais constituent une exception notoire, et nous consacrerons un paragraphe à leur étude.

Si l'on se limite seulement à la période qui suit immédiatement la conférence de l'île de Shelter (disons jusqu'à la fin de 1948) on est étonné de la profusion des travaux en électrodynamique quantique, de la multiplicité des approches (même si, dans certains cas, elles s'avèrent équivalentes) de l'ardeur des théoriciens à se lancer dans des calculs effrayants. Pour le calcul du déplacement de Lamb une compétition (amicale) s'établit entre les tenants des anciennes méthodes de calcul comme BETHE, WEISSKOPF et LAMB et ceux d'une reformulation de la théorie comme SCHWINGER et FEYNMAN. L'étude des travaux de TOMONAGA, SCHWINGER, FEYNMAN et DYSON qui sont à l'origine

* Citons par exemple Weisskopf au MIT, Bethe et Feynman à Cornell, Schwinger à Harvard, Lamb à Columbia, Oppenheimer à l'Université de California.

de la version moderne de l'électrodynamique quantique se prolongerait bien au-delà des limites de ce travail. Nous n'aborderons ici que les calculs effectués dans le formalisme qui existait dès 1934. Seule manquait l'idée de renormalisation de masse, et notre but est de montrer ici comment cette idée de Kramers prit tardivement corps pour extraire de l'ancienne théorie des perturbations des résultats finis, au moment où l'expérience l'exigeait.

8.6.a Les déplacements de niveau :

Le premier calcul de BETHE⁽⁷⁶⁾ : (27 juin 1947)

Selon la tradition, BETHE fut en mesure d'interpréter les 1000 MHz du déplacement de Lamb dans le train qui le ramenait de l'île de Shelter. Cette première évaluation de l'effet est non relativiste. Elle diffère surtout* de la première estimation divergente d'Oppenheimer⁽¹⁹⁾ (1930) par l'utilisation d'une renormalisation de la masse de l'électron.

Le hamiltonien du système atome + rayonnement s'écrit $H = H_0 + H_I$

$$\text{avec } H_0 = \frac{p^2}{2m} + eV(\vec{r}) + \frac{1}{2} \int (\vec{E}_I^2 + B^2) d^3x \quad \text{et } H_I = -\frac{e}{m} \vec{p} \cdot \vec{A}(\vec{r}) \left(+ \frac{e^2}{2m} A^2(\vec{r}) \right)$$

L'état non perturbé noté $|n\rangle$ correspond à l'état atomique de nombres quantiques n et au vide de rayonnement; il a l'énergie $E_n^{(0)}$. Le deuxième terme du hamiltonien d'interaction donne un déplacement constant, on peut donc le négliger.

La correction au deuxième ordre de l'énergie $E_n^{(0)}$ est

$$E_n^{(2)} = -\frac{2}{3\pi} \frac{\alpha}{m^2} \int_0^\infty k dk \sum_m \frac{\langle m | \vec{p} | m \rangle \langle m | \vec{p} | n \rangle}{E_m + k - E_n}$$

si l'on néglige l'effet de retard c'est-à-dire les exponentielles dans le développement de \vec{A} . Ce terme correspond à l'émission suivie de l'absorption d'un photon virtuel, le couplage étant proportionnel à $\langle m | \vec{p} | m \rangle \frac{1}{\sqrt{k}} \frac{e}{m}$. Comme nous l'avons vu au chapitre 3 cette expression de $E_n^{(2)}$ diverge linéairement dans la coupure K .

* L'autre différence, utilisation du hamiltonien de Dirac de l'électron au lieu du hamiltonien classique est inessentielle, d'autant que sans la théorie des trous Oppenheimer ne pouvait prétendre à un traitement correct des effets relativistes.

Mais, souligne Bethe, $E_n^{(2)}$ contient l'effet divergent de la masse électromagnétique ; pour soustraire cet effet Bethe admet qu'il doit être le même pour un électron libre de même énergie cinétique moyenne que l'état $|n\rangle$, c'est-à-dire d'impulsion \vec{p} telle que $\frac{p^2}{2m} = \langle n | \frac{p^2}{2m} | n \rangle$

L'énergie au deuxième ordre d'un tel électron libre est, compte-tenu du fait que l'impulsion \vec{p} est diagonale dans la base des $|\vec{p}\rangle$

$$E_p^{(2)} = -\frac{2}{3\pi} \frac{\alpha}{m^2} \int_0^{\infty} k dk \frac{p^2}{k^2} \quad ; \text{ utilisant la relation de fermeture } p^2 = \sum_m \langle m | \vec{p} | m \rangle \langle m | \vec{p} | m \rangle \quad \text{et la partie observable}$$

du déplacement de niveau s'écrit :

$$W_m = E_m^{(2)} - E_p^{(2)} = \frac{2}{3\pi} \frac{\alpha}{m^2} \sum_m k_m |\langle \vec{p} | m \rangle|^2 \ln \frac{K}{|E_m - E_m|}$$

l'intégration sur k une fois effectuée.

Le résultat diverge encore, mais seulement logarithmiquement. Le caractère non relativiste du calcul conduit à couper la fréquence des photons virtuels à $K \approx m$. Bethe pensait (à juste titre) que le calcul entièrement relativiste devait être convergent, puisque l'énergie propre d'un électron, comme l'avait démontré Weisskopf, diverge logarithmiquement et non plus linéairement dans la théorie des trous, et par conséquent la contribution des photons virtuels de longueur d'onde inférieure à celle de Compton ($\frac{h}{mc}$) devait être négligeable en première approximation. La coupure $K = mc^2$ donne numériquement pour $W_{2S_{1/2}} - W_{2P_{1/2}}$ de l'hydrogène la valeur 1040 MHz en bon accord avec l'expérience de Lamb-Retherford.

On peut dire que Bethe a eu de la chance que l'accord soit aussi bon : il n'était pas évident que le calcul relativiste ne modifierait que de quelques 10 MHz (1051 MHz) son estimation. (La polarisation du vide contribuait déjà de - 27 HHz). Mais ses conclusions, entièrement exactes, marquaient un progrès définitif dans la compréhension de l'électrodynamique quantique :

"(1) le déplacement de niveaux dû à l'interaction avec le rayonnement est un effet réel et de grandeur finie.

(2) l'effet de la masse électromagnétique infinie d'un électron ponctuel peut être éliminé par une identification convenable des termes dans la théorie du rayonnement de Dirac.

(3) une investigation précise, expérimentale et théorique, du déplacement de niveaux peut établir des effets relativistes (par exemple la théorie des trous de Dirac). Ces effets seront de l'ordre de l'unité en comparaison au $\ln K$.

Avant d'aborder les calculs relativistes du déplacement de Lamb, nous jugeons utile de présenter les différentes méthodes de renormalisation de masse. Le lecteur averti a déjà pu constater que la prescription de Bethe diffère de celle enseignée aujourd'hui le plus couramment. Dorénavant nous utiliserons les notations suivantes pour les différentes masses : m_0 désigne la masse nue (dite "mécanique") qui intervient dans le hamiltonien de départ, la masse physique notée m se compose de m_0 et d'une contribution δm électromagnétique due aux particules virtuelles créées par l'interaction électromagnétique au voisinage de l'électron. Les différentes prescriptions de soustraction ainsi que la raison de leur équivalence sont clairement exposées par FEYNMAN dans son article sur la coupure relativiste⁽⁷⁷⁾ dont nous reparlerons bientôt. L'énergie au deuxième ordre (en e) du niveau n calculée avec le hamiltonien usuel de l'électrodynamique quantique où intervient la masse nue m_0 s'écrit $\mathcal{E}_m = E_m^{(0)}(m_0) + E_m^{(2)}(m_0)$; cette énergie doit être exprimée à l'aide de la masse expérimentale pour être comparée à sa valeur empirique. Supposons δm connu jusqu'au deuxième ordre en e , et faible par rapport à m_0 (ce qui suppose l'utilisation d'une coupure suffisamment basse, $K \ll m_0 c^2$ dans la théorie relativiste) alors

$$\mathcal{E}_m = E_m^0(m) - \delta m \frac{\partial E_m^{(0)}}{\partial m} + E_m^{(2)}(m) + \text{ordres supérieurs}^*$$

Le déplacement est donné par $\mathcal{E}_m - E_m^0(m) = W_m = E_m^{(2)}(m) - \delta m \frac{\partial E_m^{(0)}}{\partial m}$

Si tout se passe bien, la divergence en K de $E_n^{(2)}(m)$ est compensée par celle de δm .

Pour déterminer δm il suffit de remarquer que l'énergie propre d'un électron libre d'impulsion \vec{p} est entièrement due à un effet de masse électromagnétique :

* Dans cette expression et dans celles qui suivent nous avons utilisé ou sous-entendu m comme argument à la place de m_0 chaque fois que l'ordre d'approximation le permettait.

$$E_p^{(0)}(m_0) + E_p^{(2)}(m_0) = E_p^{(0)}(m) = E_p^{(0)}(m_0) + \delta m \frac{\partial E_p^{(0)}}{\partial m}$$

d'où $\delta m = E_p^{(2)}(m_0) / \frac{\partial E_p^{(0)}}{\partial m}$ **

(α) - La prescription utilisée le plus couramment aujourd'hui*** consiste à écrire le hamiltonien sous la forme $H(m_0) = H(m) - \frac{\partial H}{\partial m} \delta m$ le deuxième terme étant appelé "contre-terme de masse" ($-\beta \delta m$ dans la théorie relativiste) et δm étant déterminé par la formule précédente : le contre-terme de masse est considéré comme une perturbation (au premier ordre puisque δm est du deuxième ordre en e) dans le calcul de $E_n^{(2)}$. D'après la théorie des perturbations du premier ordre $\frac{\partial E_n^{(0)}}{\partial m} = \langle m | \frac{\partial H}{\partial m} | m \rangle$ et l'on retrouve bien l'expression souhaitée de W_n .

(β) - La prescription de Bethe revient à soustraire l'énergie propre "d'un paquet d'ondes libres de même forme que l'état $|n\rangle$ " c'est-à-dire

$$W_m = E_m^{(2)}(m) - \sum_{\vec{p}} |\langle m | \vec{p} \rangle|^2 E_p^{(2)}(m_0).$$

Utilisant

$$E_p^{(2)}(m_0) = \delta m \frac{\partial E_p^{(0)}}{\partial m} = \delta m \langle \vec{p} | \frac{\partial H}{\partial m} | \vec{p} \rangle$$

il vient $W_m = E_m^{(2)}(m) - \delta m \langle m | \frac{\partial H}{\partial m} | m \rangle$, en accord avec la prescription précédente.

(γ) - On peut aussi, suivant Epstein⁽⁷⁸⁾**** (novembre 1947) calculer directement $\frac{\partial E_m^{(0)}}{\partial m}$. (Epstein utilise cette méthode pour les sections efficaces de diffusion ; alors une modification de masse δm induit une variation $\delta m \frac{\partial \sigma^{(0)}}{\partial m}$ de la section efficace).

* Précisons que \vec{p} représente l'impulsion de l'électron m avant branchement de l'interaction électromagnétique, mais aussi l'impulsion totale du système électron + rayonnement après branchement de l'interaction, puisque le hamiltonien total commute avec l'impulsion totale et que l'état non perturbé était un état propre de l'impulsion totale.

** Pour un électron au repos dans la théorie relativiste $E_0^{(0)}(m) = m$ et donc $\delta m = E_0^{(2)}$

*** Il est difficile d'en trouver l'origine exacte. Sa première apparition explicite et claire est dans une lettre de Koba et Tomonaga datant du 30 décembre 1947.

**** Le même procédé fut utilisé en mars 1948 indépendamment par Koba et Tomonaga⁽⁹⁷⁾ (P. 296)

(δ) - Enfin on peut donner au contre-terme de masse $-M = -\frac{\partial H}{\partial m} \delta m$ une autre forme en remarquant que δm est déterminé au deuxième ordre par $\langle \vec{p}, s | \delta m \frac{\partial H}{\partial m} | \vec{p}, s \rangle = \langle \vec{p}, s | H_I \frac{1}{E_p - H_0} H_I | \vec{p}, s \rangle$ et que si cette relation est vérifiée on a aussi $\langle \vec{p}, s' | M | \vec{p}, s \rangle = \langle \vec{p}, s' | H_I \frac{1}{E_p - H_0} H_I | \vec{p}, s \rangle$. s dénote la variable de spin (démontré par Feynman⁽⁷⁷⁾). Autrement dit l'action de M sur un état $|\vec{p}, s\rangle$ est la même que celle de $H_I \frac{1}{E_p - H_0} H_I$. L'avantage de cette méthode est de permettre de modifier l'intégrand de $E_n^{(2)}$ terme à terme avant d'effectuer l'intégration sur les moments*. Ainsi, formellement tout au moins, le nouveau hamiltonien H-M donne des résultats finis au deuxième ordre de la série des perturbations. Ce genre de contre-terme M correspond justement à ceux introduits par Schwinger dans ses recherches en 1947. (Voir plus loin).

Calculs relativistes du déplacement de Lamb

L'extension relativiste du premier calcul de Bethe présentait deux types de difficulté ; la soustraction de l'effet de la renormalisation de masse était ambiguë, comme chaque fois que l'on soustrait deux quantités infinies. Par ailleurs, on avait besoin dans la plupart des premiers calculs de raccorder la contribution au déplacement des photons virtuels de petite longueur d'onde et de ceux de grande longueur d'onde. La première de ces contributions ne tenait compte que de termes linéaires dans le potentiel coulombien, dont l'effet domine si l'énergie des photons virtuels est très grande devant l'énergie de liaison de l'électron ($k > K \gg \frac{e^2}{a}$), la seconde ($k < K$) était obtenue par le calcul non relativiste de Bethe tenant compte de l'effet de l'interaction coulombienne à tous les ordres. Les approximations dans les deux domaines $k > K$ et $k < K$ étaient justifiées à condition de prendre $\frac{e^2}{a} \ll K \ll mc^2$. Le raccord des deux contributions ne présentait pas de difficultés à condition que la limitation des deux domaines soit obtenue effectivement par une même coupure K. Ce n'était pas le cas dans l'article de Feynman de juillet 48 (Voir à propos de la diffusion) où

* Mais la définition de M est ambiguë : elle dépend du type de coupure choisi, d'une éventuelle transformation canonique dans l'espace des états ..., puisque sa limite pour une valeur infinie de la coupure diverge.

la suppression de l'effet divergent des photons mous dans la première contribution était obtenue en donnant une masse au photon. Feynman donna une correspondance incorrecte entre K et cette masse, d'où une légère erreur dans le résultat final.

Dans les premiers travaux de BETHE, DYSON et WEISSKOPF et SCHWINGER^(80 à 82) annoncés aux rencontres annuelles de la Société Américaine de physique du 29 au 31 janvier 1948 on ne trouve pas cette dernière erreur, mais la première difficulté, celle de l'ambiguïté des soustractions était présente. L'énergie perturbée trouvée par ces auteurs n'était pas covariante relativiste, ce qui laissait penser que la méthode de soustraction naïve dans le formalisme hamiltonien conventionnel était incorrecte. La contribution du moment magnétique anormal de l'électron au déplacement de niveau devait être artificiellement corrigée. (Voir Schwinger⁽⁷⁹⁾ note 5). Probablement pour cette raison et aussi par manque de temps ces calculs ne furent pas publiés in extenso. Le résultat était théoriquement suspect malgré le bon accord avec l'expérience. Le calcul de Bethe⁽⁸⁰⁾ utilisait la prescription (β). Bethe remarquait que les niveaux $L = 0$ comme $2P_{1/2}, 2P_{3/2}$ de l'hydrogène étaient très légèrement déplacés dans la théorie relativiste.

($\Delta\nu(2P_{1/2}) = -17\text{MHz}$, $\Delta\nu(2P_{3/2}) = +8\text{MHz}$, suivant French et Weisskopf). DYSON⁽⁸¹⁾, alors étudiant de Bethe, reprit le calcul dans le cas où les électrons sont décrits (de manière purement formelle) par des scalaires de Pauli et Weisskopf ; il trouva un ordre de grandeur voisin de celui du déplacement de Lamb. SCHWINGER et WEISSKOPF⁽⁸²⁾, utilisant deux méthodes différentes obtinrent des résultats analogues à ceux de Bethe. Weisskopf utilisait la théorie des perturbations conventionnelle assortie de la prescription (δ) et Schwinger opérait d'après "une modification du hamiltonien pour ramener la masse électromagnétique de l'électron dans la masse mécanique". Nous préciserons la nature de cette modification dans le paragraphe sur l'école japonaise (sans décrire son adaptation ultérieure à la théorie covariante des perturbations).

Nous avons dit combien l'expérience de Lamb-Retherford avait précipité les nouveaux développements de la théorie quantique des champs. Cependant il serait faux de croire, qu'aucun théoricien avant 1947 ne songeait à calculer le déplacement des niveaux atomiques par des méthodes d'électro-

dynamique pure. Kramers était très près du résultat. WEISSKOPF, après la guerre chercha à calculer cet effet avec son élève Bruce FRENCH, en cherchant à "isoler l'énergie propre infinie de l'électron". Weisskopf a donné lui-même un compte-rendu de ces recherches dans "My life as a Physicist"⁽⁸³⁾ : "C'était un calcul difficile parce que les techniques de renormalisation n'étaient pas encore développées ; nous devions calculer la différence entre l'énergie de l'électron libre et celle de l'électron lié, les deux étant infinies. Nous devions être très prudents, car prendre la différence de grandeurs divergentes est risqué. Mais nous nous frayâmes un chemin à travers les difficultés, lentement car il n'y avait pas de résultats expérimentaux vraiment bons. Alors un jour Lamb et Retherford firent une bonne expérience, et nous travaillâmes plus intensément, ainsi que beaucoup d'autres gens. Finalement nous obtînmes un résultat en excellent accord avec l'expérience. J'en fis part à Julian Schwinger et à Dick Feynman. Schwinger était de l'autre côté de la rue à Harvard et Feynman était à Cornell à ce moment-là : ils reprirent nos calculs mais n'obtinrent pas le même nombre que nous Je dis à French "Eh bien la probabilité est grande qu'ils aient raison et nous tort". Nous reportâmes notre publication afin de trouver l'erreur et cherchâmes pendant six mois. Entre-temps Lamb et Kroll publièrent leurs calculs du même effet qui étaient à peu près identiques aux nôtres. Alors je reçus un coup de téléphone d'Ithaca; c'était Feynman "Vous avez raison, j'ai tort !".

