

Géométrie
de la gravitation et du modèle standard

D. Philipona
Directeur de stage : D. Bennequin

Paris, 30 juillet 2004

Remerciements

Je tiens tout particulièrement à remercier Daniel Bennequin pour sa grande patience et sa disponibilité, j'ai conscience du débordement en temps par rapport au projet initial.

Pour les personnes, dont je suis, qui se heurtent implacablement à la frontière de leurs connaissances, la compréhension d'une théorie physique semble procéder à l'inverse de la compréhension d'une théorie mathématique : par la perte de vue de ses fondations. D'une certaine manière, on comprend l'utilisation physique des probabilités parce qu'on ne se demande plus pourquoi la fréquence des piles et faces d'un lancer de pièce devrait tendre vers une limite. Aussi avais-je besoin de remonter assez loin en amont de la théorie des interactions fondamentales pour avoir l'impression de pouvoir la comprendre. Le panorama dont je rends compte, pour ridiculement succinct qu'il soit, est une tentative de réponse à ce besoin. Encore une fois, je voudrais remercier Daniel Bennequin de m'avoir dirigé dans cette recherche un peu absurde d'une *compréhension*, plutôt que d'un *savoir*. Une majorité de physiciens a en effet l'habitude de couper cours aux discussions sur la compréhension en affirmant "la nature est comme ça". Pourtant, sans prétendre aller contre l'argument de fond, il est certain que l'interprétation n'est pas *dictée*, seulement contrainte, par les résultats d'une expérience : le renversement, induit par la relativité et la physique quantique, de l'interprétation Newtonienne de la physique à l'échelle classique en est un exemple. L'histoire de la science est faite de ces revirements, tel l'affrontement entre les deux modèles du mouvement des planètes, celui de Ptolémée et celui de Copernic, lourdement opposés conceptuellement mais assez similaires au niveau des résultats. Que les nouvelles interprétations prétendent s'imposer par un plus grand domaine d'application soulève la question suivante : est-il sensé d'imaginer qu'on puisse jamais disposer d'un corpus de résultats qui contraigne sans ambiguïté leur interprétation ? Il me semble que le comment de la nature, i.e. un sens, ne peut émerger que d'une approche mathématico-physique : seule compte *la partie irréductible du modèle*, et saisir cette part inévitable est proprement mathématique. J'irais même presque jusqu'à dire : géométrique, si l'on considère que la géométrie est bâtie sur cette idée viscérale qu'un objet peut être vu de différents points de vues, qui plus est partiels, et que pourtant il existe quelque chose à la fois en amont de ces points de vues et leur somme synthétique.

Je voudrais également remercier les organismes qui me soutiennent financièrement, le Sony CSL, l'ANRT, le projet européen Enactive, ainsi que les personnes qui m'encadrent et me font une confiance qui m'honore en me laissant la liberté de mener mes recherches par des chemins peu standards : Olivier Coenen, Kevin O'Regan, Jean-Pierre Nadal.

Table des matières

1	Introduction	4
2	Mécanique classique : espace, temps, interactions conservatrices	6
2.1	Le point de vue Lagrangien	6
2.1.1	Lagrangien d'une particule	6
2.1.2	Symétries et conservation des grandeurs	7
2.1.3	Approche statistique et limite continue	7
2.2	Géométrie symplectique	8
2.2.1	Point de vue Hamiltonien	8
2.2.2	Structure symplectique et quantification	9
2.3	Mécanique quantique	10
3	Relativité générale : l'espace-temps	12
3.1	Cadre physique	12
3.1.1	Le modèle local d'espace-temps	12
3.1.2	Concepts de la théorie	13
3.1.3	Les équations d'Einstein	14
3.2	Existence et unicité des solutions	15
3.2.1	Problème de Cauchy	15
3.2.2	Existence locale, coordonnées harmoniques	15
3.2.3	Unicité globale	16
3.3	Solutions physiques	16
3.3.1	Espace-temps de Schwarzschild	16
3.3.2	Espace-temps stationnaires et statiques	17
3.4	Singularités	18
3.4.1	Effondrement gravitationnel	18
4	Théories des champs : espace-temps, création et annihilation de particules	19
4.1	Point de départ : le modèle relativiste de l'électron	19
4.1.1	Équation de Klein-Gordon	19
4.1.2	Équation de Dirac	20
4.2	Théorie des champs	20
4.2.1	Champs classiques : théorème de Noether	20
4.2.2	Champs quantiques	21

4.2.3	Brisure spontanée de symétrie, particules sans-masses et particules massives . . .	22
4.3	Le modèle standard : équations des champs de jauge	23
4.3.1	Dynamique libre des fermions et des hadrons	23
4.3.2	Dynamique libre des bosons : équation de Yang-Mills	23
4.3.3	Interactions des fermions et des mesons avec les bosons de jauge	24
4.3.4	Interactions entre fermions et mesons	24
4.4	Modèle standard : synthèse et couplage avec la gravitation	26
4.4.1	Synthèse, théorèmes d'existence	26
4.4.2	Couplage avec la gravitation	27
5	Annexes	29
5.1	Géométrie différentielle	29
5.1.1	Rappels et notations	29
5.1.2	Variétés Lorentziennes	29
5.1.3	Classes caractéristiques	30
5.2	Spineurs	30
5.2.1	Algèbre de Clifford	30
5.2.2	Groupes Pin et Spin	31
5.2.3	Représentation	32
5.2.4	Structures Spin, variétés spinorielles, spineurs, opérateur de Dirac	32
5.3	Notes physiques	33
5.3.1	La cohérence quantique et les probabilités complexes	33
5.3.2	Causalité et localité	34

Chapitre 1

Introduction

“Only connect.” - EM Forster, Howard’s End.

Notre vision actuelle du monde est celle d’un ensemble d’objets dont les interactions sont régies par quatre forces : la gravitation, l’électromagnétisme, les forces nucléaires fortes et faibles. La décroissance de ces forces détermine leurs domaines d’influence : partout l’électromagnétisme, avec en plus aux grandes distances la gravitation, et aux courtes distances les forces fortes et faibles.

L’idée centrale de la relativité générale est l’unification des notions d’espace, temps, et force gravitationnelle dans une même entité : une variété Lorentzienne.

Le modèle standard décrit quant à lui les interactions entre particules élémentaires, et ses idées centrales sont celles de renormalisation et de symétrie, au travers des fibrés principaux. Dans ce modèle, dit des champs de jauge, les *forces* électromagnétiques, nucléaires fortes et faibles, sont décrites comme des connexions sur un certain fibré principal. On considère qu’elles correspondent à des particules : la force électromagnétique par les photons (particules γ), la force faible par les particules W^+ , W^- et Z , la force forte par les gluons (8 sortes). L’ensemble de ces particules sont des *bosons*, i.e. des particules de spin entier.

