UNIVERSITE PARIS-SUD

LICENCE ET MAGISTERE DE PHYSIQUE

LES IDEES & LEUR HISTOIRE

EN PHYSIQUE

HISTOIRE, ACTUALITE & HORIZONS

DU MODELE STANDARD

GILLES COHEN-TANNOUDJI

http://perso.club-internet.fr/gicotan

TABLE DES MATIERES

1. Introduction

Vue panoramique du Modèle standard

Les objectifs du cours

2.Les deux premiers modèle standards, de Galilée et Newton à Maxwell et Boltzmann

3.1894-1930: la révolution scientifique du 20ème siècle

4.Élaboration et mise en œuvre du Modèle standard

5.Les nouveaux horizons

6.Conclusion: bilan épistémologique et méthodologique

Introduction

Vue panoramique du Modèle Standard (MS)

Le modèle standard de la physique des particules, dont l'histoire est indiquée dans le tableau ci-joint, est la théorie de référence qui décrit les constituants élémentaires de la matière (les fermions) et les interactions fondamentales (gravitationnelle, électromagnétique, faible et forte) auxquelles ils participent.

Les constituants élémentaires de la matière (voir le tableau) se partagent en deux grandes catégories déterminées par la participation aux interactions fondamentales, les leptons d'une part, qui ne participent pas à l'interaction forte et les quarks d'autre part, qui participent à toutes les interactions. Dans la catégorie des leptons, les leptons chargés participent à l'interaction électromagnétique et à l'interaction faible et les leptons neutres ou neutrinos ne participent qu'à l'interaction faible. La participation des constituants élémentaires aux interactions fondamentales est conditionnée par leur nombres quantiques conservés, ou charges d'interaction. À chaque constituant de la matière est associée son antiparticule, une particule de même masse et de charges opposées.

Le tableau des constituants élémentaires de la matière montre une autre classification indépendante de la participation aux interactions, en trois répliques, ou générations, de particules qui ne diffèrent que par leurs masses. L'électron, le quark u et le quark d, qui appartiennent à la première génération sont les particules massives les plus légères et elles sont stables ; elles sont les constituants exclusifs de la matière ordinaire. À l'exception des neutrinos, les particules des deux autres générations sont plus lourdes et instables ; elles ne peuvent être observées ou détectées que dans les expériences auprès des accélérateurs ou dans le rayonnement cosmique. Cette structuration en générations des constituants de la matière n'est pas expliquée par le modèle standard ; elle reste l'un des problèmes les plus intriguants de la physique contemporaine.

D'après la mécanique quantique, pour qu'il y ait une interaction il faut qu'au moins une particule élémentaire, un boson, soit émise, absorbée ou échangée. Le photon est le boson de l'interaction électromagnétique, les bosons intermédiaires W^+, W^- et Z sont les bosons de l'interaction faible et les gluons sont les bosons de l'interaction forte au niveau des quarks, la chromodynamique quantique ou QCD (pour Quantum Chromodynamics).

Le cadre théorique du modèle standard est la théorie quantique des champs qui permet de décrire quantitativement les interactions fondamentales des particules élémentaires en respectant les principes de la relativité restreinte et ceux de la mécanique quantique. D'après la mécanique quantique, pour observer une structure microscopique à haute résolution temporelle et spatiale, il est nécessaire de lui transférer une énergie-impulsion, d'autant plus élevée que la résolution souhaitée est élevée. Mais d'après la théorie de la relativité, ce transfert d'énergie-impulsion peut se transformer en l'apparition de particules qui n'étaient pas présentes dans l'état initial : les fermions peuvent être produits ou annihilés par paires particule/antiparticule, les bosons peuvent l'être en nombre arbitraire. Tous les processus relevant d'une même interaction fondamentale sont reliés les uns aux autres. La démarche de la théorie quantique des champs, dans laquelle les propriétés de symétries jouent un rôle fondamental, est une démarche d'unification des processus au sein de chaque interaction fondamentale et des interactions fondamentales au sein de grandes synthèses théoriques, comme la théorie électrofaible de Glashow, Salam et Weinberg ou la grande unification de toutes les interactions fondamentales.

Le modèle standard consiste en un ensemble d'algorithmes, appelé développement perturbatif, permettant de calculer, par approximations successives, à l'aide d'un nombre fini et fixé de paramètres déterminés expérimentalement, les probabilités des réactions des leptons et des quarks, du photon et des bosons intermédiaires en interactions électromagnétique et faible (théorie électrofaible), et les probabilités des réactions des quarks et des gluons en interaction forte à grand transfert d'énergie-impulsion (QCD).

Le modèle standard a passé avec succès tous les tests expérimentaux auxquels il a été soumis. Il a permis d'anticiper de très nombreuses découvertes expérimentales décisives :

- réactions d'interaction faible à courants neutres en 1973 ;
- quark charmé en 1975 ;
- gluon en 1979 ;
- bosons intermédiaires de l'interaction faible W^+, W^- et Z en 1983 ;
- quark top en 1995.

Toutes les prédictions du modèle standard ont été confirmées par l'expérience, à l'exception d'une seule, l'existence du boson de Higgs, une particule que l'on espère bien découvrir au LEP ou au LHC.

Les lignes de force du modèle standard sont :

- un nouveau niveau d'élémentarité, celui des quarks, les constituants élémentaires qui participent à toutes les interactions fondamentales ;
- une explication simple de la violation de l'invariance par parité dans l'interaction faible ;
- un principe unificateur, l'invariance de jauge, selon lequel toutes les interactions fondamentales découlent de l'invariance dans des opérations de symétrie qui dépendent du point d'espace-temps où elles ont appliquées ;
- un critère de fiabilité, la renormalisabilité qui permet d'éliminer les infinis qui pourraient compromettre le développement perturbatif, et de discerner, à chaque échelle d'énergie, les degrés de liberté et les paramètres pertinents ;
- un mécanisme efficace de brisure de symétrie, le mécanisme de Higgs, à l'origine de la différentiation des masses des particules élémentaires, qui permet l'unification, sans les confondre, des interactions électromagnétique et faible au sein de la théorie électrofaible ;
- une perspective cosmogonique, sur laquelle il débouche en se rapprochant du modèle du big bang, le modèle standard de la cosmologie.

Le modèle standard comporte cependant des points faibles qui nécessitent des recherches théoriques et expérimentales, et qui font penser qu'il doit être considéré comme une théorie effective, susceptible d'être améliorée ou approfondie :

- certains des paramètres dont il dépend, comme les masses des constituants élémentaires, sont en très grand nombre, leurs valeurs sont très éparpillées et toutes n'ont même pas encore été déterminées ;
- il laisse complètement ouvert le problème de la masse des neutrinos ;
- il ne fournit pas d'explication à la classification des constituants de la matière en trois générations de particules, alors que c'est précisément l'existence de ces trois générations qui permet de rendre compte de la violation de l'invariance CP qui est vraisemblablement à l'origine du déséquilibre matière/antimatière dans l'univers primordial ;
- il est en échec face au traitement quantique de la gravitation ;
- il ne fournit pas, dans le cadre du développement perturbatif de la chromodynamique quantique, d'explication à la propriété fondamentale du confinement qui interdit aux quarks de se propager à l'état libre hors des hadrons, les particules dont ils sont les constituants ;
- le mécanisme de Higgs, qui en est la clé de voûte, n'a pas encore été testé expérimentalement et il apparaît comme trop « ad hoc » pour pouvoir être généralisé à

l'unification de toutes les interactions fondamentales, car il impliquerait des ajustements fins de paramètres qui le rendraient peu crédible.

C'est justement à propos de ce mécanisme de Higgs et de la particule qui lui est associée, le boson de Higgs, que se concentrent l'essentiel des efforts théoriques et expérimentaux.

Les recherches théoriques en vue du dépassement du modèle standard empruntent principalement deux grandes voies :

- la voie de la supersymétrie qui, en corrigeant certaines des insuffisances du mécanisme de Higgs permettrait, tout en conservant tous les acquis du modèle standard, de relancer les tentatives de grande unification, et fournirait peut-être des clés pour la quantification de la gravitation ; dans le modèle standard supersymétrique minimal (MSSM), il y aurait plusieurs bosons de Higgs ; la supersymétrie associerait à chaque particule (boson ou fermion) du modèle standard un partenaire de l'autre statistique (fermion ou boson) ; le partenaire supersymétrique de plus petite masse serait une particule massive n'interagissant que très faiblement, qui pourrait permettre de résoudre le problème de la masse cachée de l'univers.
- la voie de la sous-structure selon laquelle il existerait un nouveau niveau d'élémentarité sous-jacent aux particules du modèle standard (ou à certaines d'entre elles), et qui déboucherait sur toute une floraison de nouvelles particules composites, analogues aux hadrons, mais à des masses deux à trois mille fois plus élevées.

Mais seule l'expérience permettra de trancher entre ces deux options théoriques, ou de mettre en évidence une nouvelle voie. Telle est la perspective dans laquelle se prépare le programme du LHC. Grâce au modèle standard, grâce aux résultats expérimentaux que vont continuer à fournir les installations existantes comme le LEP au CERN, HERA à Hambourg et le TEVATRON aux Etats-Unis, ce programme du LHC semble doté d'un très haut potentiel de découverte.

L'HISTOIRE DU MODELE STANDARD

Dates	Cadre théorique	Gravitation	Électro-magnétisme	Interaction faible	Interaction forte
17ème siècle	Galilée, Newton	**Newton**			
19ème siècle	Euler, Lagrange, Jacobi, Hamilton		**Maxwell**		
1896-1898				Découverte de la radioactivité	
1900-1930	Mécanique quantique				
1905	Relativité restreinte				
1915	Relativité générale	**Einstein**			
1930-1950	Théorie quantique des champs		**QED**	Fermi	Yukawa
1960-1997	Le Modèle Standard	**Big bang**	**théorie électrofaible de Glashow, Salam et Weinberg**		**QCD**
	Les problèmes ouverts	Le boson de Higgs, le confinement, les masses des fermions, les générations, la masse cachée de l'univers, la brisure de la symétrie matière/antimatière, la quantification de la gravitation.			
1997- ?	Demain, les superthéories ? Une sous-structure ?....	La grande unification de toutes les interactions fondamentales, la supersymétrie, la **supergravité**, la **supercorde**, un nouveau niveau d'élémentarité...			

Les objectifs du cours

- Mettre en perspective historique et prospective l'ensemble des cours de tronc commun
- Sensibiliser des étudiants de second cycle à l'importance de l'interprétation
- Les sensibiliser aux enjeux fondamentaux de la physique contemporaine
- Leur faire apprécier l'importance de la rétroaction des avancées de la physique contemporaine sur l'interprétation et la ré-interprétation du formalisme

Les deux premiers modèles standards, de Galilée et Newton à Maxwell et Boltzmann

Le MS de Newton vu par Einstein

« La première tentative en vue de poser un fondement théorique unifié est représenté par l'œuvre de Newton. Dans son système tout est ramené au concepts suivants : 1/ des points matériels dont la masse est invariable ; 2/ action à distance entre deux points matériels ; 3/ lois du mouvement pour le point matériel. (...) De plus, Newton se rendait parfaitement compte que le temps et l'espace étaient des éléments essentiels en tant que facteurs physiques en fait de son système, quoi que seulement d'une manière implicite. »

« Règles nécessaires pour philosopher » de Newton

Règle I

« Les causes de ce qui est naturel ne doivent pas être admises en nombre supérieur à celui des causes vraies ou de celles qui suffisent à expliquer les phénomènes de ce qui est naturel. »

Règle II

« Ainsi, il faut assigner les mêmes causes aux effets naturels de même genre autant que faire se peut. Telles sont la respiration dans l'homme et la bête ; la chute des pierres en Europe et en Amérique ; la lumière d'un feu de cuisine et celle du soleil ; la réflexion de la lumière sur la terre et sur les planètes. »

Règle III

"Les qualités des corps qui ne peuvent être ni augmentées ni diminuées, et qui appartiennent à tous les corps sur lesquels on peut faire des expériences, doivent être considérées comme les qualités de tous les corps en général. »

Règle IV

« En philosophie expérimentale les propositions que l'on réunit par induction à partir des phénomènes doive être devenue poncerait, puisque que des hypothèses contraires de leur sont pas obstacles, soit précisément soit approximativement jusqu'à ce que se présente d'autres phénomènes qui soient les rangs de plus précise soient les affranchissent exception. »

Les principales étapes du développement de la mécanique

- Newton et les mathématiques
    - Le calcul différentiel et intégral,en même temps que et indépendamment de Leibniz (1646-1716)
    - Formule du binôme, développement en série, calcul des variations
- Les quatre lois
- Euler (1707-1783) fondateur de l'analyse et fondateur de la mécanique du solide
- Euler et le principe de moindre action: formulation rigoureuse du principe heuristique de Maupertuis (1698-1759)
- Euler et la mécanique des fluides. Travaux complétés par Navier (1785-1836)
- Principe de d'Alembert (1717-1783): toute dynamique peut être rapportée aux lois de Newton.
- Les nouvelles mécaniques: Lagrange (1736-1813), Laplace (1749-1827) et Hamilton (1805-1865).
- La mécanique analytique de Lagrange :

« On ne trouvera pas de figures dans cet ouvrage. Les méthodes que j'y expose ne demandent ni constructions, ni raisonnements géométrique ou mécanique, mais seulement des opérations algébriques assujetties à une marche régulière et uniforme. »

La physique de Laplace. Le « démon de Laplace ».

- Déterminisme et probabilités.
- Fondation de la Société d'Arcueil dont le but est de « ramener la vision newtonienne du monde aux dimensions de la physique moléculaire en appliquant, par les mathématiques, les lois de Newton aux phénomènes tels que la lumière, la chaleur, le son, l'électricité ou le magnétisme. »
- La mécanique de Hamilton : point de départ des mécaniques post-classiques.

Électricité, magnétisme, électromagnétisme

- Coulomb (1736-1806) et sa loi en $1/r^2$. Loi très difficile à vérifier expérimentalement mais admise à cause de son aspect newtonien.
- L'électricité « animale » de Galvani (1737-1798). Controverse avec Volta (1745-1827) qui pense qu'il n'y a qu'une seule sorte d'électricité.

- La pile de Volta (1799) et la naissance de l'électrocinétique. Recherche de nouveaux éléments en chimie grâce à l'électrolyse.
- Découverte en 1819, par Oersted (1777-1851) de l'effet magnétique produit par un courant électrique.
- Ampère (1775-1836) et la formalisation de l'électrodynamique
    - Introduction de la notion de courant électrique
    - Action mutuelle de deux courants électriques
- Faraday (1791-1867) expérimentateur autodidacte
    - Rédaction des notes de cours de Davy
    - Travaux de chimie (découverte de deux chlorures de carbone)
    - Électromagnétisme: production d'électricité avec du magnétisme
    - Travaux sur la liquéfaction des gaz, le benzène et les verres lourds
    - Découverte de l'induction 29 août 1831: courant « induit » qui n'apparaît que lors des ouvertures ou fermetures du circuit. Neumann (1798-1895) établit une théorie mathématique de l'induction
- Introduction de la notion de champ
    - « La théorie électromagnétique de la lumière, telle qu'il la propose, est la même en substance que celle que j'ai commencé à développer dans cet article de 1846, sauf qu'en 1846 il n'y avait pas de données permettant de calculer la vitesse de propagation » Commentaire de Maxwell à propos des travaux de Faraday.
    - « Pour nous, qui avons pour ainsi dire sucé les idées de Faraday avec le lait de notre mère, il est difficile d'apprécier leur grandeur et leur audace. » Einstein op. cit. p. 82
- Faraday et les lois de l'électrolyse
    - « Si nous empruntons la terminologie de la théorie atomique, des corps qui sont équivalents entre eux ont associé à eux des quantités égales d'électricité. »
- Diélectriques et cage de Faraday
- La décharge dans les gaz
- Effet Faraday (effet d'un champ magnétique sur la polarisation de la lumière)
- Paramagnétisme et diamagnétisme

Maxwell, la synthèse

- La loi d'Ohm (1789-1854). Théorie des « courants galvaniques » analogue à la théorie analytique de la chaleur de Fourier
- Gauss (1777-1855), « prince des mathématiciens »
    - Premières réflexions sur des géométries non-euclidiennes (à 15 ans!)
    - Travaux sur l'arithmétique et les statistiques
    - Directeur de l'observatoire de Göttingen
    - A partir de 1829 travaux sur l'électricité et le magnétisme ainsi que sur l'optique.

L'œuvre de Maxwell (1831-1879)

- Inspirations :
    - Travaux expérimentaux de Faraday
    - Formalisme mathématique de William Thomson (futur Lord Kelvin), lui même imprégné des travaux de Fourier
- 1855-1868: cinq articles fondant l'électromagnétisme:
    - On Faraday's lines of force
    - On physical lines of force
    - Introduction de la constante c (proche de la vitesse de la lumière)
    - A dynamical theory of the electrodynamic field
    - On a method of making a direct comparison of electrostatic with elctromagnetic force; with a noteon the electromagnetci theory of light
- « Du point de vue de l'histoire de l'humanité, il est très probable que la découverte, par maxwell, des lois de l'électrodynamique sera considérée comme l'événement le plus marquant de tout le 19ème siècle. Même la guerre civile américaine passera pour une histoire d'intérêt local à côté de cet événement autrement décisif des années 1860. » (Feynman)
- Hertz (1857-1894), et les transmissions hertziennes
    - 1884: validation expérimentale de la théorie de Maxwell: émission et détection d'ondes électromagnétiques se propageant à la vitesse de la lumière.
    - Naissance des transmissions hertziennes : Lodge (1851-1940), Branly (1844-1940), Marconi (1874-1937)

Fresnel et le développement de l'optique

- De l'antiquité à la première mesure de la vitesse de la lumière
    - Euclide (330-270 av. JC): notion de rayon lumineux, principes de propagation rectiligne et de retour inverse
    - Naissance de la lunette
    - Giambattista della Porta (1538-1615)
    - Lunette astronomique de Galilée
    - Loi de la réfraction: Snell (1580-1626) et Descartes (1596-1650)
    - La double réfraction
    - Römer (1644-1710): « le mouvement de la lumière n'est pas instantané, ce qui se fait voir par l'inégalité des immersions et émersions du premier satellite de Jupiter. »
- Huygens vs Newton
    - L'optique de Huygens qui lui permet d'expliquer la double réfraction: la lumière est une onde qui se propage sphériquement dans un milieu de référence, l'éther, en obéissant à deux principes:
        - Le principe de superposition ou de composition
        - Le principe de Huygens qui stipule que le front d'onde se construit comme l'enveloppe des ondelettes émises par tous les points d'un front antérieur.
    - L'optique de Newton:
        - modèle corpusculaire de la lumière
        - Travaux sur les couleurs
        - Le télescope à miroir
- Young (1773-1829) et les interférences
- Malus (1775-1812) et la polarisation
- Fresnel (1788-1826) et Arago (1786-1853): le triomphe de la théorie ondulatoire de la lumière
    - 1818-1819: travaux conjoints d'Arago et Fresnel sur les interférences en lumière polarisée
    - Fresnel et la diffraction
    - Fresnel et la polarisation
- Fizeau, Foucault et la vitesse de la lumière

- Fizeau: « dispositif qui fournit une valeur de la vitesse de la lumière peu différente de celle qui est admise par les astronomes. »
- Foucault et l'expérience cruciale de l'optique: la vitesse de la lumière est moindre dans les liquides que dans l'air.

Du point matériel à l'atome

- Carnot et le développement de la thermodynamique
  - La mesure de la température
    - Galilée 1603
    - Newton 1701
    - Celsius (1701-1744)
  - La nature a-t-elle horreur du vide?
    - Expérience de Toricelli 1608-1647)
    - Expérience de Pascal (1623-1662) (au sommet du Puy-de-Dôme)
    - Définition de la pression (Pascal, unité de pression)
  - Denis Papin (1647-1714) et la machine à vapeur
- Carnot et les deux principes de la thermodynamique
  - « Les physiciens sont partagés sur la nature de la chaleur. Plusieurs d'entre eux la regardent comme un fluide répandu dans toute la nature et dont les corps sont plus ou moins pénétrés, à raison de leur température et de leur disposition à le retenir. (…) D'autres physiciens pensent que la chaleur n'est que le résultat des mouvements insensibles des molécules de la matière. (…) C'est ce mouvement intestin qui, suivant les physiciens dont n,ous parlons, constitue la chaleur. » (Laplace 1780)
  - Fourier (1768-1830) et sa théorie analytique de la chaleur. Invention de la transformation de Fourier
- Carnot (1796-1832) et ses « réflexions sur la puissance motrice du feu »:
  - « Il ne suffit pas, pour donner naissance à la puissance motrice de produire de la chaleur: il faut encore se procurer du froid; sans lui, la chaleur serait inutile. »
  - « Lorsqu'une hypothèse ne suffit plus à l'explication des phénomènes, elle doit être abandonnée. C'est le cas où se trouve l'hypothèse par laquelle on considère le calorique comme une matière, comme un fluide subtil. (…) la

chaleur n'est autre que la puissance motrice ou plutôt le mouvement qui a changé de forme, c'est un mouvement. (…) la production d'une unité de puissance motrice nécessite la destruction 2,70 unités de chaleur. »

- Joule, l'équivalent mécanique de la chaleur et la conservation de l'énergie.
- Clausius (1822-1888) et la formalisation de la thermodynamique:
  - Introduction du concept d'entropie: « je proposerai donc d'appeler la quantité S l'entropie du corps d'après le mot grec τροπη, transformation. »
  - Premier principe: conservation de l'énergie
  - Second principe: non décroissance de l'entropie

- Boltzmann (1844-1906) et la naissance de la thermodynamique statistique
  - Premiers travaux sur la théorie cinétique des gaz
    - Bernouilli: $pV=1/3Nm\langle v^2\rangle$
    - Loi d'Avogadro Ampère: $pV=RT$
    - Joule: énergie cinétique moyenne proportionnelle à la température
  - 1872: Boltzmann relie la thermodynamique à la mécanique par l'intermédiaire du « théorème H »: $S=k \log W$, où k est la constante baptisée par Planck, constante de Boltzmann et W le nombre d'états microscopiques permettant de réaliser un état macroscopique donnée.
  - La chimie et la théorie atomique

La formulation lagrangienne de la mécanique classique

Le principe de moindre action

- Degrés de liberté, trajectoire dans l'espace de configuration

$$\mathbf{Q}(t)=\left\{\mathbf{q}(t);\dot{\mathbf{q}}(t)\right\}\equiv\left\{q_1(t),q_2(t),...q_N(t);\dot{q}_1(t),\dot{q}_2(t),...\dot{q}_N(t)\right\}$$

$$\mathbf{Q}(t)=\left\{\mathbf{q}(t);\dot{\mathbf{q}}(t)\right\}\equiv\left\{q_1(t),q_2(t),...q_N(t);\dot{q}_1(t),\dot{q}_2(t),...\dot{q}_N(t)\right\}$$

$$\dot{q}_i(t)\equiv\frac{dq_i(t)}{dt}$$

- Intégrale d'action, fonctionnelle de la trajectoire Q(t)

$$I = \int_{t_i}^{t_f} dt\ \mathbf{L}\left\{\mathbf{q}(t);\dot{\mathbf{q}}(t)\right\}$$

- Lagrangien **L** = différence de l'énergie cinétique et de l'énergie potentielle **L=T-V**
- Variation infinitésimale

$$\delta q(t), \delta \dot{q}(t) = \frac{d}{dt}\delta q(t)$$

- en fixant les deux extrémités de la trajectoire

$$\delta q(t_i) = \delta q(t_f) = 0$$

- Variation de l'action

$$\delta I = \int_{t_i}^{t_f}\left[\frac{\partial \mathbf{L}}{\partial \mathbf{q}(t)}\delta\mathbf{q}(t) + \frac{\partial \mathbf{L}}{\partial \dot{\mathbf{q}}(t)}\delta\dot{\mathbf{q}}(t)\right]dt$$

où

$$\frac{\partial \mathbf{L}}{\partial \mathbf{q}(t)}\delta\mathbf{q}(t) = \sum_i \frac{\partial \mathbf{L}}{\partial q_i(t)}\delta q_i(t)$$

En intégrant par parties on obtient

$$\delta I = \left[\frac{\partial \mathbf{L}}{\partial \dot{\mathbf{q}}(t)}\delta\mathbf{q}(t)\right]_{t_i}^{t_f} + \int_{t_i}^{t_f}\left[\frac{\partial \mathbf{L}}{\partial \mathbf{q}(t)} - \frac{d}{dt}\frac{\partial \mathbf{L}}{\partial \dot{\mathbf{q}}(t)}\right]\delta\mathbf{q}(t)$$

- comme le terme tout intégré est nul, à cause des conditions aux extrémités, on obtient les équations d'Euler-Lagrange, les équations fondamentales de la mécanique classique

$$\frac{\partial \mathbf{L}}{\partial q_i(t)} = \frac{d}{dt}\frac{\partial \mathbf{L}}{\partial \dot{q}_i(t)}; i = 1,2,...N$$

Les équations de Hamilton

- Le hamiltonien, **H**=**T**+**V** et les moments conjugués

$$p_i(t) = \frac{\partial \mathbf{L}}{\partial \dot{q}_i(t)}$$

- Espace de phases, degrés de liberté et moments conjugués, qui ne sont rien d'autre que les impulsions

$$\mathbf{T} = \frac{1}{2}\sum_i p_i \dot{q}_i; \mathbf{H} = 2\mathbf{T} - \mathbf{L} = \sum p_i \dot{q}_i - \mathbf{L}(\mathbf{q}, \dot{\mathbf{q}})$$

- Les équations de Hamilton

$$\begin{cases} \dot{q}_i = \dfrac{\partial \mathbf{H}}{\partial p_i} \\ \dot{p}_i = -\dfrac{\partial \mathbf{H}}{\partial q_i} \end{cases}$$

(Remarque: la deuxième équation n'est autre que l'équation de Newton $f=m\gamma$)

- Les **crochets de Poisson**. Soit *f(p,q,t)* une fonction quelconque des coordonnées, des impulsions et du temps. On a

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_k \left(\frac{\partial f}{\partial q_k} \dot{q}_k + \frac{\partial f}{\partial p_k} \dot{p}_k \right) = \frac{\partial f}{\partial t} + \{H, f\}$$

$$\text{où } \{H, f\} = \sum_k \left(\frac{\partial H}{\partial p_k} \frac{\partial f}{\partial q_k} - \frac{\partial H}{\partial q_k} \frac{\partial f}{\partial p_h} \right)$$

- f est une **intégrale première** si

$$\frac{\partial f}{\partial t} + \{H, f\} = 0$$

si elle ne dépend pas explicitement du temps, alors

$$\{H, f\} = 0$$

Symétries et lois de conservation: le théorème de Nœther.