Bien qu'elle n'ait été publiée qu'en décembre 1948, nous présentons avant celle de Kroll et Lamb, l'étude de FRENCH et WEISSKOPF⁽⁸⁴⁾ du déplacement des raies spectrales. Le hamiltonien utilisé se composait du hamiltonien non perturbé $H_0 = \int \psi^\dagger (\vec{\alpha} \cdot \vec{p} + \beta m) \psi d^3x + \frac{1}{2} \int (E_1^2 + B^2) d^3x + e \int \psi^\dagger V \psi d^3x$ et de la perturbation $H_I = \int \psi^\dagger (-e \vec{\alpha} \cdot \vec{A}) \psi d^3x + \frac{e^2}{8\pi} \int \frac{\psi^\dagger(\vec{x}) \psi^\dagger(\vec{x}')}{|\vec{x} - \vec{x}'|} d^3x d^3x'$

Il était encore d'usage en 1947 de séparer la contribution des modes longitudinaux du champ électromagnétique sous forme d'interaction coulombienne, ainsi que l'avait fait Oppenheimer, puis Weisskopf dans leurs études d'énergies propres. Cette façon de procéder introduisit une lourdeur inutile dans les calculs, sans compter la confusion qu'elle entraîna dans le calcul de Dancoff dont nous reparlerons. Néanmoins French et Weisskopf vinrent à

bout de toutes les difficultés de calcul et ils trouvèrent même une méthode pour lever l'ambiguïté sur la définition de l'opérateur M (de la prescription (δ)) et sur la soustraction : ils exigèrent pour cela que l'énergie perturbée en présence de champs statiques électrique et magnétique ait une forme covariante. Dans ce cadre plus général où B est non nul, les auteurs retrouvèrent la valeur $\frac{g-2}{2} = \frac{\alpha}{2\pi}$ donnée par Schwinger.

Le calcul de KROLL et LAMB⁽⁸⁵⁾, datant d'octobre 1948, fut le premier calcul relativiste publié du déplacement de Lamb. Il présente le défaut par rapport au précédent de ne pas lever complètement l'ambiguïté de la soustraction, il reste une liberté dans le choix de g-2 qui en dernier recours doit être donné par l'expérience. La méthode utilisée par Kroll et Lamb pour soustraire la partie infinie de l'énergie perturbée diffère un peu de la méthode (α). La perturbation à l'ordre e² des valeurs propres E_n⁽²⁾(m₀) (calculées sans renormalisation de masse) s'écrit :

$$E_n^{(2)}(m_0) = \frac{\alpha}{\pi} \langle m | [(\ln \infty) \beta + (\ln \infty) eV + \frac{\vec{\alpha} \cdot \vec{p}}{6} - \frac{i}{\hbar} \frac{e}{m} \vec{\alpha} \cdot \vec{\nabla} V + \frac{cte}{m^2} \Delta V] | m \rangle$$

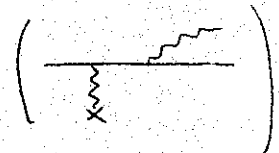
Une renormalisation naïve de masse et de charge élimine les deux premiers termes divergents. Mais l'expression résultante pour E_n⁽²⁾(m₀) n'est pas covariante en raison du terme $\vec{\alpha} \cdot \vec{p} / 6$. En effet pour V = 0 et |n> = |p̄> ce terme donne $\frac{\alpha}{6\pi} \vec{p} \cdot \vec{v}$ et la relation impulsion énergie E² + p² = cte indépendante de p ne peut être préservée même en modifiant la masse. Il faut en conclure que la prescription (α) jointe à une théorie des perturbations non covariante donne des résultats erronés. Pour sortir de l'impasse Kroll et Lamb décident de soustraire à la valeur E_n⁽²⁾ dans un potentiel, la valeur libre obtenue pour V = 0 (on revient à la méthode de Bethe(β)). Alors il ne reste plus que les deux derniers termes : celui en ΔV revient à introduire un potentiel effectif, celui égal à $-\frac{i}{\hbar} \alpha \frac{e}{m} \vec{\alpha} \cdot \vec{\nabla} V$ correspond à l'introduction d'un moment de Pauli $\alpha/2\pi \mu_B$ (l'interaction étant en $\sigma_{\mu\nu} F^{\mu\nu}$). Malheureusement on est encore libre de modifier le hamiltonien électronique libre de telle sorte que sa valeur moyenne dans un état libre reste la même. Cette liberté comme le montrent Kroll et Lamb revient à modifier le moment magnétique anormal. Il faut donc supposer ce dernier fixé par l'expérience (où par une méthode de calcul plus convenable) pour lever l'ambiguïté. En revanche le calcul de Lamb et Kroll présentait l'avantage de justifier l'utilisation du calcul de Bethe pour les photons

virtuels mous : l'expression de Bethe était réétablie en prenant une approximation non relativiste de la forme relativiste exacte de la contribution des photons mous.

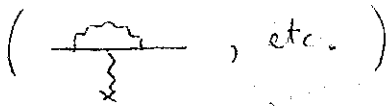
Parmi les autres calculs du même effet il faut citer celui exposé par Kramers⁽¹²⁾ au congrès Solvay 1948 (voir § 2). Comme nous l'avons dit la méthode de Kramers permettait, dans le cas non relativiste de travailler avec un hamiltonien contenant les paramètres physiques dès le départ. FEYNMAN⁽⁷⁷⁾ en juillet 1948 publia une étude des corrections radiatives à la diffusion coulombienne dont nous reparlerons bientôt. Pour la contribution des photons de grande impulsion, il suffit de substituer l'interaction effective obtenue en étudiant le problème de diffusion, à l'interaction coulombienne pure, pour obtenir le déplacement d'énergie d'un état lié. Le déplacement total est obtenu en raccordant au résultat de Bethe. Bien que Feynman se soit trompé en effectuant ce dernier raccord, sa méthode présentait l'avantage d'éliminer les ambiguïtés par une procédure différente de celle de French et Weisskopf : disons déjà qu'il utilisait une méthode de régularisation invariante relativiste , toujours toutefois dans l'ancienne théorie des perturbations. Le pas suivant fut franchi par SCHWINGER (et un peu plus tard par Feynman lui-même) : il fallait condamner l'ancienne théorie des perturbations en faveur d'une autre, covariante relativiste. Schwinger publia en janvier 1949⁽⁸⁶⁾ l'application de sa nouvelle méthode au calcul du déplacement.

8.6.b Les corrections radiatives à la diffusion coulombienne :

Parmi les exemples de calculs divergents en électrodynamique quantique nous avons cité (au chap.3 §3) les calculs de DANCOFF⁽⁸⁷⁾ de corrections à la diffusion coulombienne en e^2 . Le calcul de Dancoff était essentiellement motivé par les premières explorations du problème infra-rouge (voir chap.5 §2) par Bloch, Nordsieck, Pauli et Fierz. On savait avant Dancoff que la divergence infra-rouge des processus réels



au cours desquels un photon mou est présent dans l'état final, est

compensée par celle des processus virtuels ( , etc.)

où un photon mou virtuel est émis puis réabsorbé. Pour des photons d'énergie supérieure à celle de l'électron incident, il n'existe pas de processus réel du même type ; en revanche les processus virtuels sont toujours possibles et doivent apporter une correction à la section efficace de diffusion. En 1939 Dancoff put démontrer que cette correction divergeait logarithmiquement. A la fin de 1947 Oppenheimer demanda à H. W. LEWIS⁽⁸⁸⁾ d'examiner si une telle divergence pouvait être absorbée par renormalisation de masse*. Si l'on faisait confiance au calcul de Dancoff, aucun nouveau calcul n'était nécessaire pour répondre à cette question : il suffisait suivant une idée d'EPSTEIN d'appliquer la prescription (γ), c'est-à-dire de comparer l'effet de la variation de masse (déduite du calcul relativiste de Weisskopf de l'énergie propre d'un électron libre) dans la section efficace relativiste sans corrections radiatives, à l'effet des corrections radiatives calculées par Dancoff. Bien que les divergences fussent logarithmiques dans les deux cas, il n'y avait pas de compensation, car le coefficient du logarithme calculé par Dancoff était faux : il avait oublié tous les termes associés aux photons virtuels longitudinaux**. Rajoutant ces termes, Lewis établit la compensation des divergences. Le premier calcul de la partie finie des corrections qui restent après cette compensation revient à FEYNMAN⁽⁷⁷⁾ (juillet 1948). La difficulté essentielle tenait toujours à l'ambiguïté des soustractions. Feynman, reprenant une méthode qu'il venait d'utiliser pour la théorie électromagnétique classique⁽⁸⁹⁾ et que BOPP avait inspirée

* Il est très douteux que Dancoff se fût posé la même question dès 1939, quoi qu'en ait dit Tomonaga (Voir note au chap. 3}3)

** Sans être fin psychologue, on peut avancer que Dancoff a dû croire que le terme $e \int \psi^\dagger \gamma \psi d^3x$ d'interaction des électrons avec le champ coulombien extérieur représentait toute l'interaction coulombienne. Il n'en est rien si l'on a éliminé à la manière d'Oppenheimer 1930 les photons longitudinaux au profit d'un terme coulombien supplémentaire

$\frac{e^2}{8\pi} \int \frac{\psi^\dagger(\vec{x}) \psi(\vec{x}')}{|\vec{x} - \vec{x}'|} d^3x d^3x'$ décrivant l'interaction coulombienne du champ électronique avec lui-même.

en 1942 (voir chap. 7 § 1), introduisit un procédé de coupure relativiste des grandes impulsions des photons virtuels dans la théorie des perturbations conventionnelle. La méthode consiste à considérer les impulsions virtuelles k comme indépendantes des fréquences ω intervenant dans les dénominateurs d'énergie (introduisant ainsi comme dans la future théorie des perturbations covariantes des photons virtuels en dehors de la couche de masse) et à remplacer les sommations sur k $\int \frac{d^3k}{k} \dots$ par $2 \int d\omega d^3k g(\omega^2 - k^2) \dots$ où g est une régularisation de la distribution δ^* . Avec une telle "coupure relativiste" Feynman obtint une énergie propre de l'électron libre covariante : la partie électromagnétique était de la forme $\delta E = \frac{m \delta m}{E}$, δm étant la masse électromagnétique encore égale à l'accroissement de l'énergie d'un électron au repos. C'est bien la forme à laquelle on s'attend si l'on effectue une transformation de Lorentz du référentiel de repos au référentiel où l'électron a le moment \vec{p} : dans tout référentiel on a la relation $E^2 + p^2 = m^2$ avant et après branchement de l'interaction électromagnétique, si bien que la variation d'énergie est reliée à la variation de masse suivant $E \delta E = m \delta m$. Avec la coupure ordinaire Weisskopf dans son étude de 1934, n'avait pu obtenir de forme covariante**. Pour l'électron lié, Feynman était maintenant en mesure de donner un sens bien défini à la perturbation du deuxième ordre de l'amplitude de diffusion par le champ de rayonnement et au contre-terme de masse calculé en suivant la prescription (α). Feynman interpréta son résultat sous la forme d'un hamiltonien effectif d'interaction avec le champ extérieur électromagnétique (lentement variable) :

$$\Delta H \approx - \frac{\alpha}{3\pi m^2} \left(\ln \frac{m}{\mu} - \frac{3}{8} - \frac{1}{5} \right) \square e \delta^0 A_{ext} + \frac{\alpha}{2\pi} \frac{e}{2m} \frac{1}{2} \delta^0 \sigma_{\mu\nu} F_{ext}^{\mu\nu}$$

* Plus précisément $g(\omega^2 - k^2) = \int_0^\infty [\delta(\omega^2 - k^2) - \delta(\omega^2 - k^2 - \lambda^2)] G(\lambda) d\lambda$ avec $G(\lambda)$ fonction "lisse" telle que $\int_0^\infty G(\lambda) d\lambda = 1$. Ce type de régularisation est un ancêtre de la régularisation de PAULI-VILLARS.

** Cette difficulté fut mentionnée par Weisskopf⁽⁹¹⁾ dans une note (p. 77) de son article de 1939 sur l'énergie propre de l'électron libre. DIRAC fut sans doute le premier à trouver, lors de la quantification explicitement covariante de sa théorie de 1938, une énergie propre de l'électron libre covariante (voir chap. 7 § 2).

μ désignant une coupure infrarouge contenue dans $g(\omega^2 - k^2)$. Le terme $-\frac{1}{5}$ n'est pas présent chez Feynman qui omet délibérément la polarisation du vide. La forme covariante du résultat (correspondant à une modification $\Delta L = -\Delta H$ scalaire du lagrangien) plaidait en faveur de la méthode de coupure relativiste ; mais celle-ci n'était qu'une recette puisque les étapes intermédiaires de calcul utilisant l'ancienne théorie des perturbations n'étaient pas manifestement covariantes. Bien que cet article ne soit qu'une pâle anticipation des travaux futurs de Feynman, il faut souligner l'élégance des calculs : par exemple Feynman travaillait directement avec les photons virtuels longitudinaux au lieu de les refouler dans une auto-interaction électrostatique compliquée. Nous avons déjà dit la légère erreur commise dans le calcul du déplacement de Lamb ; elle n'invalidait pas la méthode suivie par Feynman. Nous ne traiterons pas ici des développements ultérieurs par SCHWINGER, FEYNMAN et DYSON du même problème dans une théorie covariante à toutes les étapes du calcul.

8.6.c Le moment magnétique anormal de l'électron :

Le 30 décembre 1947 Julian SCHWINGER⁽⁹²⁾ envoya une lettre à la Phys. Rev. intitulée "L'électrodynamique quantique et le moment magnétique de l'électron". Schwinger y exposait brièvement sa méthode et le résultat obtenu pour le moment magnétique anormal de l'électron $\frac{g\mu}{\mu} = \frac{\alpha}{2\pi}$ (l'article original⁽⁹²⁾ contient une erreur d'impression malheureuse). Nous avons vu à propos des travaux de Kramers comment une transformation canonique permettait de séparer un terme de masse électromagnétique, le hamiltonien effectif résiduel donnant des résultats finis pour les corrections radiatives. Chez Kramers une telle transformation n'était effectuée que dans le cas non relativiste. Schwinger généralisa cette idée dans le cas relativiste, en déterminant ordre par ordre la transformation canonique et le hamiltonien effectif. (Voir § 8.6.e et appendice)

Cette introduction d'un hamiltonien effectif ne contenant que la masse observable de l'électron constitue la première étape importante dans l'élaboration de la méthode de Schwinger. A ce stade il restait une difficulté importante, tout comme dans le calcul de Lamb et Kroll on aboutissait à des coefficients de $\vec{\alpha} \cdot \vec{E}$ et de $\vec{\sigma} \cdot \vec{B}$ incompatibles avec la

covariance relativiste, ce qui n'avait rien d'étonnant dans un formalisme hamiltonien non covariant. La deuxième étape fut l'introduction d'une théorie covariante des perturbations.

L'article de FEYNMAN (juillet 48) sur la coupure relativiste, ainsi que le calcul de FRENCH et WEISSKOPF (décembre 48) confirmèrent le résultat. Quelques mois plus tôt en mai 1948 LUTTINGER⁽⁹³⁾ (travaillant en Suisse avec Pauli et Jost) avait montré comment le $g-2$ se calculait en utilisant les techniques de Weisskopf pour le calcul des énergies propres, c'est-à-dire en calculant l'énergie propre d'un électron en présence d'un champ magnétique extérieur B dans l'ancienne théorie des perturbations. Seul le terme linéaire en B était calculé. Luttinger put même éviter toute renormalisation de masse grâce à l'astuce suivante : il existe pour un électron de Dirac plongé dans un champ B homogène un état quantique d'énergie égale à l'énergie de repos : $E_0^{(0)}(m_0) = m_0$. L'énergie orbitale compense alors exactement l'énergie de spin. Si l'interaction avec le champ de rayonnement fait apparaître un terme linéaire en B dans $E_0^{(2)}$ celui-ci contient le vrai moment magnétique anormal, puisque le terme de renormalisation de masse ne dépend pas de B . ($\frac{\partial E_0^{(0)}}{\partial m} = 1$ indépendant de B dans la prescription (γ)).