Ces forces agissent sur, ou plutôt interagissent avec, des particules “chargées” nommées *fermions*, qui sont des particules de spin fractionnaire : les *quarks*, sensible à la force forte et composés de *quarks*, et les *leptons*, qui n’y sont pas. Les fermions ont chacun leur antiparticules de même masse et de charge opposée, et se classent en trois familles. La première est constituées par deux leptons : l’électron e et l’électron neutrino ν_e , et deux quarks : haut u et bas d . Ce sont les particules les plus courantes. La seconde famille est constituée du muon μ et du muon neutrino ν_μ , et des quarks charmés c et étranges s . La troisième famille est constituée du tau τ et tau neutrino ν_τ , et des quarks top t et bottom b . Les particules composées de quarks, les *hadrons*, sont par ailleurs divisés en deux groupes, selon le nombre de quarks qui les composent : les *mesons* (pions, kaons, J, D, ad Upsilon), constitués d’un quark et d’un antiquark, et les *baryons* (proton, neutron, lambda, et leurs anti-particules, ainsi que sigma-plus, sigma-moins, lambda charme), constitués de trois quarks. Comme ils sont composés d’un nombre pair de fermions, les mesons sont donc une sorte particulière de boson, qui interagissent avec les forces fondamentales. Finalement, il existe une dernière particule chargée fondamentale, la particule de Higgs, qui n’est ni un lepton ni un quark, conjecturée pour permettre de relier les forces électromagnétiques et faibles.

Les deux théories ont nécessité une réinterprétation profonde de notre conception du temps, de l’espace, de la particule, de la causalité, de l’observation. Un principe général a servi, a priori ou a posteriori, de

lumière rationnelle pour éclairer ces terrains obscurs : l'interprétation géométrique. L'objectif de ce texte est de synthétiser les points de départ des relations les plus communes entre physique et structure géométriques.

Les courtes introductions de chaque chapitre sont inspirées très directement des notes de cours de Gilles Cohen-Tannoudji disponibles sur internet.

Chapitre 2

Mécanique classique : espace, temps, interactions conservatrices

Selon Einstein [3], "la première tentative en vue de poser un fondement théorique unifié [de la mécanique] est représentée par l'oeuvre de Newton. Dans son système tout est ramené aux 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 qu'il soit seulement d'une manière implicite."

Les cadres du Lagrangien et Hamiltoniens, développés fin du *XVIII^{ème}* début du *XIX^{ème}*, représentent la réponse mathématique à ces idées et mettent en lumière les liens entre les symétries de l'action et la conservation de certaines grandeurs.

2.1 Le point de vue Lagrangien

2.1.1 Lagrangien d'une particule

Le principe de moindre effort en mécanique classique consiste à décrire la trajectoire suivie par une particule comme celle minimisant une certaine grandeur qu'on nomme action. Étant donné un espace de configurations \mathcal{M} , cette grandeur prend la forme (\mathbf{q} dénote la trajectoire $\tau \mapsto q(\tau)$) :

$$S_t(\mathbf{q}) = \int_0^t L(q(\tau), \dot{q}(\tau), \tau) dt \quad (2.1)$$

où le Lagrangien L est donc une fonction scalaire sur $T\mathcal{M} \times \mathbb{R}$, qui doit être bornée inférieurement. Cette formulation, indépendamment de l'expression effective de L , permet de tirer les conséquences de l'idée que l'action (quoique cela veuille dire a priori) est la somme au cours du temps d'actions *instantanées*. Plus généralement, elle permet d'accroître la complexité du modèle par simple ajout de termes : le Lagrangien d'un produit de systèmes non couplés est la somme des Lagrangiens de chaque système, et le couplage se traduit par l'addition d'un terme d'interaction.

On montre facilement que la différentielle (fonctionnelle) de S en q s'annule pour les variations δq de trajectoires s'annulant aux limites si la trajectoire q satisfait à tout temps τ à l'équation d'Euler-Lagrange :

$$\frac{\partial L}{\partial q}(q(\tau), \dot{q}(\tau), \tau) = \frac{d}{d\tau} \left[\frac{\partial L}{\partial \dot{q}}(q(\tau), \dot{q}(\tau), \tau) \right]$$

qui établissent le lien entre l'approche fonctionnelle et le point de vue plus Newtonien des équations différentielles.

2.1.2 Symétries et conservation des grandeurs

L'un des intérêts principaux de la formulation de la mécanique en termes de Lagrangien est de permettre l'utilisation des symétries physiques dans la déduction de l'action (une inférence très élégante des formules de la mécanique du point est présentée dans [7], basée uniquement sur des arguments de symétrie et de structure du Lagrangien), et plus généralement d'étudier leurs conséquences vis-à-vis de la conservations de certaines grandeurs sur les trajectoires extrémales. Par exemple, les notions d'énergie et de moment sont naturellement liées à la conception Lagrangienne. Dans le cas d'un Lagrangien symétrique par translation en espace on obtient :

$$\frac{\partial L}{\partial q} = 0 = \frac{d}{d\tau} \frac{\partial L}{\partial \dot{q}}$$

et donc la quantité $p = \frac{\partial L}{\partial \dot{q}}$ nommée moment, est conservée pendant le mouvement. Dans le cas d'une invariance par translation en temps, on obtient :

$$\frac{dL}{d\tau} = \frac{\partial L}{\partial q} \dot{q} + \frac{\partial L}{\partial \dot{q}} \ddot{q} + 0 = \dot{q} \frac{d}{d\tau} \frac{\partial L}{\partial \dot{q}} + \ddot{q} \frac{\partial L}{\partial \dot{q}} = \frac{d}{d\tau} \left(\dot{q} \frac{\partial L}{\partial \dot{q}} \right)$$

et donc la quantité $E = p\dot{q} - L$, qu'on nomme énergie, est conservée.

La conservation de ces grandeurs au cours du temps, lorsque le Lagrangien est invariant en temps et donc non dissipatif, légitime leur statut de grandeur physique : les hypersurfaces de niveau de la fonction énergie ont une signification directe en termes de contraintes sur les configurations accessibles.

2.1.3 Approche statistique et limite continue

Loin d'être limitée à un cadre déterministe fort, la maximisation de l'action intervient naturellement dans un cadre probabiliste après passage à la limite continue. Considérant par exemple le cas d'un processus de Markov dans $E = \mathbb{R}^d$ avec un noyau de transition ρ invariant par translation, on a :

$$P_n(q) = \int_E \rho(q - q') P_{n-1}(q') dq'$$

et par récurrence, dans le cas $P_0(q) = \delta(q)$,

$$P_n(q) = \int_{E^n} \exp \left(- \sum_{j=1}^n w(q_j - q_{j-1}) \right) \prod_{j=1}^{n-1} dq_j, \quad q \equiv q_n$$

où $w(q) = -\ln \rho(q)$. En considérant la trajectoire continue définie par

$$\tau_j = \frac{j}{n} t \quad \text{et} \quad q^n(\tau) = q_{j-1} + \frac{\tau - \tau_{j-1}}{\tau_j - \tau_{j-1}} (q_j - q_{j-1}) \quad \text{pour } \tau \in [\tau_{j-1}, \tau_j]$$

il est alors facile de voir que

$$\sum_{j=1}^n w(q_j - q_{j-1}) \frac{t}{n} = \int_0^t w\left(\frac{t}{n} \dot{q}^n(\tau)\right) d\tau$$