Le théorème de Nœther articule

- une propriété de relativité
- une invariance par une transformation de symétrie
- une loi de conservation

Noter la complémentarité (qui sera précisée en physique quantique) entre temps et énergie, position et impulsion, orientation angulaire et moment cinétique

Inobservable	**Symétrie**	**Loi de conservation**
Origine du temps	Translation dans le temps	Energie
Origine de l'espace	Translation dans l'espace	Impulsion
Direction privilégiée	Rotation	Moment cinétique

1894-1930

LA REVOLUTION SCIENTIFIQUE DU XX° SIECLE

Vie et œuvre de Max Planck

Jeunesse et vocation scientifique

- Max Planck naît le 23 avril 1858 à Kiel
- Père professeur de jurisprudence, meurt en 1900
- Mère, issue de famille de pasteurs, meurt en 1914
- 1867, lycée à Munich
- «j'aurais pu devenir philologue ou musicien »
- 1874, université de Munich
- 1878, université de Berlin
- Direction scientifique de Helmholtz et Kirchoff
- 1879, thèse sur le deuxième principe de la thermodynamique
- 1880, thèse d'habilitation
- 1885, professeur à Kiel
- 1887, mariage avec Marie Merck
- 1888, premier enfant
- 1889, professeur à Berlin, en remplacement de Kirchoff
- 1892, attribution d'une chaire

La physique en 1892

« La physique est définitivement constituée dans ses concepts fondamentaux ; tout ce qu'elle peut désormais apporter, c'est la détermination précise de quelques décimales supplémentaires. Il y a bien deux petits problèmes : celui du résultat négatif de l'expérience de Michelson et celui du corps noir, mais ils seront rapidement résolus et n'altèrent en rien notre confiance… »

(Lord Kelvin)

Oui mais…

« On entend souvent dire que les physiciens dans la dernière partie du dix-neuvième siècle estimaient connaître toutes les lois de physique et que la seule chose importante qui leur

restait à faire était de calculer quelques décimales de plus. Quelqu'un a pu dire cela une fois, et d'autres l'ont copié. Mais une lecture attentive de la littérature de cette époque montre que quelque chose les préoccupait tous. »
R.P. Feynman

Quel était le problème ?

Maxwell à propos de la chaleur spécifique des gaz :
« Finalement, en établissant une relation nécessaire entre les mouvements de translation et de rotation de toutes les particules non sphériques, nous avons montré qu'un système de telles particules ne peut pas satisfaire les relations connues entre les deux chaleurs spécifiques. »
« Je viens maintenant de vous présenter ce que je considère comme étant la difficulté la plus grande, rencontrée jusqu'à présent par la théorie moléculaire. »

Planck et le corps noir

- 19 octobre 1900, présentation et confirmation expérimentale
- 14 décembre 1900, la loi du corps noir et les quanta d'énergie

Le temps des épreuves

- 17 octobre 1909, mort de Marie Planck
- 14 mars 1911, remariage avec Marga von Hoesslin
- 14 mars 1912, naissance de Hermann
- 1913, l'atome de Bohr
- 2 août 1914, déclaration de guerre
- 26 mai 1916, mort de Karl, fils de Max Planck devant Verdun
- 15 mai 1917, mort de Grete, l'une des filles jumelles de Planck
- 21 novembre 1919, mort de la sœur jumelle de Grete
- Mais … Novembre 1919, prix Nobel de physique au titre de l'année 1918

1929, le jubilé

- Le 28 juin 1929 est fêté le Goldene Docktorjubiläum de Max Planck à l'occasion du cinquantenaire de son doctorat est aussi crée la Médaille Max Planck
- Ce jour là sont attribuées les deux premières médailles Planck à Albert Einstein et Max Planck

Les années sombres

- 1930, Planck nommé président du Kaiser Wilhem Institute
- 1933, Hitler au pouvoir
- 1935, lois antisémites de Nuremberg
- 1936, attaques des antisémites Stark et Lenard
- 1938, Planck démissionne de ses fonctions à l'académie de Prusse sous la pression nazie
- 1939, début de la seconde guerre mondiale

Encore un drame

- 20 juillet 1944 : attentat manqué contre Hitler
- 23 juillet 1944 : arrestation de son fils Erwin Planck (au moment de la prise de pouvoir par Hitler en 1933, Erwin Planck avait démissionné de son poste de secrétaire d'État à la chancellerie du Reich)
- 23 janvier 1945 Erwin Planck est exécuté

La fin de la vie de Planck

« Planck était maintenant dans sa quatre-vingt-huitième année. Son corps était recroquevillé par l'arthrose ; du marcheur et grimpeur sportif qu'il fut, il était devenu un vieillard courbé, ne pouvant plus se déplacer qu'avec peine, en s'appuyant sur sa canne. Lui qui avait été le maître d'une maison cultivée habitait maintenant dans un meublé. Il avait perdu tous ses biens. Les quatre enfants de son premier mariage étaient tous morts » (Armin Herman)

Max Planck s'éteint à 89 ans le 4 octobre 1947

« C'est Planck, et nul autre, qui a osé faire le premier pas en introduisant une nouvelle constante universelle. Le courage génial qu'il a manifesté sera encore pour les siècles à venir un exemple exaltant aux yeux des hommes de sciences »

Max Von Laue

Le rayonnement du corps noir

« Si donc, ainsi qu'il est facile, on regarde au travers de l'ouverture, on ne pourra distinguer dans l'enceinte aucun détail de forme et on aura l'impression singulière d'un gouffre de lumière.

… Cette ouverture doit être appelée parfaitement noire si nous pensons que le caractère essentiel d'un corps noir est de ne rien renvoyer de la lumière qu'il reçoit. »

Jean Perrin, Les atomes, Champs Flammarion, Paris 1991, pp. 212 213

Loi de Stéfan-Boltzmann

$$U = uV$$

U = Energie interne du corps noir

U = Densité du rayonnement (intégrée sur toutes les fréquences).

$$dQ = dU + dE = dU + pdV$$

$$p = u\int_0^{\pi/2} \cos^2\theta \sin\theta d\theta = \frac{u}{3}$$

$$dQ = d(uV) + 1/3udV = 4/3udV + Vdu$$

U ne dépend que de T (universalité), donc

$$dQ = \frac{4}{3}udV + V\frac{du}{dT}dT$$

$$dS = \frac{dQ}{T} = \frac{4}{3}\frac{u}{T}dV + \frac{V}{T}\frac{du}{dT}dT$$

dS= différentielle exacte, donc

$$\frac{d}{dT}\left(\frac{4}{3}\frac{u}{T}\right) = \frac{1}{T}\frac{du}{dT}$$

$$\frac{du}{dT} = 4\frac{u}{T} \quad ; \quad u = cT^4$$

Modèle des résonateurs de Planck

N « résonateurs » :

$$U_N = NU$$

$$S_N = NS$$

Hypothèse fondamentale (permettant de calculer l'entropie comme le logarithme du nombre de complexions)

$$U_N = P\varepsilon$$

où ε est une « énergie élémentaire », et P un entier (*on trouve cependant dans l'article de Planck la note en bas de page suivante, citée par O. Darrigol : « Lorsque le rapport U/e n'est pas un nombre entier, on peut prendre pour P le nombre entier le plus voisin »*)

Calcul du nombre de complexions W

Exemple :

1 2 3 4 5 6 7 8 9 10 : N= 10
7 38 11 0 9 2 20 4 4 5 : P=100

$$W = \frac{(N+P-1)!}{(N-1)!P!} \underset{M!\approx M^M}{\approx} \frac{(N+P)^{N+P}}{N^N P^P}$$

$$S_N = k\,\mathrm{Log}\,W = k\left\{(N+P)\mathrm{Log}(N+P) - N\,\mathrm{Log}\,N - P\,\mathrm{Log}\,P\right\} =$$

$$= kN\left\{\left(1+\frac{U}{\varepsilon}\right)\mathrm{Log}\left(1+\frac{U}{\varepsilon}\right) - \frac{U}{\varepsilon}\mathrm{Log}\frac{U}{\varepsilon}\right\}$$

$$S = k\left\{\left(1+\frac{U}{\varepsilon}\right)\mathrm{Log}\left(1+\frac{U}{\varepsilon}\right) - \frac{U}{\varepsilon}\mathrm{Log}\frac{U}{\varepsilon}\right\}$$

<u>*S* est donc une fonction de</u> $\frac{U}{\varepsilon}$

Loi de Wien

Densité spatiale d'énergie

$$E_\lambda d\lambda = T^5 \psi(\lambda T) d\lambda ; \int E_\lambda d\lambda = cT^4$$

En fréquences

$$E_\lambda d\lambda \to u_\nu d\nu$$

$$\lambda \to c/\nu$$

$$d\lambda \to \frac{c d\nu}{\nu^2}$$

$$u_\nu = T^5 \frac{c}{\nu^2}\psi\left(\frac{cT}{\nu}\right) = \frac{\nu^3}{c^3} f\left(\frac{T}{\nu}\right)$$

Travaux de Planck de 1897 à 1900

- Série de cinq articles consacrés aux « processus de rayonnement irréversible »
- Loi de Wien-Planck

$$u_\nu \approx \frac{8\pi b \nu^3}{c^3}\exp\left(-\frac{a\nu}{T}\right)$$

- Les échelles de Planck (1899)

$$b = h\,;\, a = \frac{h}{k}\,;\, f = \text{constante de la gravitation}$$

$$l_p = \left(\frac{bf}{c^3}\right)^{\frac{1}{2}}\,;\, M_p = \left(\frac{bc}{f}\right)^{\frac{1}{2}}\,;\, t_p = \left(\frac{bf}{c^5}\right)^{\frac{1}{2}}\,;\, T_p = a\left(\frac{c^5}{bf}\right)^{\frac{1}{2}}$$

La stratégie de Planck

« Comme la signification du concept de l'entropie n'était point parvenue à être pleinement dégagée, personne n'accorda la moindre attention à la méthode que j'avais adoptée et je pus développer mes calculs complètement à ma guise, avec une rigueur absolue, sans avoir à redouter de troubles ou de compétition. »

La voie de l'entropie

- Spectre d'énergie d'un résonateur en équilibre avec le champ (formule non démontrée dans l'article de décembre 1900) (Voir Feynman Mécanique 2) :

$$u_\nu = \frac{8\pi\nu^2}{c^3}U$$

- Lorsque $U = kT$, on trouve la loi de Rayleigh
- Et, à partir de la loi de Wien on a

$$U = \nu f\left(\frac{T}{\nu}\right)$$

$$U = \nu f\left(\frac{T}{\nu}\right) \rightarrow T = \nu f\left(\frac{U}{\nu}\right)$$

$$\frac{1}{T} = \frac{dS}{dU} = \frac{1}{\nu} f\left(\frac{U}{\nu}\right) \rightarrow S = f\left(\frac{U}{\nu}\right)$$

- Mais, comme d'après le calcul de l'entropie à partir du modèle des résonateurs, on avait obtenu que l'entropie est une fonction de U/ε, on trouve que ε et ν sont nécessairement proportionnels :

$$\varepsilon = h\nu$$

La formule de Planck

$$S = k\left\{\left(1+\frac{U}{h\nu}\right)\mathrm{Log}\left(1+\frac{U}{h\nu}\right) - \frac{U}{h\nu}\mathrm{Log}\left(\frac{U}{h\nu}\right)\right\}$$

$$\frac{1}{T} = \frac{dS}{dU} = \frac{k}{h\nu}\mathrm{Log}\left(1+\frac{h\nu}{U}\right)$$

$$U = \frac{h\nu}{\exp\left(\frac{h\nu}{kT}\right)-1}$$

$$u_\nu = \frac{8\pi h\nu^3}{c^3}\frac{1}{\exp\left(\frac{h\nu}{kT}\right)-1}$$

Quelques points historiques

- Loi de Planck découverte le dimanche 7 octobre 1900. Formule envoyée à Rubens sur carte postale
- 19 octobre 1900 présentation de la formule devant la société de physique de Berlin et demande de vérification expérimentale
- 14 décembre présentation de la dérivation théorique de la formule
- Article soumis le 14 décembre:Annalen der Physik 4, 553-563 (1901) Origine de h: « hilfe grösse »

L'apport d'Einstein

- Hypothèse fondamentale: la quantification de l'énergie est une propriété universelle du champ électromagnétique lui-même

$$S = k \operatorname{Log} W$$

- W interprété comme une probabilité temporelle

$$W = \left(\frac{V}{V_0} \right)^{\frac{E}{h\nu}}$$

Elucidation de l'effet photoélectrique

« Dès lors qu'un rayonnement monochromatique (de densité suffisamment faible) se comporte, relativement à la dépendance en volume de son entropie, comme un milieu discontinu constitué de quanta d'énergie hn, on est conduit à se demander si les lois de la production et de la transformation de la lumière n'ont pas également la même structure que si la lumière était constituée de quanta d'énergie de ce type. Telle est la question dont nous allons maintenant nous occuper. »

La chaleur spécifique des gaz

- L'échec de la théorie moléculaire pour rendre compte de la chaleur spécifique des gaz était le problème soulevé par Maxwell
- Einstein part de la formule de Planck et il écrit que l'énergie de vibration des ions d'un solide est, comme celle des résonateurs de Planck un multiple de $h\nu$ et il obtient l'expression de la chaleur spécifique

$$C = 5.94 \sum \exp\left(\frac{h\nu_i}{kT} \right) \left[\exp\left(\frac{h\nu_i}{kT} \right) - 1 \right]^{-2}$$

1. *Einstein Ann. Der Physik, 22, 180, 1907)*

Émission induite

Einstein propose en 1916 une nouvelle interprétation de la formule de Planck

$$u_\nu = \frac{A_{mn}}{B_{nm} \exp\left(\frac{h\nu}{kT}\right) - B_{mn}}$$

où

- A_{mn} représente l'émission spontanée
- B_{nm} représente l'absorption
- A_{nm} représente l'émission induite

La compatibilité avec la formule de Planck impose que

$$B_{mn} = B_{nm} \; ; \frac{A_{mn}}{B_{mn}} = \frac{8\pi h\nu^3}{c^3}$$

Qu'est-ce que la cohérence quantique ?

La cohérence quantique est une propriété du formalisme de la théorie quantique qui permet de réconcilier les descriptions ondulatoire et corpusculaire irréductiblement contradictoires en physique classique. La méthodologie qui repose sur l'articulation de cette cohérence et d'une autre notion, l'indiscernabilité, s'est avérée d'une efficacité impressionnante dans le développement des implications théoriques et pratiques de la physique quantique, de l'effet laser au condensat de Bose-Einstein en passant par la superfluidité.

À la fin du XIXème siècle, la controverse entre les interprétations de la lumière, et plus généralement de l'interaction électromagnétique, en termes d'ondes ou en termes de corpuscules, avait été tranchée en faveur des premières. C'est le succès de l'interprétation ondulatoire pour décrire les phénomènes de diffraction et d'interférence qui avait emporté la décision. Or les travaux de Planck en 1900, puis ceux d'Einstein en 1905, avaient mis en évidence des aspects nettement corpusculaires dans le domaine de l'interaction électromagnétique (rayonnement du corps noir et effet photoélectrique). La physique quantique qui s'est développée depuis lors n'a cessé de confirmer la coexistence des aspects ondulatoires et des aspects corpusculaires dans tous les phénomènes de la microphysique, qu'ils concernent les interactions ou la structure de la matière. Au cœur du formalisme de la physique quantique, la propriété de cohérence est adaptée à la prise en compte de cette coexistence. C'est cette propriété qui est à l'origine des paradoxes les plus déroutants de l'univers quantique, mais c'est aussi elle qui est à l'origine de la découverte de nouveaux états

de la matière comme le condensat de Bose-Einstein et de nouveaux effets comme la superfluidité ou l'effet laser débouchant sur de multiples applications.

L'expérience d'Young

Pour comprendre comment fonctionne la cohérence quantique, il convient de revisiter la fameuse expérience des interférences d'Young, celle qui avait fait pencher la balance en faveur de l'interprétation ondulatoire de la lumière, mais en la réalisant dans des conditions où peuvent se manifester des aspects corpusculaires. Dans l'expérience d'Young, la lumière produite par une source ponctuelle passe au travers d'un cache percé de deux trous et est recueillie sur un écran situé à quelque distance du cache. C'est le formalisme des amplitudes de champ qui permet de rendre compte mathématiquement des figures d'interférence observées dans cette expérience. L'intensité de l'éclairement en un point de l'écran est proportionnelle au flux de l'énergie électromagnétique qui parvient en ce point, elle-même proportionnelle à la somme des carrés du champ électrique et du champ magnétique. Cette intensité est nécessairement représentée par un nombre réel, non négatif. Si, pour obtenir l'intensité totale en un point de l'écran il fallait additionner les deux intensités correspondant, chacune au passage par un des deux trous, aucune figure d'interférence ne serait jamais obtenue, puisque ces deux intensités sont des nombres réels positifs ou nuls. Le nom d'amplitude de champ est donné à un nombre complexe, caractérisé par un module et une phase, ou une partie réelle et une partie imaginaire, dont le module au carré est proportionnel à l'intensité. Dans l'expérience d'Young, des interférences sont obtenues si, pour calculer l'intensité de l'éclairement en un point de l'écran, les deux amplitudes de champ correspondant, chacune au passage par l'un des deux trous, sont additionnées comme des nombres complexes pour obtenir l'amplitude totale dont le carré du module est proportionnel à l'intensité d'éclairement au point considéré. Cette propriété d'additivité complexe des amplitudes de champ ne fait que traduire la propriété, fondamentale en théorie électromagnétique de la lumière, de l'additivité vectorielle du champ électrique et du champ magnétique.

Les amplitudes de probabilité

Des aspects corpusculaires peuvent commencer à se manifester dans l'expérience d'Young si elle est tentée à la limite des très faibles intensités. À cette limite, le champ électromagnétique devient fluctuant ; ses fluctuations sont des processus élémentaires dans lesquels un photon émis par la source, parvient au détecteur en passant par l'un des deux trous. En remplaçant

l'écran par un détecteur très sensible, les impacts de photons peuvent être enregistrés un à un. Au bout d'un certain temps, ces impacts reproduisent la figure d'interférence : accumulation d'un grand nombre d'impacts dans les zones éclairées et petit nombre d'impacts dans les zones sombres. Comme fluctuations quantiques du champ électromagnétique, les processus élémentaires ne sont pas descriptibles à l'aide de la mécanique corpusculaire classique. La mécanique classique repose en effet sur la continuité de l'action (produit de l'énergie par le temps), alors que la mécanique quantique résulte précisément de la découverte d'une limite à la divisibilité de l'action. Le passage à la théorie quantique suppose le renoncement à une description déterministe des processus élémentaires mettant en jeu une action égale au quantum d'action. Ces processus ne sont descriptibles que de manière statistique. Dans le cas qui nous intéresse, il est clair que l'intensité recueillie en un point du détecteur est proportionnelle au nombre d'impacts de photons et donc à la probabilité d'impact au point considéré. Pour rendre compte de l'apparition d'une figure d'interférence, il sera nécessaire, en complète analogie avec la description ondulatoire, d'introduire le concept *d'amplitude de probabilité* : pour calculer la probabilité de l'impact d'un photon en un point du détecteur ; il faudra additionner comme des nombres complexes les deux amplitudes de probabilité correspondant au passage par chacun des deux trous et prendre le module au carré de l'amplitude résultante.

L'expérience d'Young avec des électrons ou des atomes

D'autres généralisations de l'expérience d'Young permettent d'avancer dans l'exploration des bizarreries de l'univers quantique. Il est par exemple possible de remplacer les photons par des électrons. Le même phénomène est observé : les impacts des électrons construisent, au bout d'un certain temps, le même type de figure d'interférence. Plus récemment l'expérience d'Young a pu être réalisée avec des atomes à la place des photons. Et là aussi peut être observée la fascinante formation de la figure d'interférence. Le formalisme des amplitudes de probabilité permet de rendre compte de ces observations. Dès les années vingt, de Broglie avait proposé d'étendre aux constituants de la matière la dualité onde/corpuscule, et donc d'associer des ondes aux électrons, aux noyaux ou aux atomes. On sait que l'une des principales limitations traduites par le quantum d'action concerne l'impossibilité d'attribuer à un corpuscule microscopique une trajectoire bien déterminée : il est impossible de mesurer avec des précisions arbitrairement élevées et simultanément la position et la vitesse d'une particule comme un électron. Toute l'information concernant par exemple l'état d'un électron

est contenue dans la fonction d'onde de cet état, qui n'est rien d'autre qu'une amplitude de probabilité : le carré du module de la fonction d'onde est proportionnel à la probabilité que l'électron se trouve dans l'état considéré.

Cohérence et indiscernabilité

Le phénomène des interférences dans le cadre de l'interprétation corpusculaire est extrêmement paradoxal : alors qu'il est facile de comprendre comment une onde peut se scinder en deux parties qui passent chacune par un trou et se recombinent en parvenant sur l'écran, il est difficile d'imaginer comment une particule pourrait se scinder en deux parties ou bien passer par les deux trous à la fois ! La nouveauté radicale qu'introduit la mécanique quantique est que pour des conditions expérimentales données, il peut y avoir des questions sans réponses. Par exemple, dans l'expérience d'Young avec des électrons, et si l'on se contente d'enregistrer les impacts sur le détecteur, la question de savoir par quel trou est passé l'électron qui a donné un impact est sans réponse. Dans ces conditions expérimentales, les voies de passage par chacun des deux trous sont dites *indiscernables*. Il convient immédiatement de noter que cette notion n'est pas absolue, définitive : l'indiscernabilité est relative à certaines conditions d'observation ; elle est donc provisoire ; il est possible de la lever ; mais alors il faut faire une expérience différente. Dans l'expérience d'Young avec des électrons, il est possible de lever l'indiscernabilité et de monter un nouveau dispositif expérimental permettant de surveiller (par exemple avec un faisceau de photons en « lumière rasante ») le passage par l'un des deux trous. Dans cette nouvelle expérience, ô surprise, les figures d'interférences sont absentes !

La méthodologie quantique

Il devient maintenant possible d'expliciter les étapes de la méthodologie quantique qui repose sur l'articulation de l'indiscernabilité et de la cohérence. Comme toute mesure aux échelles microscopiques est une interaction soumise à la contrainte de mise en jeu d'au moins un quantum d'action, il devient impossible, même en principe, de faire totalement abstraction des conditions de l'observation.

Comme les processus élémentaires mettant en jeu une action égale au quantum d'action ne sont pas descriptibles au moyen des équations de la physique classique, et comme les conditions de l'observation ne peuvent en général pas être mieux déterminées que de manière statistique, il faudra renoncer à une prédictibilité déterministe au profit d'une prédictibilité probabiliste.

En physique quantique, l'évaluation de la probabilité d'un certain événement ou processus fait intervenir des amplitudes de probabilité qui prennent en compte l'indiscernabilité relative aux conditions de l'observation.

Pour évaluer la probabilité d'un processus, il convient de déterminer toutes les voies indiscernables par lesquelles il peut se produire dans les conditions expérimentales considérées, d'associer à chacune de ces voies une certaine amplitude de probabilité, de sommer comme des nombres complexes (somme cohérente) toutes ces amplitudes pour obtenir l'amplitude totale dont le carré du module donne la probabilité recherchée.

Mode d'emploi des amplitudes de probabilité

Du point de vue mathématique, la cohérence quantique s'exprime dans le cadre de l'espace de Hilbert, une espace vectoriel de fonctions complexes dans lequel est défini un produit scalaire qui est un nombre complexe. Les états quantiques sont représentés par des vecteurs de l'espace de Hilbert, et les observables par des opérateurs linéaires agissant sur ces vecteurs. La cohérence quantique traduit le *principe de superposition* : une combinaison linéaire de vecteurs est un vecteur ; une combinaison linéaire d'opérateurs est un opérateur.