8.6.d Les calculs intuitifs de corrections radiatives :

Un appel à l'intuition dans le genre de problème que nous venons de traiter peut paraître suspect. La notion intuitive de base utilisée généralement est celle de fluctuation quantique. Le mot de fluctuation lui-même semble se référer à une image classique de grandeurs sous-jacentes variant aléatoirement dans le temps autour d'une valeur fixe, et ces grandeurs seraient en principe exactement mesurables. Au contraire, dans la théorie quantique si une observable "fluctue" cela implique qu'on ne peut la mesurer à mieux que l'amplitude de ces fluctuations, en somme il y a fluctuation sans fluctuant, et même pis, il se peut que ces fluctuations - c'est le cas de l'oscillateur harmonique m du champ électromagnétique - existent dans l'état fondamental du système, et qu'on n'en puisse tirer aucune énergie. Bien qu'elles n'aient pas de substrat "physique", et bien qu'elles n'aient d'autre conséquence expérimentale directe qu'une certaine dispersion

statistique des mesures éventuelles de l'observable qu'elles affectent, ces "fluctuations" ont des conséquences physiques observables indirectes intéressantes, quand elles se manifestent par l'interaction avec un autre système. Le champ électromagnétique est couplé avec le champ d'électron par une interaction $j_\mu A^\mu$ linéaire dans le potentiel électromagnétique. De ce fait la probabilité de transitions réelles ou virtuelles du vide de rayonnement à un quelconque autre état

$$\sum_\alpha \langle 0 | j \cdot A | \alpha \rangle \langle \alpha | j \cdot A | 0 \rangle = \langle 0 | j_\mu j_\nu A^\mu A^\nu | 0 \rangle$$

contient en facteur les fluctuations du vide $\langle 0 | A^\mu A^\nu | 0 \rangle$. Par ce biais les fluctuations sont responsables de l'instabilité des états excités atomiques : elles engendrent alors des transitions réelles. De plus elles déplacent les niveaux correspondants par une interaction virtuelle du second ordre.

Allant au-delà de ces considérations générales, WELTON⁽⁹⁴⁾ (juin 1948), encouragé par Weisskopf, décrit l'effet des fluctuations du vide dans un modèle semi-classique plus palpable et donna une justification formelle de son modèle, expliquant pourquoi il semblait s'affranchir de renormalisation de masse. Le point de départ de Welton est le suivant : "... nous supposons que la fluctuation de position (de l'électron) est un concept réel, tandis que nous n'accorderons pas de signification physique à l'énergie d'interaction entre les fluctuations du champ et de la particule". La quantification du champ électromagnétique libre donne pour les fluctuations du vide

$$\langle E^2 \rangle_0 = \langle B^2 \rangle_0 = \frac{2}{\pi} \int k^3 dk$$

L'équation du mouvement d'un électron non relativiste dans le champ \vec{E} , \vec{B} est $m\ddot{\vec{r}} = e\vec{E}$. * Pour la composante de fréquence k $m(-k^2)\vec{r}_k = e\vec{E}_k$. Les fluctuations de \vec{E}_k forcent des fluctuations de \vec{r}_k avec $\langle (\delta r)^2 \rangle = \frac{e^2}{m^2} \frac{2}{\pi} \int_k^K \frac{dk}{k}$ K est pris égal

à m pour tenir compte du caractère non relativiste du calcul. La coupure infrarouge est déterminée par les conditions particulières du problème

* Notons que la masse m désigne ici la masse physique, puisque le champ fluctuant du vide introduit dans le second membre ne contient pas le champ propre de l'électron. Cette remarque donne une première explication de l'absence de renormalisation de masse dans le calcul de Welton.

qui invalident l'équation du mouvement $m\ddot{\vec{r}} = e\vec{E}$ pour les composantes de basse fréquence.

Les fluctuations forcées se superposent au mouvement moyen de l'électron dans un potentiel extérieur V : à un instant donné, le potentiel auquel est soumis l'électron vaut

$$V(\vec{r} + \delta\vec{r}) \simeq V(\vec{r}) + \delta\vec{r} \cdot \vec{\nabla}V + \frac{1}{2} \delta r_i \delta r_j \frac{\partial^2 V}{\partial r_i \partial r_j}$$

En moyenne les fluctuations sont isotropes donc

$$\langle V \rangle \simeq V(\vec{r}) + \frac{1}{6} \langle (\delta r)^2 \rangle \Delta V \simeq V(r) + \frac{\alpha}{3\pi} \left(\ln \frac{m}{\mu} \right) \frac{1}{m^2} \Delta V$$

Cette dernière expression donne le coefficient exact du logarithme pour les corrections radiatives de vertex, et par conséquent une bonne estimation du déplacement de Lamb. Pour que les fluctuations de position restent faibles devant le rayon atomique a , il faut prendre $\mu \simeq \frac{1}{a}$. La meilleure valeur de μ , c'est-à-dire l'énergie d'excitation moyenne introduite par Bethe ne peut être prédite dans cette approche.

Welton put aussi évaluer les corrections à la diffusion de la lumière par un électron : la section efficace d'interaction doit diminuer puisque la charge de l'électron est étendue par les fluctuations. Plus quantitativement on peut évaluer l'effet des variations aléatoires de la phase de l'onde incidente au lieu de l'électron, et de la phase de l'onde émergente :

$$\langle e^{i\vec{k} \cdot (\vec{r} + \delta\vec{r})} \rangle \simeq \left(1 - \frac{1}{6} k^2 \langle \delta r^2 \rangle \right) e^{i\vec{k} \cdot \vec{r}} \quad . \text{La variation}$$

$\frac{\delta\sigma}{\sigma}$ qui en résulte vaut deux fois la variation relative d'amplitude de l'onde incidente plus deux fois celle de l'émergente, d'où $\frac{\delta\sigma}{\sigma} = -\frac{2}{3} k^2 \langle \delta r^2 \rangle$

On prend $\mu \simeq k$.




Mais pour le moment magnétique anormal de l'électron, Welton échoua ; il trouvait le bon ordre de grandeur mais le mauvais signe, et il accusait la nature relativiste du spin.

La justification formelle de la méthode intuitive de Welton se trouvait déjà dans la théorie de KRAMERS, comme Welton le vit lui-même. Rappelons que Kramers et ses disciples avaient introduit un hamiltonien

$$H = \frac{p^2}{2m} + V(\vec{R} - \frac{e}{mc} \vec{Z}) + \frac{1}{2} \int (E_{\perp}^2 + B^2) d^3x \quad \text{où } m \text{ désignait la masse physique et } \vec{Z} \text{ le vecteur de Hertz tel que } \text{rot rot } \vec{Z} = \vec{E}_{\perp} \text{ et } \text{div } \vec{Z} = 0. \text{ Ce}$$

hamiltonien se déduisait en fait du hamiltonien de Dirac habituel (avec toutes les approximations ici considérées) par transformation canonique, à un terme de masse électromagnétique près. Le vecteur $\delta\vec{r}$ de Welton coïncide

juste avec $-\frac{e}{mc} \vec{Z}$. Lors des calculs de corrections radiatives ce vecteur intervient quadratiquement exactement de la manière envisagée par Welton et sa moyenne quadratique dans le vide restitue les fluctuations introduites a priori par Welton.

WEISSKOPF⁽⁹⁵⁾ fit en 1945 un exposé vulgarisé de l'évolution récente de la théorie de l'électron, où il fit usage des arguments de Welton, dont l'intérêt essentiel, pensons-nous, est de se bien prêter à la vulgarisation. La difficulté mentionnée pour généraliser la méthode au calcul du moment magnétique anormal de l'électron fut résolue par le physicien japonais KOBA⁽⁹⁶⁾ (Avril 49). D'après Koba, l'échec de Welton venait d'une image classique erronée du spin de l'électron. Au lieu d'être représenté par un dipôle rigide dont seule l'orientation pouvait fluctuer, il fallait prendre l'image classique de la "Zitterbewegung" de Schrödinger (voir chap. 2 § 2) c'est-à-dire représenter le spin par un mouvement hélicoïdal rapide ($v \sim m$) autour de l'orbite moyenne de l'électron. C'est bien cette image qu'il faut retenir si l'on croit que la théorie classique qui précède une théorie quantique est décrite par les équations du mouvement des opérateurs en représentation de Heisenberg. Nous ne répèterons pas les calculs semi-quantitatifs de Koba et la correspondance qu'il établit avec la théorie "rigoureuse", mais nous donnerons l'image qualitative expliquant le phénomène. Le moment magnétique de l'électron, représenté par une boucle de courant  est soumis à deux effets de signe opposé : - le pivotement autour du centre de l'anneau de courant  a pour effet de diminuer le moment magnétique, - les fluctuations de l'anneau dans son plan couplées à celles du centre  ont pour effet d'augmenter le moment magnétique et ce dernier effet domine le précédent.

8.6.e L'Ecole Japonaise :

Les distinctions de nationalité dans l'ensemble, ne nous ont pas semblé jouer un rôle important dans l'étude historique de la théorie quantique des champs. Les particularités nationales s'estompent dans une communauté scientifique devenue internationale. Dès le début de ce siècle,

la plupart des grands physiciens furent de grands voyageurs, achevant fréquemment leurs études loin de leur patrie, collaborant avec des collègues d'autres nationalités, assistant à de grandes conférences internationales. Mais la seconde guerre mondiale a empêché cette libre circulation des idées, en isolant une partie de la communauté de physiciens allemands. L'isolement du Japon pendant la guerre du Pacifique, et la lenteur des échanges d'informations avec les Etats-Unis dans l'immédiat après-guerre a pris une dimension historique tout aussi réelle : les recherches japonaises en physique théorique de cette période se sont déroulées indépendamment ou presque. Certains travaux importants de Tomonaga et de ses collaborateurs ne furent publiés en anglais* qu'avec plus de trois ans de retard ! Les idées de Schwinger qui servirent de fondement à la nouvelle électrodynamique quantique : l'usage d'une transformation canonique éliminant le champ électromagnétique libre et celui d'une formulation relativiste manifestement covariante, virent le jour indépendamment au Japon à peu près en même temps pour la première, et bien avant pour la seconde. Il nous semble légitime de traiter séparément ces contributions des physiciens japonais.

Rappelons d'abord (voir chap. 5 et 6) comment l'Ecole japonaise s'était déjà illustrée dans des problèmes de théorie des champs. Parmi les premières applications du concept de champ quantique à la physique nucléaire nous avons dit la nouveauté et la portée des travaux de YUKAWA, SAKATA, TAKETANI et TOMONAGA ... l'interaction nucléaire était interprétée pour la première fois comme un échange virtuel de mésons. Un autre problème, cette fois purement formel trouva d'amples développements au Japon : celui d'une formulation manifestement covariante de l'électrodynamique quantique. Partis de considérations originales de Dirac, les japonais YUKAWA, WATANABE, et surtout TOMONAGA cherchèrent à les généraliser. En substituant au temps une surface quelconque du genre espace comme argument de la fonction d'onde et en utilisant la représentation d'interaction (voir chap. 5 § 4). Tomonaga

* Nous ne donnerons que les références des traductions anglaises des textes. Celles des publications japonaises originales y sont rapportées.

donna la forme générale d'un formalisme quantique relativiste manifestement covariant dès 1943 et son application à l'électrodynamique quantique en 1946. Il manquait encore l'idée de renormalisation pour que l'édifice s'achevât, mais comme le fit remarquer Tomonaga (6 novembre 1946)⁽⁹⁸⁾, "le développement d'une telle théorie nous semble intéressant non seulement parce que nous obtenons ainsi une théorie formellement plus satisfaisante, mais aussi parce que de cette manière nous pouvons espérer que certains aspects des difficultés de base de la théorie quantique des champs usuelle se révéleront d'eux-mêmes."

Stratégie historique

Pour le problème qui nous intéresse de plus près dans ce chapitre, l'extraction de résultats finis pour les corrections radiatives à des processus d'électrodynamique, il faut distinguer plusieurs approches japonaises. La première en date trouve son origine dans une discussion historico-philosophique, lancée par Mituo TAKETANI et surtout développée lors de plusieurs discussions à l'Université de Nagoya. L'intervention d'une analyse historique dans l'histoire de la théorie quantique des champs n'est pas un fait très commun. On se souvient des analogies historiques utilisées par Heisenberg (voir chap. 4 § 6) pour justifier l'introduction d'une longueur universelle (\hbar , c , puis r_0). Dans les années 40 les théories alternatives à la théorie de Heisenberg-Pauli devenaient si nombreuses que certains grands théoriciens ressentaient le besoin de les comparer dans une perspective historique (voir les articles de Pauli, Wentzel, Kramers cités au chapitre 7). Dans une conférence prononcée en 1943⁽⁹⁹⁾ TAKETANI procéda à une revue de toutes les tentatives d'élimination des divergences, en insistant sur les rapports entre ce problème et l'évolution de la physique nucléaire théorique à laquelle les Japonais avaient tant participé (voir chap. 6). Le problème central selon Taketani est le "conflit entre matière et champ" (par exemple particules chargées versus rayonnement). L'incompatibilité des concepts de base de la théorie quantique, reflétée par le problème des divergences, doit être identifiée à ce conflit : "Les deux concepts fondamentaux de matière et de champ sont ceux dont l'opposition me paraît aujourd'hui la plus essentielle ; ils contiennent le problème

le plus essentiel de la théorie contemporaine - de même que les concepts incompatibles d'onde et de particule donnèrent naissance à la mécanique quantique". Sous cet aspect les théories récentes se classifient en théories " monistes " contenant un seul champ en auto-interaction comme celles de MIE, BORN, INFELD et BOPP, et les théories dualistes dont l'électrodynamique quantique usuelle est un exemple. Taketani donne sa préférence aux théories monistes, débarassées du conflit matière-champ, mais présentant d'autres difficultés : introduction du spin, équation du mouvement des singularités représentant les particules, quantification. Ceux qui cherchaient une origine orientale aux idées de Taketani seront déçus, car c'est bien à la dialectique hégélienne que le théoricien japonais se réfère : "nous croyons qu'il résulte de l'analyse logique dialectique qui précède, que le conflit entre les deux concepts fondamentaux de matière et de champ est la contradiction en présence, celle qui nous montre dans quelle direction nous orienter". Shoichi SAKATA publia en avril 1947 un article⁽¹⁰⁰⁾ sur la "théorie de l'interaction des particules élémentaires" résumant les réflexions du groupe de Nagoya. On trouve la "Wissenschaft der Logik" de Hegel citée entre Taketani et Pauli. Son application à l'histoire naissante de la théorie quantique des champs se trouve précisée. Les trois stades d'évolutions historiques s'ordonnent suivant le tableau :

phénoménologique	an sich	Tycho Brahé	coupures U.V., λ -limite soustractions
substantialiste	für sich	Kepler	Mie, Born et Bopp et théories mixtes
essentialiste	an und für sich	Newton	? (longueur universelle ?)

En termes naïfs : "si la loi du développement de la physique, gouvernant la direction dans laquelle la théorie doit progresser, pouvait être découverte et être appliquée consciencieusement à l'analyse des problèmes

concrets, alors nous pourrions certainement atteindre notre fin beaucoup plus facilement et plus sûrement qu'avec les succès accidentels dus à l'intuition donnée aux chercheurs". Sakata craint bien qu'une telle tentative aboutisse à des conclusions "post hoc" plutôt que "propter hoc" et doute ainsi de la justification historique de la longueur universelle de Heisenberg. Les théories monistes constituent un progrès dans l'optique de Sakata même si elles ne peuvent être présentées comme le stade ultime de la théorie. La théorie de BOPP, comme l'a montré BOPP lui-même (voir chap. 7 § 1) est équivalente à une théorie mixte où la matière est couplée à la fois au champ électromagnétique et à un champ vectoriel massif de Proca. Les théories mixtes, issues du point de vue moniste doivent être explorées en priorité : "c'est la méthode la plus efficace pour développer la théorie, et c'est une voie pleine de promesses pour la résolution des difficultés actuelles : il faut examiner les corrélations internes des interactions entre tous les champs interagissant avec la même particule et toutes les particules agissant comme source du même champ". Ces considérations philosophiques de Sakata peuvent nous paraître suspectes, mais elles dirigèrent toute une série de recherches fructueuses du groupe de Nagoya.

Les théories mixtes : le champ de cohésion

La théorie de BOPP⁽¹⁰¹⁾ (voir chap. 7 § 1) illustrant le point de vue moniste cher aux théoriciens de Nagoya se trouvait être équivalente à une théorie mixte. Sakata et Hara s'intéressèrent à une théorie mixte du même genre où la matière se trouvait simultanément couplée au champ électromagnétique et au champ scalaire neutre de Yukawa. L'énergie propre de l'électron comportait alors deux contributions : l'une venant des processus électromagnétiques virtuels, l'autre de l'émission virtuelle de vecteurs neutres. Ces deux contributions étaient de signe opposé et de même forme si bien qu'un choix convenable du rapport des deux constantes de couplages annulait la dépendance dans la coupure de l'énergie propre. PAIS avait fait en 1945 une constatation analogue (voir chap. 7 § 1). Le

8 avril 1947 SAKATA et HARA annoncèrent dans une lettre⁽¹⁰³⁾ la même élimination de la divergence de l'énergie propre (à l'ordre α). Le calcul complet fut publié par HARA l'année suivante⁽¹⁰²⁾. Pouvait-on éliminer de la même manière et avec le même rapport de constantes de couplages la divergence des corrections radiatives à la diffusion coulombienne ? La première réponse à cette question, donnée par ITO, Koba et TOMONAGA⁽¹⁰⁴⁾ le 1er novembre 1947, fut négative. Ces trois auteurs avaient répété l'erreur de Dancoff, ils avaient négligé certains termes dans le calcul perturbatif. Le 30 décembre 1947 un erratum⁽¹⁰⁵⁾ apporta les corrections nécessaires, révéla que le champ de cohésion assurait bien (à l'ordre α) la finitude des corrections radiatives (à condition toutefois d'opérer simultanément les soustractions habituelles de divergence de la polarisation du vide). Les calculs complets furent publiés plus tard^(106,107). Il n'était pas encore question de renormalisation de masse, mais pour la première fois une valeur finie des corrections radiatives pouvait être obtenue par un procédé naturel, étroitement apparenté à la renormalisation de masse, comme allait bientôt le constater Tomonaga : pour passer d'un procédé à l'autre il fallait substituer un contre-terme de masse au contre-terme de cohésion.