En general, la suite de fonctions $(q \mapsto \frac{n}{t} w(\frac{t}{n} q))_n$ n'admet pas de limite, du moins non triviale, et l'on cherche donc une renormalisation de l'espace-temps $E \times \mathbb{R}$ qui lui en donne une, i.e., dans ce cas particulier, une fonction λ telle que

$$\frac{n}{t} w\left(\frac{t}{n} \frac{q}{\lambda(t, n)}\right) \xrightarrow{n \rightarrow \infty} w^*(q)$$

En considérant la trajectoire $\mathbf{x}^n = \lambda(t, n) \mathbf{q}^n$, il vient alors

$$P_n(q) \sim \int_{\mathcal{C}_{n,t}} \exp(-S_t(\mathbf{x})) d\mathbf{x} \quad \text{avec} \quad S_t(\mathbf{x}) = \int_0^t w^*(\dot{x}(\tau)) d\tau$$

où $\mathcal{C}_{n,t}$ est l'ensemble des trajectoires continues allant de 0 à $\lambda(t, n)q$ en un temps t , l'intégrale sur les chemins étant l'intégrale de Wiener. Il apparaît ainsi que les trajectoires minimisant l'action S_t vont jouer un rôle dominant pour le calcul des valeurs moyennes.

2.2 Géométrie symplectique

Il est naturel de s'interroger sur la structure géométrique impliquée dans cette approche. En particulier, le passage du Lagrangien aux équations d'Euler-Lagrange (utilisées pour dériver les propriétés de conservations) est problématique car s'il est aisé de donner un sens intrinsèque à

$$p = \frac{\partial L}{\partial \dot{q}}(q, \dot{q})$$

comme la dérivée de l'application

$$\dot{q} \in T\mathcal{M}|_q \mapsto L(q, \dot{q})$$

(et donc $p \in T^*\mathcal{M}|_q$), \dot{q} fixe lui la valeur de q et empêche la définition de

$$\frac{\partial L}{\partial q}(q, \dot{q})$$

par un même principe. De plus, les équations d'Euler-Lagrange ne sont valides qu'à l'intérieur du domaine d'une carte de \mathcal{M} . Il faut donc chercher un autre objet que le Lagrangien ; cet objet apparaît lors du passage au point de vue Hamiltonien.

2.2.1 Point de vue Hamiltonien

Si le Lagrangien est tel qu'on puisse inverser $(q, \dot{q}) \mapsto (q, p)$, on considère la fonction

$$H(q, p, \tau) = p\dot{q} - L$$

Les équations d'Euler-Lagrange deviennent alors

$$\dot{p} = \frac{\partial L}{\partial q}$$

et de l'expression de la différentielle de H :

$$dH = dp\dot{q} + p d\dot{q} - \dot{p}dq - pd\dot{q} - \frac{\partial L}{\partial \tau} d\tau = -\dot{p}dq + \dot{q}dp - \frac{\partial L}{\partial \tau} d\tau$$

on déduit les équations de Hamilton, ou équations canoniques :

$$\dot{q} = \frac{\partial H}{\partial p} \quad , \quad \dot{p} = -\frac{\partial H}{\partial q} \quad , \quad \frac{dH}{d\tau} = \frac{\partial H}{\partial \tau}$$

où l'on lit directement que l'invariance par translation spatiale du Lagrangien implique la conservation du moment et son invariance par translation temporelle la conservation de l'énergie.

Dans le cas où l'hamiltonien est non dissipatif, on lit également le champ de vecteur $X_H \in T(T^*\mathcal{M})$ décrivant la dynamique :

$$X_H(q, p) = \left(\frac{\partial H}{\partial p}(q, p), -\frac{\partial H}{\partial q}(q, p) \right)$$

et en notant le crochet de Poisson de deux fonction de (q, p)

$$\{f, g\} = \sum_i \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i}$$

on constate alors que les équations canoniques s'écrivent

$$\dot{q}_i = \{q_i, H\} \quad , \quad \dot{p}_i = \{p_i, H\} \quad \text{et} \quad \{q_i, p_j\} = \delta_{ij} \quad , \quad \{p_i, p_j\} = \{q_i, q_j\} = 0$$

où p_i et q_i sont des fonctions scalaires sur $T^*\mathcal{M}$. Ainsi apparaît une structure d'algèbre sur les grandeurs mesurables, les grandeurs conservées satisfaisant $\{f, H\} = 0$, et la mise en valeur du rôle crucial de l'espace des phases : le fibré cotangent $T^*\mathcal{M}$.

D'une manière plus générale, on note que pour une fonction de (q, p, τ) :

$$\{f, H\} = \sum_i \frac{\partial f}{\partial q_i} \dot{q}_i + \frac{\partial f}{\partial p_i} \dot{p}_i = X_H \cdot f = \dot{f} - \frac{\partial f}{\partial \tau}$$

2.2.2 Structure symplectique et quantification

Tout fibré cotangent est doté de manière canonique d'une 1-forme ζ , la forme de Liouville :

$$\zeta_{|(x, \alpha)} : v \in T(T^*\mathcal{M})_{|(x, \alpha)} \mapsto \alpha(\pi_*(x, \alpha)(v))$$

où π est la projection de $T^*\mathcal{M}$ sur \mathcal{M} , et π_* son application tangente. On en déduit une 2-forme $\omega = d\zeta$ fermée, la forme symplectique. A toute fonction scalaire lisse H sur $T^*\mathcal{M}$, on montre alors qu'il est possible d'associer canoniquement un champ X_H (champ de vecteur hamiltonien) tel que

$$dH = \omega(X_H, \cdot)$$

Le flot d'un tel champ X_H est un symplectomorphisme, i.e. un difféomorphisme qui conserve la forme ω . On retrouve la structure d'algèbre de Poisson des observables avec :

$$\{f, g\} = \omega(X_f, X_g)$$

qui vérifie bien $\{f, H\} = \omega(X_f, X_H) = df(X_H) = X_H \cdot f$. Enfin, comme $X_{\{f, g\}} = [X_f, X_g]$, on trouve que cette algèbre est isomorphe à l'algèbre de Lie du groupe des symplectomorphismes.

A l'opposé des points de vue précédents, cette construction géométrique ne privilégie pas un système de coordonnées a priori, et permet une approche globale de la situation. On peut ensuite revenir a posteriori dans des coordonnées, par exemple avec le théorème d'Arnold-Mineur-Liouville qui pointe une propriété très contraignante satisfaite par les systèmes entièrement intégrables dont les champs de vecteur Hamiltoniens sont complets :

Théorème 1. *Tout point régulier $a \in \mathbb{R}^d$ d'une application moment¹ A défini dans son voisinage \mathcal{U} une fibration triviale, dont la fibre typique est une union de cylindres. De plus, il existe des coordonnées $(x, \xi) : \mathcal{U} \rightarrow \mathbb{T}^d \times \mathbb{R}^d$ (coordonnées "angle-action") telles que l'Hamiltonien est indépendant de x .*

Dans ce cadre, la question de la quantification géométrique d'une variété symplectique (M, ω) est celui de la représentation de ses sous-groupes de symplectomorphismes dans le groupe linéaire des sections d'un fibré préquantique au dessus de M , i.e. un fibré en ligne complexe L au dessus de M muni d'une connexion de courbure $i\omega$ et d'un produit hermitien qui lui est compatible.