Le mode d'emploi des amplitudes de probabilité tient en quatre règles :

1. S'il existe plusieurs voies de transitions indiscernables, l'amplitude totale de transition est la somme des amplitudes des différentes voies : $\langle f|i\rangle = \sum_k \langle f|i\rangle_k$ où $\langle f|i\rangle$ désigne l'amplitude totale de transition entre l'état initial $|i\rangle$ et l'état final $\langle f|$, et $\langle f|i\rangle_k$ désigne l'amplitude associée à la voie k
2. S'il existe plusieurs états finals discernables, c'est la probabilité totale qui est la somme des probabilités de transition vers les états finals individuels : $\left|\langle f|i\rangle\right|^2 = \sum_k \left|\langle f_k|i\rangle\right|^2$
3. Si la transition s'effectue via un état intermédiaire v, l'amplitude se factorise : $\langle f|i\rangle = \langle f|v\rangle\langle v|i\rangle$
4. Lorsque deux particules effectuent chacune et indépendamment une transition, l'amplitude se factorise : $\langle f_1 f_2|i_1 i_2\rangle = \langle f_1|i_1\rangle\langle f_2|i_2\rangle$.

Bosons et fermions

L'une des implications les plus importantes du quantum d'action concerne l'indiscernabilité des particules identiques. En physique classique, des particules de même

espèce restent discernables, « étiquetables », puisqu'il est toujours possible (au moins en principe) de suivre la trajectoire d'une particule. Ce n'est plus vrai en physique quantique car tout étiquetage coûte au moins un quantum d'action. Formellement, cette nouvelle propriété d'indiscernabilité s'exprime au moyen d'une propriété de symétrie : la symétrie ou l'invariance par permutation de particules identiques. Soit un système de deux particules identiques, appelées 1 et 2, observées, l'une au point x et l'autre au point y. L'invariance par permutation signifie simplement que la probabilité d'observer 1 en x et 2 en y est égale à la probabilité d'observer 2 en x et 1 en y :

$$\Pr\{1 \quad en \quad x; 2 \quad en \quad y\} = \Pr\{2 \quad en \quad x; 1 \quad en \quad y\}$$

Comme la probabilité est proportionnelle au carré du module de la fonction d'onde, l'invariance par permutation implique l'égalité *à une phase près* des deux fonctions d'ondes :

$$\Psi(1 \quad en \quad x; 2 \quad en \quad y) = \exp(i\phi)\Psi(1 \quad en \quad y; 2 \quad en \quad x)$$

mais comme une deuxième permutation des deux particules ramène à l'état initial, la phase est nécessairement de carré égal à +1 ; ϕ ne peut valoir que 0 (modulo 2π) ou π (modulo 2π) ; la phase ne peut valoir que +1 ou –1.

- Lorsque la phase vaut +1, c'est-à-dire si

$$\Psi(1 \quad en \quad x; 2 \quad en \quad y) = +\Psi(1 \quad en \quad y; 2 \quad en \quad x)$$

on dira que les particules identiques obéissent à la statistique de Bose-Einstein et ces particules seront appelées des bosons ;

- lorsque la phase vaut –1, c'est-à-dire si :

$$\Psi(1 \quad en \quad x; 2 \quad en \quad y) = -\Psi(1 \quad en \quad y; 2 \quad en \quad x)$$

on dira que les particules identiques obéissent à la statistique de Fermi Dirac et ces particules seront appelées des fermions.

Le théorème de la connexion spin/statistique qu'il est possible d'établir à partir d'hypothèses extrêmement générales, assigne des spins nuls ou entiers aux bosons et des spins demi-entiers aux fermions.

Le terme de statistique est utilisé parce que c'est en physique quantique statistique que le comportement des fonctions d'onde sous la permutation de particules identiques a le plus de conséquences. C'est plus de quinze ans après le travail d'Einstein que le physicien indien Bose s'est aperçu que les photons n'obéissent pas à la statistique classique, et qu'il faut les traiter comme des particules indiscernables, et en quelque sorte "superposables", ce que permet la statistique qui porte son nom associé à celui d'Einstein. Le fait que les particules

associées aux interactions par la théorie quantique soient des bosons permet de rendre compte de l'additivité vectorielle des champs de force, qui était valide en physique classique.

Les fermions obéissent au principe d'exclusion de Pauli : si, en effet, on essaie de placer deux fermions au même point, leur fonction d'onde sera égale à son propre opposé : elle ne pourra que s'annuler. Cette propriété convient tout à fait aux particules constituantes de la matière, dont elle rend compte du caractère impénétrable, intuitif en physique classique.

Mais s'il est vrai que les statistiques de particules quantiques indiscernables rendent compte de caractéristiques intuitives en physique classique, elles ont aussi des implications tout à fait inattendues qu'il faut mettre à l'actif de la physique quantique. C'est la cas de l'effet laser ou du condensat de Bose-Einstein.

La relativité restreinte

Les racines de la théorie de la relativité

Le bilan de l'œuvre de Newton

(D'après Einstein, "Considerations concerning the fundamentals of theoretical physics", Science N° 91, 24 mai 1940, traduit dans "Einstein, conceptions scientifiques, Champs Flammarion pp 77-96)

- Les fondements théoriques unifiés de l'œuvre de Newton se ramènent à trois principes:
    - Des points matériels dont la masse est invariable
    - Action à distance entre points matériels
    - Loi du mouvement pour le point matériel
- La seule loi d'action à distance explicitée est la loi de la gravitation. Il n'y a pas de lois a priori de l'action entre points matériels, à l'exception de l'égalité de l'action et de la réaction.
- Le temps et l'espace, considérés comme absolus, sont implicitement des éléments essentiels de la théorie de Newton
- Les succès de la théorie de Newton et de ses prolongements :
    - Mécanique des masses discrètes et des milieux continus
    - Explication du principe de conservation de l'énergie
    - Théorie complète de la chaleur, si l'on assimile les atomes ou molécules à des points matériels
- Les points faibles:

- L'électrodynamique
- La théorie corpusculaire de la lumière s'expliquant par la réticence à envisager, pour une théorie ondulatoire, un milieu remplissant tout l'espace "l'éther", constitué de points matériels, et propageant les ondes lumineuses sans manifester la moindre propriété mécanique
- L'action à distance: échec de toutes les tentatives d'interpréter la gravitation par les collisions d'hypothétiques particules matérielles.

La révolution de Faraday-Maxwell-Hertz-Lorentz

- Triomphe de la conception ondulatoire de la lumière
  - Vitesse déterminée de propagation
  - Phénomènes d'interférences
  - Diffraction
  - Polarisation
- récusation par Faraday de l'explication de l'électromagnétisme par l'action à distance de masses électriques:"comment dans la limaille de fer répandue sur une feuille de paier, chaque grain pourrait-il connaître les particules électriques circulant dans un conducteur placé à côté? - Einstein-"
- Le concept de champ introduit par Faraday: les forces électrique et magnétique sont décrites au moyens du champ électrique E et du champ magnétique H, qui se manifestent quand on y introduit des corps chargés. Le concept de champ commence à acquérir le statut de réalité physique intrinsèque. On continue néanmoins à invoquer un milieu porteur des champs, l'éther.
- Les **équations de Maxwell** relient les dérivées spatiales et temporelles du champ électrique et du champ magnétique :

$$\begin{aligned}
\operatorname{rot}\mathbf{E} &= -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} \\
\operatorname{div}\mathbf{H} &= 0 \\
\operatorname{rot}\mathbf{H} &= \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \frac{1}{c}\mathbf{j} \\
\operatorname{div}\mathbf{E} &= \rho
\end{aligned}$$

- Dans le vide ($\mathbf{j} = \rho = 0$), les équations de Maxwell admettent des solutions en forme d'ondes, mises en évidence expérimentalement et interprétées théoriquement par Hertz.

- Les ondes de lumière :
    - dans le vide, ($\mathbf{j} = \rho = 0$), les équations de Maxwell s'écrivent

$$\operatorname{rot}\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t},\ \operatorname{div}\mathbf{H} = 0$$
$$\operatorname{rot}\mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t},\ \operatorname{div}\mathbf{E} = 0$$

Pour résoudre ces équations il est commode d'introduire le potentiel scalaire ϕ et le potentiel vecteur **A**. Ces potentiels ne sont pas fixés de manière univoque, on peut leur imposer des conditions. On peut fixer ϕ à 0 et ajouter au potentiel vecteur le gradient d'une fonction quelconque; on peut donc imposer div**A** = 0. On a alors

$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t},\ \mathbf{H} = \operatorname{rot}\mathbf{A}$$
$$\operatorname{rot}\ \operatorname{rot}\mathbf{A} = -\Delta\mathbf{A} + \operatorname{grad}\ \operatorname{div}\mathbf{A} = -\frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2}$$
$$\operatorname{div}\mathbf{A} = 0 \Rightarrow \Delta\mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2}$$

Qui est l'équation de d'Alembert des ondes se propageant à la vitesse *c*

- Cas particulier où le champ ne dépend que d'une seule coordonnée, soit *x*. Les équations du mouvement prennent alors la forme suivante

$$\frac{\partial^2 f}{\partial t^2} - c^2\frac{\partial^2 f}{\partial x^2} = 0$$

où *f* est n'importe quelle composante de E ou H. Pour résoudre cette équation on la réécrit

$$\left(\frac{\partial}{\partial t} - c\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x}\right)f = 0$$
$$\xi = t - \frac{x}{c},\ \eta = t + \frac{x}{c},$$
$$t = \frac{1}{2}(\eta + \xi),\ x = \frac{c}{2}(\eta - \xi)$$

L'équation pour f prend alors la forme

$$\frac{\partial^2 f}{\partial\xi\partial\eta} = 0 \Rightarrow f = f_1\left(t - \frac{x}{c}\right) + f_2\left(t + \frac{x}{c}\right)$$

qui décrit des ondes de vitesse *c* le long de l'axe *x*

- Les avancées dues à **Lorentz**
    - Un pas supplémentaire est franchi dans la voie de l'émancipation du concept de champ: la théorie de Maxwell était en difficulté à l'intérieur de la matière, où il

fallait introduire deux vecteurs électriques et deux vecteurs magnétiques. Pour Lorentz, « partout (y compris à l'intérieur des corps pondérables) les siège du champ est l'espace vide; La participation de la matière aux phénomènes électromagnétiques repose uniquement sur le fait que les particules élémentaires de la matière portent des charges inaliénables et sont, pour cette raison sujettes d'une part aux actions des forces pondéromotrices et possèdent d'autre part la propriété d'engendrer un champ. Les particules élémentaires obéissent à la loi de Newton du point matériel » - Einstein –

- Les équations de Maxwell sont invariantes sous la transformation de Lorentz

$$x = \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}},\ y = y',\ z = z',\ t = \frac{t' + \frac{v}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Poincaré et Einstein en 1905: la structure fine de la théorie de la relativité

- La situation en 1905
  - On a découvert l'invariance des équations de Maxwell par les transformations de Lorentz, dont Poincaré a montré qu'elles forment un groupe.
  - On a accepté le résultat négatif des expériences de Michelson et de Michelson et Morley
- La relativité de Poincaré, qui considère comme acquise l'invariance des équations de Maxwell par les transformations de Lorentz
  - Principe de relativité: équivalence des systèmes d'inertie
  - Hypothèse de la dualité temps vrai/temps local pour accommoder l'existe de t et t' dans la transformation de Lorentz
  - Hypothèse de l'électron déformable, pour accommoder la contraction des longueurs dans la transformation de Lorentz
- Cette théorie, presque simultanée avec celle d'Einstein en est indépendante, et de nature très différente: elle conserve l'idée de l'éther
- Les prémisses de la relativité chez Einstein
  - Jusque en 1905 les travaux d'Einstein concernent davantage la thermodynamique que l'électrodynamique
  - 1905, "l'annus mirabilis d'Einstein" avec les trois articles, à quelques semaines d'intervalles qui devaient révolutionner toute la physique

        - 18 mars 1905: "Sur un point de vue heuristique concernant la production et la conversion de la lumière" (quanta de lumière, effet photoélectrique)
        - 11 mai 1905: "Mouvement des particules en suspension dans un fluide au repos impliqué par la théorie moléculaire cinétique de la chaleur" (théorie du mouvement brownien conduisant à la mise en évidence des atomes)
        - "Sur l'électrodynamique des corps en mouvement" (Théorie de la relativité restreinte).
    - Tous ces travaux sont sous tendus par
        - Le rejet de l'idée de l'éther
        - L'idée de la "matérialité" du champ, réalité intrinsèque, portant de l'énergie et de la quantité de mouvement
        - La notion d'événement
        - L'ambition de refonder toute la mécanique, en remettant en cause sa partie la plus fondamentale, la cinématique.

La refondation relativiste de la mécanique

- Un nouveau statut de la constante *c*
    - L'absence d'action instantanée à distance implique l'existence d'une borne supérieure à toute vitesse de propagation, de quoi que ce soit, dans tout l'univers.
    - Si une certaine interaction se propage à une vitesse égale à cette borne supérieure, sa propagation doit violer la loi de composition des vitesses.
    - C'est ce qui se produit avec l'interaction électromagnétique qui se propage à la vitesse *c* dans tous les référentiels d'inertie
    - La vitesse de la lumière dans le vide *c* est donc interprétée comme la constante universelle traduisant l'absence d'interaction instantanée à distance, et représentant la borne supérieure de toute vitesse de propagation
- La remise en cause de la cinématique classique
    - Alors que Poincaré élabore un modèle dynamique pour expliquer l'invariance des équations de Maxwell par les transformations de Lorentz, Einstein donne à cette invariance une signification beaucoup plus vaste: c'est la partie la plus

fondamentale de la mécanique, la cinématique, c'est-à-dire "la doctrine de l'espace et du temps" qu'il remet en cause.

- Il interprète la transformation de Lorentz qui modifie le temps et la métrique spatiale, comme un changement de référentiel d'inertie dans un continuum à quatre dimensions **l'espace-temps** (un concept introduit, après l'article d'Einstein, par Minkowski)
- La constante *c* s'interprète alors comme la constante de structure de l'espace-temps, une constante qui, partout dans l'univers, et pour tout phénomène relie l'espace au temps.
- Les remises en causes impliquées par cette théorie sont de grande ampleur: le temps, la métrique spatiale, la simultanéité ne sont plus absolus. Dans une transformation de Lorentz, le temps et la métrique spatiale varient, et ils varient de telle sorte que la métrique d'espace-temps est invariante.

- Intervalles, genres, temps propre
  - La transformation de Lorentz

$$x = \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}},\ y = y',\ z = z',\ t = \frac{t' + \frac{v}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}$$

appliquée à deux **événements** (points d'espace-temps) 1 et 2, laisse invariant **l'intervalle d'espace-temps**

$$s_{12}^2 = c^2 t_{12}^2 - l_{12}^2,\ t_2 - t_1 = t_{12},\ (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 = l_{12}^2$$

$$c^2 t_{12}^2 - l_{12}^2 = c^2 t'^2_{12} - l'^2_{12}$$

- Trois cas sont possibles:
  - $s^2_{12} > 0$; l'intervalle est dit du **genre temps** : il existe un référentiel d'espace-temps où les deux événements ont eu lieu au même point; la succession temporelle est absolue
  - $s^2_{12} < 0$; l'intervalle est dit du **genre espace** : il existe un référentiel d'espace-temps où les deux événements sont simultanés; la séparation spatiale est absolue
  - $s^2_{12} = 0$; l'intervalle est dit du **genre lumière** : dans tout référentiel, les deux événements sont reliés à la vitesse *c*
- Le temps propre (Landau & Lifchitz, Théorie des champs, pp. 18 - 19)

- En plus du principe de relativité et du principe d'invariance de la vitesse *c*, la théorie de la relativité d'Einstein suppose un troisième principe celui de l'identité des unités de mesure (tiges rigides identiques et rythmes identiques). Ce principe ne se justifie pleinement que dans un cadre quantique (c'est pourquoi la théorie d'Einstein est "pré-quantique"). C'est ce principe qui permet d'attacher rigidement une horloge à un système de coordonnées. (Une telle horloge pourra être par exemple un atome de Césium).
- Supposons que nous observions, dans un système d'inertie une telle horloge en mouvement, que nous pouvons considérer comme uniforme sur un intervalle infinitésimal de temps. Pendant *dt*, l'horloge parcourt $(dx^2+dy^2+dz^2)^{1/2}$. Le temps *dt'* indiqué par cette horloge, au repos dans son référentiel (*dx'*=*dy'*=d*z'*=0), sera égal, en vertu de l'invariance de l'intervalle :

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 = c^2dt'^2,$$

$$dt' = dt\sqrt{1 - \frac{dx^2 + dy^2 + dz^2}{c^2dt^2}} = \frac{ds}{c} = dt\sqrt{1 - \frac{v^2}{c^2}}$$

- Une horloge mobile marche plus lentement qu'une horloge au repos.
  - Soit une horloge en mouvement rectiligne uniforme dans un référentiel d'inertie K. Le référentiel K' de l'horloge est aussi d'inertie. L'horloge de K' retarde sur celle d'un observateur dans K. Inversement du point de vue de K', c'est l'horloge dans K qui retarde! Comment est-ce possible? Pour pouvoir comparer la marche d'horloges dans des référentiels différents il faut avoir plusieurs horloges dans un référentiel de comparaison : ce processus n'est pas symétrique relativement aux deux référentiels.
  - Si l'on considère deux horloges et que l'une revienne vers l'autre (immobile) après avoir décrit une trajectoire fermée, c'est l'horloge mobile qui retardera par rapport à l'horloge immobile. Le raisonnement inverse d'après lequel les rôles des horloges seraient intervertis est faux car l'horloge décrivant la trajectoire fermée n'a pas un mouvement rectiligne uniforme, et son référentiel ne peut être un référentiel d'inertie.

Champs et particules en Relativité

- L'absence d'interaction instantanée à distance amène à délaisser le concept de force (action d'un point matériel sur un autre) au profit du concept de **champ** : une particule

crée un champ et toute particule se trouvant dans ce champ sera soumise à une certaine force

- C'est seulement au bout d'un certain temps que le changement de la position de l'une des particules produit un effet sur les autres. Une interaction ne peut avoir lieu à tout instant qu'entre deux points voisins de l'espace (théorie locale des champs).
- En mécanique classique, on peut utiliser des corps rigides absolus. De tels corps ne peuvent pas exister en Relativité : ils pourraient transmettre instantanément des actions à distance. Seules interviennent en Relativité des particules élémentaires, (sans structure et sans extension spatiale). La Relativité conserve le concept essentiel, en mécanique classique, de point matériel.

La mécanique relativiste du point matériel

- Intervalle d'espace-temps
  - Soient x et v la position et la vitesse d'un point matériel. L'intervalle d'espace-temps sur la trajectoire de ce point matériel est noté

$$\Delta s^2 = c^2 \Delta t^2 - \Delta \mathbf{x}^2 = c^2 \Delta \tau^2$$

$$d\tau = \left(1 - \frac{\mathbf{v}^2}{c^2}\right)^{1/2} dt = \gamma^{-1} dt$$

τ est le temps propre le long de la trajectoire

- Notations quadri-vectorielles

$$x^\mu \equiv (ct, \mathbf{x}) \quad \mu = 0,1,2,3$$
$$\Delta s^2 = g_{\mu\nu} \Delta x^\mu \Delta x^\nu = \Delta x_\mu \Delta x^\mu$$

où le **tenseur métrique** est défini par

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad g_\mu{}^\nu = \delta_\mu{}^\nu$$

qui fait monter et descendre les indices.

- Invariances de Poincaré et de Lorentz

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu \quad \Lambda^T g \Lambda = g$$

Où Λ est une transformation de Lorentz qui peut s'écrire comme le produit d'une rotation ordinaire par un "**boost**" le long d'une direction **n**

$$\mathbf{x}' = \mathbf{x} + \mathbf{n}\left[(\cosh\alpha - 1)\mathbf{x}.\mathbf{n} - ct\sinh\alpha\right]$$
$$ct' = ct\cosh\alpha - \mathbf{n}\mathbf{x}\sinh\alpha$$

où **v**= c thα **n** est la vitesse du repère en mouvement. La **quadrivitesse**

$$u^\mu = \frac{dx^\mu}{d\tau} = \left(c\frac{dt}{d\tau}, \frac{d\mathbf{x}}{d\tau}\right) = \frac{dt}{d\tau}(c, \mathbf{v})$$

est un quadrivecteur du genre temps de longueur constante égale à *c*. Sa dérivée par rapport au temps propre

$$\frac{d^2x^\mu}{d\tau^2} = \frac{du^\mu}{d\tau}$$

lui est orthogonale et est donc du genre espace.

- Le quadrimoment
  - Pour une particule libre de masse *m*, le quadrimoment s'écrit

$$p^\mu = mu^\mu = \left(\frac{E}{c}, \mathbf{p}\right) = \frac{m}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}}(c, \mathbf{v})$$

$$p^2 = m^2c^2 = \frac{E^2}{c^2} - \mathbf{p}^2 \quad \mathbf{p} = \frac{E\mathbf{v}}{c^2}$$
$$\mathbf{p} = 0 \Rightarrow E_0 = mc^2$$
$$m = 0 \Rightarrow E = |\mathbf{p}|c$$
$$\frac{v}{c} \to 0 \Rightarrow E \approx mc^2 + \frac{1}{2}m\mathbf{v}^2$$

Cinématique des réactions particulaires

- Désintégration d'une particule en deux corps Une particule *A* de masse *M* se désintègre en deux particules *1* et *2* de masses m_1 et m_2. Le système où *A* est au repos est le système du centre de masse de la désintégration: $\boldsymbol{p}_1 + \boldsymbol{p}_2 = 0$; $E_1 + E_2 = M$ (**Attention! : à partir de maintenant c=1)**

$$P_A.p_1 = ME_1 = (p_1 + p_2).p_1 = m_1^2 + p_1.p_2 = \frac{M^2 - m_1^2 - m_2^2}{2} + m_1^2$$
$$E_1 = \frac{M^2 + m_1^2 - m_2^2}{2M};\ E_2 = \frac{M^2 + m_2^2 - m_1^2}{2M}$$

- Les énergies des deux particules produits de la désintégration sont bien déterminées. Dans une désintégration à trois corps, les énergies ne sont pas déterminées; il y a un spectre d'énergie. L'observation d'un spectre d'énergie pour l'électron dans la radioactivité β a conduit à l'hypothèse du neutrino formulée par Pauli.

- Réactions à deux corps
  - Conservation de l'énergie-impulsion: $p_1+p_2=p_3+p_4$ (quatre équations)

$$s=(p_1+p_2)^2=(p_3+p_4)^2$$
$$t=(p_1-p_3)^2=(p_2-p_4)^2 \quad s+t+u=\sum_i m_i^2$$
$$u=(p_1-p_4)^2=(p_2-p_3)^2$$

$$S_{ij}=\left\{\left[s-(m_i+m_j)^2\right]\left[s-(m_i-m_j)^2\right]\right\}^{1/2}$$

$$E_i=\frac{s+m_i^2-m_j^2}{2\sqrt{s}},\ i=1,2\leftrightarrow j=2,1;\ i=3,4\leftrightarrow j=4,3$$

$$p_{ij}=\frac{S_{ij}}{2\sqrt{s}},\ i=1,2\leftrightarrow j=2,1;\ i=3,4\leftrightarrow j=4,3$$

$$w_\mu=\frac{-2\varepsilon_{\mu\nu\rho\sigma}p_1^{\ \nu}p_2^{\ \rho}p_3^{\ \sigma}}{\left[\Phi(s,t)\right]^{1/2}} \quad (w_\mu w^\mu=-1) \quad \mathbf{w}=\mathbf{p_1}\times\mathbf{p_3}$$

$$\Phi(s,t)=stu-s(m_2^2-m_4^2)(m_1^2-m_3^2)-t(m_1^2-m_2^2)(m_3^2-m_4^2)$$
$$-(m_1^2+m_4^2-m_3^2-m_2^2)(m_1^2m_4^2-m_2^2m_3^2)$$

Théorie relativiste du champ électromagnétique

Le champ électromagnétique comme un système dynamique infini

- Equations de Maxwell dans les unités de Heaviside. Dans ces unités, la force de Coulomb s'écrit $QQ'/4\pi r^2$ et nous prenons c=1.

$$\operatorname{div}\mathbf{E}=\rho \quad \operatorname{div}\mathbf{B}=\mathbf{0}$$
$$\operatorname{rot}\mathbf{B}-\frac{\partial\mathbf{E}}{\partial t}=\mathbf{j} \quad \operatorname{rot}\mathbf{E}+\frac{\partial\mathbf{B}}{\partial t}=0$$

- Interprétation physique, sous forme intégrale
  - Loi de Gauss: le flux de **E** à travers une surface est égal à la charge enfermée

$$\int_S \mathbf{E}.d\mathbf{S}=\int_V \rho d^3x$$

  - La circulation de B le long d'une courbe fermée C bordant une surface S est égale au flux à travers S de la somme du courant usuel et du courant de déplacement de Maxwell

$$\oint_C \mathbf{B}.d\mathbf{x}=\int_S\left(\mathbf{j}+\frac{\partial\mathbf{E}}{\partial t}\right)d\mathbf{S}$$

  - **Loi d'induction de Faraday** : un flux magnétique variable engendre une force électromotrice

$$-\oint_C \mathbf{E}.d\mathbf{x} = \int_S d\mathbf{S}.\frac{\partial \mathbf{B}}{\partial t}$$

- Mise sous forme covariante relativiste

$$x^\mu = (t, \mathbf{x}) \quad j^\mu = (\rho, \mathbf{j})$$

$$F^{\mu\nu} = -F^{\nu\mu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}$$

  - Symbole antisymétrique de Lévi-Civita $\varepsilon\mu\nu\rho\sigma$ égal à +1 ou -1 selon que (μνρσ) est une permutation paire ou impaire de (*0123*) et 0 dans tout autre circonstance

$$\tilde{F}^{\mu\nu} = -\tilde{F}^{\nu\mu} = \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} = \begin{pmatrix} 0 & -B^1 & -B^2 & -B^3 \\ B^1 & 0 & E^3 & -E^2 \\ B^2 & -E^3 & 0 & E^1 \\ B^3 & E^2 & -E^1 & 0 \end{pmatrix}$$

$$\tilde{F} = F(E \to B, B \to -E)$$

  - Qui se transforment comme des tenseurs antisymétriques

$$F_{\mu\nu}F^{\mu\nu} = -\tilde{F}_{\mu\nu}\tilde{F}^{\mu\nu} = -2\left(\mathbf{E}^2 - \mathbf{B}^2\right)$$

$$F_{\mu\nu}\tilde{F}^{\mu\nu} = -4\mathbf{E}.\mathbf{B}$$

  - Les équations de Maxwell deviennent alors :

$$\partial_\mu F^{\mu\nu} = j^\nu \quad \partial_\mu \tilde{F}^{\mu\nu} = 0$$

  - Et la conservation du courant :

$$\partial_\mu j^\mu = 0$$

  - **E** et **B** ne peuvent pas être considérées comme des variables dynamiques puisque les équations de Maxwell ne font intervenir que des dérivées du premier ordre. D'autre part, la covariance de Lorentz n'est pas évidente.