Cette étape fut franchie par TOMONAGA et Koba⁽⁹⁷⁾. Dans une lettre aussi datée du 30 décembre 1947, ils rapportèrent les résultats récents de l'expérience de Lamb-Retherford et l'explication de Bethe dont ils avaient pris connaissance*. Les deux auteurs annoncèrent des calculs en cours utilisant une formulation nouvelle de l'idée de renormalisation : ils préconisaient d'introduire un contre-terme de masse $-\delta m \int \psi^\dagger \beta \psi d^3x$ par une méthode dite de " self-consistent subtraction " puisque l'opposé de ce terme était rajouté au hamiltonien électronique libre de départ pour restituer la masse expérimentale ($m = m_0 + \delta m$; prescription α). Un programme à plus long terme était annoncé : utiliser conjointement cette

* Ces informations étaient déjà connues lors du symposium de Kyoto des 24 et 25 novembre 1947 sur la théorie des particules élémentaires.

méthode de renormalisation et la formulation "super-multi-temps" de Tomonaga. Le calcul des corrections radiatives à la diffusion coulombienne avec renormalisation de masse dans l'ancienne théorie des perturbations fut effectué indépendamment des travaux de Lewis et d'Epstein par Koba Takeda et Tomonaga (mars 1948).

Au congrès de Kyoto des 24 et 25 novembre 1947 - d'où furent issues les premières publications japonaises faisant état des découvertes américaines, concernant le déplacement de Lamb - TATI et TOMONAGA⁽¹¹⁰⁾ exposèrent une nouvelle méthode de calcul du déplacement de niveau, identique à celle utilisée indépendamment* par Schwinger aux Etats-Unis. La méthode consiste à "effectuer une transformation canonique des variables séparant le champ des photons libres du champ lié à l'électron". Lors d'une première exploration des conséquences de cette méthode (mai 1948) Tati et Tomonaga en précisèrent l'origine : "Bien que cette hypothèse (la renormalisation de masse) ne fût exprimée clairement et explicitement qu'à la lumière de l'évidence expérimentale, une telle idée avait été utilisée implicitement dans des considérations antérieures de Bloch et Nordsieck ainsi que de Pauli et Fierz sur le champ propre de l'électron"⁽¹¹¹⁾. Nous avons vu (voir chap. 5 § 2) comment Bloch et Nordsieck définirent de nouveaux états électroniques habillés de photons mous et comment Pauli et Fierz explicitèrent dans le cas non relativiste la transformation canonique engendrant ces états à partir des états "nus". L'effet précis de cette transformation était d'associer à tout état propre du hamiltonien libre (sans couplage -j.A matière-rayonnement) un état propre du hamiltonien couplé. Ce hamiltonien couplé avec en sus le potentiel extérieur coulombien pouvait être réexprimé dans les nouvelles variables canoniques et les différents termes recevaient une interprétation immédiate : terme de masse électromagnétique venant de l'habillage des électrons, termes de matière et de rayonnement de structure identique au hamiltonien libre, et terme d'interaction avec le champ coulombien extérieur. Une fois connu le

* La première mention publiée de la méthode de Schwinger date du 30 décembre 1947 (lettre sur le g-2).

principe de renormalisation selon lequel toute divergence UV devait être absorbée par une redéfinition de la masse, l'utilité d'une généralisation de la transformation de Pauli et Fierz qui isolât au moins perturbativement le terme de masse électromagnétique fut claire aux esprits de Tomonaga et de Schwinger.

Suivant une propriété générale du branchement adiabatique des perturbations, les états propres du hamiltonien variable restent des états propres lors du branchement, si bien que l'opérateur d'évolution du système matière-rayonnement, lors du branchement de leur interaction relie comme la transformation de Fierz et Pauli les états "nus" aux états habillés. En fait nous montrons en appendice que ces deux opérateurs ne diffèrent que d'un facteur de phase*. Grâce à cette nouvelle interprétation, la généralisation de la transformation de Fierz et Pauli au cas relativiste devient évidente. En représentation d'interaction son expression est $U_I = T \exp -i \int_{-\infty}^t H_I dt$ avec $H_I = -e \int \psi^\dagger \vec{\alpha} \cdot \vec{A} \psi d^3x$ ψ et \vec{A} désignant les champs libres, car dans cette même représentation l'équation d'évolution des états s'écrit $i \frac{\partial}{\partial t} |\chi\rangle = H_I |\chi\rangle$

En présence d'un champ extérieur agissant par le hamiltonien d'interaction Λ (par exemple $\Lambda = e \int \psi^\dagger v \psi d^3x$ pour un potentiel coulombien v) cette équation devient $i \frac{\partial}{\partial t} |\chi\rangle = (H_I + \Lambda) |\chi\rangle$

Un nouveau changement de représentation $|\chi\rangle = U_I |\varphi\rangle$ donne $i \frac{\partial}{\partial t} |\varphi\rangle = [(U_I^{-1} H_I U_I - H_I) + U_I^{-1} \Lambda U_I] |\varphi\rangle$. Une analyse de la structure du premier terme** montre qu'il s'interprète comme le terme de masse électromagnétique ; le deuxième terme donne un résultat fini pour la diffusion par Λ et ses corrections radiatives (ou pour les

* Ces considérations clarifiant le lien entre la transformation de Pauli et Fierz et l'évolution par branchement adiabatique sont absentes des articles de Tomonaga et de Schwinger.

** Il est aisé de vérifier qu'au deuxième ordre $U_I^{-1} H_I U_I - H_I = -i [H_I^{(1)}, \int_{-\infty}^t H_I dt]$ coïncide avec le contre-terme de masse introduit dans la prescription (δ) (une fois revenu dans la représentation de Schrödinger).

déplacements de niveau). Bien sûr, les premiers calculs n'étaient effectués qu'à l'ordre e^2 et U_I n'admet pas dans le cas relativiste d'expression non perturbative explicite contrairement au cas classique. Nous avons exposé la méthode dans le formalisme hamiltonien conventionnel, mais dans la communication de Tati et Tomonaga le formalisme "super-multi-temps", sans doute moins familier au lecteur, était utilisé.

L'application aux déplacements de niveau fut précisée plus tard par HARA et TOKANO⁽¹¹²⁾ et l'analyse de la structure de différents termes poursuivie dans plusieurs articles dont nous avons cité le premier⁽¹¹¹⁾. Ces développements sortent du cadre de notre étude. Parmi les travaux japonais de 1948 signalons seulement l'étude des corrections radiatives aux diffusions autres que la coulombienne, comme la diffusion Compton, ou la diffusion électron-électron. Il apparut que les contre-termes introduits dans le hamiltonien pour rendre finie la polarisation du vide comprenaient un terme en $\int A^2 d^3\alpha$ non invariant de jauge, ni invariant relativiste^(113,114,115). OPPENHEIMER, dans une note ajoutée à une lettre de TOMONAGA⁽¹¹⁶⁾ qu'il présentait à Phys. Rev. souligna cet aspect des travaux japonais. SCHWINGER avait déjà pu obtenir la valeur zéro pour l'énergie propre du photon (comme l'exige l'invariance de jauge), mais la situation ne fut vraiment clarifiée qu'avec la régularisation de PAULI-VILLARS respectant à tous les stades du calcul l'invariance de jauge et l'invariance relativiste.

Si l'on ne doit retenir des travaux de l'Ecole Japonaise que les plus essentiels il faudra distinguer d'une part l'élaboration avant 1946 d'une version manifestement covariante des théories quantiques relativistes qui devint après 1947 un terrain de choix pour explorer les conséquences du principe de renormalisation, d'autre part la constitution de l'électrodynamique quantique moderne par TOMONAGA, parallèlement aux travaux de SCHWINGER. La théorie des champs de cohésion n'allait certes pas trouver de prolongements historiques importants (elle n'aurait sans doute pas pu à elle seule régler tout le problème des divergences). Mais elle fut un espoir d'obtenir des résultats finis pour des corrections radiatives, avant même que le principe de renormalisation de masse fût connu au Japon, et avant que les théoriciens occidentaux (autres que Kramers) lui accordent toute son importance.

Appendice : De la transformation canonique de Pauli et Fierz
à la représentation de Tomonaga-Schwinger

Le hamiltonien de Dirac dans l'approximation non relativiste utilisée par Pauli et Fierz, s'écrit

$$H = H_0 + H_1 + eV(\vec{R}) \quad \text{avec} \quad \begin{cases} H_0 = \frac{p^2}{2m_0} + \frac{1}{2} \int (E^2 + B^2) d^3x \\ H_1 = -\frac{e}{m_0} \vec{p} \cdot \vec{A} \\ V \text{ potentiel coulombien extérieur} \end{cases}$$

La transformation canonique de Pauli et Fierz est engendrée par la transformation unitaire $U_{PF} = \exp(i \frac{e}{m_0} \vec{p} \cdot \vec{Z})$, \vec{Z} étant le vecteur de Hertz défini par $\vec{E} = \text{rot rot } \vec{Z}$ et $\text{div } \vec{Z} = 0$. A un état $|\psi\rangle$ dans l'ancienne représentation correspond un état $|\varphi\rangle = U_{PF}^{-1} |\psi\rangle$ dans la nouvelle. Les nouveaux champs canoniques conjugués (\vec{A}', \vec{E}') sont définis par $\vec{A}' = U_{PF} \vec{A} U_{PF}^{-1}$ et $\vec{E}' = U_{PF} \vec{E} U_{PF}^{-1}$, de même pour (\vec{r}', \vec{p}') ; le hamiltonien dans ces nouvelles variables est

$$\text{(à l'ordre } e^2) \quad H' = \frac{p'^2}{2m_0} + \frac{1}{2} \int (E'^2 + B'^2) d^3x + V(\vec{r}' - \frac{e}{m_0} \vec{Z}) + \delta H$$

$$\text{avec } \delta H = \frac{\mu}{m_0} \frac{p'^2}{2m_0} \quad \text{et } \mu = \frac{4}{3} \frac{e^2}{4\pi} \int_0^\infty d\omega$$

En l'absence de V , H' est identique à H_0 à la substitution près $m_0 \rightarrow m = m_0 + \mu$ (toujours à l'ordre e^2): μ doit être interprétée comme la masse électromagnétique comme le remarquaient Pauli et Fierz (ce qui manquait à ces deux auteurs pour atteindre la théorie renormalisée c'était l'idée que toutes les divergences ultraviolettes pouvaient être absorbées en exprimant les processus physiques calculés à l'aide de la masse totale $m_0 + \mu$. Ils pensaient au contraire que la validité de leur calcul nécessitait une masse électromagnétique μ négligeable ($\mu \ll m_0$)).

Ainsi à un état propre du hamiltonien libre $H_0(m_0 + \mu)$, U_{PF} associe un état propre de $(H_0 + H_1)(m_0)$ de même valeur propre. Ou encore à un état propre $H_0(m_0)$ U_{PF} associe un état propre de $H_0 + H_1(m_0)$ de valeur propre différente. Il est donc instructif de comparer U_{PF} à l'opérateur d'évolution du système matière-rayonnement lors du branchement adiabatique de H_1 . Pour ce faire on substitue dans H_1 $e(t)$ à e avec $e(t)$ croissant lentement de $e(-\infty) = 0$ à $e(\tau) = e$. En représentation d'interaction (Réf. I p 177) cet opérateur d'évolution s'écrit $U_I = \mathcal{T} \exp -i \int_0^\tau H_{I-1} dt$

Dans ce cas particulier la T-exponentielle ne diffère de l'exponentielle ordinaire que par un facteur de phase (Réf. I p 179), du fait que les valeurs de H_{II} à deux instants différents ont un commutateur non opératoire (puisque $[a, a^\dagger] = 1, [a, a] = 0, [\vec{p}, \vec{A}] = 0$ etc.). Donc

$$U_I = \exp -i \int_{-\infty}^t H_{II} dt \quad \text{avec } H_{II} = -\frac{e(t)}{m} \vec{p} \cdot \vec{A}_I$$

Or $\vec{A}_I = -\frac{\partial \vec{Z}_I}{\partial t}$ comme il résulte de $\begin{cases} \vec{E}_I = -\partial \vec{A}_I / \partial t \\ \vec{E}_I = \text{rot rot } \vec{Z}_I = -\Delta \vec{Z}_I = -\frac{\partial^2 \vec{Z}_I}{\partial t^2} \end{cases}$

$$\int_{-\infty}^t H_{II} dt = -\frac{1}{m} \vec{p} \int_{-\infty}^t e(t) \vec{Z}_I dt = -\frac{1}{m} \vec{p} \left[e(z) \vec{Z}_I(z) - \int_{-\infty}^z e \vec{Z}_I dt \right]$$

Le deuxième terme est négligeable dans le cas d'un branchement adiabatique,

d'où $U_I = \exp \frac{e}{m} \vec{p} \cdot \vec{Z}(z)$ ou encore en revenant en représentation

de Schrödinger $U = \exp \frac{e}{m} \vec{p} \cdot \vec{Z}$ c.à.d. $\boxed{U = U_{PF}}$

Cette identité permet de généraliser la transformation de Pauli-Fierz au cas relativiste : on prendra $U_I = T \exp -i \int j_\mu A^\mu d^4x$ (les conditions aux limites doivent bien sûr être précisées). Cette fois $[j \cdot A(x), j \cdot A(y)]$ est un opérateur et il n'est plus possible de donner à U_I une forme explicite non perturbative. C'est uniquement la simplicité du terme d'interaction $j \cdot A$ dans une approximation non relativiste qui avait permis à Bloch et Nordsieck et à Pauli et Fierz de donner une solution non perturbative au problème infrarouge.

REFERENCES

Abréviations :

ZP : Zeitschrift für Physik

PRS : Proceedings of the Royal Society of London (série A)

AP : Annalen der Physik

PR : Physical Review

1. Born, Heisenberg, Jordan ZP 39, 557 (1926)
2. Planck Berliner Ber. 1-700 (1900)
3. Einstein AP 17, 132 (1905)
4. Ehrenfest Phys. Zeits. 7, 528 (1906)
5. Debye A.P. 33, 1427 (1910)
6. Einstein Phys. Zeits. 10, 185, 817 (1909)
7. Dirac PRS 114, 243 (1927)
8. Bose ZP 26, 178 (1924)
9. G.N. Lewis Nature 118, 874 (1926)
10. Jordan ZP 44, 473 (1927)
11. Jordan et Wigner ZP 47, 631 (1928)
12. Jordan et Klein ZP 45 (1927) 751
13. Jordan et Pauli ZP 47 (1928) 151
14. Pauli ZP 43, 601 (1927)
15. Darwin PRS 116, 227 (1927)
16. Dirac PRS 117 (1928) 610
17. Darwin PRS 118, 654 (1928)
18. Gordon ZP 48, 11 (1928)
19. Klein ZP 53 (1929) 157
20. Dirac PRS 126, 360 (1930)
21. Oppenheimer PR 35, 562 (1930)
22. Weyl "Gruppentheorie und Quantenmechanik" 2^o ed. p. 234 (1931)
23. Dirac PRS 133, 60 (1931)
24. Gray et Tarrant Nature 128 (1931) 100
25. Heisenberg et Pauli ZP 56, 1 (1929)
26. Heisenberg et Pauli ZP 59, 168 (1930)
27. Fermi Rend. Lincei 5, 881 (1929)
28. Fermi Rend. Lincei 12, 431 (1930)
29. Oppenheimer PR 35, 461 (1930)
30. Waller ZP 62, 673 (1930)
31. Wigner et Weisskopf ZP 63, 54 (1930)
32. Møller ZP 70, 786 (1931)
33. Mott PRS 124, 425 (1929)

Dans ce cas particulier la T-exponentielle ne diffère de l'exponentielle ordinaire que par un facteur de phase (Réf. I p 179), du fait que les valeurs de H_{II} à deux instants différents ont un commutateur non opératoire (puisque $[a, a^\dagger] = 1, [a, a] = 0, [\vec{p}, \vec{A}] = 0$ etc.). Donc

$$U_I = \exp -i \int_{-\infty}^z H_{II} dt \quad \text{avec } H_{II} = -\frac{e(t)}{m} \vec{p} \cdot \vec{A}_I$$

Or $\vec{A}_I = -\frac{\partial \vec{Z}_I}{\partial t}$ comme il résulte de $\begin{cases} \vec{E}_I = -\partial \vec{A}_I / \partial t \\ \vec{E}_I = \text{rot rot } \vec{Z}_I = -\Delta \vec{Z}_I = -\frac{\partial^2 \vec{Z}_I}{\partial t^2} \end{cases}$

$$\int_{-\infty}^z H_{II} dt = -\frac{1}{m} \vec{p} \int_{-\infty}^z e(t) \vec{Z}_I dt = -\frac{1}{m} \vec{p} \left[e(z) \vec{Z}_I(z) - \int_{-\infty}^z e \vec{Z}_I dt \right]$$

Le deuxième terme est négligeable dans le cas d'un branchement adiabatique, d'où $U_I = \exp \frac{e}{m} \vec{p} \cdot \vec{Z}(z)$ ou encore en revenant en représentation de Schrödinger $U = \exp \frac{e}{m} \vec{p} \cdot \vec{Z}$ c.à.d. $U = U_{PF}$

Cette identité permet de généraliser la transformation de Pauli-Fierz au cas relativiste : on prendra $U_I = T \exp -i \int_{j_\mu A^\mu} d^4x$ (les conditions aux limites doivent bien sûr être précisées). Cette fois $[j \cdot A(x), j \cdot A(y)]$ est un opérateur et il n'est plus possible de donner à U_I une forme explicite non perturbative. C'est uniquement la simplicité du terme d'interaction $j \cdot A$ dans une approximation non relativiste qui avait permis à Bloch et Nordsieck et à Pauli et Fierz de donner une solution non perturbative au problème infrarouge.