2.3 Mécanique quantique

L'origine de l'intérêt porté à cette construction provient bien sur des méthodes développées par la physique quantique pour rendre compte des expériences menées d'abord sur la lumière, puis sur les électrons et les atomes. La fructueuse recette de la "première quantification" est 1) de représenter les observables du modèle classique dans l'algèbre des opérateurs hermitiens d'un espace de Hilbert, avec un produit et une dynamique hérités de l'analyse classique (les facteurs i et i/\hbar étant nécessaires pour rendre le produit hermitien et la dérivée homogène) :

$$[\mathcal{Q}(f), \mathcal{Q}(g)] = i\mathcal{Q}(\{f, g\}) \quad \text{et} \quad \frac{d}{dt}\mathcal{Q}(f) = \frac{i}{\hbar} [\mathcal{Q}(H), \mathcal{Q}(f)] + \frac{\partial \mathcal{Q}(f)}{\partial t}$$

et 2) de décrire le phénomène de mesure de f comme un "effondrement de la fonction d'onde" sur les vecteurs propres de $\mathcal{Q}(f)$, i.e. la transition de l'état $|\psi\rangle$ à l'un des états propres $|\lambda\rangle$ selon la probabilité

$$P(|\psi\rangle \rightarrow |\lambda\rangle) = \frac{|\langle \psi | \lambda \rangle|^2}{\langle \psi | \psi \rangle}$$

Comme seule la probabilité est véritablement accessible à l'expérience, on se restreint à la sphère unité de l'espace, ce qui justifie alors que les observables soient hermitiennes puisqu'elles sont isomorphes à l'algèbre de Lie de la représentation des symplectomorphismes dans le groupe unitaire.

¹une application $A : T^*\mathcal{M} \rightarrow \mathbb{R}^d$ (d dimension de \mathcal{M}) est une application moment si les A_i sont lisses, vérifient $\{A_j, H\} = \{A_i, A_j\} = 0$ et les différentielles dA_i sont indépendantes presque partout

L'un des avantages de cette formulation est qu'elle permet de résoudre formellement la trajectoire des observables. En effet, dans le cas d'observable qui ne dépendent pas explicitement du temps, en posant

$$U(t) = e^{-\frac{i}{\hbar} Q(H)t}$$

et

$$Q(f)(t) = U(t)^\dagger Q(f)(0)U(t) \quad (2.2)$$

on trouve par un calcul direct

$$\frac{d}{dt} Q(f) = \frac{i}{\hbar} (Q(H)Q(f) - Q(f)Q(H)) = \frac{i}{\hbar} [Q(H), Q(f)]$$

Comme l'équation 2.2 peut par ailleurs s'interpréter non pas comme une évolution de l'observable sur des états fixes, mais l'évolution des états avec une observable fixe, on trouve une dualité entre le point de vue de Heisenberg (état fixes, observables dynamiques), et celui de Schrödinger (états dynamiques, observables fixes). Dans cette dernière approche, l'équation de l'évolution de l'état est donc :

$$\frac{d}{dt} |\psi(t)\rangle = \frac{d}{dt} U(t) |\psi(0)\rangle = -\frac{i}{\hbar} Q(H)U(t) |\psi(0)\rangle = -\frac{i}{\hbar} Q(H) |\psi(t)\rangle$$

Chapitre 3

Relativité générale : l'espace-temps

La théorie de la relativité générale est sous-tendue par une critique profonde de l'approche Newtonienne, en particulier le refus de l'idée d'actions instantanées, et l'hypothèse d'une borne supérieure à toute vitesse de propagation, de quoi que ce soit, dans tout l'univers. Il s'en suit que 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 le cas de l'interaction électromagnétique, qui se propage à la vitesse c dans tous les référentiels d'inertie (expérience de Michelson). Pour fonder théoriquement l'origine de cette invariance, la relativité remet en cause, au début du XX^{eme} siècle, le temps, la métrique spatiale et la simultanéité en tant que cadres absolus, pour les remplacer par un espace-temps Lorentzien unifié.

Les sources principales pour ce chapitre sont [10, 1].

3.1 Cadre physique

Après une rapide motivation de la modélisation de l'espace-temps par une variété Lorentzienne, on présente (non moins rapidement) les objets élémentaires de la théorie de la relativité.

3.1.1 Le modèle local d'espace-temps

L'espace-temps de Minkowski est le modèle local de la notion d'espace-temps : un espace vectoriel $M = \mathbb{R}^4$, muni d'une pseudo-norme g de signature $(+, +, +, -)$. Ce modèle dérive de l'idée à la base de la théorie de la relativité : la lumière se déplace à la même vitesse, dans toutes les directions et dans tout repère inertiel.

Considérons un physicien A dans un monde unidimensionnel, doté d'un chronomètre et d'un outil de mesure instantané des grandeurs spatiales. Cherchant à décrire la trajectoire d'un front d'onde électromagnétique, il relève au cours de son histoire propre les mesures (x, t) délivrées par son mètre et son chronomètre, et trace le graphe correspondant dans \mathbb{R}^2 (c.f. Fig. 3.1). Dans ce plan, les vecteurs $(1, 0)$ et $(0, 1)$ représentent la mesure unité des instruments, les lignes horizontales représentent l'instantanéité et les lignes verticales la localisation. Normalisant correctement les unités, A va obtenir un cône $t = |x|$.

En supposant qu'au cours de cette expérience un deuxième physicien B se trouvait, avec son propre équipement de mesure, dans un repère se déplaçant à vitesse constante par rapport au repère de A , ce

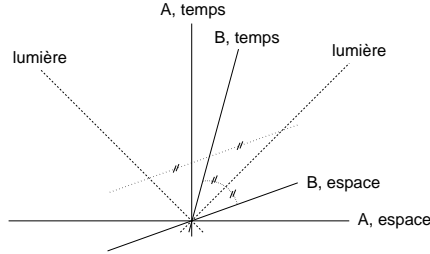


FIG. 3.1 – Mesures de A , et la représentation des notions d'instantanéité et de localité de B .

dernier peut également tracer sur son graphe les positions de B et obtenir ainsi une demi-droite $x = vt$. On cherche alors à trouver la relation entre les mesures trouvées par A et par B à propos du front d'onde. La notion de localité de B , représentée dans le graphe de A , est constituée des lignes parallèles à sa trajectoire : il doit en être ainsi pour qu'il se pense immobile. Sa notion d'instantanéité, toujours dans le graphe de A , doit être telle que le front d'onde lui semble 1) se déplacer à vitesse 1 et 2) de manière symétrique à droite et à gauche.