- Quadri-potentiel $A\mu$ et arbitraire de jauge

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$$

$$i.e.\ \ \mathbf{E} = -\nabla A^0 - \frac{\partial \mathbf{A}}{\partial t} \quad \mathbf{B} = \mathrm{rot}\,\mathbf{A}$$

  - Arbitraire de jauge : $A^\mu(x) \to A^\mu(x) + \partial^\mu\phi$ ne change pas F.
  - En termes du quadri-potentiel, les équations de Maxwell s'écrivent

$$\Box A^{\mu} - \partial^{\mu}\left(\partial_{\nu}A^{\nu}\right) = j^{\mu}$$

$$\Box \equiv \partial_{\mu}\partial^{\mu} \equiv \left(\partial/\partial t\right)^2 - \Delta$$

forme qui n'est pas affectée par une transformation de jauge

$$\rho = \Box A^0 - \frac{\partial}{\partial t}\left(\frac{\partial A^0}{\partial t} + \operatorname{div}\mathbf{A}\right) = -\Delta A^0 - \frac{\partial}{\partial t}\operatorname{div}\mathbf{A}$$

$$\mathbf{j} = \Box\mathbf{A} + \nabla\left(\frac{\partial}{\partial t}A^0 + \operatorname{div}\mathbf{A}\right)$$

- Jauge de Coulomb, div**A**=0, d'où $A^0(t,\mathbf{x}) = \int \frac{d^3x'}{4\pi}\frac{\rho(t,x')}{|\mathbf{x}-\mathbf{x}'|}$ qui est le potentiel coulombien instantané.

Formulation lagrangienne

La formulation lagrangienne, en prenant comme variables dynamiques les composantes du quadri-potentiel va permettre de mettre en relation invariance de jauge et covariance relativiste.

Pour assurer l'invariance de Lorentz, on écrit la fonction de Lagrange comme une intégrale sur l'espace d'une densité, le lagrangien, dont l'intégrale sur l'espace-temps est l'intégrale d'action $I = \int d^4x\mathcal{L}(x)$

On suppose que l'intégrale s'étend sur tout l'espace-temps et que les champs s'annulent suffisamment vite à l'infini pour que l'on puisse faire des intégrales par parties.

Nous considérons $A^{\mu}(x)$ comme un **champ quadri-vectoriel**, et nous essayons d'écrire un lagrangien dépendant de ce champ et de ses premières dérivées, dont dérivent les équations de Maxwell à partir des équations d'Euler Lagrange généralisées à des systèmes dynamiques infinis

$$\delta I = \int d^4x\left\{\frac{\partial\mathcal{L}(x)}{\partial\varphi_i(x)}\delta\varphi_i(x) + \frac{\partial\mathcal{L}(x)}{\partial\left[\partial_{\mu}\varphi_i(x)\right]}\delta\left[\partial_{\mu}\varphi_i(x)\right]\right\}$$

$$= \int d^4x\left\{\frac{\partial\mathcal{L}(x)}{\partial\varphi_i(x)} - \partial_{\mu}\frac{\partial\mathcal{L}(x)}{\partial\left[\partial_{\mu}\varphi_i(x)\right]}\right\}\delta\varphi_i(x)$$

$$\frac{\delta I}{\delta\varphi_i(x)} \equiv \frac{\partial\mathcal{L}(x)}{\partial\varphi_i(x)} - \partial_{\mu}\frac{\partial\mathcal{L}(x)}{\partial\left[\partial_{\mu}\varphi_i(x)\right]} = 0$$

On détermine les coefficients *a,b,c,d,* et *e* pour retrouver les équations de Maxwell à partir de

$$\mathcal{L}(x) = \frac{1}{2}\left[a\partial_\mu A^\nu \partial^\mu A_\nu + b\partial_\mu A^\nu \partial_\nu A^\mu + c\left(\partial_\mu A^\mu\right)^2 dA_\mu A^\mu + eA_\mu j^\mu \right]$$

on trouve après un calcul simple :

$$\mathcal{L}(x) = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - j_\mu A^\mu + \frac{c}{2}\left[\left(\partial_\mu A^\mu\right)^2 - \partial_\mu A^\nu \partial_\nu A^\mu \right]$$

Mais le terme multipliant *c* est une divergence :

$$\left(\partial_\mu A^\mu\right)^2 - \partial_\mu A^\nu \partial_\nu A^\mu = \partial_\mu \left[A_\nu \left(g^{\mu\nu}\partial_\rho A^\rho - \partial^\nu A^\mu \right) \right]$$

Il peut donc être omis, et on a finalement :

$$I = -\int d^4x \left(\frac{1}{4}F^2 + j.A \right) = \int d^4x \left(\frac{\mathbf{E}^2 - \mathbf{B}^2}{2} - \rho A^0 + \mathbf{j}.\mathbf{A} \right)$$

A cause du terme en *A,* présent si il y a un courant externe, ce lagrangien n'est pas invariant de jauge, puisque sous une transformation de jauge, le terme en *j.A* induit un terme supplémentaire. Mais ce terme supplémentaire est une quadri-divergence si le courant est conservé

$$j_\mu \partial^\mu \phi = \partial^\mu \left(\phi \;\; j_\mu \right)$$

La conservation du courant est la condition nécessaire et suffisante pour l'invariance de jauge de la théorie

Relativité générale et théorie relativiste du champ gravitationnel

Champ de gravitation en mécanique non relativiste

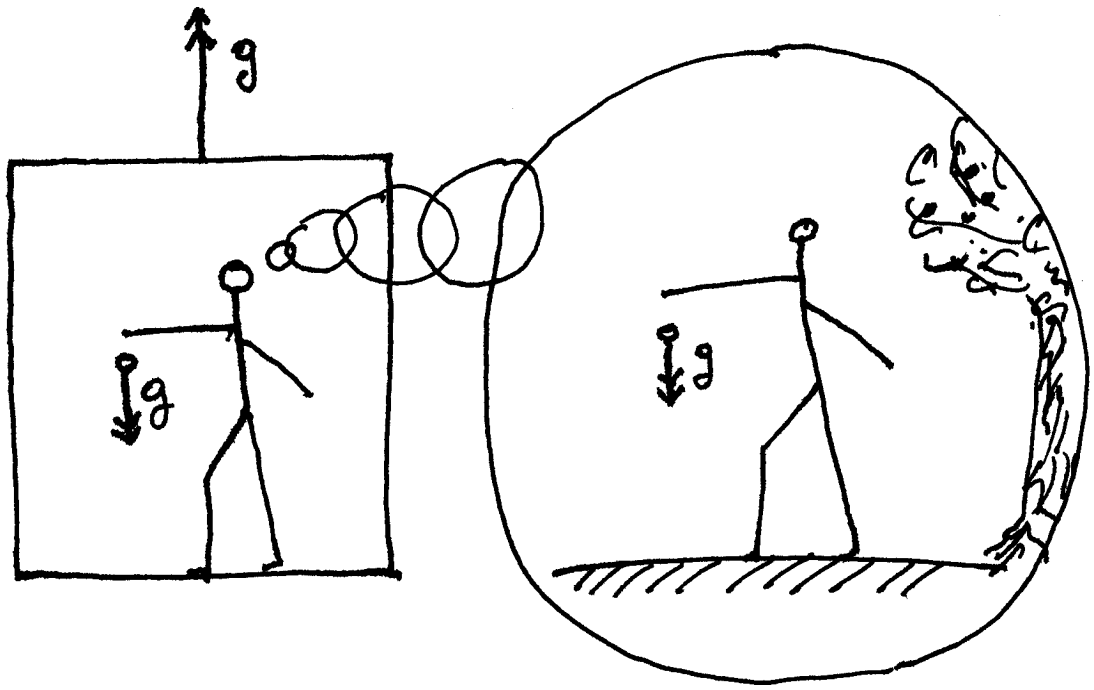
Principe d'équivalence

Les champs de gravitation jouissent de la propriété fondamentale suivante: tous les corps s'y meuvent, indépendamment de leur masse de la même manière (voir chute libre)

Analogie entre le mouvement dans un champ gravitationnel et le mouvement sans champ mais du point de vue d'un référentiel non inertiel.

Principe d'équivalence:un référentiel non inertiel est équivalent à un champ gravitationnel

Voir la première expérience de pensée d'Einstein (référentiel uniformément accéléré)



Limites du principe d'équivalence

Les champs équivalents à un référentiel non inertiel disparaissent dès que l'on passe à un référentiel non inertiel. Ces champs (fictifs) ne s'annulent pas à l'infini. Les champs gravitationnels réels (qui s'annulent à l'infini) ne peuvent être éliminés (partout) par un changement de référentiel. On ne peut les éliminer qu'en un point de l'espace-temps et dans un voisinage petit.

Fonction de Lagrange pour une particule dans un champ de gravitation, en mécanique non relativiste

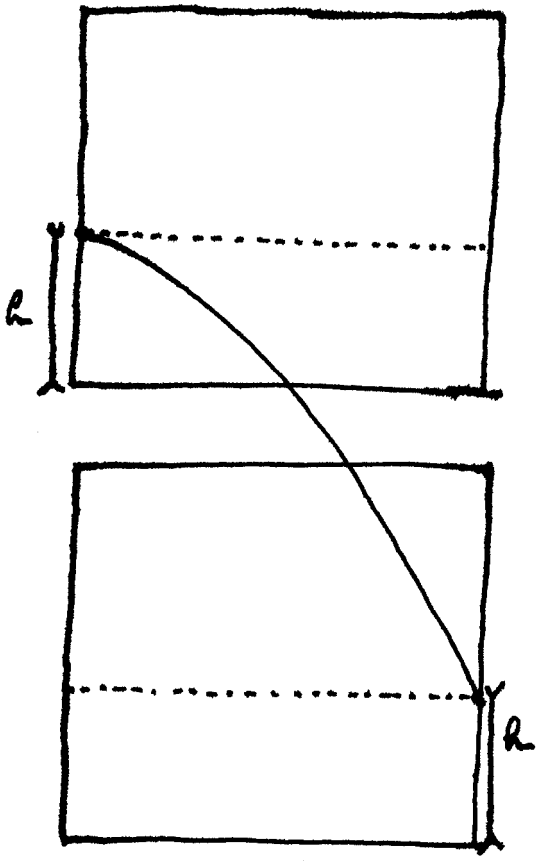
$$L = \frac{mv^2}{2} - m\phi$$

où ϕ est le potentiel gravitationnel. Les équations du mouvement sont

$$\dot{\mathbf{v}} = -\operatorname{grad}\phi$$

Champ de gravitation en mécanique relativiste

Deuxième expérience de pensée d'Einstein



Le principe d'équivalence reste valable. Dans un référentiel d'inertie, l'intervalle d'espace-temps

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$$

est invariant quand on passe à un autre référentiel inertiel. Mais il change si on passe à un référentiel non inertiel. Par exemple quand on passe à un référentiel en rotation uniforme autour de l'axe z,

$$x = x'\cos\omega t - y'\sin\omega t \quad y = x'\sin\omega t + y'\cos\omega t \quad z = z'$$

l'intervalle d'espace-temps devient

$$ds^2 = \left[c^2 - \omega^2\left(x'^2 + y'^2\right)\right]dt^2 - dx'^2 - dy'^2 - dz'^2$$
$$+2\omega y'dx'dt - 2\omega x'dy'dt$$

qui, quelle que soit la loi de transformation du temps ne peut être mise sous la forme d'une somme de carrés des différentielles des quatre coordonnées. La forme la plus générale de l'intervalle dans un référentiel non inertiel est

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu$$

qui définit la **métrique** de l'espace-temps. Le champ de métrique est le champ de force équivalent au référentiel non inertiel.

D'après le principe d'équivalence, le champ de gravitation est donc associé à une modification de la métrique de l'espace-temps: **la gravitation est le phénomène physique qui détermine la géométrie de l'espace-temps (sa métrique).**

En l'absence de gravitation, dans un référentiel inertiel la métrique se réduit à la forme dite galiléenne

$$g_{00} = 1,\ g_{11} = g_{22} = g_{33} = 0,\ g_{ik} = 0,\ i \neq k$$

Tout comme en mécanique non relativiste, le principe d'équivalence ne fonctionne que localement: en présence d'un champ de gravitation, l'espace-temps est tel que les $g_{\mu\nu}$ déterminant sa métrique ne peuvent être ramenés par aucune transformation des coordonnées à la forme galiléenne dans tout l'espace. Un tel espace-temps est dit **courbe**, par opposition à un espace-temps plat où cette réduction serait possible.

La seule chose qui puisse être faite est de réduire, par un choix adéquat de coordonnées la métrique à sa forme galiléenne, en un point donné, et dans son voisinage immédiat.

Notons que le signe du déterminant de la métrique est toujours négatif

$$g = \det g_{\mu\nu} = \left| g_{\mu\nu} \right| < 0$$

La relativité générale

En mécanique non relativiste, la notion de référentiel s'obtenait à partir de corps rigides. C'est d'ailleurs à partir de corps rigides que l'on peut aller à une conception de l'espace (voir Einstein, "Conceptions Scientifiques" pp. 29-33). Cette association corps rigide /référentiel est encore possible en relativité restreinte qui ne considère que des référentiels inertiels. Elle n'est plus possible en relativité générale.

Exemple du référentiel tournant: contraction tangentielle et non contraction radiale: le rapport de la circonférence au diamètre n'est plus égal à π ! Il en est de même en présence d'un champ de gravitation. Métrique spatiale non euclidienne.

Expérience du seau d'eau de Newton.

En présence d'un champ de gravitation variable, non seulement la métrique spatiale n'est pas euclidienne, mais en plus, elle varie avec le temps.

Le "**mollusque de référence**": une sorte de milieu emplissant tout l'espace, avec en chaque point de l'espace, une horloge attachée donnant un temps arbitraire.

Relativité générale: **les lois de la nature doivent être écrites sous une forme qui conviennent formellement dans tout référentiel Attention!**: équivalence formelle ne veut pas dire équivalence concrète; l'expression concrète des phénomènes physiques dépend du référentiel.

Coordonnées curvilignes

Quadrivecteur contravariant se transforme comme les coordonnées

$$x^{\mu} = f^{\mu}\left(x'^{0}, x'^{1}, x'^{2}, x'^{3}\right)$$

$$dx^{\mu} = \frac{\partial x^{\mu}}{\partial x'^{\nu}} dx'^{\nu}$$

$$A^{\mu} = \frac{\partial x^{\mu}}{\partial x'^{\nu}} A'^{\nu}$$

Quadrivecteur covariant se transforme comme le gradient d'un scalaire

$$\frac{\partial \phi}{\partial x^{\mu}} = \frac{\partial \phi}{\partial x'^{\nu}} \frac{\partial x'^{\nu}}{\partial x^{\mu}}$$

$$A_{\mu} = \frac{\partial x'^{\nu}}{\partial x^{\mu}} A'_{\nu}$$

Quadritenseurs contravariants, covariants, mixtes, à nombre arbitraires d'indices. Par exemple

$$A^{\mu\nu} = \frac{\partial x^{\mu}}{\partial x'^{\rho}} \frac{\partial x^{\nu}}{\partial x'^{\sigma}} A'^{\rho\sigma}$$

Le quadritenseur unité

$$\delta^{\mu}{}_{\nu} = 0, \ \mu \neq \nu \quad \delta^{\mu}{}_{\nu} = 1, \ \mu = \nu$$

est un tenseur car pour tout quadrivecteur contravariant on a

$$A^{\nu}\delta^{\mu}{}_{\nu} = A^{\mu}$$

La métrique est un tenseur covariant et symétrique, son inverse est le tenseur contravariant

$$ds^2 = g_{\mu\nu} dx^{\mu} dx^{\nu}$$
$$g_{\mu\nu} = g_{\nu\mu}$$
$$g_{\mu\nu} g^{\nu\rho} = \delta^{\rho}{}_{\mu}$$

Les tenseurs métriques (covariant et contravariant) sont les tenseurs qui font monter et descendre les indices

$$A^{\mu} = g^{\mu\nu} A_{\nu}, \ A_{\mu} = g_{\mu\nu} A^{\nu}$$

Métrique galiléenne

$$\eta_{\mu\nu} = \eta^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Le tenseur unité complètement antisymétrique. En électromagnétisme nous avions introduit le symbole de Lévi Civita $\varepsilon^{\mu\nu\rho\sigma}$. A partir de ce symbole nous pouvons construire le tenseur correspondant. Dans des coordonnées galiléennes *(x')*, le tenseur se réduit au symbole. A partir de la loi générale de transformation des tenseurs, on a ($E^{\mu\nu\rho\sigma}$ désigne le tenseur)

$$E^{\mu\nu\rho\sigma} = \frac{\partial x^{\mu}}{\partial x'^{\mu'}} \frac{\partial x^{\nu}}{\partial x'^{\nu'}} \frac{\partial x^{\rho}}{\partial x'^{\rho'}} \frac{\partial x^{\sigma}}{\partial x'^{\sigma'}} \varepsilon^{\mu'\nu'\rho'\sigma'}$$

$$E^{\mu\nu\rho\sigma} = J\varepsilon^{\mu\nu\rho\sigma}$$

où J est le déterminant des dérivées des nouvelles coordonnées par rapport aux anciennes, c'est-à-dire le *jacobien* du changement de coordonnées. Ce jacobien peut s'exprimer à partir du déterminant de la métrique; En effet

$$g^{\mu\nu} = \frac{\partial x^{\mu}}{\partial x'^{\rho}} \frac{\partial x^{\nu}}{\partial x'^{\sigma}} \eta^{\rho\sigma}$$

$$\det g^{\mu\nu} = \frac{1}{g},\ \det \eta^{\mu\nu} = -1,$$

$$\frac{1}{g} = -J^2,\ J = \frac{1}{\sqrt{-g}}$$

De manière générale les tenseurs font toujours apparaître la racine carrée du déterminant de la métrique. En coordonnées curvilignes, l'intégrale d'un scalaire est aussi un scalaire. Lorsque l'on passe des coordonnées galiléennes aux coordonnées curviligne l'élément d'intégration $d\Omega'$ devient

$$d\Omega' \to \frac{1}{J} d\Omega = \sqrt{-g}\, d\Omega$$

Distances et durées

Temps propre

$$dx^1 = dx^2 = dx^3 = 0$$
$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = c^2 d\tau^2 = g_{00} (dx^0)^2$$
$$d\tau = \frac{1}{c} \sqrt{g_{00}}\, dx^0$$

Elément de distance spatiale

Soit un signal lumineux partant d'un point *B* de l'espace et allant à un point *A* infiniment voisin et réfléchi instantanément en sens inverse. Le temps nécessaire à ce trajet est égal, après multiplication par *c* au double de la distance entre les deux points

$$ds^2 = g_{ij} dx^i dx^j + 2 g_{0i} dx^0 dx^i + g_{00} (dx^0)^2 \quad (i, j = 1, 2, 3)$$

Résolvant l'équation $ds^2 = 0$ relativement à dx^0, on trouve deux solutions

$$dx^{0(\pm)} = \frac{1}{g_{00}} \left(-g_{0i} dx^i \mp \sqrt{\left(g_{0i} g_{0j} - g_{ij} g_{00} \right) dx^i dx^j} \right)$$

le laps de "temps" écoulé est

$$dx^{0(-)} - dx^{0(+)} = \frac{2}{g_{00}} \sqrt{\left(g_{0i} g_{0j} - g_{ij} g_{00} \right) dx^i dx^j}$$

d'où, en multipliant par

$$\sqrt{g_{00}} / c$$

pour avoir le temps propre et par *c/2* pour avoir la distance

$$dl^2 = \gamma_{ij} dx^i dx^j,\ \gamma_{ij} = -g_{ij} + \frac{g_{0i} g_{0j}}{g_{00}}$$

qui est le tenseur métrique tridimensionnel, déterminant les propriétés géométriques de l'espace. Il faut se rappeler que les g_{ij} dépendent de x^0 et donc que la notion de distance entre deux corps perd tout sens en relativité générale, sauf localement. Les déterminants g et γ sont reliés par $-g=g_{00}\gamma$.

La simultanéité en relativité générale

Dans l'expérience d'aller retour de signal lumineux évoquée plus haut, il convient de considérer comme simultanée avec l'instant x^0 au point A l'indication de l'horloge en B à l'instant milieu des deux solutions

$$x^0 + \Delta x^0 = x^0 + \frac{1}{2}\left(dx^{0(+)} + dx^{0(-)}\right),\ \Delta x^0 = -\frac{g_{0i}dx^i}{g_{00}} \equiv g_i dx^i$$

Cette relation permet de synchroniser des horloges dans n'importe quel volume infiniment petit de l'espace.

Dérivation covariante

En coordonnées galiléennes, la différentielle d'un vecteur est un vecteur. Ce n'est plus vrai en relativité générale parce que la différentielle est la différence de deux vecteurs ne se trouvant pas au même point. Pour définir la différentielle, il faut "transporter" l'un des deux vecteurs infiniment voisins au point ou se trouve le second. L'opération de transport doit être telle qu'en coordonnées galiléennes, la différence coïncide avec la différentielle. D'où la notion de transport parallèle: quand on compare deux vecteurs voisins, on doit transporter parallèlement l'un d'eux au point où se trouve l'autre.

Coefficients de Christoffel

Alors la différence entre deux vecteurs se trouvant maintenant au même point s'écrit

$$DA^\mu = dA^\mu - \delta A^\mu$$

l'accroissement dû au transport parallèle dépend linéairement des composantes (pour que la loi de transformation soit la même pour la somme de deux vecteurs que pour chacun des deux)

$$\delta A^\mu = -\Gamma^\mu{}_{\nu\rho} A^\nu dx^\rho$$

où les coefficients $\Gamma^{\mu}{}_{\nu\rho}$, appelés **coefficients de Christoffel**, sont des fonctions des coordonnées, dépendant du choix du système de coordonnées. Ces coefficients sont tous nuls dans un système galiléen. Ces coefficients ne forment pas un tenseur, car un tenseur nul dans un système de coordonnées est nul dans tout autre système. Nous verrons que l'on peut choisir un référentiel tel que les Christoffel s'annulent en un point donné à l'avance. Il est utile d'introduire

$$\Gamma_{\mu,\nu\rho} = g_{\mu\sigma}\Gamma^{\sigma}{}_{\nu\rho} \quad \Gamma^{\mu}{}_{\nu\rho} = g^{\mu\sigma}\Gamma_{\sigma,\nu\rho}$$

le produit scalaire de deux vecteurs ne variant pas dans un transport parallèle, on a

$$\delta(A_\mu B^\mu) = 0 \rightarrow B^\mu \delta A_\mu = -A_\mu \delta B^\mu = \Gamma^{\mu}{}_{\nu\rho} B^\nu A_\mu dx^\rho \rightarrow \delta A_\mu = \Gamma^{\nu}{}_{\mu\rho} A_\nu dx^\rho$$

Les expressions entre parenthèses dans

$$DA^\mu = \left(\frac{\partial A^\mu}{\partial x^\rho} + \Gamma^{\mu}{}_{\nu\rho} A^\nu \right) dx^\rho$$

$$DA_\mu = \left(\frac{\partial A_\mu}{\partial x^\rho} - \Gamma^{\nu}{}_{\mu\rho} A_\nu \right) dx^\rho$$

sont des tenseurs, puisque leur produit par un vecteur est un vecteur, on les appelle **les dérivées covariantes** et on les note

$$DA^\mu = A^\mu{}_{;\rho} dx^\rho, \; DA_\mu = A_{\mu;\rho} dx^\rho$$

En coordonnées galiléennes, les Christoffel s'annulent et les dérivées covariantes coïncident avec les dérivées ordinaires. Il est facile de définir les dérivées covariantes d'un tenseur, on a, par exemple pour un tenseur contravariant

$$A^{\mu\nu}{}_{;\rho} = \frac{\partial A^{\mu\nu}}{\partial x^\rho} + \Gamma^{\mu}{}_{\sigma\rho} A^{\sigma\nu} + \Gamma^{\nu}{}_{\sigma\rho} A^{\mu\sigma}$$

On vérifie facilement que la dérivée covariante d'un produit obéit aux mêmes règles que la dérivation ordinaire

$$(A_\mu B_\nu)_{;\rho} = A_{\mu;\rho} B_\nu + A_\mu B_{\nu;\rho}$$

Symétrie des Christoffel par rapport aux indices inférieurs

$$A_\mu = \frac{\partial \phi}{\partial x^\mu}$$

$$\frac{\partial A_\mu}{\partial x^\nu} = \frac{\partial^2 \phi}{\partial x^\mu \partial x^\nu} = \frac{\partial A_\nu}{\partial x^\mu} \rightarrow A_{\nu;\mu} - A_{\mu;\nu} = \left(\Gamma^\rho{}_{\mu\nu} - \Gamma^\rho{}_{\nu\mu}\right)\frac{\partial \phi}{\partial x^\rho}$$

en coordonnées galiléennes, les dérivées covariantes coïncident avec les dérivées ordinaires, et le premier membre de l'égalité est nul. Mais comme la différence des deux dérivées covariantes est un tenseur, son annulation dans un référentiels est vraie dans tout référentiel.