REFERENCESAbréviations :

- ZP : Zeitschrift für Physik
 PRS : Proceedings of the Royal Society of London (série A)
 AP : Annalen der Physik
 PR : Physical Review

1. Born, Heisenberg, Jordan ZP 39, 557 (1926)
2. Planck Berliner Ber. 1-700 (1900)
3. Einstein AP 17, 132 (1905)
4. Ehrenfest Phys. Zeits. 7, 528 (1906)
5. Debye A.P. 33, 1427 (1910)
6. Einstein Phys. Zeits. 10, 185, 817 (1909)
7. Dirac PRS 114, 243 (1927)
8. Bose ZP 26, 178 (1924)
9. G.N. Lewis Nature 118, 874 (1926)
10. Jordan ZP 44, 473 (1927)
11. Jordan et Wigner ZP 47, 631 (1928)
12. Jordan et Klein ZP 45 (1927) 751
13. Jordan et Pauli ZP 47 (1928) 151
14. Pauli ZP 43, 601 (1927)
15. Darwin PRS 116, 227 (1927)
16. Dirac PRS 117 (1928) 610
17. Darwin PRS 118, 654 (1928)
18. Gordon ZP 48, 11 (1928)
19. Klein ZP 53 (1929) 157
20. Dirac PRS 126, 360 (1930)
21. Oppenheimer PR 35, 562 (1930)
22. Weyl "Gruppentheorie und Quantenmechanik" 2° ed. p. 234 (1931)
23. Dirac PRS 133, 60 (1931)
24. Gray et Tarrant Nature 128 (1931) 100
25. Heisenberg et Pauli ZP 56, 1 (1929)
26. Heisenberg et Pauli ZP 59, 168 (1930)
27. Fermi Rend. Lincei 5, 881 (1929)
28. Fermi Rend. Lincei 12, 431 (1930)
29. Oppenheimer PR 35, 461 (1930)
30. Waller ZP 62, 673 (1930)
31. Wigner et Weisskopf ZP 63, 54 (1930)
32. Møller ZP 70, 786 (1931)
33. Mott PRS 124, 425 (1929)
34. Klein-Nishima ZP 52, 853 (1929)

CHAPITRE 0

35. Bethe et Heitler PRS 146, 83 (1934)
36. Heitler et Sauter Nature 132, 892 (1933)
37. Anderson Science 76 (1932) 238
38. Blackett et Occhialini PRS 139, 699 (1933)
39. Oppenheimer PR 50 (1936) 389
40. Bhabha et Heitler Nature 138 (1936) 401
41. Carlson et Oppenheimer PR 51 (1937) 220
42. Heisenberg ZP 101 (1936) 533
43. Ellis et Wooster PRS 117, 109 (1927)
44. Meitner et Orthmann ZP 60, 143 (1930)
45. Pauli lettre du 4 déc. 1930 à Geiger et Meitner
46. Chadwick Nature 129, 312 (1932)
47. Joliot et Curie C.R. Ac. Sc. Paris 194, 273 (1932)
48. Heisenberg ZP 77 (1932) 1
49. Majorana ZP 82, 137 (1933)
50. Fermi Ric. Sci. 4 (2), 491 (1933)
51. Tamm Nature 133, 981 (1934)
52. Iwanenko Ibid.
53. Yukawa Proc. phys. math. Soc. Japan 17, 48 (1935)
54. Street et Stevenson PR 51, 1005 (1937)
55. Oppenheimer et Serber PR 51, 1113 (1937)
56. Stueckelberg PR 52, 41 (1937)
57. Anderson et Neddermeyer Conf. de Londres de 1934 Hermann (1936)
58. Heisenberg AP 9, 273 (1931)
59. Dirac PRS 136, 453 (1932)
60. Dirac-Fock-Podolsky Phys. Zeits. d. Sow. 1, 801 (1932)
61. Rosenfeld ZP 76, 729 (1932)
62. Tomonaga Bull IPCR Riken-iho 22, 545 (1943)
63. Dirac Phys. Zeits. d. Sow. 3, 64 (1933)
64. Pauli et Weisskopf Helv. Phys. Acta 7, 709 (1934)
65. Fock C.R. Leningrad 6, 267 (1933)
66. Pauli PR 58, 716 (1940)
67. Weisskopf ZP 89, 27 et 90, 817 (1934)
68. Landau-Peierls ZP 69, 56 (1931)
69. Bohr Rosenfeld Kgl. Danske. Vid. Sels. fys. Medd. 12, 8 (1933)

CHAPITRE 0

70. Heisenberg ZP 110, 251 (1938)
71. Heisenberg ZP 120 (1943) 513
72. Références au chapitre 7
73. Dirac PRS 165 (1938) 199
74. Dirac Comm. Dublin Inst. Adv. Stud. ser A n° 1 (1943)
75. Pais PR 68, 227 (1945)
76. Hara et Sakata Prog. th. Phys. 2, 30 (1947)
77. Pauli et Villars Rev. Mod. Phys. 21 (1949) 434
78. Furry et Oppenheimer PR 45, 245 (1934)
79. Dirac Proc. Camb. 30, 150 (1934)
80. Heisenberg ZP 90, 209 (1934)
81. Uehling PR 48, 55 (1935)
82. Serber PR 48, 49 (1935)
83. Euler et Kockel Naturw. 23, 246 (1935)
84. Weisskopf Kong. Danske Vid. Sels. Medd. 14 n° 6 (1936)
85. Block et Nordsieck PR 52, 54 (1937)
86. Pauli et Fierz Nuov. Cim. 15, 167 (1938)
87. Kramers : Hd und Jb. d. th. Phys. I, 2 (1937)
"Quantenth. des Elektrons und der Strahlung" §§ 89, 90
88. Bethe PR 72, 339 (1947)
89. Rosenfeld "On Q.E.D." In Niels Bohr and the devt of phys. Pauli Ed.
Pergamon 1955
90. Weinberg "The Search for Unity ..." Daedalus 106 (1977).

CHAPITRES 1 et 2

Abréviations :

ZP : Zeitschrift für Physik

PRS : Proceedings of the Royal Society of London (série A)

AP : Annalen der Physik

PR : Physical Review

- I. Itzykson et Zuber : "Quantum Field Theory" Mc Graw-Hill (1980).
 - II. Cohen, Laloë, Diu : "Mécanique Quantique" Hermann (1977).
 - III. A. Hermann : "The Genesis of Quantum Theory" (1899-1913) MIT Press
 - IV. T.S. Kuhn "Black-body theory and the quantum discontinuity" (1894-1912) Oxford Un. Press (1978).
 - V. M. Jammer : "The Conceptual development of Quantum Mechanics" NY 1966
-
1. Planck Berliner Ber. (Déc. 14, 1900) 1-700
 2. Rayleigh.Phil. Mag. 49, 539 (1900)
 3. M. Planck "Vorlesungen über die theorie der Wärmestrahlung" Leipzig 1906
 4. Rayleigh Nature 72, p. 54 et p. 243 (1905)
 5. Jeans Phil. Mag. 10 p. 91 (1905)
 6. Ehrenfest Phys. Z. 7, 528 (1906)
 7. Einstein AP 17 S132-148 (1905)
 8. Debye AP 33, 1427 (1910)
 9. Einstein Phys. ZS 10, 185, 817 (1909)
 10. H.A. Lorentz "Les Théories Statistiques en Thermodynamique" (Leipzig 1916) P. 59
 11. Ehrenfest Comptes Rendu publié dans ZP 34, 362 (1925)
 12. Born Heisenberg Jordan ZP 35, 557 (1926)
 13. DIRAC PRS 114, p. 243 (1927)
 - 13'. Dirac PRS 114, 710 (1927)
 14. Mehra in Aspects of Quantum Theory Ed. by Salam and Wigner Cambridge Un. Press p. 17 à 59
 15. Eden and Polkinghorne Ibid. p. 2
 16. Dirac "The Principles of Quantum Mechanics" 3^e ed. (1947) Oxford
 17. HEISENBERG-PAULI I ZP 56, 1 (1929)
 18. HEISENBERG-PAULI II ZP 59, 168 (1930)
 19. FERMI Rend. Lincei 5, 881-997 (1929)

CHAPITRES 1 et 2

20. FERMI Rend. Lincei 12, 431 (1930)
21. Fermi Rev. of Modern Phys. 4, 125 (1932)
22. H. Weyl Gruppen_theorie und Quantenmechanik Leipzig Hirzel (1928)
Traduction anglaise par H.P. Robertson : The Theory of Groups and
Quantum Mec. Dover
23. H. Weyl Sitz. d. Preuss Ak. d. Wiss. P. 465 (1918)
24. Schrödinger ZP 12, 13 (1922)
25. London ZP 42, 375 (1927)
26. Oppenheimer PR 35, 461 (1930)
27. Wentzel "Einführung in die Quantentheorie der Wellenfelder"
Franz Deuticke, Wien, 1943
28. S.T. Ma PR 75, 535 (1949)
29. Belinfante PR 76, 226 (1949)
30. Gupta PRS A 63, 46 (1950)
31. Bleuler Helv. Phys. Acta 23, 567 (1950)
32. Rosenfeld AP 5, 113 (1930)
33. W. Heitler "The Theory of Radiation" Oxford (1936)
34. Landau ZP 45, 430 (1927)
35. Bloch Phys. Zeits. XXIX (58) (1928)
36. Ehrenfest Naturwiss 11, 543 (1923)
37. Weisskopf et Wigner ZP 63, 54 (1930)
38. Weisskopf et Wigner ZP 65, 18 (1930)
39. Dirac ZP 44, 485 (1927)
40. A. Unsöld Zs f. Astrophys. 2, 199 (1931)
41. W. Schütz ZP 45, 30 (1927)
42. Rosenfeld ZP 71, 273 (1931)
43. Casimir ZP 81, 496 (1933)
44. Heisenberg AP 9, 338 (1931)
45. Fermi Rend. Lincei 10, 72 (1929)
46. Racah Rend. Lincei 11, 837 et 1100 (1930)
47. G. Breit PR 34, 553 (1929)
48. G. Breit PR 39, 616 (1932)
49. C. Møller ZP 70, 786 (1931)
50. C. Møller AP 14, 531 (1932)
51. Dirac PRS A114, 710 (1927)

CHAPITRES 1 et 2

52. Kramers et Heisenberg ZP vol. 31 p. 681 (1925)
53. Mott PRS A126, 259 (1930)
54. Rosenfeld ZP 73, 253 (1931)
55. Bethe et Fermi ZP 77, 256 (1932)
56. Raman et Krishnan Nature 121, 501 (1928)
57. G. Landsberg et L. Mandelstamm Naturev. 16, 557 et 772 (1928)
58. Amaldi Rend. Lincei (1929)
59. Weisskopf AP 9, 23 (1931)
60. Weisskopf ZP 85, 451 (1933)
61. Heitler ZP 84, 145 (1933)
62. Bethe et Heitler PRS 146, 83 (1934)
63. Mott Proc. Camb. Phil. Soc. 27, 255 (1931)
64. Sommerfeld AP 11, 257 (1931)
65. Waller ZP 58, 75 (1929)
66. Waller ZP 61, 837 (1930)
67. Einstein, Ehrenfest ZP 19, 301 (1923)
68. Mayer-Göppert Naturw. 17, 932 (1929)
69. Mayer-Göppert AP 9, 273 (1931)
70. S.N. Bose ZP 26, 178 (1924)
71. Einstein Sitzungb. d. Preuss Ak. de Wiss. 3 (1925)
72. Fermi Atti. Accad. Lincei 3, 145 (1926) ; ZP 36, 902 (1926)
73. Dirac PRS A112, 661 (1926)
74. Jordan et Klein ZP 45, 751 (1927)
75. Schrödinger AP 82, 265 (1927)
76. Jordan ZP 45, 766 (1927)
77. Fock ZP 49, p. 339 (1928)
78. Jordan ZP 44, 473 (1927)
79. Jordan ZP 44, 1 (1927)
80. Jordan-Wigner ZP 47, 631 (1928)
81. Fock ZP 75, 622 (1932)
82. Landau et Peierls ZP 62, 188 (1930)
83. Heisenberg ZP 90, 209 (1939)
84. Landau-Peierls ZP 69, 56 (1931)
85. Landau-Peierls ZP 62, 188 (1930)

CHAPITRES 1 et 2

86. Lorentz "The Theory of Electrons" p. 217 Leipzig 1909
87. Compton Journ. Frankl. Inst, 192, 145 (1921)
88. Uhlenbeck et Goudsmit Naturw. 13, 953 (1925)
89. Pauli ZP 43, 601 (1927)
90. Darwin PRS A116, 227 (1927)
91. Darwin Nature 119, 282 (1927)
92. Klein ZP 41, 407 (1927)
93. Gordon ZP 40, 117 (1926)
94. Dirac "The Development of Quantum Theory" Gordon and Breach(1971)
95. Mehra in Aspects of Quantum Theory Salam and Wigner Ed.
96. Dirac PRS A117, 610 (1928)
97. Dirac PRS A118, 351 (1928)
98. Darwin PRS A118, 654 (1928)
99. Darwin PRS A120, 621 (1928)
100. Gordon ZP 40, 11 (1928)
101. Sommerfeld "Atombau und Spektrallinien" Braunschweig 1919 § 5
(trad. française existe)
102. Klein Nishima ZP 52, 853 (1929)
103. Nishima ZP 52, 869 (1929)
104. Tamm ZP 62, 545 (1930)
105. Compton PR 21, 483 (1923)
106. Breit PR 27, 362 (1926)
107. Dirac PRS 111, 405 (1926)
108. Klein Nishima Nature 122, 398 (sept. 1928)
109. Schrödinger Sitzungsberichte d. Berlin, Ak. 418 (1930)
110. Dirac Phys. Zeits. 29, 561 (1928)
111. Tamm ZP 62, 545 (1930)
112. Breit Proc. Nat. Ac. Am . 14, 553 (1925)
113. Fock ZP 55, 127 (1929)
114. Klein ZP 53, 157 (1929)
115. Dirac PRS 126, 360 (1930)
116. Weyl ZP 56, 332 (1929)
117. Dirac Proc. Camb. Phil. Soc. 26, 361 (1930)
118. Oppenheimer PR 35, 939 (1930)
119. Oppenheimer PR 35, 562 (1930)
120. H. Weyl "Gruppentheorie und Quantenmechanik" 2° ed. p. 234 (1931)
ou traduction anglaise "The Theory of Groups and Quantum Mechanics"
Dover (p. 262, 263, 264)

CHAPITRES 1 et 2

121. Dirac PRS 133, 60 (1931)
122. Pauli Handb. d. Phys. Geiger Scheel, Bd. XXIV /I (p. 246)
123. Schrödinger Berl. Ber. 63 (1931)
124. Anderson PR 43, 491 (1933)
125. Anderson Science 76, 238 (1932)
126. Blackett et Occhialini PRS 139A, 699 (1933)
127. Curie et Joliot JOurn. de Phys. 4, 21 (p. 31) (1933)
128. Curie et Joliot Comptes Rendus de l'Ac. des Sc. 196 p. 1105
129. Blackett Nature 132, 917 (1933)
130. Oppenheimer et Plesset PR 44, 53 (1933)
131. Heitler et Sauter Nature 132, 892 (1933)
132. Rapport du Congrès Solvay 1933 - Gauthier Villars 1934
133. Jordan Pauli ZP 47, 151 (1928)
134. Rosenfeld et Solomon Journal de Phys. 2, 139 (1931)
135. Möglich Ann. der Phys. 2, 676 (1929)
136. Amaldi The collected papers of Enrico Fermi p. 305 Chicago Press
137. Fermi Nuovo Cimento 8, 121 (1931)
138. Eddington PRS 121, 524 (1929) ; 122, 358 (1929)
139. Tetrode ZP 49, 858 (1928)
140. Möglich 48, 852 (1928)
141. Mie Ann. d. Phys. 85, 711 (1928)
142. Rosenfeld ZP 70, 454 (1931)
143. Weisskopf ZP 89, 27 (1934)
144. Dirac PRS A136, 453 (1932)
145. Frenkel Phys. Zeits. d. Sow. 1, 99 (1932)
146. Dirac Fock et Podolsky Phys. Z. d. Sow. 2, 6 (1932)
147. Rosenfeld ZP 76, 729 (1932)
148. Fock et Podolsky Phys. Z. d. Sow. 1, 301 (1932)
149. F. Bloch Phys. Zeits. d. Sow. 5, 301 (1934)
150. Fock et Podolsky Phys. Z. d. Sow. 2, 275 (1932)
151. Nikolsky Phys. Z. d. Sow. 2, 447 (1932)
152. Tomonaga Progr. of Th. Phys. 1, 27 (1946)
153. Heisenberg ZP 110, 251 (1938)
154. Sauter ZP 69, 742 (1931)