De cette simple construction géométrique on peut déduire, à un facteur multiplicatif global près, la représentation des mesures unités de B et donc les relations de transformations entre les mesures. Ce facteur est fixé par une hypothèse supplémentaire, celle de l'identité des unités de mesure, qui ne se justifie que dans un cadre quantique, faute de quoi on a en fait une théorie conforme. En étendant à un espace tridimensionnel, on trouve que ces transformations forment le groupe de Lorentz, groupe orthogonal d'une forme quadratique de signature $(+, +, +, -)$: la quantité physique, indépendante du système de coordonnées, est donc cette pseudo-norme.

3.1.2 Concepts de la théorie

Espace-temps La relativité générale cherche à appréhender la force de gravitation comme l'effet d'une courbure de l'espace-temps. Un espace-temps (M, g, D) est une variété connexe M munie d'une métrique Lorentzienne et de la connexion de Levi-Civita D associée. Un champ gravitationnel est une classe d'équivalence d'espace-temps, à isométrie près préservant l'orientation et l'orientation du temps (c.f. annexe).

Particules, observateurs Les objets d'intérêt dans un espace-temps sont les particules, les observateurs et les cadres de référence. Une particule ponctuelle (γ, m) est une trajectoire γ dans M et m est un réel positif caractérisant la trajectoire en chute libre de cette particule par l'équation¹ $g(\dot{\gamma}, \dot{\gamma}) = -m^2$.

Comme on l'a vu dans la section "modèle local", la trajectoire d'un observateur en translation uniforme détermine la manière dont il mesure le temps². On définit donc un observateur comme une trajectoire z dans M vérifiant $g(\dot{z}, \dot{z}) = -1$, où cette contrainte est purement conventionnelle pour que le temps mesuré par l'observateur soit

$$\int_a^u \|\dot{z}(s)\| ds = u - a$$

¹ Le lien avec la trajectoire classique se fait comme suit : considérons un espace-temps de Minkowski \mathbb{R}^2 , et une particule Newtonienne (x, m) , cette dernière est une particule ponctuelle (γ, m) au sens de la relativité en posant $\gamma(u) = (x \circ s^{-1}(u), s^{-1}(u))$ où $s(t) = (1/m) \int_0^t (1 - \dot{x}^2(u))^{1/2} du$. On vérifie que $\gamma^1 = x \circ \gamma^2$ et que $g(\dot{\gamma}, \dot{\gamma}) = -m^2$

² En ce qui concerne la mesure de l'espace, un observateur doit également spécifier l'évolution d'un gyroscope.

Les notions d'énergie et de moment d'une particule sont unifiés et simplifiés par cette approche : étant donné un observateur instantané (z, \dot{z}) et une particule instantanée $(\gamma, \dot{\gamma})$ au même point de l'espace-temps ($z = \gamma$), on a la décomposition

$$\dot{\gamma} = e\dot{z} + p$$

L'analogie avec l'approche Newtonienne (c.f. note de bas de page 1) montre que $e = m/(1 - \|v\|^2)^{1/2}$ et $p = mv/(1 - \|v\|^2)^{1/2}$, donc dans le cas où $\|v\| \ll 1$, $e \approx m + \frac{1}{2}m\|v\|^2$ et $p \approx mv$, ce qui légitime d'appeler e et p l'énergie et les moments de γ mesurés par z . Grâce à la formule de Pythagore, et en revenant en unités, on trouve :

$$-m^2c^4 = -e^2 + \|p\|^2c^2$$

Flots, faisceaux, tenseur énergie-impulsion, tenseur de charge-courant Plus généralement, on s'intéresse aux flots et faisceaux de particules, et aux modèles de matières. Un flot de particules (P, η) de masses m est une fonction de densité η et un champ de vecteur P , dit de moment-énergie, tel que chaque courbe intégrale de P est une particule de masse m . On lui associe un tenseur, dit énergie-impulsion : $\hat{T} = \eta P \otimes P$ (on notera dans la suite $\hat{\cdot}$ et $\tilde{\cdot}$ la dualité sur les tenseurs induite par la métrique g , c.f. annexe). Un faisceau de particules est un flot de particule avec en plus une charge électrique e , auquel on associe la densité de charge-courant : $J = e\eta P$. Un modèle de matière est une collection $\mathcal{M} = \{(m_A, e_A, P_A, \eta_A)\}$ de faisceaux de particules, son tenseur énergie-impulsion est $\hat{T} = \sum_A \eta_A P_A \otimes P_A$ et sa densité de charge-courant est $J = \sum_A e_A \eta_A P_A$.

Fluides relativistes Dans le cas où il existe des interactions entre les particules telles qu'elles se comportent comme un fluide et non pas comme un ensemble de particules indépendantes, le tenseur énergie-impulsion est de la forme $T_{jk} = (\rho + p)u_j u_k + p g_{jk}$, où u est la 4-vitesse du fluide (telle que $g(u, u) = -1$), p est la pression du fluide et ρ sa densité telles que mesurées par un observateur de vitesse u .

Modèle relativiste, champ électromagnétique Un modèle relativiste est un triplet (M, \mathcal{M}, F) , où $F \in \Omega^2(M)$ est une 2-forme décrivant un champ électromagnétique³. On dit que ce modèle satisfait les équations de Maxwell si F est fermée et $div \hat{F} = 4\pi J$, J la charge-courant du modèle de matière \mathcal{M} . Le tenseur énergie-impulsion du champ électromagnétique est $E_{ij} = \frac{1}{4\pi} [F_{im} F_j^m - \frac{1}{4} g_{ij} F^{mn} F_{mn}]$.

3.1.3 Les équations d'Einstein

Le modèle relativiste (M, \mathcal{M}, F) satisfait l'équation d'Einstein du champ si

$$G = 8\pi\kappa(T + E)$$

où G est le tenseur d'Einstein $G = Ric - \frac{1}{2}Sg$, T est le tenseur énergie-impulsion de la matière et E et le tenseur énergie-impulsion du champ électromagnétique.

Le problème mathématique général est de résoudre en (M, g, F) , c'est-à-dire trouver les variétés Lorentzienne avec champ électromagnétique (de Maxwell) satisfaisant cette équation. Plus précisément, il s'agit de trouver les *champs gravitationnels*, qui seuls ont une réalité physique puisque la généralité de l'équation d'Einstein implique l'existence d'une infinité de solutions à partir d'une seule : le problème ne peut donc se poser qu'à difféomorphisme près.

³Une 2-forme de courbure, c.f. le chapitre sur le modèle standard.