Formule de transformation des Christoffel

$$\Gamma^\mu{}_{\nu\rho} = \Gamma'^\sigma{}_{\tau\upsilon}\frac{\partial x^\mu}{\partial x'^\sigma}\frac{\partial x'^\tau}{\partial x^\nu}\frac{\partial x'^\upsilon}{\partial x^\rho} + \frac{\partial^2 x'^\sigma}{\partial x^\nu \partial x^\rho}\frac{\partial x^\mu}{\partial x'^\sigma}$$

Annulation des Christoffel dans un référentiel particulier. Soit un point donné choisi à l'origine des coordonnées, et soient $(\Gamma\mu_{\nu\rho})_0$ les Christoffel en ce point. On effectue en ce point la transformation

$$x'^\mu = x^\mu + \frac{1}{2}\left(\Gamma^\mu{}_{\nu\rho}\right)_0 x^\nu x^\rho$$

pour laquelle on a

$$\left(\frac{\partial^2 x'^\sigma}{\partial x^\nu \partial x^\rho}\frac{\partial x^\mu}{\partial x'^\sigma}\right)_0 = \left(\Gamma^\mu{}_{\nu\rho}\right)_0$$

Ce qui implique que tous les Γ' sont nuls. Ce référentiel est dit **localement géodésique.** La transformation considérée est telle que

$$\left(\frac{\partial x'^\mu}{\partial x^\nu}\right)_0 = \delta^\mu{}_\nu$$

de sorte qu'elle ne change pas les composantes de n'importe quel tenseur: l'annulation des Christoffel peut être faite en même temps que la réduction du tenseur métrique à la forme galiléenne: c'est l'expression mathématique du principe d'équivalence.

Lien entre les Christoffel et le tenseur métrique

La dérivée covariante du tenseur métrique est nulle. En effet

$$DA_\mu = g_{\mu\nu} DA^\nu$$
$$A_\mu = g_{\mu\nu} A^\nu$$
$$DA_\mu = D\left(g_{\mu\nu} A^\nu\right) = g_{\mu\nu} DA^\nu + A^\nu Dg_{\mu\nu}$$
$$\forall A^\mu \rightarrow Dg_{\mu\nu} = 0$$

on peut alors en déduire l'expression des Christoffels en fonction du tenseur métrique. On trouve

$$\Gamma_{\mu,\nu\rho} = \frac{1}{2}\left(\frac{\partial g_{\mu\nu}}{\partial x^\rho} + \frac{\partial g_{\mu\rho}}{\partial x^\nu} - \frac{\partial g_{\nu\rho}}{\partial x^\mu}\right)$$
$$\Gamma^\mu{}_{\nu\rho} = g^{\mu\sigma}\Gamma_{\sigma,\nu\rho} = \frac{1}{2} g^{\mu\sigma}\left(\frac{\partial g_{\sigma\nu}}{\partial x^\rho} + \frac{\partial g_{\sigma\rho}}{\partial x^\nu} - \frac{\partial g_{\nu\rho}}{\partial x^\sigma}\right)$$

Il est possible de généraliser la notion de divergence et de généraliser la théorème de Gauss. On trouve

$$\oint A^\mu \sqrt{-g}\, dS_\mu = \int A^\mu{}_{;\mu} \sqrt{-g}\, d\Omega$$

Mouvement d'une particule dans un champ de gravitation

En relativité restreinte, le mouvement d'une particule libre est déterminé par le principe de moindre action

$$\delta S = -mc\delta \int ds = 0$$

En relativité générale, on a formellement la même équation pour le mouvement d'une particule dans un champ gravitationnel, qui n'est rien d'autre que la modification de la métrique de l'espace-temps, qui modifie le *ds*.

Dans un champ gravitationnel la particule se meut sur une extrémale (**géodésique**) de l'espace-temps.

Le mouvement d'une particule libre, en relativité restreinte s'écrit

$$u^\mu = \frac{dx^\mu}{ds},\ du^\mu = 0$$

qui se généralise en coordonnées curvilignes en $Du\mu=0$, soit

$$du^\mu + \Gamma^\mu{}_{\nu\rho} u^\nu dx^\rho = 0$$
$$\frac{d^2 x^\mu}{ds^2} + + \Gamma^\mu{}_{\nu\rho} \frac{dx^\nu}{ds}\frac{dx^\rho}{ds}$$

Le champ de gravitation est un champ d'accélération, et

$$-m\Gamma^\mu{}_{\nu\rho} u^\nu u^\rho$$

peut être appelée la **quadriforce** agissant sur la particule dans le champ de gravitation. Le tenseur métrique joue le rôle de potentiel gravitationnel. La possibilité du choix du référentiel annulant les Christoffel est l'expression du principe d'équivalence.

En mécanique non relativiste

$$L = -mc^2 + \frac{mv^2}{2} - m\phi$$

$$S = \int L dt = -mc \int \left(c - \frac{v^2}{2c} + \frac{\phi}{c} \right) dt = -mc \int ds$$

$$\rightarrow ds = \left(c - \frac{v^2}{2c} + \frac{\phi}{c} \right) dt$$

Soit en élevant au carré, en négligeant les termes s'annulant pour c tendant vers l'infini

$$ds^2 = \left(c^2 + 2\phi \right) dt^2 - d\mathbf{r}^2 \quad (\mathbf{v}dt = d\mathbf{r})$$

$$g_{00} = 1 + \frac{2\phi}{c^2}$$

Champ gravitationnel constant

$$\tau = \frac{1}{c} \sqrt{g_{00}} x^0$$

$$\tau = \frac{x^0}{c} \left(1 + \frac{\phi}{c^2} \right)$$

Equations du champ de gravitation

Définition covariante de la courbure

Il est possible de généraliser à des variétés quadri-dimensionnelles la notion de courbure qui est intuitive pour une surface à deux dimensions plongée dans un espace euclidien à trois dimensions. Dans ce cas en effet, il suffit pour se faire une idée de la courbure, de suivre la rotation d'une normale à la surface lorsque l'on se déplace sur la surface.

Dans le cas de l'espace-temps courbe de la relativité générale, la notion même de "normale" n'est pas évidente.

L'idée de base est à nouveau d'avoir recours au transport parallèle de vecteurs: on transporte parallèlement à lui-même un quadri-vecteur le long d'une courbe infinitésimale fermée; la courbure se manifestera par le fait que dans ce transport parallèle, le vecteur subira une rotation: à l'arrivée le vecteur transporté fera un angle avec le vecteur initial.

Exemple du transport parallèle sur une sphère.

Transport parallèle = angle constant avec une géodésique.

A partir des Christoffel et de leur dérivées, il est ainsi possible de construire un tenseur à quatre indices, **le tenseur de courbure de Riemann,**

$$R_{\mu\nu\rho\sigma}$$

qui permet de caractériser, de manière covariante, la courbure de l'espace-temps de la relativité générale.

Le tenseur de Riemann a des propriétés de symétrie ou d'antisymétrie par permutation de certains de ses indices et il satisfait l'importante **identité de Bianchi**

La contraction sur les indices qui ne donne pas zéro par antisymétrie permet de définir un tenseur de courbure à deux indices, **le tenseur de Ricci**

$$R_{\mu\nu} = g^{\rho\sigma} R_{\rho\mu\sigma\nu}$$

Lequel donne, par contraction, **la courbure scalaire**

$$R = g^{\mu\nu} R_{\mu\nu}$$

L'action pour un champ de gravitation

Pour trouver les équations du champ de gravitation il nous faut écrire l'action dont elles dérivent par principe de moindre action. Cette action est écrite sous forme d'une intégrale scalaire

$$S_g = \int K\sqrt{-g}d\Omega$$

où K est une quantité scalaire ne dépendant que du tenseur métrique et de ses premières dérivées, c'est-à-dire du tenseur métrique et des Christoffel.

La courbure scalaire R dépend du tenseur métrique, de ses dérivées premières mais aussi de ses dérivées secondes. Mais, comme la dépendance dans les dérivées secondes est linéaire, on peut ramener l'intégrale sur l'espace-temps de la courbure scalaire à une intégrale d'un scalaire ne dépendant que du tenseur et de ses dérivées premières plus l'intégrale d'un divergence, sans effet sur les équations du mouvement.

L'action pour un champ de gravitation s'écrit donc

$$S_g = \kappa\int R\sqrt{-g}d\Omega$$

où κ est une constante que l'on détermine en imposant de retrouver la théorie de Newton à l'approximation non relativiste.

Les équations d'Einstein

On dérive l'équation de la gravitation en ajoutant à l'action de la gravitation l'action de la matière et en appliquant le principe de moindre action. On trouve, tous calculs faits, les équations suivantes, appelées équations d'Einstein

$$\delta S_M + \delta S_g = 0$$

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu}$$

Le tenseur du membre de gauche, appelé **tenseur d'Einstein**, a une divergence qui s'annule identiquement. L'annulation de la divergence du membre de droite (tenseur d'énergie impulsion de la matière, y compris du champ électromagnétique) exprime les équations du mouvement de la matière (conservation de l'énergie et de l'impulsion).

La constante G qui intervient dans le second membre de l'équation n'est autre que … la constante de Newton.

Noter que les équations d'Einstein sont non linéaires dans le champ de gravitation.

En l'absence de matière, les équations d'Einstein se réduisent à l'annulation du tenseur de Ricci, qui n'implique pas nécessairement que l'espace-temps soit plat.

La loi de Newton

Rappelons la relation entre la métrique et le potentiel gravitationnel à l'approximation non relativiste

$$g_{00} = 1 + \frac{2\phi}{c^2}$$

A l'approximation non relativiste, (mouvement lent de toutes les particules de matière) le tenseur d'énergie impulsion de la matière se réduit à la seule composante

$T_{00} = \mu c^2$ où μ est la densité de masse.

La seule équation d'Einstein restante devient

$$R_{00} = \frac{4\pi G}{c^2} \mu$$

$R_{00} \approx \frac{1}{c^2} \Delta\phi$ Dont la solution est

$\Delta\phi = 4\pi G\mu$

$$\phi = -G \int \frac{\mu dV}{r}$$

Confirmations observationnelles de la théorie de la relativité générale

Avance du périhélie de Mercure

La théorie de la relativité générale permet de calculer l'avance du périhélie de Mercure. Elle donne

$$+\frac{24\pi^3 a^2}{T^2 c^2 \left(1-e^2\right)}$$

où a est le demi grand axe de l'orbite, T la période de révolution, et e l'excentricité de l'orbite. Soit, pour Mercure, la planète la plus proche du Soleil, 43 seconde d'arc par siècle, en contradiction avec les prédictions de la théorie de Newton mais en accord avec les observations. En 1916, lorsque la théorie de la relativité a été publiée, les écarts entre la théorie de Newton et le mouvement des autres planètes étaient trop petits pour être décelés. De nos jours ces écarts ont été observés pour toutes les planètes, et ils sont expliqués par la RG.

Déflexion de la lumière par un champ de gravitation

La RG permet de calculer la déflexion de la lumière par un champ de gravitation. On trouve

$$\alpha = \frac{1.7\,\text{sec}}{\Delta}$$

où Δ est le paramètre d'impact. Remarquons que la théorie de Newton prédit aussi une déflexion, mais deux fois moindre. Les observations lors de l'éclipse totale de 1919, rapportées par Eddington, confirment la prédiction de la RG.

Confirmations modernes

- Les observations du mouvement de pulsars doubles (analogues à l'avance du périhélie de Mercure) sont en excellent accord (avec une précision comparable aux tests de l'électrodynamique quantique) avec les calculs basés sur la RG
- Des effets de lentilles gravitationnelles, et micro-lentilles gravitationnelles ont été observés en très grand nombre. L'effet de micro-lentille gravitationnelle devient un nouveau moyen de recherche astrophysique (recherche de "naines brunes")

Les trous noirs et le rayon de Schwarzschild

Un raisonnement heuristique, non rigoureux basé sur la théorie de Newton permet de calculer le rayon de Schwarzschild. La vitesse de libération est alors donnée par

$$\frac{1}{2}mv^2 = G\frac{mM}{r}$$

$$v_l^{\,2} = \frac{2GM}{r}$$

où M est la source du champ. Le rayon de Schwarzschild est la distance du centre de la source à l'objet animé d'une vitesse de libération égale à la vitesse de la lumière.
Une étoile de deux masses solaires peut avoir un **effondrement gravitationnel** produisant un **trou noir**

.

La théorie quantique des champs

La seconde quantification

Il est possible d'emprunter deux voies pour parvenir à la théorie quantique des champs. La première voie consiste à partir d'une théorie classique relativiste des champs (comme la théorie relativiste de l'interaction électromagnétique) et à la quantifier ; la seconde consiste à partir de la mécanique quantique non relativiste, gouvernée par l'équation de Schrödinger, et à la rendre relativiste. Cette deuxième voie est celle de la *seconde quantification*, que nous allons maintenant discuter brièvement.

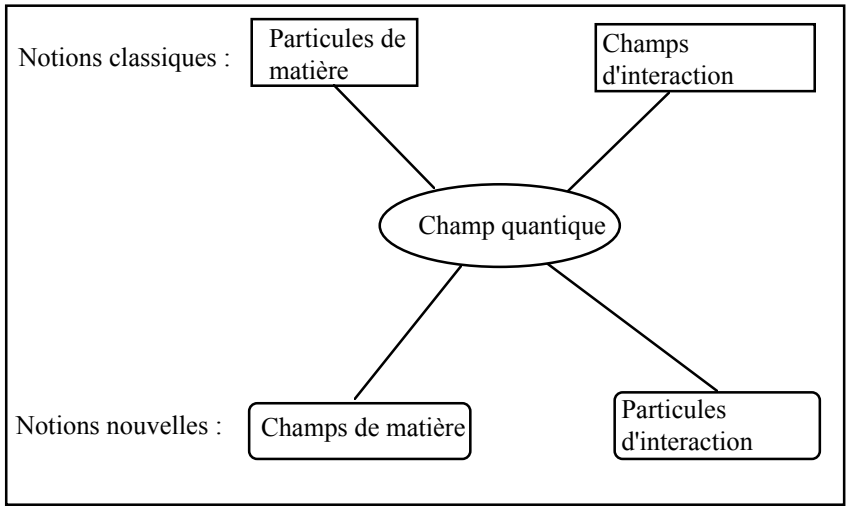
La « première quantification » de la mécanique non relativiste de la particule ponctuelle consiste à remplacer la position $\boldsymbol{x}$ de la particule et son impulsion $\boldsymbol{p}$ par des *opérateurs* agissant sur un *espace de Hilbert.* Les règles de *commutation* de ces opérateurs sont établies par analogie avec la formulation de la mécanique rationnelle à l'aide des *crochets de Poisson*. Les éléments ou vecteurs de l'espace de Hilbert décrivent les configurations possibles ou *états* du système à une particule. La représentation dans l'espace des coordonnées d'un tel état est ce que l'on appelle une *fonction d'onde*. Cette fonction d'onde est une *amplitude de probabilité*, c'est-à-dire une fonction complexe de la position $\boldsymbol{x}$ et du temps t , $\psi(\boldsymbol{x},t)$, dont le carré du module est la probabilité de trouver la particule au point $\boldsymbol{x}$, à l'instant t. La fonction d'onde est solution de l'*équation de Schrödinger*, qui est une équation aux dérivées partielles.

Le processus de seconde quantification consiste à traiter $\psi(\boldsymbol{x},t)$, l'état du système une première fois quantifié, non plus comme une fonction mais comme un opérateur relatif au système qui est alors quantifié une « seconde fois ». $\boldsymbol{x}$ et $\boldsymbol{p}$ ne sont plus traités comme des opérateurs mais comme des indices continus du nouvel opérateur.

Pourquoi cette théorie de seconde quantification peut-elle être qualifiée de théorie quantique des champs ? Parce que l'on peut considérer la fonction d'onde de la théorie de première quantification, qui n'est rien d'autre qu'une fonction ordinaire définie sur tout l'espace-temps, comme un champ classique, qui, au travers du processus de seconde quantification est devenu un *champ quantique,* solution d'une équation de champ opératorielle. L'intérêt essentiel de la seconde quantification est que cette équation de champ garde la même forme quand on passe d'un système à une particule à un système à deux, trois ou un nombre quelconque de particules. Comme nous le verrons plus bas, cette propriété est très utile quand on veut rendre relativiste la théorie, puisqu'en relativité, le nombre de

particules n'est pas conservé. Dans la théorie quantique relativiste des champs, les opérateurs décrivant des champs quantiques sont des opérateurs de *création* ou d'*annihilation* de particules. L'espace de Hilbert sur lequel agissent ces opérateurs est ce que l'on appelle un espace de *Fock*, c'est-à-dire un empilement infini d'espaces de Hilbert, communiquant par l'intermédiaire des opérateurs champs et comportant le *vide*, espace à zéro particule, l'espace à une particule, l'espace à deux particules, etc.

La théorie quantique des champs permet de réconcilier les deux grandes approches de la physique classique, qui semblaient totalement incompatibles, celle du *point matériel* et celle du *champ*. Comme le montre la figure 2, le concept de champ quantique réconcilie les concepts classiques de particules de matière et de champs d'interaction et introduit les concepts nouveaux de *champs de matière* et de *particules d'interaction*.



Causalité et localité

La théorie de la relativité, avons nous dit, résulte de la prise en compte de la contrainte d'absence d'interaction à distance. En théorie quantique des champs, cette contrainte va entrer en conflit avec une propriété directement liée à la flèche du temps, la *causalité*. Cette propriété signifie tout simplement que la cause doit précéder l'effet : le temps qui s'écoule d'une cause à son effet comporte nécessairement une flèche. Le principe de causalité sera sans doute l'un des derniers auxquels les sciences renonceront un jour.

En mécanique quantique non relativiste, la causalité est inscrite dans l'équation de Schrödinger qui fait jouer au *hamiltonien*, l'opérateur que la première quantification associe à l'énergie totale du système, le rôle de générateur infinitésimal des translations dans le temps. Le temps est traité, en mécanique quantique non relativiste, non pas comme un opérateur mais comme un paramètre continu ; la description des interactions y est *locale* dans le temps. Rien

par contre ne nous impose de traiter de la même façon l'espace : la projection sur un axe de coordonnée de la position spatiale n'est pas une variable continue mais plutôt un opérateur. Les interactions n'ont pas à être locales dans l'espace : comme on ne se préoccupe pas du temps pris par la propagation des interactions, on peut avoir des interactions *locales* dans le temps et *non locales* dans l'espace, c'est-à-dire des actions instantanées à distance.

On comprend alors en quoi la seconde quantification répond à l'objectif de rendre relativiste la mécanique quantique : en traitant la position spatio-temporelle $x=(\boldsymbol{x}, t)$ comme un ensemble de quatre variables continues, indices continus des opérateurs champs, on rétablit la symétrie de traitement de l'espace et du temps nécessaire à la théorie de la relativité. Mais on voit alors surgir un problème inattendu: des champs quantiques ne peuvent être couplés les uns aux autres qu'en des *points* d'espace-temps ; en théorie quantique des champs les seules interactions possibles sont des *interactions de contact spatio-temporel.* Un retour à Descartes en somme! Pour comprendre les implications considérables de la localité spatio-temporelle, il nous faut revenir aux inégalités de Heisenberg : d'après ces inégalités, on voit qu'il est impossible de faire tendre Δt et $\Delta \boldsymbol{x}$ vers zéro (si l'on veut limiter la région d'interaction à un point d'espace-temps), sans faire tendre ΔE et $\Delta \boldsymbol{p}$ vers l'infini. Les inégalités de Heisenberg prennent alors une signification nouvelle, qui était déjà celle de la première inégalité en première quantification : dans une petite région spatio-temporelle définie par Δt et $\Delta \boldsymbol{x}$, les lois de conservation de l'énergie et de l'impulsion peuvent être violées, avec des erreurs ΔE et $\Delta \boldsymbol{p}$ reliées à Δt et $\Delta \boldsymbol{x}$ par les inégalités de Heisenberg. De même, le nombre de particules n'est pas conservé. Les processus pour lesquels ces lois de conservation seraient violées mettraient en jeu une action inférieure au quantum d'action et ils ne peuvent donc être réels ; on les qualifie de *virtuels*. Comme le point d'espace-temps est nécessairement une idéalisation, impossible à réaliser pratiquement, il nous faut l'approcher avec une certaine résolution, et envisager d'un point de vue théorique, l'ensemble des processus virtuels qui peuvent intervenir dans la limite de la résolution spatio-temporelle. Plus cette résolution est élevée (plus est petite la région d'espace-temps explorée) plus grandes peuvent être les violations des lois de conservation de l'énergie-impulsion, plus virtuels sont les processus qu'il faut prendre en considération. D'un point de vue expérimental on peut explorer les très courtes distances spatio-temporelles par exemple en provoquant des collisions entre particules à très haute énergie. A l'aide de ces réactions particulaires on peut étudier, de manière statistique, l'actualisation de certains de ces processus virtuels. Tel est le programme de la

physique des particules : la théorie quantique des champs fournit les probabilités des processus élémentaires que l'on provoque expérimentalement dans les réactions particulaires à haute énergie. Vaste programme dont on devine les énormes difficultés : la causalité implique la localité mais la localité fait surgir le problème d'une *singularité ponctuelle*, assez analogue à celle du big bang en cosmologie.

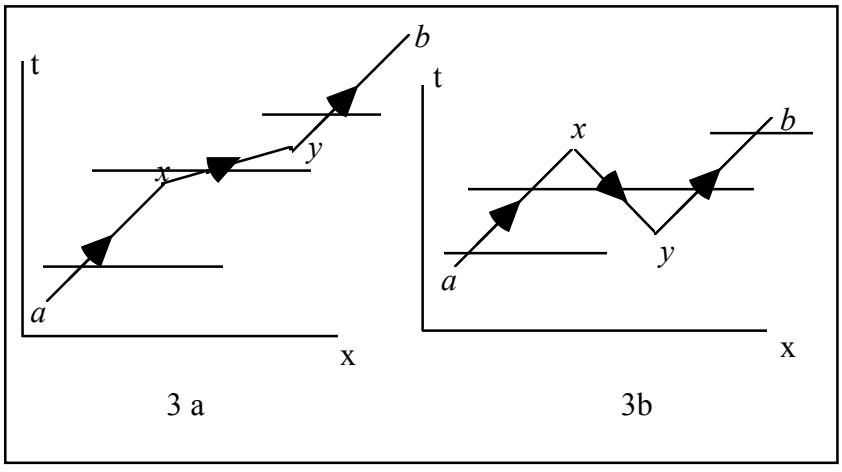
Causalité et antiparticules

Avant de discuter les méthodes qui permettent d'affronter ces difficultés, nous devons examiner d'autres implications de la causalité qui nécessitent des adaptations du cadre général de la théorie quantique des champs.

Les contraintes de la causalité s'expriment au moyen des règles de commutation des opérateurs champs. Un opérateur de création $\phi^*(x)$ d'une particule au point d'espace-temps x et l'opérateur d'annihilation de cette même particule $\phi(y)$ au point d'espace-temps y doivent commuter pour une séparation de x et y du genre espace et ne pas commuter pour une séparation du genre temps : ces règles empêchent une particule de se propager sur une ligne du genre espace (ce qui voudrait dire que la particule se propagerait plus vite que la lumière) et, pour la propagation sur une ligne du genre temps, que la création de la particule a précédé son annihilation. Ces règles ne peuvent pas être satisfaites si la décomposition en onde planes (analyse de Fourrier) des opérateurs champs ne comporte pas de *modes de fréquence négative.* Mais fréquence négative signifie, en mécanique quantique, énergie négative. Que faire alors de ces états d'énergie négative ? Si on considère une onde plane d'énergie négative, on remarque que par renversement du sens du temps ($t \rightarrow -t$), les états d'énergie négative se mettent à ressembler à des états d'énergie positive. On résout donc le problème des états d'énergie négative en supposant que ces états ne se propagent *qu'en remontant le temps,* et en ré interprétant une particule d'énergie négative qui remonte le temps comme *une antiparticule d'énergie positive qui le descend.* Particule et antiparticule doivent avoir la même masse, des charges opposées et de manière générale, tous les nombres quantiques opposés. Ainsi, pour que le formalisme de la théorie quantique des champs soit compatible avec la relativité et la causalité, il a fallu inventer un concept nouveau, celui d'antiparticule. De fait il apparaît bien que toute particule connue a un partenaire de même masse et de nombres quantiques opposés que l'on peut assimiler à son antiparticule. C'est déjà un succès considérable à mettre à l'actif de cette théorie.