CHAPITRES 1 et 2

155. Pauli Phys. Zeits. 20, 457 (1919)
156. Einstein Physica 5, 330 (1925)
157. Weisskopf, Gregory Lectures 1979 "The beginning of Field Theory - A personal Recollection" CERN 1^o Nov. 1979
158. Weisskopf "Physics in the 20th century" MIT pr. (1972)
159. Jordan in "The physicist's conception of nature" ed. by Mehra p. 296 (1972)
160. Van der Waerden - Ibid. p. 429
161. Born et Jordan ZP 34, 858 (1925)
162. Pauli Wissenschaftlicher Briefwechsel Hermann, Meyenn, Weisskopf ed.
163. J. Bromberg HSPS 7, 161 (1976)
164. Peierls in Aspects of Quantum Theory Ed. by Salam and Wigner Cambridge
165. Gray Nature 129, 241 (1929)
166. Mott PRS 124, 425 (1929)
167. Rich and Wesley "Current Status of the lepton g-factors" Rev. Mod. Phys. 44, 250 (1972)
168. Moyer Am. J. of Phys. 49, n^o 10, 11 et 12 (1981)
169. Dymond PRS 145, 657 (1934)
170. Rupp ZP 79, 644 (1932)
171. Congrès International d'Electricité - Volume II Gauthier Villars 1932
172. Klein ZP 53, 157 (1929)
173. Natanson Phys. Zeits. 12, 659 (1911)
174. Ehrenfest Annalen d. Phys. 36, 91 (1911)
175. Dirac Proc. Camb. 35, 416 (1939)
176. Smekal Naturw. 11, 873 (1923)

CHAPITRE 3

1. Pauli : Die Naturwissensch.48, 55 (1933)
2. Dirac : PRS 114, 243 (1927)
3. Heitler : "The quantum theory of radiation", Oxford 1936
4. Wiggaker, Sir Edmond : "A history of the theories of aether and electricity"
Vol I : Thomas Nelson and Sons Ltd., London 1951
Vol II : Phil. Library, N.Y. 1954
5. Rohrlich : "Classical charged particles" Addison-Wesley chap. 2
6. Pais : "Theory of the electron" in "Aspects of quantum theory"
Salam and Wigner Ed. Cambridge Univ. Press.
7. Thomson : Phil. Mag. (5), 11, 229 (1881)
8. Abraham : AP 10, 105 (1903)
9. Lorentz : "The theory of electrons and its applications to the phenomena
of light and radiant heat", 2° Ed. Dover Pub. N.Y. 1952
(d'après des cours de 1906)
10. Kaufmann : AP 19, 487 (1906)
11. Poincaré : Comptes rendus 140, 1504 (1905)
12. Goudsmit et Uhlenbeck : Physica 6, 273 (1926)
13. Heisenberg Pauli I : ZP 56, 1 (1929)
14. Heisenberg Pauli II : ZP 59, 168 (1930)
15. Oppenheimer : PR 35, 461 (1930)
16. Dirac : ZP 44, 585 (1927)
17. Bethe : PR 72, 339 (1947)
18. Kroll et Lamb : PR 75, 388 (1949)
19. Waller : ZP 62, 673 (1930)
20. Rosenfeld : ZP 70, 454, (1931)
21. Rosenfeld : ZP 65, 589 (1930)
22. Solomon : ZP 71, 162 (1931)
23. Weisskopf : ZP 89, 27 (1939)
24. Weisskopf : ZP 90, 817 (1934)
25. Weisskopf : PR 56, 72 (1939)

CHAPITRE 3

26. Schrödinger : Berl. Ber. 1930, 418 (1930)
27. Pauli et Weisskopf : Helv. Phys. Acta 7, 709
28. Heisenberg : ZP 90, 209 (1934)
29. Wentzel : Th. Phys. in the 20th cent., Pauli-Weisskopf, Ed. (1960), p. 60
30. Dancoff : PR 55,959 (1939)
31. Dirac : Théorie du positron - Rapport du 7^e Congrès Solvay de Physique Structure et propriétés des noyaux atomiques p. 203 (1934)
32. Dirac : "The principles of quantum mechanics" 3^e Ed. Oxford (1947) p. 308
33. Pauli : Handbuch de Phys. Geiger Scheel Vol 24 T. 1, 272 (1933)
34. Villars : Theoretical Phys. in the 20th cent. Fierz and Weisskopf Ed. Interscience Pub. p. 80
35. Wentzel : Réf. précédente (34) p. 54-55
36. Heisenberg : ZP 65, 4 (1930)
37. Jordan : Ergebnisse der ex.Naturw. 7, 207 (1928)
38. Pauli : Wissenschaftlicher Briefwechsel Hermann, Meyenn, Weisskopf Ed. Springer (1979)
39. Serber : PR 45, 545 (1936)
40. Dirac : PRS 133, 60 (1931)

CHAPITRE 4

Abréviations :

- ZP : Zeitschrift für Physik
- PRS : Proceedings of the Royal Society of London (Série A)
- A.P. : Annalen der Physik
- PR : Physical Review

I Itzykson et Zuber "Quantum Field Theory" (1980) Mc. Graw-Hill

II Cohen, Laloë, Diu "Mécanique Quantique" Hermann 1977

1. Landau-Peierls Zf. Ph. 69 (1931), 56
2. Dirac Proc. Roy. Soc. London (A) 117, 610 (1927)
3. Heisenberg-Pauli Zf.Phys. 56,1 (1929)
4. Heisenberg-Pauli Zf. Phys. 59,168 (1930)
5. Dirac Proc. Roy. Soc. A-117 613 (1928)
6. N.Bohr. Die Naturwissenschaften 16,245(1928)
7. Landau-Lifchitz. Cours de phys. th. Tome III §44
8. Landau-Lifchitz. Cours de phys. th. Tome I §1
9. Par ex. Réf. I p67
10. Landau-Peierls Zf. Phys. 62,188 (1930)
11. Heisenberg Ann. d.Phys. 32,20 (1938)
12. Wu in th. Phys. in the 20th cent. Fierz and Weisskopf ed. p252
13. " " " " " " p245
14. Bohr-Rosenfeld. Kgle Danske Vid. Sels. fys. Medd. 12,8 (1933)
15. Traduction anglaise du papier précédent par Aage Petersen in Selected papers of Leon Rosenfeld p.357 à 400
16. Rosenfeld "Niels Bohr and the development of physics" - Pergamon 1955- p.70 Pauli ed.
17. id p.73
18. Jordan Pauli Zf. Physik 47 (1928) 151
19. Thirring et Touschek : Phil. Mag.62,244 (1951)
20. Peierls Proc.Roy. Soc. A214,143 (1952)
21. Rosenfeld Zf. Phys. 70,454 (1931)
22. Bohr-Rosenfeld Phys. Rev. 78,6 (1950)794
23. Heisenberg Sachs. Ak. der Wiss. Vol 86 p.317 (1934)

CHAPITRE 4

25. Heisenberg Zf. Phys. 90,205 (1934)
26. Pais, Development in the theory of the electron (Princeton Un.Press, Princeton, N.J. 1948)
27. Bohr et Rosenfeld. Phy. Rev 78,6 (1950) 794
28. Heisenberg Ann. der Phys. 5-32 (1938)
29. Heisenberg Zeits. f. Phys. 110-251 (1938)
30. Heisenberg Zeits. f. Phys. 101-533 (1936)
31. Pauli Handb d. Phys. Geiger-Scheel Vol 24 part I p. 246 (note) (1933)
32. Heisenberg. Zf. Phys. 120 (1943)
33. Heisenberg, conference at the Zürich Seminar for theoretical physics (1944)
34. Dirac. Quantum Mechanics. 2° édition p.197 et suiv. Oxford 1935
35. Wentzel Rev. Mod. Phys. 19,1 (1947)
36. Heisenberg Zf. Phys. 120,673 (1943)
37. Heisenberg. Der Teil and das Ganze (Piper Munich 1969) p.92
trad. Physics and Beyond (Harper and Row, N.Y., 1971) P.63
38. Aharonov et Albert Phys. Rev 21D 12 (1980) 3316
39. Aharonov et Albert Phys. Rev 24D 2 (1981) 359
40. Niels Bohr "Atomic theory and the description of nature" Cambridge 1934
41. W. Heisenberg "Die Physikalischen Prinzipien der Quantentheorie" Leipzig 1930
42. Dirac "The principles of quantum Mechanics" Oxford 1930
43. Landau. Mécanique quantique - Editions de Moscou
44. D'Espagnat. "Conceptual foundations of quantum mechanics" 2nd Ed. p.250
Benjamin 1976
45. Congrès de Rome 1931
46. Congrès de Solvay 1933
47. Moyer
48. Bohr
49. Cassidy. HSPS 12:1 (1981)
50. Heisenberg Zf. Phys. 65 (1930) 5
51. Holton "Heisenberg, Oppenheimer et l'émergence de la physique moderne"
La Recherche. N°130 p.190 (1982)

CHAPITRE 5

Abréviations Physical Review : PR
 Zeitschrift für Physik : ZP
 Annalen der Physik : AP
 Proceedings of the Royal Society : PRS (Série A)

1. Mott Proc. Camb. Phil. Soc. 27, 255 (1931)
2. Sommerfeld AP 11, 257 (1931)
3. Bethe et Heitler PRS A 146, 83 (1934)
4. Knipp et Uhlenbeck Physica 3, 425 (1936)
5. Bloch PR 50, 272 (1936)
6. Bloch et Nordsieck PR Vol. 52 p. 54 (1937)
7. Braunbeck et Weinmann ZP 110, 360 (1938)
8. Pauli et Fierz Nuovo Cim. 15, 167 (1938)
9. Dancoff PR 55, 959 (1939)
10. Bethe et Oppenheimer PR 70, 451 (1946)
11. Eliezer PRS 187 (1946)
12. Pauli et Jauch PR 65, 255 (A) (1944)
13. Jost PR 72, 815 (1947)
14. Heitler The Quantum theory of radiation 2nd ed. Oxford (1944)
15. Jauch et Rohrlich Helv. Phys. Acta 27, 613 (1954)
16. Morpurgo Nuovo Cim. 8, 109 (1950)
17. Dirac "Théorie du Positron" Rapport du 7ème Conseil Solvay de Physique 203 (1934)
18. Dirac Proc. Camb. Phil. Soc. 25, 62 (1929) ; 26, 376 (1930) ; 27, 240 (1931)
19. Furry et Oppenheimer PR 45, 245 (1934)
20. Dirac Proc. Camb. Phil. Soc. 30, 150 (1934)
21. Peierls PRS A 146, 420 (1934)
22. Heisenberg ZP 90, 209 (1934)
23. Heisenberg ZP 92, 692 (1934)
24. Euler et Heisenberg ZP 98, 714 (1936)
25. Valatin PRS A 222, 93 (1954)
26. Wightman in "Aspects of Quantum theory" Salam and Wigner Ed. Cambridge
27. Heisenberg AP 10, 888 (1931)

CHAPITRE 5

28. Wentzel "Einführung in die Quantentheorie der Wellenfelder"
Wien (1943)
29. Uehling PR 48, 55 (1935)
30. Serber PR 48, 49 (1935)
31. Pauli et Rose PR 49, 462 (1936)
32. Breit et Wheeler PR 46, 108 (1934)
33. Sauter ZP 69, 742 (1931)
34. Halpern PR 44, 855 (1933)
35. Delbrück ZP 84, 144 (1933)
36. Euler et Kockel Natura 23, 246 (1935)
37. Euler AP 26, 398 (1935)
38. Furry PR 51, 125 (1937)
39. Weisskopf Kong. Danske Vid. Sels. Medd. XIV N°6 (1936)
40. Wigner "Group theory and its application to the Q. M. of atomic spectra"
41. Pauli ZP 31, 765 (1925)
42. Heisenberg-Pauli ZP 56, 1 p. 29, 30 (1929)
43. Jordan Ergebnisse der exacten Naturw. 7 p. 205-206
44. Pauli "Handbuch der Physik Geiger Scheel" 24 p. 258
45. Rapport du congrès de Solvay 1933 p. 213
46. Fermi Ric. Sci. 2. N. 12 (1933)
47. Yukawa Proc. Phys. Math. Soc. Japan 17, 48
48. Proca J. Phys. Radium 7, 347 (1936)
49. Stueckelberg Helv. Phys. Acta XI 299 (1938)
50. Kemmer PRS A 166, 154 (1938)
51. Yukawa, Sakata, Taketani Proc. Phys. Math. Soc. Japan 20, 319
52. Bhabha PRS A 166, 501
53. Fröhlich, Heitler, Kemmer PRS A 166, 154
54. Sakata, Yukawa Proc. Phys. Math. Soc. Japan 19, 91 (1937)
55. Darwin PRS A 118, p. 657 (1928)
56. Wigner et von Neumann ZP 47, 203 (1927)
57. van der Waerden, "in theoretical physics in the 20th cent."
Fierz-Weisskopf ed.
58. H. Weyl Gruppentheorie und Quantenmec. Leipzig (1928)
59. van der Waerden Nach. Ges. Wiss. Göttingen mat. phys. 1929, 100
60. Laporte et Uhlenbeck PR Vol. 37 1380 (1931)

CHAPITRE 5

61. Dirac PRS A 117, 615 (1928)
62. Majorana Nuovo Cim. 9, 335 (1932)
63. Dirac PRS A 155, 447 (1936)
64. Wigner, Ann. of math. 40, 149 (1939)
65. Wigner ZP 124, 665 (1947)
66. Pauli et Weisskopf, Helv. Phys. Acta 7, 709 (1974)
67. Iwanenko et Sokolov, Phys. Zeit. Sov. 11, 590 (1937)
68. Pauli, Ann. Inst. Poincaré 6, 137, (1936)
69. De Wet PR 57, 646 (1940)
70. Fierz, Helv. Phys. Acta 12, 3, 1939
71. Pauli et Fierz, Helv. Phys. Acta 12 (297) 1939
72. Pauli PR 59 (1940) 716
73. H. Weyl "Gruppen Theorie und Quanten mechanik" 2° ed. p. 234 (1931)
Traduction anglaise Dover p. 225 et 264
74. Pauli, Ann. Inst. Poinc. 6 (1936) 130
75. Kramers, Proc. Roy. Ac. Amsterdam 40 (1937) 814
76. Majorana Nuov. Cim. 14, 171 (1937)
77. Belinfante, Physica 6 (1939) 849
78. Belinfante, Physica 6 (1939) 870
79. Pauli et Belinfante, Physica 7 (1940) 177
80. Pauli, Prog. Th. Phys. Kyoto 5 (1950) 526
81. Pauli in "Niels Bohr and the development of physics" Pauli ed. p. 30
82. Kemmer et Weisskopf, Nature 137 (1936) 659
83. Kemmer, Helv. Phys. Acta 10 (1937) 112
84. Weisskopf "Physics in the 20th Cent" MIT press 1972
85. Tomonaga, Conf. Nobel (6 mai 1966) réadité par Mehra in "the physicist's
conception of nature", 1973
86. Fock CR Leningrad 6 (1933) 267
87. Kemmer, Helv. Phys. Acta 10 (1937) 112
88. Pauli, Wissenschaftlicher Briefwechsel. Hermann, Meyenn, Weisskopf
Ed. Springer (1980)
89. Pauli, Nature 17 (1929) 259
90. Serber PR 49 (1936) 545
91. Dirac PR5 136 (1932) 453
92. Dirac, Fock, Podolsky, Phys. Zeits. der Sow. 2 (1932) 468
93. Dirac, Phys. Zeit. der Sow. 3, (1933) 64
94. Rosenfeld ZP 76 (1932) 729
95. Bloch, Phys. Zeits. der Sow. 5 (1934) 301

CHAPITRE 5

96. Stückelberg, Helv. Phys. Acta 11 (1938) 225 § 5
97. Dirac, Rev. Mod. Phys. 17 (1945) 195
98. Feynman, Rev. Mod. Phys. 20 (1948) 267
99. Tomonaga, Nobel lecture in "The Physicists Conception of Nature", Mehra Ed. Reidel Pub. Co.
100. Yukawa, Kagaku 12 (1943) 251, 282 et 322
101. Tomonaga, Bull. IPCR Riken-iho 22 (1943) 545
102. Tomonaga, Prog. Theor. Phys. 1 (1946) 27
103. Koba Tati Tomonaga Prog. Theor. Phys. 2 (1947) 101
104. Koba Tati Tomonaga Prog. Theor. Phys. 2 (1947) 198
105. Kanesawa, Tomonaga Prog. Theor. Phys. 3 (1948) 1
106. Kanesawa, Tomonaga Prog. Theor. Phys. 3 (1948) 101
107. Watanabe PTP 2 (1947) 71 et PTP 3 (1948) 378
108. Tanikawa PTP 2 (1947) 71
109. E. Cartan Bull Soc Math de France 41 (1913) 53
110. Pauli Rev. Mod Phys 13 (1941) 203 (p. 223)
111. Pauli Rev. Mod Phys 15 (1943) 175 (p. 175 note 2)

CHAPITRE 6

Abréviations :

ZP : Zeitschrift für Physik

PRS : Proceedings of the Royal Society of London (série A)