3.2 Existence et unicité des solutions

3.2.1 Problème de Cauchy

La question du développement d'une solution est le suivant : étant donnée une variété S de dimension 3, les données initiales

$$\overset{\circ}{g}_{ij} : S \rightarrow \mathbb{R} \quad \text{et} \quad \overset{\circ}{k}_{ij} : S \rightarrow \mathbb{R} \quad \text{pour} \quad i, j = 1 \dots 3$$

on recherche un espace-temps (M, g) , un plongement de S dans M en une sous-variété de genre espace (on identifiera S et son plongement), un système de coordonnées $(x_0, x') \in \mathbb{R} \times \mathbb{R}^3$ au voisinage de S tel que $S = \{x_0 = 0\}$ et

$$g_{ij}|_S = \overset{\circ}{g}_{ij} \quad \text{et} \quad \frac{\partial g_{ij}}{\partial x_0}|_S = \overset{\circ}{k}_{ij} \quad (3.1)$$

Le calcul de $G_k^0|_S$ montre que ces termes sont fonction de $g_{ij}|_S$, $\frac{\partial g_{ij}}{\partial x_0}|_S$, et leur dérivées en espace. Il s'en suit une condition nécessaire de cohérence sur les données initiales :

$$\mathcal{G}_k^0(\overset{\circ}{g}_{jk}, D_S^2 \overset{\circ}{g}_{jk}, \overset{\circ}{k}_{jk}, \nabla_S \overset{\circ}{k}_{jk}) = 0$$

3.2.2 Existence locale, coordonnées harmoniques

Dans le cas d'un espace-temps vide, $G = 0$ implique

$$Tr(Ric) - \frac{1}{2}S \cdot Tr(g) = (1 - \frac{1}{2}Tr(g))S = 0$$

et donc les équations d'Einstein se réduisent à $Ric = 0$. Or on a

$$Ric_{jk} = -\frac{1}{2}g^{lm}\partial_l\partial_m g_{jk} + \frac{1}{2}(g_{jl}\partial_k + g_{kl}\partial_j)\lambda^l + H_{jk}(g, \nabla g)$$

où

$$\lambda^l = \frac{1}{2}g^{jk}g^{lm}(\partial_k g_{jm} + \partial_j g_{km} - \partial_m g_{jk}) = -\square x_l$$

avec l'opérateur de Laplace-Beltrami

$$\square u = g^{jk}\partial_j\partial_k u - \lambda^l\partial_l u$$

Si l'on considère des coordonnées solutions de $\square x = 0$, dites harmoniques, on trouve donc que la métrique satisfait au système hyperbolique quasi-linéaire

$$g^{lm}\partial_l\partial_m g_{jk} - 2H_{jk}(g, \nabla g) = 0$$

avec les données initiales 3.1.

On construit alors un développement à partir de la métrique

$$\tilde{g}_{jk}(x_0, x') = \overset{\circ}{g}_{jk}(x') + x_0 \overset{\circ}{k}_{jk}(x')$$

définie au voisinage de S . La clef est que les coordonnées harmoniques pour \tilde{g} sont harmoniques pour toute métrique g telle que

$$g_{jk} - \tilde{g}_{jk} = O(y_0^2)$$

Par conséquent, il suffit de chercher les coordonnées harmoniques pour \tilde{g} , puis de résoudre en g un système hyperbolique.

3.2.3 Unicité globale

Il est possible de montrer un théorème d'unicité pour les éventuelles solutions satisfaisant certaines conditions de régularité.

Un espace-temps est dit : *régulier* s'il est muni d'une orientation temporelle globale et si la métrique est régulière au voisinage de tout point, *complet vers le futur* si toutes les géodésiques temporelles peuvent être prolongées indéfiniment vers le futur, *globalement hyperbolique* si l'ensemble des chemins joignant deux points quelconques est compact. Un sous-ensemble d'un espace-temps est dit *compact vers le passé* si l'ensemble des chemins temporels qui en sortent et aboutissent à un point x est compact quelque soit x .

On a alors [1] :

Théorème 2. *Si (V, g) et (V', g') sont deux espaces-temps solutions de l'équation d'Einstein, réguliers et globalement hyperboliques, (V', g') complet vers le futur, et tels que (\bar{A}) dénotant le complémentaire*

$$P \subset V, P' \subset V' \text{ compacts vers le passé} \Rightarrow \exists f : \bar{P} \rightarrow \bar{P}' \text{ telle que } f_*g = g'$$

alors il existe une isométrie de (V, g) sur (D', g') ou D' est un ouvert de V' .

3.3 Solutions physiques

3.3.1 Espace-temps de Schwarzschild

L'espace-temps de Schwarzschild est une solution à symétrie sphérique des équations d'Einstein. Une variété Lorentzienne est dite à symétrie sphérique s'il existe une action efficace de $SO(3)$ comme groupe d'isométries de la variété. En supposant que les orbites de cette action sont de genre espace, M est difféomorphe localement à $]a, b[\times]c, d[\times S^2$ et la métrique prend la forme :

$$ds^2 = -C(r, t)dt^2 + D(r, t)dr^2 + 2E(r, t)drdt + F(r, t)d\omega^2$$

où $d\omega^2$ est la métrique standard sur la sphère unité de \mathbb{R}^3 . L'expression de cette dernière et la diagonalisation de la forme quadratique ds^2 mène à la formulation :

$$ds^2 = -e^{\nu(r, t)}dt^2 + e^{\lambda(r, t)}dr^2 + r^2(d\varphi^2 + \sin^2\varphi d\theta^2)$$

Dans le cas d'un espace vide, i.e. $G = 0$, le calcul de la courbure de Ricci et de la courbure scalaire permettent d'établir que la métrique doit être de la forme (pour un certain K)

$$ds^2 = -\left(1 - \frac{K}{r}\right)dt^2 + \left(1 - \frac{K}{r}\right)^{-1}dr^2 + r^2d\omega^2$$

La singularité qui semble exister pour $K = r$ est en fait le résultat d'un mauvais choix de coordonnées et n'apparaît pas avec un choix plus judicieux, celui des coordonnées de Eddington-Finkelstein. Si $K > 0$, les géodésiques de cette métrique sont une approximation (pour r grand et une petite vitesse) des solutions de l'équation de Newton et pour $K < 0$ de l'équation de Coulomb.

3.3.2 Espace-temps stationnaires et statiques

Un espace-temps (M, g) est dit stationnaire s'il existe un champ de Killing⁴ Z de genre temps sur M . Ce champ permet de construire une fibration $p : M \rightarrow S$ sur une variété de dimension 3 dont les fibres sont les courbes intégrales de Z . Ce fibre a une structure de \mathbb{R} -fibre principal, et sa base S hérite naturellement d'une structure Riemannienne. On peut enfin doter p d'une connexion par :

$$\omega(X, Y) = P_0 [X, Y]$$

ou P_0 est la projection orthogonale en chaque point x sur le sous-espace de $T_x M$ engendré par Z_x .