Le concept d'antiparticule avait été inventé par Dirac lorsqu'il avait essayé de rendre relativiste la mécanique quantique de l'électron. Il avait aussi buté sur le problème des états d'énergie négative. La solution qu'il proposait consistait à redéfinir le vide, non plus comme espace à zéro particule, mais comme la configuration d'énergie minimale. Cette configuration est celle dans laquelle tous les états possibles d'énergie négative sont occupés chacun par un électron. En fait, l'énergie du vide n'est pas nulle elle vaut plutôt $-\infty$. Mais le principe d'exclusion de Pauli auquel obéissent les électrons interdit d'ajouter à un tel vide aucun électron d'énergie négative, puisque tous les états possibles sont déjà occupés. Il devient alors possible de considérer le vide comme un espace à zéro particule et d'énergie nulle. Tout se passe donc comme si les états d'énergie négative n'existaient pas : ils ne sont que *virtuels*. Supposons alors qu'il manque au vide un électron d'énergie négative ; ce "trou" d'énergie négative représente un antiélectron ou *positon* d'énergie positive. Si un électron "tombe dans un trou", on dira qu'il y a eu annihilation d'une paire électron-positon. Si une certaine interaction ponctuelle éjecte du vide un des électrons d'énergie négative, en lui donnant une énergie positive et en laissant un « trou » à sa place, on dira qu'il y a eu création d'une paire électron-positon.

Le concept d'antiparticule permet de comprendre pourquoi lors des processus virtuels le nombre de particules n'est pas conservé. Considérons la propagation d'une particule depuis un point d'espace-temps *a* jusqu'à un point d'espace-temps *b* (voir la figure 3). Supposons que lors de sa propagation, la particule subisse des interactions avec un champ extérieur aux points d'espace-temps *x* et *y*.



Les lignes horizontales figurent des « coupes à temps constant » qui permettent de déterminer le nombre de particules en propagation. Sur la figure 3a, à chaque instant, une seule particule se propage ; les interactions en *x* et *y* ne font que modifier la trajectoire de la

particule. L'événement *x* a précédé l'événement *y*. Dans la figure 3b, on a d'abord une particule puis trois (deux particules et une antiparticule) puis une seule à nouveau. L'événement *y* (création d'une antiparticule se propageant de *y* à *x et* d'une particule se propageant de *y* à *b*) a précédé l'événement *x* (annihilation de la particule venant de *a* et de l'antiparticule venant de *y*). On voit donc que lors des interactions ponctuelles des particules et antiparticules peuvent être créées ou annihilées par paires. En réalité des particules qui, comme le photon, coïncident avec leur propre antiparticule peuvent être créées ou annihilées autrement que par paires.

Les symétries discrètes et leur violations

On appelle *conjugaison de charge*, notée C l'opération qui consiste à transformer une particule en son antiparticule. A cette opération est associé un opérateur qui agit dans l'espace de Hilbert des états. De même *T* est l'opération (ou l'opérateur associé) de *renversement du sens du temps*. Pour être complète, la résolution du problème des états d'énergie négative, nécessite une troisième opération (ou opérateur), la *parité* d'espace, notée *P* qui consiste à changer de signe les coordonnées spatiales. *C, P* et *T* sont ce que l'on appelle des *symétries discrètes*, parce qu'il s'agit d'opérations de symétries dont le carré vaut 1 (répéter l'opération ramène à l'état initial). En théorie quantique des champs, le *théorème CPT* stipule que toutes les interactions sont invariantes par l'opération *CPT* qui consiste à remplacer chaque particule par son antiparticule, à changer de signe toues les coordonnées d'espace et à changer le sens du temps. Il s'agit d'une *prédiction* que la théorie quantique des champs est conduite à faire si elle veut prétendre à rendre compte de la relativité restreinte, de la mécanique quantique et de la causalité dans les processus élémentaires. Jusqu'à présent, ce théorème n'a subi aucune contradiction expérimentale. Encore un incontestable succès de la théorie quantique des champs.

Le théorème *CPT* ne dit rien sur l'invariance par chacune des opérations prise séparément : ce théorème autorise des violations de *P, C,* ou T dans telle ou telle interaction, dès lors que le produit *CPT* n'est pas violé. Alors que les interactions électromagnétique et nucléaire forte semblent invariantes par chacune des trois symétries discrètes, l'interaction nucléaire faible se montre complètement récalcitrante à l'égard de ces symétries. Lee et Yang avaient émis l'hypothèse que cette interaction n'est pas invariante par parité d'espace, et Mme Wu a observé expérimentalement cette violation de symétrie discrète. L'interaction faible viole aussi la conjugaison de charge. Dans la théorie moderne de l'interaction faible, la

violation de la parité et celle de la conjugaison de charge s'interprètent comme une sensibilité de l'interaction à la *chiralité* des leptons, des quarks et de leurs antiparticules : seuls les leptons et quarks *gauches* ou lévogyres et les antileptons et antiquarks *droits* ou *dextrogyres* participent à l'interaction faible. Jusqu'en 1964 on croyait, faute d'indication expérimentale contraire, que l'interaction faible était invariante par le produit *CP et* donc, d'après le théorème *CPT* aussi par *T*. Mais l'expérience de Christensen, Cronin, Fitch et Turlay a révélé une violation (certes très ténue mais incontestable) de la symétrie *CP* dans la désintégration par interaction faible de certaines particules (les "mésons K neutres"). Jusqu'à présent cet effet n'est pas compris et il n'a été observé dans aucune autre situation. Comme la violation de *CP* implique, si l'on admet le théorème *CPT*, la violation de *T*, cette brisure de symétrie discrète est peut-être l'indice d'une flèche microscopique du temps. De plus, dans le cadre des théories de grande unification, et avec le rapprochement de la physique des particules et de la cosmologie, il a été envisagé, en suivant une idée proposée par A. Sakharov, que la brisure de la symétrie *CP* fût à l'origine du léger déséquilibre matière/antimatière dans l'univers primordial, nécessaire à la dominance, dans l'univers actuel, de la matière sur l'antimatière. La connexion spin-statistique

Connexion spin-statistique

Pour terminer cette revue des implications de la causalité sur le cadre général de la théorie quantique des champs, nous mentionnons l'important théorème de la *connexion spin-statistique*. Selon ce théorème, dont les seules hypothèses sont la relativité, la mécanique quantique et la causalité, les fermions, les particules de matière qui obéissent au principe d'exclusion de Pauli, sont des particules de *spin demi-entier*, alors que les bosons, les particules d'interaction qui peuvent se trouver à plusieurs dans le même état quantique, sont des particules de *spin entier ou nul.* Comme le spin d'une particule est son *moment cinétique intrinsèque* (une notion purement quantique, puisque, classiquement on ne voit pas comment le moment cinétique intrinsèque d'une particule ponctuelle peut ne pas être identiquement nul), la conservation du moment cinétique interdit aux particules de matière (des fermions de spin demi-entier) d'être produites ou annihilées autrement qu'associées à leurs antiparticules (qui sont aussi des fermions de spin demi-entier), alors que les bosons peuvent être produits ou annihilés en nombre arbitraire.

L'intégrale de chemins, la mise en œuvre concrète de la théorie quantique des champs

Le programme de la théorie quantique des champs, que nous ne pouvons que résumer d'une manière extrêmement schématique est connu sous le nom de *méthode de l'intégrale de chemins de Feynman.*

Dans une première étape,

avant seconde quantification, il s'agit, pour chaque interaction fondamentale, de déterminer

quels sont les champs de matière participant à cette interaction (masses et nombres quantiques des particules associées à ces champs),

- quels sont les champs d'interaction qui la véhiculent, (masses et nombres quantiques des particules associées à ces champs),
- quels sont les caractéristiques des interactions élémentaires entre ces champs de matière et champs d'interaction (constantes de couplages, lois de conservations).

Toutes les informations concernant cette première étape sont consignées (on pourrait dire « encodées ») dans le *lagrangien* de l'interaction fondamentale concernée. Le lagrangien est une fonctionnelle des champs de matière et d'interaction qui exprime toutes les symétries et lois de conservations supposées pour l'interaction considérée. Avant la seconde quantification, tous les champs intervenant dans le lagrangien sont des champs classiques, et le principe classique de moindre action permet de dériver à partir du lagrangien, les équations classiques du mouvement de ces champs.

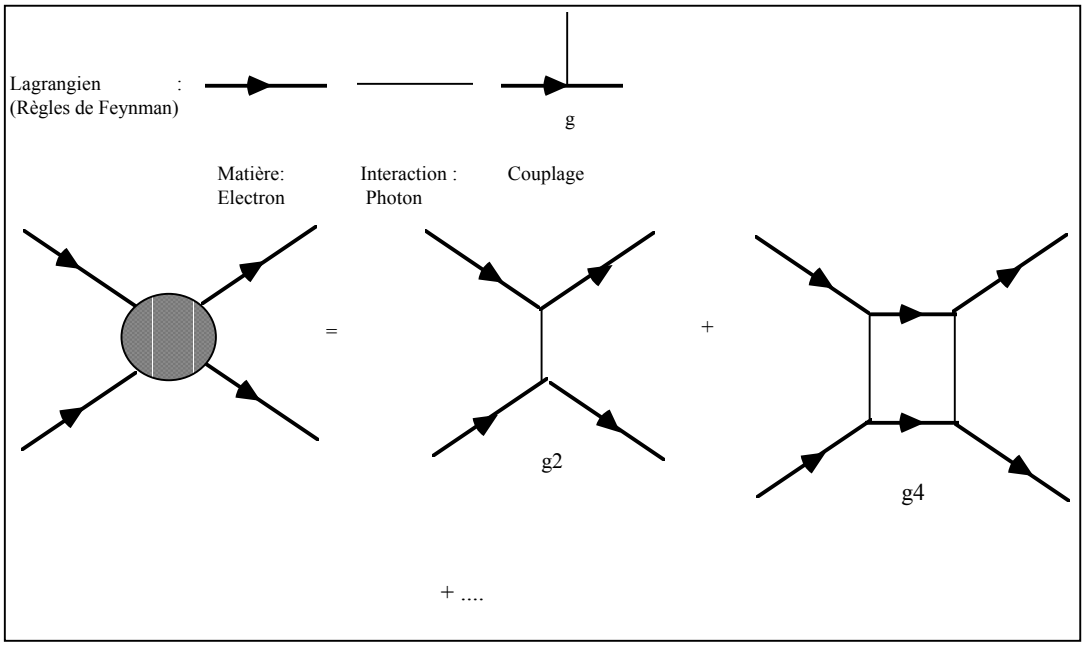
La deuxième étape

est celle de la seconde quantification dont l'intégrale de chemins de Feynman est une méthode particulière (il en existe d'autres que nous n'aborderons pas ici). Très sommairement décrite, cette méthode consiste, pour chaque processus élémentaire relevant de l'interaction fondamentale considérée, à déterminer l'ensemble des voies *quantiquement indiscernables* que peut emprunter le processus considéré, et à déterminer l'amplitude probabilité correspondant à chacune de ces voies. L'amplitude de probabilité totale du processus élémentaire considéré est alors l'intégrale fonctionnelle de toutes ces amplitudes de voies indiscernables. Le carré du module de cette amplitude totale est la probabilité du processus mesurable par des expériences répétables.

Même sommairement décrit, le programme de l'intégrale de chemins semblera, à juste titre, d'une extraordinaire difficulté : les voies indiscernables forment une infinité continue, et les intégrales fonctionnelles ne sont jamais que des intégrales à une infinité continue de variables ! Bien que dans quelques cas, malheureusement trop rares, il soit possible d'effectuer analytiquement ces intégrales fonctionnelles, ou tout au moins de déduire à partir de raisonnements généraux certaines propriétés des amplitudes qui s'expriment à partir d'elles, il en général nécessaire d'avoir recours à des méthodes d'approximation pour évaluer les amplitudes de probabilité des processus élémentaires.

La troisième étape

est donc celle de la mise en œuvre d'une telle méthode d'approximation, *la méthode des perturbations*. Le langage dans lequel s'exprime cette méthode est celui des *diagrammes de Feynman.* Un diagramme de Feynman est une représentation diagrammatique de l'amplitude de probabilité associée à une voie indiscernable que peut emprunter virtuellement un processus élémentaire. Les éléments constitutifs des diagrammes de Feynman sont les *propagateurs* symbolisant la propagation des particules et des *vertex* symbolisant les interactions élémentaires. Le programme de l'intégrale de chemins conduit donc, pour chaque processus relevant d'une interaction fondamentale dont le lagrangien contient les règles de Feynman élémentaires (définitions de tous les propagateurs et vertex possibles), à dessiner tous les diagrammes de Feynman possibles, à calculer l'amplitude de probabilité associée à chaque diagramme et à sommer toutes ces amplitudes pour obtenir l'amplitude totale. Les règles de Feynman (voir la figure)



permettent de calculer l'amplitude de probabilité représentée par chaque diagramme. L'une de ces règles indique que l'amplitude de probabilité représentée par un diagramme comportant *n* vertex d'interactions élémentaires, est proportionnelle à la constante de couplage élevée à la puissance *n*. Supposons alors que la constante de couplage de l'interaction considérée soit petite. Il est clair que plus un diagramme est compliqué, plus il comporte de vertex d'interactions, plus la puissance de la constante de couplage en facteur de l'amplitude de probabilité est élevée, plus est négligeable sa contribution à l'amplitude totale. Telle est la signification de la méthode des perturbations : à l'aide des contributions des quelques diagrammes les plus simples, il est possible, pour peu que la constante de couplage soit petite, d'obtenir une bonne approximation des amplitudes de probabilité des processus élémentaires.

Une remarque s'impose d'emblée à propos de cette méthode : on ne comprend pas bien ce qu'elle peut signifier si la constante de couplage n'est pas un nombre sans dimension. Si en effet la constante de couplage est dimensionnée, on peut changer sa valeur par changement d'unités, et alors quelle fiabilité accorder à des calculs qui ne sont valables que pour certains choix d'unités ? Au moment où a été mise au point la méthode des perturbations, seule l'interaction électromagnétique semblait satisfaire les conditions permettant son application. La constante de couplage de l'interaction électromagnétique est le carré de la charge électrique de l'électron, qui a le contenu dimensionnel du produit d'une action par une vitesse. Comme en physique quantique et relativiste, le quantum d'action et la vitesse de la lumière sont des constantes universelles que l'on peut poser à 1, on peut considérer la

constante de couplage de l'interaction électromagnétique comme un nombre sans dimension. De plus, il se trouve que cette constante est petite, elle vaut 1/137.

Comment se présente la situation des autres interactions fondamentales du point de vue de l'application de la méthode des perturbations ? Comme son nom l'indique, l'interaction forte a une grande constante de couplage, et elle semble donc peu disposée à se laisser traiter par la méthode des perturbations. L'interaction faible a bien une petite constante de couplage, mais ce n'est pas un nombre sans dimension.

Le paradigme de l'électrodynamique quantique (QED)

Les diagrammes de Feynman

Dans l'intégrale de chemins, toute l'information concernant une interaction fondamentale est encodée dans le lagrangien : les termes d'énergie cinétique correspondent à la propagation des champs et les termes d'énergie potentielle à leurs couplages. Propagation et couplages se retrouvent dans les éléments de base des diagrammes de Feynman, les propagateurs et les vertex. Le lagrangien de l'électrodynamique s'écrit :

$$\mathfrak{L}_{QED} = \overline{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - e\overline{\psi}\gamma^{\mu}A_{\mu}\psi$$

où le premier terme est le lagrangien de Dirac décrivant la propagation du champ spinoriel ψ, (un champ électronique par exemple), le second le lagrangien de Maxwell donnant l'énergie cinétique du champ électromagnétique $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$, où le dernier terme représente le couplage du courant porté par le champ spinoriel au quadrivecteur potentiel A_{μ}, et où la charge électrique *e* de l'électron mesure l'intensité de l'interaction électromagnétique au niveau élémentaire.

Au moyen des diagrammes de Feynman, le programme de l'intégrale de chemins prend une forme relativement simple dans son principe : il s'agit, pour toute réaction en QED, de dessiner tous les digrammes de Feynman possibles (en réalité, comme ces diagrammes sont en nombre infini, il s'agit plutôt de les caractériser de manière topologique) en les classant par le nombre de vertex qu'ils comportent. L'amplitude associée à un digramme de Feynman comportant *N* vertex est proportionnelle à e^N. Comme *e* est un nombre petit (la constante de structure fine, $\alpha = e^2/\hbar c$, est un nombre sans dimension qui vaut environ 1/137), le développement en perturbation a des chances d'être efficace en QED : la contribution des digrammes les plus complexes, qui comportent le plus de vertex, et qui

correspondent aux ordres élevés de ce développement ont des chances d'être rapidement négligeables.

Les diagrammes d'ordre le plus bas sont des digrammes « en arbre », c'est-à-dire qu'ils ne comportent pas de boucles ; ils représentent ce que l'on appelle le terme de Born, qui n'est rien d'autre que l'approximation classique. Les effets quantiques apparaissent avec les digrammes qui comportent des boucles, dont l'amplitude s'obtient par intégration sur les quadrimoments arbitraires qui circulent le long de ces boucles. Malheureusement, en général, ces intégrales divergent, et cette difficulté menace l'ensemble du développement en perturbations.

Le miracle de la renormalisabilité

Le premier réflexe face à une telle difficulté est de considérer les contributions qui font diverger les intégrales sont celles de processus hautement virtuels faisant intervenir des énergies arbitrairement élevées et de négliger purement et simplement ces contributions en supprimant des intégrales les contributions venant d'énergies supérieures à un paramètre de coupure Λ (on dit alors que l'on a *régularisé* théorie). Il se trouve qu'en QED les divergences rencontrées ne sont pas trop violentes ; qu'elles ne sont que logarithmiques, si bien que même avec un paramètre de coupure Λ proche de l'énergie de Planck, soit 10^{19} GeV, les corrections au terme de Born sont relativement petites. Il n'en reste pas moins que, le paramètre de coupure étant arbitraire, l'ensemble de la théorie devient suspecte.

La technique de la *renormalisation* qui a permis de contourner l'obstacle, consiste à dédoubler les paramètres de la théorie, la charge électrique et la masse de l'électron en distinguant d'une part une charge e_0 et une masse m_0 dites *nues*, qui seraient la charge et la masse s'il n'y avait pas d'interaction et d'autre part la charge e et la masse m physiques ou *renormalisées* par l'interaction : on s'attend en effet à ce que le couplage du champ électronique avec le champ électromagnétique modifie la charge et la masse de l'électron. La procédure de la renormalisation se poursuit alors de la façon suivante.

- On va calculer le développement en perturbation de e et de m, en fonction de e_0, m_0 et d'un paramètre de coupure arbitraire Λ.
- On va ensuite inverser ces relations, c'est-à-dire exprimer e_0 et m_0 en fonction de e et m :

$$e_0 = e + 1/2\beta_2 e^3 \,\mathrm{Log}\left(\frac{\Lambda}{m}\right) + O(e^5)$$

$$m_0 = m + \gamma_1 m e^2 \,\mathrm{Log}\left(\frac{\Lambda}{\mathrm{m}}\right) + O(e^4)$$

Survient alors le « miracle » de la *renormalisabilité* : soit F l'amplitude d'un processus physique quelconque de QED. Si, dans le développement en perturbation de F on élimine, à l'aide des équations (5), e_0 et m_0 au profit de e et de m, alors, la limite $\Lambda \to \infty$ de l'expression ainsi obtenue *est finie*, et il en est ainsi pour toutes les amplitudes physiques et à tous les ordres du développement en perturbation. Des théories qui partagent cette propriété avec QED, on dit qu'elles sont *renormalisables*. Il faut bien voir que cette circonstance permet de sauver l'ensemble du programme de l'intégrale de chemins : au moyen de deux paramètres physiques, expérimentalement mesurables, il devient possible d'exprimer l'ensemble des observables physiques sans qu'il soit besoin d'introduire aucun paramètre de coupure arbitraire. De fait, comme l'électrodynamique quantique n'intervient pas seulement à très haute énergie mais aussi à basse énergie, dans un domaine où des expériences de très haute précision ont pu être menées dès les années quarante, les prédictions permises par la renormalisabilité ont pu être comparées avec l'expérience avec des succès qui ont dépassé toutes les espérances.

Le groupe de renormalisation

Nous remettons à plus tard la discussion de la signification physique de la procédure de la renormalisation qui n'a pas manqué de susciter nombre d'interrogations. En attendant nous voulons aborder une circonstance qui, en quelque sorte relativise le « miracle » de la renormalisabilité. Si, dans le lagrangien de QED, on met à zéro la masse de l'électron, on obtient une expression dans laquelle aucun paramètre n'est dimensionné. Or, avons-nous dit, les divergences dans les intégrales de boucles sont logarithmiques, si bien que pour régulariser la théorie, il est nécessaire d'introduire, outre le paramètre de coupure Λ, un autre paramètre ayant la dimension d'une masse ou d'une énergie, que l'on appelle l'énergie de renormalisation μ. Ainsi la première des équations (5) devient

$$e_0 = e + 1/2\beta_2 e^3 \,\mathrm{Log}\left(\frac{\Lambda}{\mu}\right) + O(e^5)$$

si bien que la charge renormalisée e n'est plus une constante mais une fonction de μ. On parlera alors de la charge renormalisée *à l'énergie de renormalisation* μ. Or cette énergie de renormalisation est arbitraire, et la physique ne devrait pas en dépendre. Cela signifie que la charge renormalisée $e(\mu)$ doit obéir à une équation différentielle qui exprime l'invariance de la physique par changement de l'énergie de renormalisation, ce que l'on appelle équation du groupe de renormalisation :

$$\mu \frac{de^2(\mu)}{d\mu} = \beta\left(e^2(\mu)\right)$$

où la fonction β peut être obtenue à l'aide d'un développement en perturbation indépendant de tout paramètre de coupure

$$\beta(e^2) = \beta_2 e^4 + O(e^6)$$

où β_2 est un nombre pur, positif en QED.

Lorsque la masse de l'électron n'est pas nulle, la dépendance des paramètres physiques dans l'énergie de renormalisation est présente, bien que masquée parce qu'elle a en fait été implicitement posée égale à la masse de l'électron avec

$$e(\mu = m) = e$$
$$m(\mu = m) = m$$

L'invariance par le groupe de renormalisation relativise le miracle de la renormalisabilité : les paramètres physiques comme la charge et la masse de l'électron ne sont plus des paramètres fondamentaux, mais plutôt des paramètres *effectifs* qui dépendent, certes d'une façon contrainte par les équations du groupe de renormalisation, d'une échelle arbitraire d'énergie. Il nous faut donc admettre que la constante de structure fine $\alpha = e^2/\hbar c$ n'est plus une véritable constante ; ce qui vaut 1/137 c'est la « constante » de structure fine évaluée à une énergie de renormalisation égale à la masse de l'électron ; à la masse du boson Z, cette « constante » vaut environ 1/128.

Les équations du groupe de renormalisation contiennent finalement toute l'information concernant l'effet des fluctuations quantiques. Le vide de la théorie quantique des champs doit être considéré comme un milieu, siège des fluctuations quantiques (créations et annihilations de paires virtuelles particule/antiparticule). Dans le cas de QED, ces fluctuations polarisent le milieu qui devient une sorte de diélectrique, dont la constante diélectrique dépend de la résolution (l'inverse de l'énergie de renormalisation) avec laquelle il est observé. Le fait que le coefficient β_2 soit positif implique que les paires virtuelles

électron/positon « écrantent » la charge nue, si bien que la charge renormalisée croît quand la distance diminue, c'est-à-dire quand l'énergie croît. Cette circonstance est d'ailleurs le signe d'une grave pathologie de QED, à savoir l'existence dans le propagateur du photon, d'un pôle à une valeur très élevée de l'énergie, avec un résidu négatif de telle sorte qu'il n'est susceptible d'aucune interprétation physique (c'est pourquoi on l'appelle le « fantôme » de Landau du nom du physicien qui a découvert cet effet).

L'invariance de jauge

Quels que soient les problèmes et interrogations que suscite la renormalisabilité, il est extrêmement important de comprendre les caractéristiques de QED qui en sont à l'origine si l'on veut construire pour les autres interactions des théories qui soient elles aussi renormalisables. Il semble bien que l'*invariance de jauge* soit précisément une telle caractéristique de QED susceptible de généralisation à d'autres interactions.

En électromagnétisme classique, l'invariance de jauge traduit le fait que deux potentiels qui diffèrent par la quadridivergence d'une fonction arbitraire donnent lieu au même champ électromagnétique. Si on ajoute au potentiel une quadridivergence, le terme de couplage du courant au potentiel est modifié par l'ajout d'une quadridivergence (ce qui ne modifie pas les équations du mouvement) *si et seulement si le courant est conservé.* L'invariance de jauge est donc équivalente à la conservation du courant.