AP : Annalen der Physik

PR : Physical Review

I. Itzykson et Zuber : "Quantum Field Theory" Mc Graw-Hill (1980).

II. Cohen, Laloë, Diu : "Mécanique Quantique" Hermann (1977).

1. Chadwick : J. Verh. dtsch. Phys. Ges. 16, 383 (1914)

2. Ellis et Wooster : Proc. Camb. Phil. Soc. 22, 849 (1925)

3. Ellis et Wooster : PRS A117, 109 (1927)

4. Meitner : ZP 11, 35 (1922)

5. Meitner et Orthmann : ZP 60, 143 (1930)

6. Bohr : J. Chem. Soc. 349 (1932)

7. Pauli : Correspondance Springer (1981)

8. Rutherford : Bakerian lecture PRS A97, 374 (1920)

9. Glasson : Phil. Mag. 42, 596 (1921)

10. Curie et Joliot : C.R. Ac. Sc. Paris 194, 273 (1932)

11. Chadwick : Nature 129, 312 (1932)

12. Chadwick : PRS A136, 692 (1932)

13. Iwanenko : Nature 129, 798 (1932)

14. Heisenberg : ZP 77, 1 (1932)

15. Majorana : ZP 82, 137 (1933)

16. Rapports du 7ème Congrès Solvay Gauthier-Villars (1934)

17. Bethe et Becker : ZP 66, 289 (1930)

18. Curie Irène : C. R. Ac. Sci. Paris 193, 1412 (1931)

19. Fermi : Ric. Sci. 4(2), 491 (1933)

20. Fermi : Nuov. Cim. 11, 1 (1934)

21. Fermi : ZP 88, 161 (1934)

22. Gamow et Teller : PR 49, 895 (1936)

23. Carter, Reines, Wagner, Wyman : PR 113, 280 (1959)

24. L. Leprince-Ringuet : Les rayons cosmiques Albin Michel (1945)

25. Carlson et Oppenheimer : PR 38, 1787 (1931)

26. Carlson et Oppenheimer : PR 41, 763 (1932)

CHAPITRE 6

27. Bethe : ZP 76, 293 (1932)
28. Heisenberg : AP 13, 430 (1932)
29. Heitler : ZP 84, 145 (1933)
30. Blackett et Occhialini : PRS 139, 699 (1933)
31. Oppenheimer et Plesset : PR 44, 53 (1933)
32. Anderson : PR 44, 406 (1933)
33. Williams : PRS 139, 163 (1933)
34. Weizsäcker : ZP 88, 612 (1934)
35. Landau : Phys. Zeits. d. Sow. 5, 761 (1934)
36. Oppenheimer : PR 47, 44 (1935)
37. Nordheim : PR 49, 189 (1936)
38. Congrès de Londres de 1934 Hermann (1936)
39. Anderson et Neddermeyer : PR 50, 268 (1936)
40. Oppenheimer (Seattle meeting) : PR 50, 389 (1936)
41. Bhabha et Heitler : Nature 138, 401 (1936)
42. Bhabha et Heitler : PRS 159, 432 (1937)
43. Oppenheimer et Carlson : PR 51, 220 (1937)
44. Crussart et Leprince-Ringuet : C.R. 204, 240 (1937)
45. Anderson et Neddermeyer : PR 51, 884 (1937)
46. Street et Stevenson : PR 51, 1005 (1937)
47. Oppenheimer et Serber : PR 51, 1113 (1937)
48. Stückelberg : PR 52, 41 (1937)
49. Oppenheimer : Physics Today 19:11 p. 58 (1966)
50. C.G. et M.O. Montgomery : PR 50, 403 (1936)
51. Yukawa : Proc. Phys. Math. Soc. Japan 19, 712 (1937)
52. Yukawa : Proc. Phys. Math. Soc. Japan 17, 48 (1935)
53. Yukawa : Proc. Phys. Math. Soc. Japan 17, 1084 (1935)
54. Yukawa Sakata Taketani : Proc. Phys. Math. Soc. Japan 20, 319 (1936)
55. Proca : J. Phys. Radium 7, 347 (1936)
56. Kemmer : Nature 141, 116 (1938)
57. Bhabha : Nature 141, 117 (1938)
58. Stückelberg : Helv. Phys. Acta 11, 299 (1938)
59. Fröhlich Heitler Kemmer : PRS 166, 154 (1938)
60. Kemmer : PRS A166, 127 (1938)
61. Kemmer : Proc. Camb. Phil. Soc. 34, 354 (1938)

CHAPITRE 6

62. Breit et Feenberg : PR 50, 850 (1936)
63. Cassen et Condon : PR 50, 846 (1936)
64. Wick : R. C. Acc Lincei 21, 170 (1935)
65. Heisenberg : ZP 113, 61 (1939)
66. Wentzel : Helv. Phys. Acta 13, 269 (1940)
67. Brückner : PR 86, 106 (1952)
68. Rasetti in "The collected papers of E. Fermi" p. 538 à 540
69. Wu in "Theoretical physics in the 20th century" Interscience pub. NY (1960)
70. Wentzel ibid.
71. Wentzel : Rev. Mod. Phys. 19, 1 (1947)
72. Bethe, de Hoffmann : "Meson and Fields" II Row, Peterson and Co. (1955)
73. Bethe-Bacher : Rev. Mod. Phys. 8, 200 (1936)
74. Bethe : Rev. Mod. Phys. 9, 65 (1936)
75. Bethe et Livingston : Rev. Mod. Phys. 9, 245 (1937)
76. Møller Rosenfeld : Kgl. Danske Vid. Sels. Math.-Fys. Med 17, n°8 (1940)
77. Heisenberg : ZP 101, 533 (1936)

CHAPITRE 7

1. Dirac : PRS 167, 148 (1938)
2. Shankland : Nature 137, 241 (1936)
3. Dirac : Nature 137, 298 (1936)
4. Dirac : PRS 155, 447 (1936)
5. Dirac : PRS 160, 48 (1937)
6. Dirac : Nature 139, 1001 (1937)
7. Dirac : PRS 165, 199 (1938)
8. Dirac : Nature 139, 323 (1937)
9. Pais : Theory of the electron (1897-1947) in Aspects of quantum theory Salam and Wigner Ed. Camb. Univ. Press 1972
10. Frenkel : Naturw. 12, 882 (1924)
11. Frenkel : ZP 32, 518 (1925)
12. Fokker : ZP 58, 386 (1929)
13. Rohrllich : "Classical charged particles" Addison-Wesley Pub. Comp. 1965
14. Dirac : Ann. Inst. Poincaré 9,13 (1939)
15. Wentzel : ZP 86, 479 (1933)
16. Wentzel : ZP 36, 635 (1933)
17. Dirac, Fock, Podolsky : Phys. Z. d. Sov. 2, 468 (1932)
18. Waller : ZP 62, 673 (1930)
19. Weisskopf : ZP 89, 27 et 50 (1934), 817 (1934)
20. Pauli : Rev. Mod. Phys. 15, 175 (1943)
21. Riesz : Conférence de la réunion internationale des mathématiques tenue à Paris en Juillet 1937 (Paris 1939)
22. Ma S.T. : Phys. Rev. 71, 787 (1947)
23. Weisskopf : Phys. Rev. 76, 72 (1939)
24. Dirac : PRS 190, 1 (1941)
25. Dirac : Comm. Dublin Inst. Adv. Stud, ser A n° 1 (1943)
26. Pauli : Rev. Mod. Phys. 15,175 (1943)

CHAPITRE 7

27. Pauli et Weisskopf : Helv. Phys. Acta 70 , 709 (1934)
28. T. Lewis : Phil. Mag. 36, 533 (1945)
29. Dirac : Phil. Mag. 39, 31 (1948)
30. Report of an international conf. Cambridge July 1946
Vol. I Taylor and Francis (1947)
31. Dirac : Comm. Dublin Inst. Adv. Stud. ser. A n° 3
32. Eliezer : Bull. Calcutta Math. Soc. Dec. 1946
33. Eliezer, Rev. Mod. Phys. 19, 3 (1947)
34. Jauch : PR 63, 334 (1943)
36. Pauli : PR 64, 332 (1943)
37. Jauch Pauli : PR 65, 255 (1944)
38. Eliezer : PRS 191, 130 (1947)
39. Pauli : Helv. Phys. Acta 19, 234 (1946)
40. Gupta : PRS 63, 46 (1950)
41. Bleuler : Helv. Phys. Acta 23, 567 (1950)
42. Bhabha : PRS 172, 384 (1939)
43. Wigner et Weisskopf : ZP 63, 54 (1930)
44. Heitler : "Quantum theory of radiation" 2nd Ed. Oxford 1944
45. Waller : ZP 88, 436 (1934)
46. Góra : Acta Phys. Polon. 7, 159 (1938)
47. Góra : Acta Phys. Polon. 7, 374 (1939)
48. Wilson : Proc. Camb. 37, 301 (1941)
49. Heitler : Proc. Camb. 37, 291 (1941)
50. Heitler et Peng : Proc. Camb. 38, 256 (1942)
51. Góra : ZP 120, 121 (1943)
52. Pauli : "Meson theory" (1946) Interscience Pub.
53. Sokolov : J. Phys. USSR 5, 231 (1941)
54. Dirac : ZP 44, 585 (1927)

CHAPITRE 7

55. Breit et Wigner : PR 49, 519 (1936)
56. Bethe et Oppenheimer : PR 70, 451 (1946)
57. Ma et Hsüeh : Proc. Camb. 40, 167 (1944)
58. Rapport du 8° congrès Solvay - Stoops 1950
59. Pryce : PRS 168, 389 (1938)
60. Bhabha : Nature 143, 276-277 (1939)
61. Heisenberg : ZP 113, 61 (1939)
62. Booth et Wilson : Proc. Camb. 36, 446 (1940)
63. Heitler et Ma : PRS 176, 368 (1940)
64. Bhabha : Nature 145, 819 (1940)
65. Bhabha : Proc. Indian Acad. A 11, 247 (1940)
66. Wheeler : ZP 52, 1107 (1937)
67. Heisenberg : ZP 120, 513 (1943)
68. Heisenberg : ZP 120, 673 (1943)
69. Heisenberg : Zeits. für Naturforschung 1, 608 (1946)
70. Podolsky et Kikuchi : Phys. Rev. 65, 228 et 67 (1944), 184 (1944)
71. Ma : Phys. Rev. 69, 668 (1946)
72. Møller : Kongl. Danske Vid. Sels. Mat. Fys. Medd 23 n°1 (1946)
73. Møller : Kongl. Danske Vid. Sels. Mat. Fys. Medd 22 n° 19 (1946)
74. Stueckelberg : Helv. Phys. Acta 11, 225 (1938)
75. Stueckelberg : Helv. Phys. Acta 14, 51 (1941)
76. Stueckelberg : Helv. Phys. Acta 17, 3 (1944)
77. Stueckelberg : Helv. Phys. Acta 18, 195 (1945)
78. Stueckelberg : Helv. Phys. Acta 19, 242 (1946)
79. Wentzel : Rev. Mod. Phys. 19, 1 (1947)
80. Stueckelberg, Rivier : PR 74, 218 (1948)
81. Rivier : Helv. Phys. Acta 22, 265 (1949)

CHAPITRE 7

82. Dyson : PR 75, 1736 (1949)
83. Pauli et Villars : Rev. Mod. Phys. 21, 434 (1949)
84. Mie : AP 37, 511 (1912)
85. Mie : AP 39, 1 (1912)
86. Mie : AP 40, 1 (1913)
87. Born : Nature 132, 282 (1933)
88. Born, Infeld : Nature 132, 1004 (1933)
89. Born : PRS 143, 410 (1934)
90. Born : Nature 133, 63 (1934)
91. Born, Infeld : PRS 144, 425 (1934)
92. Born, Infeld : PRS 147, 522 (1934)
93. Born, Infeld : PRS 150, 141 (1935)
94. Pryce : PRS 150, 166 (1935)
95. Regener : Phys. Z. 34, 306 (1933)
96. Euler et Kockel : Naturwiss. 23, 246 (1935)
97. Schrödinger : PRS 150, 465 (1935)
98. Born, Schrödinger : Nature 135, 1342 (1938)
99. Bopp : AP 38, 345 (1940)
100. Bopp : AP 42, 575 (1943)
101. Podolsky : PR 62, 62 (1942)
102. Podolsky et Kikuchi : PR 65, 228 (1944)
103. Pais : "On the theory of elementary particles" Trans. Roy. Acad. Sc. of the Netherlands 19 (1947)
104. Pais : PR 68, 227 (1945)
105. H. Weyl : PR 46, 505 (1934)
106. Correspondance Einstein-Born 1916-55 Seuil (1972) (trad. franç.) p. 139
107. Wheeler Feynman : Rev. Mod. Phys. 17, 157 (1945)

CHAPITRE 8

Abréviations Physical Review : PR
 Zeitschrift für Physik : ZP
 Annalen der Physik : AP
 Proceedings of the Royal Society : PRS

- I. Itzykson et Zuber "Quantum Field Theory" Mc Graw-Hill (1950)
- II. Cohen, Laloë, Diu "Mécanique Quantique" Hermann (1977)
-
1. Uhlenbeck "Oude en Nieuwe Vragen der Natuurkunde" North-Holland pb. (1955)
2. Schweber, Bethe, de Hoffmann "Mesons and Fields" Vol. I (p. 273-274) Row-Peterson and Co (1955)
3. Kramers : Hd. und Jb. d. Ch. Phys. I, 2 "Theorien des Aufbaues der Materie. Quantenth. des Elektrons und der Strahlung" §§ 89, 90 (1937) Traduction anglaise North-Holl. pub. Co (1956)
4. Kramers Nuov. Cim. 15, 108 (1938)
5. Pauli et Fierz Nuov. Cim. 15, 167 (1938)
6. Serpe Physica 7, 133 (1940)
7. Serpe Physica 8, 226 (1941)
8. Weisskopf et Wigner ZP 63, 54 (1930)
9. Weisskopf et Wigner ZP 65, 18 (1930)
10. Opechowski Physica 8, 161 (1941)
11. Kramers Ned. T. Natuurk. 11, 134 (1944) traduit dans Coll. Scient. papers N.H.ed
12. Kramers Rapports du 8ème Conseil Solvay. Stoops p. 241 (1950)
13. Oppenheimer - Ibid. p. 269 à 281
14. Series "Spectrum of atomic hydrogen" Oxford (1957)
15. Lamb Reports on progress in physics 14, 19 (1951)
16. Kellogg et Millman Rev. Mod. Phys. 18, 323 (1946)
17. Gordy Rev. Mod. Phys. 20, 668 (1948)
18. Dirac ZP 44, 585 (1927)
19. Oppenheimer PR 35, 461 (1930)
20. Lamb Conf. Nobel 1955 Elsevier pub. p. 286 (1964)
21. Kusch Conf. Nobel 1955 Elsevier pub. p. 298 (1964)

CHAPITRE 8

22. Michelson et Morley Phil. Mag, 24, 46 (1887)
23. Houston et Hsieh PR 45, 263 (1934)
24. Houston PR 51, 446 (1937)
25. Williams R.C. PR 54, 558 (1938)
26. Drinkwater Richardson et Williams PRS A174, 164 (1940)
27. Giulotto PR 7, 562 (1947)
28. Kemble et Present PR 44, 1031 (1932)
29. Pasternack PR 54, 1113 (1938)
30. Bethe Handbuch der Physik 24, 273 (1933)
31. Fröhlich Heitler et Kahn PRS A171, 269 (1939)
32. Lamb PR 56, 384 (1939)
33. Fröhlich Heitler et Kahn PR 56, 961 (1939)
34. Lamb PR 57, 458 (1940)
35. Kahn Physica 8, 58 (1941)
36. Sommerfeld ZP 118, 236 (1941)
37. Blatt PR 67, 203 (1945)
38. Stern ZP 7, 21, 9 (1921)
39. Stern et Gerlach AP 74, 673 (1924)
40. Breit G. et Rabi PR 38, 2082 (1931)
41. Kellogg, Rabi, Ramsey et Zacharias PR 57, 677 (1940)
42. Kellogg, Rabi, Zacharias PR 50, 472 (1936)
43. Cohen PR 46, 713 (1934)
44. Rabi, Millman, Kusch et Zacharias PR 53, 318 (1938)
45. Rabi, Millman, Kusch et Zacharias PR 55, 526 (1939)
46. Grotrian "Graphische Darstellung der Spektren von Atomen"
Springer Berlin (1928)
47. Betz AP 15, 321 (1932)
48. Haase AP 23, 675 (1935)
49. Cleetons et Williams PR 45, 234 (1934)
50. Herzberg "Atomic Spectra and Atomic Structure" Prentice-Hall N.Y. (1937)
51. Lamb, Retherford PR 72, 241 (1947)
52. Lamb, Retherford PR 75, 1325 et 1332 (1949)
53. Lamb, Retherford PR 79, 549 (1950)
54. Lamb, Retherford PR 81, 222 (1951)
55. Lamb, Retherford PR 86, 1014 (1952)

CHAPITRE 8

56. Breit et Teller *Astrophys. J.* 91, 215 (1940)
57. Webb *PR* 43, 113 (1924)
58. Schwinger Préface de "Selected papers on Q.E.D." Dover (1958)
59. Pauli "Handbuch der Phys." Geiger-Scheel 24-1 2ème éd. p. 133 (1933)
60. Fröhlich, Heitler et Kemmer *PRS* A166, 154 (1938)
61. Kusch et Foley *PR* 74, 250 (1948)
62. Kinsler *PR* 46, 533 (1934)
63. Millman et Kusch *PR* 60, 91 (1941)
64. Fermi *ZP* 60, 320 (1930)
65. Nafe., Nelson, Rabi *PR* 71, 914 (1947)
66. Nagle, Julian, Zacharias *PR* 72, 971 (1947)
67. Nafe-Nelson *PR* 73, 718 (1948)
68. Breit *PR* 72, 984 (1947)
69. Schwinger *PR* 73, 416 (1948)
70. Breit *PR* 73, 1410 (1948)
71. Breit *PR* 74, 656 (1948)
72. A. Bohr *PR* 73, 1109 (1948)
73. Foley, Kusch *PR* 72, 1257 (1947)
74. Becker, Kusch *PR* 73, 584 (1948)
75. Foley, Kusch *PR* 73, 412 (1948)
76. Bethe *PR* 72, 339 (1947)
77. Feynman *PR* 74, 1430 (1948)
78. Epstein *PR* 73, 177 (1948)
79. Schwinger *PR* 75, 898 (1949)
80. Bethe *PR* 73, 1271 (Z8) (1945)
81. Dyson *PR* 73, 1272 (Z9) (1948)
82. Schwinger et Weisskopf *PR* 73, 1272 (Z10) (1948)
83. Weisskopf "My life as a physicist" in *Phys. in the 20th Century*
Weisskopf MIT
84. French et Weisskopf *PR* 75, 1240 (1949)
85. Kroll et Lamb *PR* 75, 388 (1949)
86. Schwinger *PR* 75, 898 (1949)
87. Dancoff *PR* 55, 959 (1939)
88. Lewis *PR* 73, 173 (1947)