On cherche alors les équations satisfaites par la projection $\sigma(s) = p \circ \gamma(s)$ dans l'espace S d'une géodésique de M . Dans ce but, on montre que

$$\nabla_Z Z = -\frac{1}{2} \text{grad} \langle Z|Z \rangle$$

et en notant β la forme anti-symétrique telle que

$$\omega(T_1, U_1) = \langle \beta(T_1)|U_1 \rangle Z, \quad \forall (T_1, U_1) \text{ orthogonaux à } Z$$

on trouve alors

$$\nabla_{\sigma'}^S \sigma' = \frac{1}{2} C_1^2 \text{grad} \Phi^{-1} - \frac{1}{2} C_1 \beta(\sigma')$$

avec $C_1 = \langle \gamma'|Z \rangle$ et $\Phi = -\langle Z|Z \rangle$

Un espace-temps stationnaire est dit statique si $\omega = 0$. Il peut s'écrire $M = S \times \mathbb{R}$ avec une métrique de la forme

$$ds^2 = -\Phi(x)dt^2 + g_S(dx, dx)$$

ou g_S est la métrique de S . Une géodésique de genre temps peut alors être reparamétrisée pour être de la forme $\gamma(t) = (x(t), t)$ avec une accélération

$$\nabla_{x'}^S x' = -\frac{1}{2} \text{grad} \Phi + \frac{1}{\Phi} \langle x'| \text{grad} \Phi \rangle x'$$

Si l'on considère sur S la métrique $g^\# = \Phi^{-1}g_S$, on peut se ramener par transformation conforme à la métrique suivante sur M

$$ds^2 = -dt^2 + g^\#(dx, dx)$$

ce qui équivaut à prendre $\Phi = 1$, et l'équation de l'accélération pour la projection des géodésiques de M devient donc

$$\nabla_{x'}^\# x' = 0$$

Comme la transformation conforme d'une métrique Lorentzienne conserve les géodésiques nulles, n'affectant que leur paramétrisation, ceci montre que les géodésiques de genre lumière se projettent sur les géodésiques espace de $(S, g^\#)$. Dans le cas des géodésiques plus généralement de type temps, on montre que

$$\nabla_{x'}^\# x' = \frac{C_2}{2C_1^2} \text{grad}^\# \Phi$$

ou $C_2 = \langle \gamma'| \gamma' \rangle$.

Remarque : l'exemple le plus classique d'espace-temps statique est l'espace-temps d'Einstein-De Sitter, avec $S = \mathbb{R}^3$ et $\Phi(x, t) = t^{-4/3}$. Son tenseur d'Einstein est $G = \frac{4}{3}t^{-2}dt \otimes dt$.

⁴On rappelle qu'un champ de vecteur X est dit champ de Killing si son fbt est une isométrie, i.e. $\mathcal{L}_X g = 0$

3.4 Singularités

Il est éclairant de présenter des cas assez naturels pour lesquels les solutions proviennent d'une singularité initiale ou en développent en un temps fini.

En ce qui concerne les singularités initiales, on ne fera que citer un théorème de Penrose-Hawking [9]

Théorème 3. *Si un espace-temps satisfait les deux conditions suivantes, alors il existe une géodésique sans extrémité vers le passé (définie sur $] - \infty, a)$) qui a une longueur finie :*

1. *M contient une hypersurface Σ de genre espace, dont les normales divergent à chaque point de Σ (i.e que le volume, le déterminant de la métrique, croît le long de chaque géodésique normale à Σ)*
2. *la matière dans l'espace-temps satisfait la condition "énergétique" : $\text{Ricc}(\dot{\gamma}, \dot{\gamma}) \leq 0$ pour toute géodésique γ de genre temps.*

En d'autres termes, s'il existe une hypersurface de genre espace telle que l'espace-temps y soit en tout point en expansion, et qu'il n'y a pas de singularité énergétique le long des géodésiques, alors il existe un big-bang : une origine, à distance finie dans le passé.

3.4.1 Effondrement gravitationnel

Considérons le cas d'un fluide dans un espace-temps $M = U \times S$, dont la métrique serait de la forme :

$$ds^2 = -dt^2 + R(t)^2 g^S$$

où g^S est une métrique de courbure constante sur S . Pour un fluide de densité et pression uniforme, on a $\rho = \rho(t)$ et $p = p(\rho)$. Les équations d'Einstein, après calcul de la courbure de Ricci et de la courbure scalaire, prennent alors la forme (K est la courbure sectionnelle supposées constante, $S_S = 6K$)

$$(R')^2 + K = \frac{8}{3}\pi\kappa\rho R^2 \quad (3.2)$$

$$2RR'' + (R')^2 + K = -8\pi\kappa p(\rho)R^2 \quad (3.3)$$

De ces équations on peut déduire les identités

$$\frac{d}{dt}(\rho R^3) = -p \frac{d}{dt}(R^3) \quad , \quad \frac{dR}{R} = -\frac{1}{3} \frac{d\rho}{\rho + p(\rho)}$$

et donc ρ comme fonction de R . Eq. 3.2 peut-être alors être vue comme une loi de conservation par rapport à la dynamique qui apparaît dans la différence de eq. 3.2 et eq. 3.3 :

$$3R'' = -4\pi\kappa(\rho + 3p)R$$

L'étude des courbes de niveau de $F(V, R) = \frac{1}{2}V - \frac{4}{3}\pi\kappa\rho R^2$ montre finalement que si $V(t_0) < 0$ pour un certain t_0 , alors $R'(t)$ est borné inférieurement et donc $R(t)$ doit devenir nul en un temps fini ultérieur à t_0 , et que si $V(t_0) > 0$, $R(t)$ à du s'annuler en un temps antérieur à t_0 .

Chapitre 4

Théories des champs : espace-temps, création et annihilation de particules

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, à la rendre relativiste puis à la quantifier une seconde fois. L'intérêt essentiel de la théorie quantique des champs est de procurer un formalisme unique quelque soit le nombre de particules, ce qui est nécessaire puisqu'en relativité, le nombre de particules n'est pas conservé.

Les sources principales pour ce chapitre sont [2, 4, 5, 8, 6].

4.1 Point de départ : le modèle relativiste de l'électron

4.1.1 Équation de Klein-Gordon

La formulation issue de la 1^{ère} quantification attribue a priori un rôle distinct à l'espace et au temps, hérité de son origine Hamiltoniennes classiques. En effet, l'étude d'un oscillateur classique libre montre que $H = \sum_{\mu} p_{\mu}^2/2m$ et la quantification donne

$$\mathcal{Q}(H) = i\hbar \frac{\partial}{\partial t} \quad \text{et} \quad \mathcal{Q}(p_{\mu}) = -i\hbar \partial_{\mu}$$

d'où l'équation de Schrödinger :

$$\sum_{\mu} -\frac{\hbar^2}{2m} \partial_{\mu}^2 \psi = i\hbar \frac{\partial}{\partial t} \psi$$

qui fait apparaître clairement une asymétrie entre les variables de temps et d'espace. Si l'on en cherche une version relativiste, il est naturel de considérer l'opérateur hamiltonien issu de l'analogie avec la formule de la relativité $E^2 = p^2 c^2 + m^2 c^4$, et donc :

$$\sqrt{\sum_{\mu} -\hbar^2 c^2 \partial_{\mu}^2 + m^2 c^4} \psi = i\hbar \frac{\partial}{\partial t} \psi \quad (4.1)$$

Cette expression étant difficile à analyser, on a d'abord travaillé avec son carré, l'équation de Klein-Gordon, qui est avec la convention $c = \hbar = 1$:

$$(\partial_t^2 - \sum_{\mu} \partial_{\mu}^2 + m^2)\psi = 0$$

où μ parcourt cette fois les indices de temps et d'espace.