En QED, le courant est exprimé au moyen du champ spinoriel ψ, et le lagrangien est invariant par une propriété de symétrie spécifiquement quantique, l'invariance par changement de la phase du champ spinoriel. Dans le lagrangien de QED, cette propriété de symétrie a la particularité d'être une invariance *locale*, c'est-à-dire que le changement de phase dépend du point d'espace-temps où on l'applique. Si on réécrit le lagrangien de QED

$$\mathfrak{L}_{QED} = \mathfrak{L}_{COV} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$
$$\mathfrak{L}_{COV} = \overline{\psi(x)} \left[i\gamma^{\mu} D_{\mu} - m \right] \psi(x)$$

où $D_{\mu} = \partial_{\mu} + ieA_{\mu}$ est appelée « dérivée covariante » par analogie à la dérivée covariante introduite en relativité générale, on constate qu'il reste invariant par l'ensemble des transformations suivantes opérées sur les champs

$$\psi(x) \to e^{-i\alpha(x)} \psi(x)$$
$$\overline{\psi(x)} \to e^{+i\alpha(x)} \overline{\psi(x)}$$
$$A_{\mu} \to A_{\mu} + \partial_{\mu} \Lambda(x)$$

à condition que $\alpha(x) = e\Lambda(x)$. L'invariance par les transformations (11) est ce que l'on appelle désormais *l'invariance de jauge* de l'électrodynamique quantique, quant au potentiel A_μ, on l'appelle désormais le *champ de jauge* de l'interaction. A partir de l'invariance de jauge on aurait pu d'ailleurs aboutir au lagrangien de QED en suivant un autre chemin : on serait parti du lagrangien de Dirac du champ spinoriel et on aurait constaté son invariance par un changement global de phase, et on aurait imposé une invariance de phase *locale* en introduisant un champ de jauge se transformant comme le potentiel de l'interaction électromagnétique. Ainsi il apparaît que l'invariance de jauge est un principe qui détermine la forme même de l'interaction. Une telle situation n'est pas sans analogie avec ce qui se produit en relativité générale : de l'équivalence locale entre un changement quelconque de référentiel et l'action de la gravitation découle la forme même de la nouvelle théorie, géométrique, ou plutôt « géométrodynamique », de la gravitation universelle. Il devient dès lors très tentant d'ériger l'invariance de jauge en un principe directeur applicable à toutes les interactions fondamentales. Et ce, d'autant plus qu'il est apparu mathématiquement que c'est l'invariance de jauge qui garantit la renormalisabilité à tous les ordres du développement en perturbations. L'élaboration du MS s'est donc poursuivie par la recherche de théories à invariance de jauge pour l'interaction forte et pour l'interaction faible.

La chromodynamique quantique (QCD)

Le modèle des quarks-partons

On sait qu'un tournant décisif est intervenu en physique des particules lorsque l'on a mis en évidence un nouveau niveau d'élémentarité, celui des quarks, les constituants élémentaires des hadrons, les particules qui participent à toutes les interactions, interaction forte comprise. Alors que la théorie quantique des champs était en difficulté face à la complexité des interactions entre hadrons, avec les quarks, on dispose de particules somme toute assez analogues à l'électron et autres leptons, que l'on pourra considérer comme les excitations de champs quantiques en interaction locale.

La structure des hadrons a été mise en évidence dans les expériences de collisions profondément inélastiques lepton-hadron. Il s'agit d'expériences dites « inclusives », dans lesquelles un lepton de haute énergie entre en collision avec un hadron cible (un proton ou un neutron dans un noyau) et produit un état final dans lequel seul le lepton final est détecté :

$$l + p \rightarrow l' + X$$

la région cinématique des collisions profondément inélastiques est celle où le transfert d'énergie entre le lepton initial et le lepton final dans le système du laboratoire $\nu = E - E'$, ainsi que la valeur absolue du carré de Lorentz du quadrimoment transféré entre ces deux leptons $Q^2 = \left|(l - l')^2\right|$ sont grands. La variable de Bjorken $x = Q^2 / 2M\nu$ (où M est la masse du hadron cible) est comprise entre 0 et 1. Comme les leptons ne participent pas à l'interaction forte, de telles expériences permettent de sonder une éventuelle structure des hadrons dont les constituants seraient élémentaires dans les interactions électromagnétique et faible. On donne le nom générique de *partons* à de tels hypothétiques constituants élémentaires des hadrons. D'après l'hypothèse fondamentale du modèle des partons, dite hypothèse de *liberté asymptotique*, dans un repère de grand moment, le hadron peut être modélisé à l'aide d'une collection d'un nombre fini de partons colinéaires et quasi libres. La collision lepton hadron peut être visualisée dans le repère de Breit, ou repère du « mur de briques » dans lequel le lepton initial et le lepton final ont la même énergie et des tri-impulsions égales et opposées $\pm Q/2$ et où le hadron a une grande impulsion égale à p. D'après l'hypothèse de liberté asymptotique, le lepton entrerait en collision avec un seul parton a portant une fraction de l'impulsion du hadron précisément égale à la variable de Bjorken x

$$p_a = x_a p = Q/2$$
$$x_a = Q/2p = x$$

et la section efficace de la collision lepton hadron serait le produit de la probabilité que le parton actif a porte une fraction d'impulsion égale à la variable de Bjorken par la section efficace de la réaction électromagnétique ou faible de la réaction lepton parton :

$$\sigma(l + p \to l' + X) = \Pr\{x_a = x\}\sigma(a + l \to a' + l')$$

La conséquence la plus significative de l'hypothèse du modèle des partons est que la probabilité qui intervient dans cette équation est une fonction dite *fonction de structure*, *invariante d'échelle*, i.e. indépendante de Q, décrivant de façon universelle le spectre d'impulsion des partons à l'intérieur du hadron. L'observation, dès la fin des années soixante de l'invariance d'échelle approximative des fonctions de structure pour des valeurs de Q à peine supérieures à 1 GeV/c a validé l'hypothèse des partons. La section efficace du processus électromagnétique ou faible dépend de Q et aussi des caractéristiques (charges électriques ou faibles) des partons impliqués, si bien que l'analyse des données expérimentales des collisions profondément inélastiques donne accès aux caractéristiques des partons. De l'ensemble des

données obtenues avec des faisceaux d'électrons, de muons et de neutrinos, on a pu induire que les quarks, qui avaient été introduits pour rendre compte de la classification des hadrons, sont des candidats à être des partons. On a ainsi trouvé que des constituants des hadrons sont d'une part les quarks de *valence*, ceux qui donnent au hadron ses nombres quantiques (par exemple deux quarks *u* et un quark *d* pour le proton) et d'autre part, comme on s'y attend en théorie quantique des champs, des paires quark-antiquark qui forment ce que l'on appelle une *mer de Fermi*. Cela étant, il est apparu que les quarks et antiquarks n'épuisent pas tout le contenu en partons des hadrons, et qu'il existe vraisemblablement d'autres partons, neutres dans les interactions électromagnétique et faible qui portent une fraction importante de l'impulsion des hadrons. On a appelé *gluons* ces hypothétiques partons supplémentaires, pour suggérer qu'il s'agit des particules responsables de la cohésion du hadron, qui « collent » ses constituants.

La chromodynamique quantique comme théorie de jauge non abélienne

Ayant identifié les partons, restait à mettre sur pied la théorie quantique des champs dans laquelle les partons sont les excitations des champs quantiques en interaction, une théorie qui, de préférence, fût renormalisable. On s'est donc tourné vers des théories à invariance de jauge, en essayant de généraliser la propriété de symétrie qui garantit la renormalisabilité de QED. La généralisation la plus naturelle consiste à considérer un groupe de symétrie plus général que le groupe commutatif $U(1)$ de la multiplication par une phase. Dès les années cinquante, Yang et Mills avaient introduit les théories de jauge non abéliennes, dans lesquelles le groupe de jauge est un groupe non commutatif ou non abélien. Dans le cadre du modèle des quarks, le schéma de la « couleur » avait été introduit par Greenberg pour donner aux quarks, particules de spin ½, la statistique de Fermi-Dirac : comme il faut trois quarks identiques, dans le même état, pour former un fermion de spin 3/2 comme le Ω^-, ou le Δ^{++} ou le Δ^-, il a eu l'idée de donner aux quarks un nouveau degré de liberté, qu'il a appelé la couleur, lui permettant d'antisymétriser la fonction d'onde du système de trois quarks, et il a postulé l'invariance par les transformations du groupe $SU(3)$ de couleur. Les quarks appartiennent à la représentation fondamentale de ce groupe, les hadrons appartiennent à la représentation triviale, et ceci est à l'origine de la terminologie de la couleur : les quarks ont les trois « couleurs » fondamentales qui se combinent pour donner des hadrons « blancs », sans couleurs. La chromodynamique quantique (le préfixe « chromo » est une référence à la terminologie de la couleur) est une théorie à la Yang et Mills dans laquelle les champs colorés

de quarks sont les champs de matière et le groupe de jauge, le groupe *SU*(3) de couleur. La symétrie de couleur devient une invariance locale. Le formalisme de la dérivée covariante utilisé en QED – voir l'équation (10) – se généralise de façon naturelle à une invariance de jauge non abélienne. Les générateurs infinitésimaux $T^a (a = 1,...,8)$ du groupe de couleur sont au nombre de huit ; ils appartiennent à la représentation adjointe du groupe :

$$\left[T^a, T^b\right] = if^{ab}{}_c T^c$$

où les $f^{ab}{}_c$ sont les *constantes de structure* du groupe. Sur les quarks $q^i (i = 1,...,3)$ l'action des générateurs est représentée par les matrices de Gell-Mann

$$\left(T^a q\right)^i = \sum_j \left(\frac{\lambda^a}{2}\right)^{ij} q^j$$

En partant du lagrangien de Dirac des quarks

$$\mathfrak{L}_q = \sum_{i=1}^{3} \overline{q_i(x)} \left(i\gamma^\mu \partial_\mu - m\right) q^i(x)$$

on peut imposer une invariance locale de couleur (g_s est une constante de couplage analogue à la charge électrique)

$$q^i(x) \to \left[\exp\left\{ig_s \sum_{a=1}^{8} T^a \Lambda^a(x)\right\}\right]^{ij} q^j(x)$$

en remplaçant dans le lagrangien la dérivée ordinaire par la dérivée covariante

$$\left(D_\mu\right)^{ij} = \partial_\mu \delta^{ij} + ig_s \sum_{a=1}^{8} \left(T^a\right)^{ij} A_\mu^a(x)$$
$$\mathbf{D}_\mu \equiv \partial_\mu + ig_s \mathbf{T}.\mathbf{A}_\mu$$

qui fait intervenir les huit champs de jauge $A_\mu^a(x)\ (a = 1,...,8)$, dont les quanta sont identifiés aux gluons, et à partir desquels on obtient par dérivation covariante les tenseurs chromodynamiques

$$\mathbf{F}_{\mu\nu} \equiv F_{\mu\nu}^a \ (a = 1,...,8)$$
$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g_s f^a{}_{bc} A_\mu^b A_\nu^c$$

Avec des notations évidentes qui rendent implicite la sommation sur les indices de couleur, on obtient alors le lagrangien de la chromodynamique quantique :

$$\mathfrak{L}_{QCD} = -\frac{1}{4}\mathbf{F}_{\mu\nu}.\mathbf{F}^{\mu\nu} + \sum_{j=1}^{N_f} \overline{\mathbf{q}_j} \left(i\mathbf{D}_\mu \gamma^\mu - m_j\right) \mathbf{q_j}$$

où $j = 1,...,N_f$ représente le nombre quantique de « saveur » (*u*, *d*, *s*, etc.) des quarks.

On ne peut qu'être frappé par l'analogie formelle entre ce lagrangien et celui de QED, une analogie que l'on retrouve dans l'expression des transformations de jauge qui laissent le lagrangien invariant

$$\mathbf{q_j}(x) \to \exp\{ig_s \mathbf{T}.\mathbf{\Lambda}(x)\}\mathbf{q_j}$$
$$A^a_\mu \to A^a_\mu - g_s f^a{}_{bc}\Lambda^b(x)A^c_\mu + \partial_\mu \Lambda^a$$

Renormalisation et liberté asymptotique

Toutefois ces analogies ne doivent masquer les différences essentielles entre QCD et QED. La présence dans l'expression des tenseurs chromodynamiques d'un terme quadratique dans les champs de jauge induit dans le lagrangien de QCD des couplages à trois et à quatre gluons, qui n'ont pas d'équivalent en QED. La nouveauté essentielle introduite par le caractère non abélien du groupe de jauge est le fait que les champs de jauge sont maintenant des champs porteurs de charge, qui sont donc autocouplés. Les règles de Feynman de QCD traduisent cette nouveauté : aux propagateurs des quarks et des gluons et au couplage du gluon au courant de quark analogues aux termes présents en QED, il convient d'ajouter les couplages à trois et à quatre gluons. Au début des années soixante-dix on a prouvé la renormalisabilité des théories de jauge non abélienne, et donc celle de QCD. L'auto couplage des gluons a des conséquences spectaculaires sur le groupe de renormalisation de QCD : les boucles de quarks produisent un effet d'écran analogue à celui des boucles d'électrons en QED, mais les boucles de gluons (qui n'ont pas d'équivalent en QED) produisent un effet d' « anti écran » qui est dominant par rapport à l'effet d'écran des boucles de quarks si le nombre de saveurs n'est pas trop grand (en fait si $N_f < 17$). Si bien que la charge forte renormalisée est une fonction décroissante de l'énergie de renormalisation, qui s'annule lorsque cette énergie tend vers l'infini. Si, par analogie avec la constante de structure fine, on définit $\alpha_s = g_s{}^2/\hbar c$, on trouve, à l'ordre d'une boucle, l'évolution suivante pour l'intensité de l'interaction au niveau élémentaire, quand on varie l'échelle de renormalisation de μ_0 à μ :

$$\alpha_s(\mu^2) = \frac{\alpha_s(\mu_0{}^2)}{1 + \frac{\alpha_s(\mu_0{}^2)}{12\pi}(33 - 2N_f)\mathrm{Log}\left(\frac{\mu^2}{\mu_0{}^2}\right)}$$

ce que l'on peut réécrire

$$\alpha_s(\mu^2) = \frac{1}{b \,\mathrm{Log}\dfrac{\mu^2}{\Lambda^2}}$$

$$b = \frac{33 - 2N_f}{12\pi}$$

de façon à faire apparaître le pôle de Landau de QCD à $\mu = \Lambda$. Le régime perturbatif de QCD s'établit à des énergies supérieures à Λ. Le fait que l'on ait observé une invariance d'échelle des fonctions de structure à une valeur relativement basse de Q suggère que ce régime perturbatif de QCD, avec une valeur faible de Λ (une centaine de MeV) est bien le cadre théorique sous jacent au modèle des partons : lorsque l'on renormalise QCD à une énergie $\mu = Q$, on retrouve bien la liberté asymptotique nécessaire pour justifier le modèle des partons. Mais le fait que l'intensité de l'interaction dépende de Q implique qu'il faille s'attendre à des violations de l'invariance d'échelle, et de fait, les équations du groupe de renormalisation appliquées aux fonctions de structure par Dokshitzer, Gribov, Lipatov, Altarelli et Parisi (équations DGLAP) conduisent à une telle prédiction qui s'est trouvée confirmée par l'expérience avec une excellente précision.

QCD sur réseau

Le comportement pathologique de QCD à basse énergie empêche toute utilisation de la méthode des perturbations dans la physique hadronique à grande distance, mais on espère cependant que c'est ce comportement qui est à l'origine du *confinement* des partons à l'intérieur des hadrons, c'est-à-dire de l'impossibilité d'observer à l'état libre un quark ou un gluon. Il existe bien plusieurs modèles phénoménologiques pour rendre compte de cette propriété (modèle des sacs, modèle des cordes hadroniques), ainsi que de puissants outils de simulation rendant compte de la transformation des partons en jets de hadrons, mais, jusqu'à présent, le confinement représente une frontière du MS, celle du « non perturbatif ».

Pour explorer ce domaine, on exploite l'analogie qui existe entre la théorie quantique des champs et la physique statistique. Si l'on passe en effet dans l'euclidien, c'est-à-dire en temps imaginaire, on s'aperçoit que l'intégrale de chemins est formellement identique à la moyenne statistique sur les configurations d'un système statistique : le lagrangien est remplacé par le hamiltonien et l'inverse de la température par le temps imaginaire. Les termes de cette correspondance se comprennent bien si on considère, comme système statistique, un ensemble de spins situés sur les nœuds d'un réseau. L'approximation du champ moyen consiste alors à remplacer chaque spin par sa valeur moyenne et à traiter en perturbations les

fluctuations autour de cette valeur moyenne. On conçoit bien qu'une telle méthodologie transforme le problème de physique statistique en un développement en perturbations pour une théorie (euclidienne) de champs quantiques en interaction locale si, dans le système statistique, l'interaction est entre plus proches voisins. Mais dans le cas d'un phénomène critique, c'est-à-dire au voisinage d'une transition de phase du second ordre, où le système présente des fluctuations importantes à toutes les échelles, cette méthodologie est en échec, tout comme, en théorie quantique des champs, la méthode des perturbations est mise en échec par les divergences dans les intégrales de boucles. Il se trouve que la méthode qui a permis de se sortir de cette difficulté en physique statistique est formellement identique à la procédure de la renormalisation dans le cadre d'une théorie des champs renormalisable. Selon cette méthode (dite méthode des « blocs de spins »), introduite par Kadanof, on effectue la moyenne sur les configurations du système par récurrence, échelle après échelle. Par exemple, partant d'un réseau carré de maille *a,* avec une énergie $\mathbf{H}_a(S)$, on moyenne les spins sur chaque carré élémentaire et on considère un réseau de maille *2a* avec une énergie effective $\mathbf{H}_{2a}(S)$. On itère l'opération qui fait passer d'un réseau de maille $2^{n-1}a$ au réseau de taille $2^n a$, que l'on appelle une transformation du groupe de renormalisation

$$\mathbf{H}_{2^n a}(S) = \mathfrak{T}(\mathbf{H}_{2^{n-1} a}(S)).$$

Pour rendre compte du fait qu'au voisinage d'une transition de phase du second ordre, le système devient *fractal*, c'est-à-dire qu'il reste similaire à lui-même quand on change la résolution (ou niveau d'agraindissement) avec laquelle on l'observe, on va rechercher un éventuel point fixe asymptotique de la transformation du groupe de renormalisation,

$$\mathbf{H}_{2^n a}(S) \underset{n\to\infty}{\to} \mathbf{H}^*(S)$$
$$\mathbf{H}^*(S) = \mathfrak{T}(\mathbf{H}^*(S))$$

dont l'état ne dépend plus de la maille du réseau, de la même façon que la physique ne dépend plus de l'échelle de coupure Λ dans une théorie renormalisable. Comme a pu le montrer Kenneth Wilson, il y a beaucoup plus qu'une simple analogie entre les invariances par le groupe de renormalisation en physique des phénomènes critiques et en physique des particules, mais plutôt la porte ouverte à une véritable unification de ces deux grands domaines de la physique. Un groupe de renormalisation définit une *classe d'universalité* à laquelle peuvent appartenir des théories de phénomènes critiques aussi bien que des théories renormalisables en physique des particules. Cette unification requérait un cadre précis que Wilson mit en place et qui fut complété par les travaux de Brézin, Le Guillou et Zinn-Justin

en 1973. Il devenait alors possible de modéliser des phénomènes critiques à l'aide de la théorie quantique des champs, et de faire des prédictions précises concernant des observables caractéristiques des phénomènes critiques. Des expériences menées, en particulier par Pierre Bergé, apportèrent des confirmations éclatantes de ces prédictions. Réciproquement, il devenait aussi possible de modéliser une théorie renormalisable comme QCD à l'aide d'un système de spins sur un réseau comportant une transition de phase du second ordre et appartenant à la même classe d'universalité. Pour de tels modèles, il existe des algorithmes informatiques efficaces permettant d'effectuer la moyenne sur toutes les configurations (algorithmes de Monte Carlo) sans avoir à faire appel à la méthode des perturbations, si bien qu'il est devenu possible d'effectuer d'authentiques *expériences informatiques* permettant d'explorer le domaine non perturbatif de QCD. Les résultats obtenus le long de cette voie (qui n'est limitée que par la puissance des ordinateurs) sont encourageants : on a obtenu quelques lumières sur le mécanisme de confinement des quarks et des gluons, on a été capable de calculer les masses de quelques hadrons (en particulier de hadrons appelés « boules de glu », purs états liés de gluons) et on a pu prédire l'existence de la transition de déconfinement.

La théorie électrofaible

Structure des courants électromagnétiques et faibles

Une fois la structure en quarks des hadrons mise en évidence, la phénoménologie des interactions faibles impliquant des hadrons et des leptons s'est attachée à comprendre la structure des courants faibles par analogie avec celle des courants électromagnétiques. Dans les années soixante, l'interaction faible était en effet décrite à l'aide du modèle phénoménologique de Fermi d'interaction courant-courant, dont le lagrangien s'écrit :

$$\mathfrak{L}_{Fermi} = -\frac{G}{\sqrt{2}} J_\lambda J^{\lambda\dagger}$$

$$G = (1.026 \pm 0.001) 10^{-5} m_p^{-2}$$

où le courant faible est la somme d'une partie leptonique

$$J_\lambda^{lept} = \overline{\nu_\mu} \gamma_\lambda (1-\gamma_5) \mu + \overline{\nu_e} \gamma_\lambda (1-\gamma_5) e + h.c.$$

et d'une partie hadronique, qui, à l'époque où l'on ne connaissait que les trois quarks *u*, *d* et *s* se réduisait à

$$J_\lambda^{had} = \overline{u} \gamma_\lambda (1-\gamma_5) (\cos\theta d + \sin\theta s) + h.c.$$

L'angle de mélange θ, introduit par Cabibbo, permet d'unifier les désintégrations faibles du neutron et celles des particules étranges.

Une des propriétés les plus caractéristiques de l'interaction faible est la violation maximale de l'invariance par parité traduite par l'égalité en valeur absolue des poids du courant vectoriel (terme en γ_μ) et du courant axial (terme en $\gamma_\mu\gamma_5$) dans l'expression des courants faibles. Cette situation est à contraster avec l'invariance par parité des interactions électromagnétique et chromodynamique qui ne font intervenir que des courants purement vectoriels en $\overline{\psi}\gamma_\mu\psi$. Si les leptons et les quarks étaient tous des particules de masse nulle, on pourrait interpréter la violation maximale de l'invariance par parité de l'interaction faible en disant que seules les particules « lévogyres » ou gauches (celles dont l'hélicité est égale à 1/2), et les antiparticules « dextrogyres » ou droites (dont l'hélicité est égale à +1/2), y sont sensibles. A l'inverse, les interactions électromagnétique et chromodynamique sont de même intensité pour les particules gauches et droites. Lorsque l'on a recherché une théorie de jauge pour l'interaction faible on a supposé que les fermions participant à cette interaction étaient, en première approximation, de masse nulle pour qu'il soit possible de les classer en gauches et droits.

Les courants faibles neutres et les générations

La théorie de Fermi n'étant pas renormalisable, on s'est tourné, comme pour l'interaction forte, vers une théorie à invariance de jauge pour l'interaction faible. Comme l'électron et son neutrino, le muon et son neutrino, la combinaison de Cabibbo du d et du s et le quark u forment des doublets, analogues aux doublets d'isospin dans l'interaction forte, on est conduit à essayer $SU(2)$ comme groupe de jauge de façon à ce que les fermions porteurs de charge faible appartiennent à la représentation fondamentale d'un groupe d' « isospin faible » que l'on prendra comme groupe d'invariance de jauge. Mais alors on s'attend à ce que, à côté des courants faibles chargés, existent aussi des courants faibles neutres comme par exemple $\overline{\nu_\mu}\gamma_\lambda(1-\gamma_5)\nu_\mu$ ou $\overline{u}\gamma_\lambda(1-\gamma_5)u$. Or de tels courants neutres sont très difficiles à mettre en évidence expérimentalement puisque, pour des particules chargées, l'interaction faible à courant neutre est dominée, de manière écrasante par l'interaction électromagnétique. Seules les interactions faibles à courant neutre des neutrinos sont épargnées par la contamination des interactions électromagnétiques. C'est pourquoi il a fallu attendre les expériences avec les faisceaux de neutrinos du CERN en 1973 (alors que tous les ingrédients théoriques du MS de

l'interaction électrofaible étaient déjà disponibles) pour avoir la confirmation expérimentale de l'existence des interactions faibles à courants neutres.

Un autre problème a surgi avec les courants faibles neutres, celui de l'absence de courants faible neutres avec changement d'étrangeté : on avait constaté en effet que, bien qu'ayant la même charge électrique, le quark étrange *s* et le quark *d* ne semblaient pas pouvoir se transformer l'un en l'autre par interaction faible. C'est alors que Glashow, Iliopoulos et Maiani ont imaginé un mécanisme, le mécanisme « GIM » permettant de résoudre ce problème. Il ont émis l'hypothèse qu'outre le doublet constitué par le quark *u* et la combinaison de Cabibbo du *d* et du *s*, existe un autre doublet constitué d'un nouveau quark de même charge que le *u* et de la combinaison orthogonale à celle de Cabibbo du *d* et du *s*. Le nom de « charme » a été donné au nouveau quark ainsi postulé, pour souligner l'élégance de l'idée théorique à l'origine de ce mécanisme. Le courant neutre hadronique s'écrirait alors :

$$J_\lambda^{N-had} = \overline{N}\gamma_\lambda(1-\gamma_5)N + \overline{N'}\gamma_\lambda(1-\gamma_5)N'$$

où

$$N \equiv \begin{pmatrix} u \\ d_\theta = d\cos\theta + s\sin\theta \end{pmatrix}; N' = \begin{pmatrix} c \\ s_\theta = -d\sin\theta + s\cos\theta \end{pmatrix}$$

où *c* désigne le quark de charme, et où l'orthogonalité des combinaisons assurerait l'annulation des contributions couplant les quarks *s* et *d*. La découverte en 1975 de la particule J/Ψ interprétée comme un « charmonium » (état lié charme-anticharme), puis des particules comportant un quark de charme a marqué un nouveau tournant décisif dans l'histoire de la physique des particules.