CHAPITRE 8

89. Feynman PR 74, 939 (1948)
90. Bopp AP 42, 573 (1942)
91. Weisskopf PR 56, 72 (1939)
92. Schwinger PR 73, 416 (1948)
93. Luttinger PR 74, 893 (1948)
94. Welton PR 74, 1157 (1948)
95. Weisskopf Rev. Mod. Phys. 21, 305 (1949)
96. Koba Prog. Th. Phys. 4, 319 (1949)
97. Koba et Tomonaga Prog. Th. Phys. 2, 218 (1947)
98. Koba, Tati, Tomonaga Prog. Th. Phys. 3, 198 (1947)
99. Taketani Prog. Th. Phys. 2, 187 (1947)
100. Sakata Prog. Th. Phys. 2, 145 (1947)
101. Bopp AP 38, 345 (1940)
102. Hara Prog. Th. Phys. 3, 188 (1948)
103. Hara et Sakata Prog. Th. Phys. 2, 30 (1947)
104. Ito, Koba et Tomonaga Prog. Th. Phys. 2, 216 (1947)
105. Ito, Koba et Tomonaga Prog. Th. Phys. 2, 217 (1947)
106. Ito, Koba et Tomonaga Prog. Th. Phys. 3, 276 (1948)
107. Ito, Koba et Tomonaga Prog. Th. Phys. 3, 325 (1948)
108. Koba et Takeda Prog. Th. Phys. 3, 98 (1948)
109. Koba et Tomonaga Prog. Th. Phys. 3, 290 (1948)
110. Tati et Tomonaga Prog. Th. Phys. 3, 211 (1948)
111. Tati et Tomonaga Prog. Th. Phys. 3, 392 (1948)
112. Hara et Tokano Prog. Th. Phys. 3, 316 (1948)
113. Koba et Takeda Prog. Th. Phys. 3, 98 (1948)
114. Koba et Takeda Prog. Th. Phys. 3, 204 (1948)
115. Koba et Takeda Prog. Th. Phys. 3, 290 (1948)
116. Tomonaga PR 74, 224 (1948)
117. Stokes Math. and Phys. papers Vol. 1 p. 17 (reprint NY and London) (1966)
118. Bechert et Meixner AP 22, 525 (1935)

INDEX ALPHABETIQUE DES NOMS DE PERSONNES

*Seules les citations les plus significatives sont indiquées.
Les pages soulignées marquent le début d'une partie du texte en majeure
partie consacrée à des travaux de la personne citée.*

ABRAHAM	:	129
AHARONOV	:	155
ALBERT	:	155
AMALDI	:	62, 65
ANDERSON	:	15, <u>89</u> , 253, <u>254</u>
AUGER	:	253
BECHERT	:	334
BELINFANTE	:	19, 115, 224
BETHE	:	14, 32, 61, 63, 90, 198, 204, 247, 251, 305, 319, 320, 334, 342 <u>343</u> , <u>353</u> , 358
BETZ	:	338
BHABHA	:	15, 19, 210, 254, 260, <u>284</u>
BLACKETT	:	15, 89, 90, 251,
BLATT	:	337,
BLEULER	:	115,
BLOCH F.	:	31, 57, 108, <u>199</u> , 318, 329
BOHR A.	:	349
BOHR N.	:	22, 49, 81, 91, 150, <u>151</u> , <u>160</u> , <u>167</u> , <u>170</u> , 209, 240, 245, 278, 332
BOOTH	:	286
BOPP	:	<u>271</u> , 309, 362, 371, 372
BORN	:	4, 26, 46, <u>268</u> , 371.
BOSE	:	44
BOTHE	:	242
BRAUNBECK	:	201
BREIT	:	60, 79, 190, 260, 304, 337, 349
BROMBERG	:	85
BROSSEL	:	331, 338
BRÜCKNER	:	262

CARLSON	:	15, 250, 254
CARTAN	:	212
CASSEN	:	260
CASIMIR	:	60, 105
CASSIDY	:	249
CHADWICK	:	15, 240, <u>242</u>
CLAT	:	249
CLETON	:	339
COHEN	:	337
COHEN-TANNOUJJI	:	56
COMPTON	:	74
CONDON	:	260
CRUSSART	:	255
CURIE	:	15, 89, 242
DANCOFF	:	141, 203, 319, 361
DARWIN	:	74, 78, 211
DEBYE	:	, 5, 44, 191
DELBRÜCK	:	191
DE WET	:	218
DIRAC	:	5, <u>6</u> , <u>9</u> , 18, 19, 26, 29, <u>49</u> , 57, 61, 63, 66, 72, <u>76</u> , 79, <u>82</u> , 83, 86, 87, <u>106</u> , 143, <u>182</u> , <u>186</u> , 214, <u>228</u> , <u>229</u> , 274, <u>278</u> , <u>287</u> , <u>289</u> , <u>291</u> , <u>292</u> , 317, 331, 345
DRINKWATER	:	333
DUHEM	:	238
DUNOYER	:	337
DYMOND	:	81
DYSON	:	32, 312, 358
EDDINGTON	:	85, 100
EHRENFEST	:	5, <u>42</u> , 44, 46, 58, 64
EINSTEIN	:	3, <u>43</u> , 46, 64, 91, 112, 178
ELIEZER	:	284, <u>298</u>
ELLIS	:	240
EPSTEIN	:	356
D'ESPAGNAT	:	151
EULER	:	30, 191, 193, 194, 264

FEENBERG	:	260
FEYNMAN	:	32, 226, 230, 280, 361, 362, 365
FERMI	:	6, <u>12</u> , <u>16</u> , 60, 61, <u>95</u> , <u>110</u> , 113, 237, <u>245</u> , 334
FIERZ	:	19, 31, <u>202</u> , 211, <u>219</u> , 220, 318, 323, 326, 377
FOCK	:	18, <u>70</u> , 71, 109, <u>185</u> , 274,
FOKKER	:	280
FOLEY	:	<u>350</u>
FOWLER	:	49
FRENCH	:	355, 365
FRENKEL	:	106, 280
FRÖHLICH	:	210, 260, 335, 346
FURRY	:	21, 29, <u>137</u> , 184
GAMOW	:	160, 247
GERLACH	:	337
GIULOTTO	:	333
GOCKEL	:	249
GÓRA	:	301
GORDON	:	75, 78, 79
GORDY	:	331
GOUDSMIT	:	74, 129
GROTRIAN	:	338
GUPTA	:	115
HALPERN	:	191
HARA	:	373, 376
HEISENBERG	:	4, 6, <u>11</u> , 15, 18, <u>23</u> , 26, 27, 29, 46, 47, 71, 93, <u>98</u> , <u>104</u> , 111, 130, 135, 140, <u>144</u> , 166, <u>172</u> , <u>175</u> , <u>187</u> , 194, 208, 237, <u>243</u> , 250, 252, 261, 284, <u>306</u> , 317
HEITLER	:	14, 15, 28, 56, 63, 90, 145, 198, 210, 250, 251, 254, 260, 286, <u>301</u> , 307, 335, 346
HERMANN	:	41
HERTZBERG	:	339
HESS	:	249
HOLTON	:	178
HOUSTON	:	332, 333
HSIEH	:	332
HSÜEH	:	304

INFELD : 26, 268, 371
IWANENKO : 216, 218, 243, 257
JAMMER : 41
JAUCH : 205, 299, 300
JEANS : 41, 85
JEFFREYS : 49
JOLIOT : 15, 89, 242
JORDAN : 4, 6, 8, 46, 67, 68, 94, 100, 136, 208
JOST : 204
JULIAN : 348
KAHN : 335, 336
KASTLER : 338
KAUFMANN : 129
KELLOG : 331, 337
KEMBLE : 334
KEMMER : 19, 192, 210, 260, 265, 346
KIKUCHI : 272, 307
KINSLER : 347
KEPLER : 371
KLEIN : 6, 8, 10, 14, 67, 75, 79, 80, 98, 100, 105
KNIPP : 198
KOBAYASHI : 233, 356, 368, 373
KOCKEL : 30, 191, 269
KRAMERS : 31, 148, 224, 278, 289, 308, 317, 319, 361, 367
KRISHNAN : 62
KROLL : 360
KUHN : 41
KUSCH : 330, 338, 347, 350
LAMB : 32, 329, 336, 339, 360
LANDAU : 22, 57, 72, 150, 156, 252
LANDSBERG : 62
LAPORTE : 214
LEE : 247
LEPRINCE-RINGUET : 249, 253 255
LEWIS H.W. : 362
LEWIS T. : 288
LEWIS G.N. : 7

LONDON : 112
LORENTZ : 74, 129, 316
LUTTINGER : 365
MA : 115, 286, 304, 308
MANDELSTAM : 62
MAJORANA : 15, 214, 224, 244
MAYER-GÖPPERT : 64
MEHRA : 49
MEITNER : 191, 240, 241
MEIXNER : 334
MIE : 26, 101, 267, 371
MILLMAN : 331, 338, 347
MÖGLICH : 100, 117

MØLLER : 14, 60, 79, 260, 308
MOYER : 85
MONTGORMERY : 254
MORPURGO : 205
MOTT : 14, 60, 53, 79, 81, 170, 198
NAFE : 348
NAGLE : 348
NATANSON : 44
NEDDERMEYER : 253, 254
NELSON : 348
NERNST : 85
VON NEUMANN : 211
NEWTON : 371
NIKOLSKY : 109
NISHIMA : 14, 79, 80
NORDHEIM : 252
NORDSIECK : 31, 199, 318, 323
OCCHIALINI : 15, 89, 251
OPECHOWSKY : 323, 326
OPPENHEIMER : 6, 10, 13, 15, 17, 21, 29, 86, 90, 103, 131, 146, 147, 184, 204,
250, 251, 252, 254, 255, 260, 305, 319, 324, 328, 331, 332, 376

ORTHMANN : 240
OVIDE : 160
PAIS : 27, 272, 372
PAULI : 6, 8, 11, 18, 19, 31, 71, 74, 93, 94, 98, 111, 119, 130, 143, 146,
184, 190, 202, 206, 215, 220, 221, 225, 226, 237, 240, 241, 244,
259, 272, 276, 293, 294, 297, 299, 305, 318, 323, 326, 345, 376, 377
PASTERNAK : 333
PETERLS : 22, 72, 150, 156, 177
PENG : 303
PLANCK : 2, 41
PLESSET : 90, 251
PODOLSKY : 18, 272, 274, 307
POINCARÉ : 129
PRESENT : 334
PROCA : 19, 210, 259
RABI : 337, 338, 339, 348, 345
RACAH : 60
RAMAN : 62
RAMSEY : 337
RAYLEIGH : 41
RETFERD : 329
RICHARDSON : 333
RIVIER : 312
ROHRICH : 205, 283
ROBINSON : 333
ROSE : 190
ROSENFELD : 18, 22, 35, 60, 61, 105, 107, 115, 117, 135, 143, 150, 160, 167,
168, 209, 260
ROSSI : 251, 253
RUPP : 81
RUTHERFORD : 241
SAKATA : 27, 210, 260, 261, 267, 369, 371, 373
SAUTER : 14, 82, 90, 190, 251
SCHRÖDINGER : 75, 80, 100, 269
SCHÜTZ : 59
SCHWEBER : 351
SCHWINGER : 32, 339, 358, 365, 375, 376, 377
SERBER : 29, 141, 190, 197, 255

SERPE : 324, 325
SHANKLAND : 278
SLATER : 278
SMEKAL : 62
SOKOLOV : 216, 218, 301
SOLOMON : 117, 135
SOMMERFELD : 63, 198, 337
STERN : 337, 346
STREATER : 226
STREET : 17, 255
STEVENSON : 17, 255
STOKES : 128, 316
STUECKELBERG : 17, 19, 27, 210
SWIFT : 99
TAKETANI : 210, 260, 261, 267, 369, 370
TAMM : 86, 257
TANIKAWA : 233
TATI : 233, 374
TELLER : 247, 342
TER HAAR : 308
TETRODE : 100
THOMSON JJ : 128, 316
TOMONAGA : 18, 32, 109, 203, 231, 356, 372, 373, 374, 376, 377
TOKANO : 376
TYCHO BRAHE : 371
UEHLING : 29, 189, 334
UHLENBECK : 74, 129, 198, 214, 319
UNSOLD : 59
VILLARS : 148, 272, 376
VOLTERRA : 95, 101
VAN DER WAERDEN : 212, 213
WAGNER : 216
WALLER : 13, 21, 62, 84, 133, 143, 302
WATANABE : 233, 369
WEBB : 343
WEINMANN : 201
WELTON : 366

WEINBERG : 37
WEISSKOPF : 14, 18, 21, 30, 58, 92, 106, 137, 192, 195, 215, 259, 276, 302,
317, 331, 346, 359, 365, 368
VAN WEIZSÄCKER : 252
WENTZEL : 141, 261, 263, 274, 287
WEYL : 22, 83, 87, 111, 212, 223
WENFLER : 190, 280, 306
WIGHTMAN : 226
WIGNER : 8, 14, 58, 69, 100, 206, 211, 214, 302, 304, 331
WILLIAMS E.J. : 252, 254
WILLIAMS R.C. : 332, 334
WILLIAMS W.E. : 333, 334
WILSON A.H. : 286, 301
WQOSTER : 240
WU : 247
YANG : 247, 257
YUKAWA : 16, 210, 231, 237, 255, 260, 261, 369
ZACHARIAS : 337, 338, 348

TABLE DES MATIERES

•	REMERCIEMENTS		
•	NOTATIONS, CONVENTIONS, FORMULES	Page	
<u>CHAPITRE</u>	0	L'EVOLUTION DES CONCEPTS ET DES METHODES DE LA THEORIE QUANTIQUE DES CHAMPS DE 1927 à 1948.	1
	•	Introduction	2
	0.1	Dans l'euphorie de la nouvelle mécanique quantique	4
	0.2	Malgré les infinis	14
	0.3	Attitudes face aux infinis	20
	0.4	A la recherche de théories alternatives	25
	0.5	L'issue	29
	0.6	Conclusion	34
	•	Bibliographie générale	38
<u>CHAPITRE</u>	1	LA QUANTIFICATION DU CHAMP ELECTROMAGNETIQUE.	40
	1.1	Les quanta de rayonnement au début de la théorie quantique	41
	1.2	Les fluctuations du corps noir à l'aube de la mécanique des matrices	46
	1.3	"La théorie quantique de l'émission et de l'absorption du rayonnement de P.A.M. Dirac	49
	1.4	Premières applications de la théorie de Dirac	56
<u>CHAPITRE</u>	2	VERS UNE THEORIE QUANTIQUE GENERALE DES CHAMPS D'ONDES.	65
	2.1	La seconde quantification	66
	2.2	La théorie relativiste de l'électron	74
	2.3	Généralisations de la théorie du rayonnement de Dirac	93
	2.4	La querelle des représentations	104
	2.5	L'invariance de jauge et la quantification de l'électrodynamique	110
	2.6	L'énergie de zéro des oscillateurs	117

<u>CHAPITRE</u>	3	LA CRISE DES DIVERGENCES ULTRAVIOLETTES.	125
	3.1	L'auto-interaction en électrodynamique classique	128
	3.2	Mise en évidence de la divergence de l'énergie propre	130
	3.3	Les divergences dans la théorie des trous de Dirac	136
	3.4	Commentaires sur la nature des infinis et sur l'avenir de la théorie	143
<u>CHAPITRE</u>	4	LE PROBLEME DE LA MESURABILITE DES GRANDEURS PHYSIQUES DANS LE DOMAINE QUANTIQUE RELATIVISTE.	149
	4.1	Introduction de la problématique	151
	4.2	"Généralisations des relations d'incertitude à la théorie quantique relativiste" (Landau et Peierls)	156
	4.3	"Du problème de la mesurabilité des champs électromagnétiques" (Bohr et Rosenfeld)	160
	4.4	Mesurabilité des charges et des courants	166
	4.5	L'électron de Dirac et la physique du noyau face à la critique de Bohr	170
	4.6	La longueur universelle de Heisenberg	172
	4.7	La matrice S	175
<u>CHAPITRE</u>	5	PROGRES FORMELS DE LA THEORIE QUANTIQUE DES CHAMPS PENDANT SA CRISE	181
	5.1	La polarisation du vide	181
	5.2	La catastrophe infrarouge	198
	5.3	La relation entre spin et statistique	206
	5.4	L'invariance relativiste manifeste	228
<u>CHAPITRE</u>	6	THEORIE QUANTIQUE DES CHAMPS ET HAUTES ENERGIES	236
	6.1	Des spectres β à la théorie de Fermi	240
	6.2	Les rayons cosmiques	248
	6.3	La théorie mésique	257
	•	Bibliographie	263

<u>CHAPITRE</u> 7	ALTERNATIVES THEORIQUES	264
7.1	Théories monistes et théories mixtes	267
7.2	A partir de théories classiques de l'électron ponctuel	274
7.3	La métrique indéfinie	291
7.4	L'amortissement par rayonnement de Heitler	301
7.5	Les théories de matrice S	306
	• Bibliographie	314
<u>CHAPITRE</u> 8	NAISSANCE DE L'IDEE DE RENORMALISATION	315
8.1	Vers un principe de renormalisation	316
8.2	Les idées de Kramers	319
8.3	Le déplacement de Lamb-Retherford	329
8.4	Le moment magnétique anormal de l'électron	345
8.5	La conférence de l'île de Shelter	351
8.6	Les premiers calculs finis de corrections radiatives	352
•	REFERENCES	
•	INDEX DES NOMS	