En fait, avant même la découverte du charme, il existait un autre argument théorique en faveur de l'existence d'un quatrième quark : on essayait de construire une théorie de jauge pour l'interaction faible, mais les corrections quantiques correspondant à des diagrammes de Feynman comportant une boucle de fermion induisaient une *anomalie*, i.e. une non-conservation des courants qui la mettait en échec. Mais il est apparu que cette anomalie pouvait s'annuler si les contributions des boucles de leptons et celles des boucles de quarks voulaient bien s'équilibrer ; on connaissait deux doublets de leptons et seulement trois quarks ; il fallait un quatrième quark pour former deux doublets et le mécanisme GIM pouvait faire l'affaire.

Une nouvelle surprise est intervenue avec la découverte, presque concomitante avec celle du charme, d'un troisième doublet de leptons, constitué du τ et de son neutrino. Mais

cette découverte faisait resurgir le problème de l'anomalie, à moins qu'il existe un troisième doublet de quarks. On a donné des noms à ces nouveaux quarks hypothétiques : *b* (pour « beauty » ou « bottom ») et *t* (pour « top » ou « truth »), et le *b* a été découvert en 1979 et le *t* en 1995. Il semble donc que fermions participant à l'interaction faible s'organisent en trois familles (on dit *générations*) comportant chacune un doublet de leptons et un doublet de quarks. Dans l'hypothèse où toutes les masses seraient nulles, les interactions faible, mais aussi électromagnétique et chromodynamique, seraient strictement identiques pour chacune de ces trois générations. La signification profonde de cette sorte de bégaiement de la nature reste une des grandes énigmes de la physique contemporaine. En tous les cas, le mécanisme de GIM pour annuler les contributions de courants neutres avec changement de saveur (on désigne ainsi les types de quarks, l'étrangeté, le charme, etc. sont des *saveurs*), peut être généralisé à trois doublets de quarks à l'aide de la *matrice* de Cabibbo, Kobayashi et Maskawa. Les éléments de cette matrice pouvant être complexes, ce mécanisme de mélange peut permettre de rendre compte de la brisure de la symétrie *CP* que nous avons évoquée plus haut, et qui fait l'objet d'autres exposés de notre colloque.

L'invariance de jauge de l'interaction faible

L'étape suivante de la construction de la théorie de jauge de l'interaction faible, celle qui a ouvert la voie à la théorie « électrofaible », a consisté en l'attribution d'isospins différents aux fermions selon leurs chiralités (caractère gauche ou droit) : ainsi tous les doublets d'isospin que nous avons rencontrés jusqu'à présent sont en fait des particules gauches alors que les particules droites de même charge ou de même saveur se voient attribuer un isospin nul. Le groupe de jauge de l'interaction faible est alors le produit du groupe $SU_L(2)$ de l'isospin faible (l'indice L rappelle que ce groupe n'agit que sur les particules gauches –left en anglais–) par le groupe de l'« hypercharge » faible, analogue de l'hypercharge introduite en physique hadronique pour rendre compte de l'étrangeté. La symétrie d'hypercharge est une invariance par un groupe $U_Y(1)$, complètement analogue à l'invariance de jauge de l'interaction électromagnétique. Pour toutes les particules, la charge électrique Q, l'hypercharge Y et la troisième composante de l'isospin faible I_3 sont reliées par la relation de Gell-Mann et Nishijima :

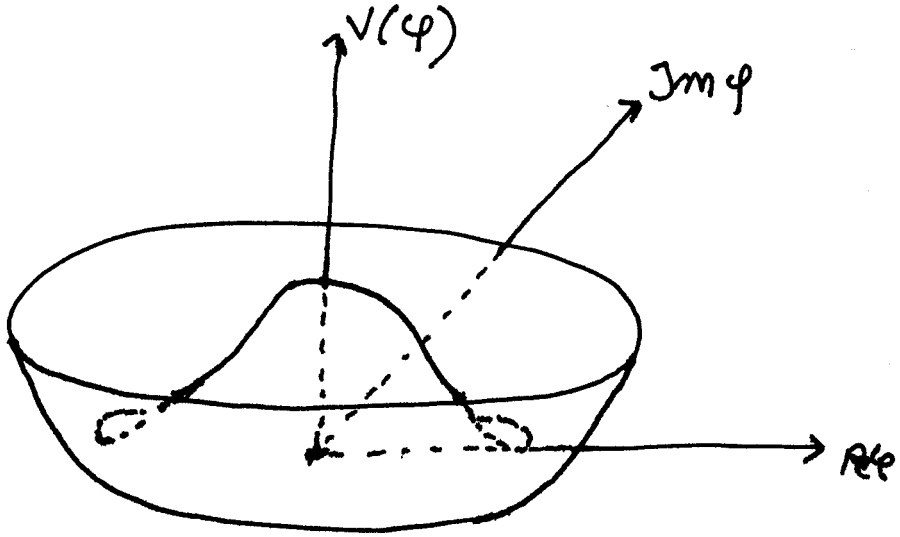
$$Q = I_3 + Y$$

Le formalisme de la dérivée covariante peut être utilisé pour écrire le lagrangien de la théorie de jauge putative de l'interaction faible avec comme groupe de jauge

$SU_L(2)\times U_Y(1)$, comme champs de jauge, un triplet d'isospin $W^i;(i=1,...,3)$ et un singulet B, comme champs de fermions porteurs de charges les doublets gauches et les singulets droits, et avec deux constantes de couplages g pour $SU_L(2)$ et g' pour $U_Y(1)$. Si on continue à supposer nulles les masses des fermions, alors la chiralité de ces fermions est conservée dans l'interaction. Tout terme de masse dans le lagrangien de la forme $m\overline{\psi}\psi = m(\overline{\psi_R}\psi_L + \overline{\psi_L}\psi_R)$ (l'indice R renvoie aux fermions droits –right en anglais–) mélangerait les deux chiralités.

Brisure spontanée de symétrie

Il fallait une bonne dose d'optimisme pour essayer, pour l'interaction faible, une telle théorie de jauge dans laquelle toutes les particules, les bosons de jauge comme tous les fermions sont nécessairement de masse nulle, alors que l'on sait bien que les bosons médiateurs de l'interaction faible sont nécessairement très massifs puisque la portée de cette interaction est très petite, et qu'il existe des fermions très massifs comme le *t* dont la masse est de 174 GeV/c^2. Il est certes possible de donner une masse au bosons médiateurs de l'interaction faible mais alors on brise explicitement l'invariance de jauge et ceci rend la théorie en construction non renormalisable. C'est pourquoi on s'est tourné vers un mécanisme dit de *brisure spontanée de symétrie*, dans lequel la dynamique est symétrique sans que les états du système le soient, en espérant que dans ce cas la renormalisabilité serait préservée. Une telle situation de brisure spontanée est relativement fréquente dans de nombreux domaines de la physique. Même en mathématiques pures, on sait bien que des équations peuvent être symétriques sans que les solutions le soient : par exemple l'équation $x^2 = 4$ est symétrique par rapport à la transformation $x \to -x$, alors que la solution $x = 2$ ne l'est évidemment pas ! En théorie quantique des champs, on a une situation de brisure spontanée de symétrie lorsque le lagrangien a une certaine propriété de symétrie sans que l'état fondamental (le vide) soit symétrique. Cette situation résulte d'un conflit entre la symétrie et la stabilité du vide : un vide potentiel qui serait symétrique serait instable, alors que seuls seraient stables des états non symétriques. Il convient de noter que, dans cette situation, le vide est dégénéré puisque plusieurs états liés les uns aux autres par la symétrie sont des vides possibles. L'archétype de la situation de brisure spontanée de symétrie est fourni par le potentiel en « fond de bouteille » (voir la figure)



qui apparaît dans le lagrangien suivant invariant par changement de la phase du champ scalaire complexe ϕ :

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - V(\phi)$$

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2); V = \frac{\lambda}{6\left(\phi^*\phi - \eta\right)^2}; \eta \text{ réel}$$

La dégénérescence du vide force à choisir comme vide un point quelconque de la « rigole ». Si on change de variable de champ, en prenant comme origine ce point, on s'aperçoit que les deux degrés de liberté du champ complexe se réarrangent en un champs scalaire de masse nulle, correspondant au mouvement, sans dépense d'énergie, « au fond de la rigole », que l'on appelle le boson de Goldstone, et un champ scalaire massif correspondant au mouvement « au dessus de la rigole ». En fait le *théorème de Goldstone* stipule que dans toute situation de brisure spontanée de symétrie existe au moins un boson de Goldstone de masse nulle.

Le mécanisme de Brout Engler Higgs

Comme l'ont montré Brout, Engler et Higgs, la brisure spontanée de symétrie est particulièrement intéressante lorsque la symétrie qui est brisée est une invariance de jauge. Rappelons que dans une théorie de jauge (abélienne par exemple, mais le raisonnement qui suit peut se généraliser à des théories non abéliennes), le boson de jauge est une particule vectorielle de masse nulle et de polarisation transverse (il n'a pas d'hélicité zéro). En cas de brisure spontanée, le boson de Goldstone se superpose aux deux modes transverses du boson de jauge pour former un boson vectoriel massif ayant trois états d'hélicité +1, –1 et 0. En

quelque sorte, le boson de Goldstone est « mangé » par le boson de jauge qui prend du poids et devient massif ! Au bilan,

- la dynamique reste invariante de jauge, le courrant est conservé, restent valables les identités de Ward qui traduisent cette conservation dans le développement perturbatif et qui sont cruciales pour la renormalisabilité
- le boson de jauge devient massif et le boson de Goldstone disparaît, la portée de l'interaction devient finie.

Le modèle standard électrofaible

La théorie électrofaible de Glashow, Salam et Weinberg découle de l'application de ce mécanisme de Brout, Engler et Higgs à la théorie de jauge $SU_L(2)\times U_Y(1)$ introduite ci-dessus. Les champs scalaires qui induisent la brisure spontanée, appelés champs de Higgs sont un doublet d'isospin faible $\left(\phi^+,\phi^0\right)$ et son conjugué $\left(\overline{\phi^0},\phi^-\right)$ (ce qui correspond à quatre degrés de liberté). Trois des quatre bosons de jauge acquièrent de la masse, c'est-à-dire que les trois bosons de Goldstone sont absorbés. Le quatrième degré de liberté des champs de Higgs devient un boson scalaire massif, le fameux *boson de Higgs*. Après brisure, on s'arrange pour avoir comme symétrie résiduelle, la symétrie $U(1)$ de QED, c'est-à-dire que l'on fait émerger *QED de la brisure de la symétrie de jauge initiale*. Les bosons de jauge W^3 et B se mélangent et leurs termes de masses forment une matrice de masse dont les états propres sont le photon A de masse nulle et le boson intermédiaire neutre Z^0 :

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W W_\mu^3$$
$$Z_\mu = -\sin\theta_W B_\mu + \cos\theta_W W_\mu^3$$

où θ_W désigne l'angle de mélange de Weinberg.

Faisons maintenant le bilan de ce qui apparaît comme une authentique unification des interactions électromagnétique et faible. Dans la théorie de jauge, avant brisure, nous avons quatre paramètres : g, g', η et λ qui se retrouvent dans M_W, M_Z, $\sin\theta_W$ et M_H (la masse du boson de Higgs). Mais la phénoménologie des interactions électromagnétiques et faibles nous fournit plusieurs autres observables, comme la charge électrique e, les constantes de couplage (de Fermi) des interactions faibles à courants chargés G et à courants neutres G_N, dont les valeurs pourront donc être prédites par le MS. Nous avons les relations suivantes :

$$e = g \sin\theta_W$$
$$\tan\theta_W = g'/g$$
$$\frac{G}{\sqrt{2}} = \frac{g^2}{8M_W^2}$$
$$M_W^2 = M_Z^2 \cos^2\theta_W$$
$$G_N = G$$

A propos de ces relations qui semblent en bon accord avec l'expérience, deux remarques s'imposent :

1. En toute rigueur, elles ne sont valables qu'à l'approximation de Born. Les corrections quantiques (ou *radiatives*) induisent dans ces relations des termes correctifs qui sont calculables par la méthode des perturbations si la théorie est renormalisable. Comme la renormalisabilité de la théorie a été prouvée (en 1971 par 'tHooft, puis par Lee et Zinn-Justin), les mesures de précision concernant ces observables et ces relations revêtent une très grande importance qui touche aux fondements mêmes de la mécanique quantique. Elles seront évoquées dans un exposé du colloque et dans un exposé plénier.
2. La relation entre les paramètres du MS et les observables est relativement directe : les constantes de couplages se retrouvent dans les constantes physiques des interactions électromagnétique et faible, η, la valeur moyenne dans le vide du champ de Higgs se retrouve dans la masse du *W*, quant à λ, il commande essentiellement, et à peu près uniquement, la masse du boson de Higgs. Il s'ensuit qu'à l'approximation de Born, le MS *ne prédit pas la masse du boson de Higgs*. Mais comme cette masse intervient comme un paramètre dans le calcul des corrections radiatives, les mesures de précisions peuvent permettre de contraindre son domaine de variation.

Les masses des fermions

Le mécanisme de Brout Engler et Higgs fournit en prime un moyen d'engendrer les masses des fermions. Compte tenu des nombres quantiques d'isospin et d'hypercharge, assignés aux particules droites et gauches ainsi qu'aux champs de Higgs, il est possible, en respectant toutes les propriété de symétrie, d'ajouter au lagrangien des termes de couplages dits de Yukawa entre un fermion gauche, son partenaire droit et un champ de Higgs du type

$$g_e\left(\overline{e_L} e_R \phi + h.c.\right)$$

Après la brisure spontanée, le champ de Higgs est remplacé par sa valeur moyenne dans le vide η qui est un c-nombre avec la dimension d'une masse (ou d'une énergie) ; il suffit alors de choisir pour la constante de couplage de Yukawa, $g_e=m_e/\eta$ pour que le terme de couplage de Yukawa devienne le terme de masse de l'électron. Comme ce mécanisme peut être utilisé pour tous les fermions massifs, on peut dire que c'est la brisure spontanée de la symétrie électrofaible qui fait émerger les masses de toutes les particules élémentaires (bosons et fermions). Ce mécanisme n'apporte rien du point de vue de l'économie des paramètres : les couplages de Yukawa sont des paramètres du MS qui commandent directement les masses observables des fermions. Trois remarques s'imposent cependant :

1. Lorsque la masse d'un fermion n'est pas connue, comme c'était le cas pour le *t* avant sa découverte, les calculs des effets des corrections radiatives sur certaines observables peuvent permettre de contraindre son domaine de variation. C'est ce qui s'est produit pour le *t,* dont la masse a été prédite avec une bonne précision grâce aux mesures effectuées au LEP.
2. Comme d'éventuels neutrinos droits auraient un isospin et une hypercharge nuls, il n'y a pas lieu de les introduire puisqu'ils ne participeraient à aucune des interactions du MS. Mais, s'il n'y a pas de neutrinos droits, il n'y a pas non plus de couplage de Yukawa ni donc de masse pour les neutrinos. C'est pourquoi on peut dire (bien que cette opinion ne soit pas unanime parmi les spécialistes) que des masses non nulles pour les neutrinos sont des indications d'une physique au-delà du MS.
3. On peut remarquer que QCD avec des quarks de masse nulle conserve la chiralité des quarks. Si donc les quarks étaient de masse nulle, les baryons, c'est-à-dire les fermions formés avec ces quarks, auraient une chiralité bien déterminée et ils seraient donc aussi de masse nulle. Il est toutefois possible, comme on sait bien que le proton et le neutron ne sont pas de masse nulle, d'invoquer un mécanisme de brisure spontanée de la symétrie chirale qui les rendrait massifs. Le pion serait alors le boson de Goldstone associé à cette brisure spontanée de symétrie. Le pion n'est pas de masse nulle (heureusement, sinon l'interaction forte serait de portée infinie !). Il faut donc supposer qu'avant la brisure spontanée de la symétrie chirale, les quarks (tout au moins les quarks *u* et *d* constituants des nucléons) ne sont pas de masse nulle mais qu'ils ont une masse très petite de telle sorte que le pion soit un *quasi* boson de Goldstone dont la masse est petite devant celle des nucléons. Autrement dit, selon un tel scénario, c'est la physique du MS électrofaible qui serait responsable de la masse

des nucléons et donc de celle des noyaux et donc de celle de toute la matière visible de l'univers et qui est responsable de la portée finie des interactions forte et faible.

GLOSSAIRE

ajustement fin (« fine tuning ») : la grande unification des interactions fondamentales fait intervenir des brisures de symétries à des échelles d'énergies très différentes (l'une à environ 10^3 GeV et l'autre à environ 10^{16} GeV) ; les mécanismes de Higgs qui rendraient compte de ces brisures devraient avoir leurs paramètres ajustés avec une précision si élevée qu'elle les rendrait peu crédibles.

antiparticule : pour rendre compatibles les principes de la mécanique quantique, de la relativité et de la causalité, la théorie quantique des champs a dû forger le concept d'antiparticule, une particule associée à chaque particule élémentaire, qui a la même masse qu'elle mais toutes les charges opposées. Les antiparticules prédites ont toutes été découvertes.

big bang [modèle du -] : modèle standard de la cosmologie selon lequel l'univers est en expansion et en refroidissement depuis une explosion primordiale, le « big bang », intervenue il y a environ quinze milliards d'années.

boson : particule obéissant à la statistique de Bose Einstein. La fonction d'onde d'un état à plusieurs bosons est symétrique par permutation de ces particules. Non seulement plusieurs bosons peuvent se trouver dans le même état, mais encore la probabilité pour qu'il en soit ainsi est beaucoup plus grande que pour des particules obéissant à la mécanique classique. Les bosons ont un spin (moment cinétique intrinsèque exprimé en unité de quantum d'action) entier ou nul.

brisure [- spontanée de symétrie] : pour une interaction décrite mathématiquement à l'aide de certaines équations, apparition de solutions ne satisfaisant pas la symétrie des équations.

cachée [masse -] : la masse de la matière visible dans l'univers est insuffisante pour rendre compte de la dynamique des étoiles dans les galaxies et des galaxies dans les amas. Comme la densité massique de l'univers est un paramètre fondamental qui conditionne le devenir de l'univers (vers l'expansion indéfinie ou l'implosion), la résolution du problème de la masse cachée de l'univers est un enjeu décisif de la physique contemporaine.

CERN : acronyme du Laboratoire Européen de Physique des Particules, situé à Genève.

charge [- d'interaction] : caractéristique, associée à un nombre quantique conservé, conditionnant la participation d'un constituant élémentaire aux interactions fondamentales. On distingue la charge électrique qui conditionne la participation à l'interaction électromagnétique, la charge faible (ou isospin faible) qui conditionne la participation à

l'interaction faible et la charge de couleur qui conditionne la participation à la chromodynamique (l'interaction forte au niveau fondamental).

confinement : impossibilité d'observer à l'état libre les quarks et les gluons.

courants neutres [interactions faibles à -] : interaction faibles n'impliquant pas d'échange de charge électrique ; dans la théorie électrofaible, interactions impliquant l'échange d'un boson intermédiaire Z^0.

effective [théorie -] : nouvelle conception de la théorie quantique des champs selon laquelle, il n'est plus nécessaire (comme pour les théories strictement renormalisables du modèle standard) de se prononcer sur le comportement des interactions à des énergies arbitrairement élevées, et qui se contente d'une description idoine des phénomènes importants à une échelle d'énergie donnée.

électrofaible [théorie -] : partie du modèle standard décrivant de façon unifiée les interactions électromagnétique et faible.

fermions : particule obéissant à la statistique de Fermi Dirac. La fonction d'onde d'un état à plusieurs fermions est antisymétrique par permutation de ces particules. Deux fermions identiques ne peuvent se trouver dans le même état quantique ; c'est le principe d'exclusion de Pauli. Les fermions ont un spin demi-entier.

génération : classification indépendante de la participation aux interactions fondamentales des constituants de la matière. D'une génération à l'autre, les quarks et les leptons chargés de mêmes charges ne diffèrent que par leurs masses. Il y a trois générations, ce qui permet au modèle standard de rendre compte de la violation de l'invariance CP.

hadron : particule composite, constituées de quarks, antiquarks et gluons, participant à toutes les interactions, y compris l'interaction forte.

HERA : collisionneur électron-proton installé à Hambourg en Allemagne

Higgs [boson de -] : particule hypothétique de spin nul, dont l'existence, associée au mécanisme de Higgs, est prédite par le modèle standard.

Higgs [mécanisme de -] : mécanisme de brisure spontanée de symétrie qui, appliqué à une théorie à invariance de jauge rend massifs les bosons de l'interaction. Dans la théorie électrofaible, ce mécanisme implique l'existence d'au moins un boson de Higgs.

invariance CP : invariance par l'opération de symétrie discrète produit de la conjugaison de charge C par la parité d'espace P.

jauge [invariance de -] : invariance par des opérations de symétries dépendant du point d'espace-temps où elle sont appliquées ; dans le modèle standard, toutes les interactions fondamentales sont décrites par des théorie à invariance de jauge.

LEP : acronyme de « Large Electron Positron collider », le grand collisionneur électron-positon installé au CERN.

lepton : fermion élémentaire qui ne participe pas à l'interaction forte.

LHC : acronyme de « Large Hadron Collider », le grand collisionneur proton-proton qui sera installé dans le tunnel du LEP au CERN, et devrait entrer en fonctionnement en 2005.

masse : dans la théorie de la relativité, la masse d'une particule est sa masse *invariante par transformation de Lorentz*. Pour une particule massive, la masse invariante coïncide avec la masse au repos. En théorie de la relativité, une particule peut avoir une masse nulle ; elle n'est alors jamais au repos, car elle se déplace à la vitesse de la lumière, dans tout référentiel. Le photon, le gluon et le graviton (l'hypothétique particule associée à la gravitation dans le domaine quantique) sont des particules de masse nulle. On ne sait toujours pas si les neutrinos sont ou ne sont pas des particules de masse nulle.

MSSM : acronyme de « Minimal Supersymmetric Standard Model », qui désigne l'extension supersymétrique minimale du modèle standard.

neutrino : lepton neutre qui ne participe qu'à l'interaction faible.

parité : opération de symétrie discrète consistant à changer de signe les coordonnées d'espace ; alors que les interactions gravitationnelle, électromagnétique et faible sont invariante par parité, l'interaction faible ne l'est pas. Le modèle standard fournit une explication simple et naturelle de cette particularité.

perturbatif [développement -] : évaluation par approximations successives des probabilités de transition dans le cadre du modèle standard.

QCD : acronyme de Quantum ChromoDynamics, ou Chromodynamique Quantique, théorie à invariance de jauge de l'interaction forte au niveau des quarks.

QED : acronyme de Quantum ElectroDynamics, ou Electrodynamique Quantique, théorie quantique et relativiste de l'interaction électromagnétique.

quantification [- de la gravitation] : à l'échelle des particules élémentaires, la gravitation, que l'on ne sait pas encore décrire en accord avec la mécanique quantique, est tellement faible qu'il est possible de la négliger. Toutefois, à des énergies de l'ordre de 10^{19} GeV, soit une distance de l'ordre de 10^{-35} cm, ou encore un temps de 10^{-43} seconde après le big bang, on

s'attend à ce que la gravitation devienne forte, voire dominante. La quantification de la gravitation est le grand défi théorique de la fin de ce siècle.

quark : fermion élémentaire participant à toutes les interactions fondamentales.

renormalisabilité : propriété des théories du modèle standard qui permet d'éliminer les infinis qui apparaissent dans le développement perturbatif et qui en garantit la fiabilité.

supercorde [théorie de la -] : théorie dans laquelle les particules élémentaires sont les excitations supersymétriques d'une corde dont la dimension avoisine la longueur de Planck (10^{-35} cm).

supergravité : extension supersymétrique de la théorie de la gravitation ; la supersymétrie y est traitée comme une invariance de jauge.

supersymétrie : symétrie imaginée pour tenter de résoudre les problèmes de la quantification de la gravitation ; une opération de supersymétrie transforme un boson en un fermion et vice et versa.

symétrie : propriété caractéristique d'une interaction fondamentale qui traduit son *invariance* par rapport à un ensemble de transformations. On distingue trois grands types de symétries, les *symétries cinématiques relativistes* (groupes de Lorentz et de Poincaré), les *symétries discrètes* (parité d'espace ou inversion des coordonnées d'espace, conjugaison de charge ou remplacement de chaque particule par son antiparticule, renversement du sens du temps), et les *symétries internes* (isospin, couleur, saveur...).

TEVATRON : collisionneur proton-antiproton installé au Laboratoire de Fermi, à Chicago aux États-Unis.

BIBLIOGRAPHIE

- *Les atomes* Jean Perrin, Champs Flammarion 1991
- *Autobiographie scientifique* Max Planck, Champs Flammarion 1960
- *Relativity, the special and general theory* Albert Einstein, Crown Trade Paperbacks 1961
- *Conceptions scientifiques* Albert Einstein, Champs Flammarion 1990
- *Le quark et le Jaguar* Murray Gell-Mann, Champs Flammarion 1997
- *Les constantes universelles* Gilles Cohen-Tannoudji, Hachette 1998
- *Max Planck et les quanta* Jean-Claude Boudenot et Gilles Cohen-Tannoudji, Editions Ellipses 2001
- *Histoire de la physique et des physiciens* Jean-Claude Boudenot, Editions Ellipses 2001
- *Symmetry in Science, an introduction to the general theory* Joe Rosen, Springer 1995
- *La révolution des quanta* Victor Weisskopf, Hachette 1989