4.1.2 Équation de Dirac

L'équation de Dirac est le résultat d'une volonté de travailler directement avec l'équation 4.1, plutôt qu'avec son carré, tout en conservant une équation linéaire. Cela est possible si le terme de gauche est un carré parfait, i.e. (toujours avec $c = \hbar = 1$)

$$\sum_{\mu} p_{\mu}^2 + m^2 = \left(\gamma_0 m + \sum_{\mu} \gamma_{\mu} p_{\mu} \right)^2$$

La résolution symbolique de cette identité mène aux équations

$$\gamma_{\mu}^2 = Id \quad \forall \mu \quad , \quad \gamma_{\mu} \gamma_{\nu} + \gamma_{\nu} \gamma_{\mu} = 0 \quad \forall \mu \neq \nu$$

La satisfaction de ces contraintes requiert de travailler au minimum avec des matrices 4×4 , avec par exemple

$$\gamma_0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \quad , \quad \gamma_{\mu} = \begin{pmatrix} 0 & \sigma_{\mu} \\ \sigma_{\mu} & 0 \end{pmatrix}$$

où les σ_i sont les matrices de Pauli, base de l'algèbre de Lie $\mathfrak{su}(2)$. L'équation de Dirac est alors

$$\left(i \sum_{\mu} \gamma^{\mu} \partial_{\mu} - m \right) \psi = 0$$

L'utilisation de matrices suppose que l'Hamiltonien n'opère plus sur une simple fonction d'onde, mais sur un vecteur d'onde. Il est dès lors difficile de continuer à faire l'économie du formalisme des fibrés, puisque les objets de base deviennent des champs vectoriels, susceptibles de présenter de riches symétries internes dont on voudrait tirer partis.

4.2 Théorie des champs

4.2.1 Champs classiques : théorème de Noether

La généralisation triviale de l'approche Lagrangienne aux champs se fait en considérant non plus des trajectoires $t \mapsto q(t) \in \mathbb{R}^d$ mais des champs $(t, q) \in \mathbb{R}^{d+1} \mapsto \varphi(t, q) \in \mathbb{R}^n$. De même que pour la trajectoire classique, cette approche suppose donnée, en chaque point de l'espace-temps, une base pour laquelle les mesures φ_i représentent les coordonnées du véritable champ physique. Un point de vue plus intrinsèque est donné par les fibrés, qui sera celui pris dans la section "champs de jauge", mais il est utile de rappeler d'abord quelques résultats physiques historiquement liés à la première approche.

Pour un Lagrangien de la forme

$$S(\varphi) = \int_U \mathcal{L}(\varphi_i(x), \partial_\mu \varphi_i(x)) dx \quad , U \subset X$$

les équations d'Euler-Lagrange deviennent :

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_i)} \right) = \frac{\partial \mathcal{L}}{\partial \varphi_i}$$

Dans le cas des particules, une symétrie était un difféomorphisme de \mathcal{M} dans lui-même dont l'application tangente laissait l'action invariante. L'action étant cependant définie sur les trajectoires, il s'agissait en fait de manière plus essentielle d'un difféomorphisme de l'ensemble des trajectoires respectant la projection sur le temps. Ce point de vue s'étend naturellement pour les champs, et l'on définit une symétrie comme une transformation :

$$S(\varphi) = [x \mapsto T(f(x))\varphi(f(x))]$$

où $T(y)$ est un automorphisme de \mathbb{R}^n .

S'il existe un groupe continu de m telles symétries, le théorème de Noether généralise alors la conservation des moments et de l'énergie. En montrant que

Théorème 4. *Les m courants*

$$j_k^\mu = \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_i)} \partial_\alpha \varphi_i - \delta_\alpha^\mu \mathcal{L} \right) X_k^\alpha - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_i)} \Phi_{ik}$$

sont conservatifs (i.e. $\partial_\mu j_k^\mu = 0$), où $X_k(x) = \partial_{\omega_k} f(x)$ et $\Phi_k(x) = \partial_{\omega_k} T(x)$ et les ∂_{ω_k} sont engendrés par l'action du groupe.

il implique en effet que les charges

$$Q_k(t) = \int_{x_0=t} dx j_k^0(x)$$

sont conservées au cours du temps.

4.2.2 Champs quantiques

Dans le cas d'un espace-temps statique, la quantification des champs peut se faire sur le même principe que la 1^{ère} quantification. Par une analogie avec la particule, on définit la généralisation des moments conjugués :

$$\pi_i = \frac{\partial \mathcal{L}}{\partial(\partial_0 \varphi_i)} \quad , \quad i = 1 \dots n$$

et on introduit le crochet de Poisson (construit par extension directe du cas fini avec l'idée que x n'est qu'un indice) :

$$\{f, g\} = \int dx \left(\sum_i \frac{\partial f}{\partial \varphi_i(x)} \frac{\partial g}{\partial \pi_i(x)} - \frac{\partial g}{\partial \varphi_i(x)} \frac{\partial f}{\partial \pi_i(x)} \right)$$

et l'on trouve, dans le cas de la charge $Q = H$ associée à la translation en temps ($X_k^\mu \partial_\mu = \partial_0$ et $\Phi_k = 0$)

$$\{H, \varphi_i(x)\} = \partial_0 \varphi_i(x) \quad , \quad \{H, \pi_i(x)\} = \partial_0 \pi_i(x)$$

et l'algèbre des courants

$$\{\varphi_i(x), \pi_j(y)\}_{ET} = \delta_{ij} \delta(x - y) \quad , \quad \{\varphi_i(x), \varphi_j(y)\}_{ET} = \{\pi_i(x), \pi_j(y)\}_{ET} = 0$$

où ET signifie "Equal Time", i.e. $x_0 = y_0$.

Si les particules sont des bosons libres, on utilise le même principe de représentation que dans la 1^{ère} quantification, à ceci près qu'on ne travaille plus sur un espace de Hilbert. Pour rendre compte de la création et annihilation de particules, on est mené à considérer un ensemble encore plus grand, nommé espace de Fock. Étant donné un espace de Hilbert E décrivant une unique particule, il s'agit du sous-espace de

$$F = \bigoplus_n E_n \quad , \quad E_n = E \otimes \dots \otimes E$$

constitué par les éléments symétriques par permutation. On note ses éléments $|x_1 \dots x_n\rangle \in E_n$, et on considère la représentation donnée par (se restreignant à 1 seul champ) :

$$\begin{aligned} \mathcal{Q}(\varphi(x)) &: |x_1 \dots x_n\rangle \mapsto \sum_i (-1)^{i+1} \delta(x - x_i) |x_1 \dots \hat{x}_i \dots x_n\rangle \\ \mathcal{Q}(\pi(x)) &: |x_1 \dots x_n\rangle \mapsto |x, x_1, \dots, x_n\rangle \end{aligned}$$

L'intérêt essentiel de cette quantification est que le calcul des valeurs moyennes (c.f. section "approche statistique") va faire intervenir l'intégrale

$$Z = \int d\psi \exp\left(-\frac{i}{\hbar} \int dx \mathcal{L}(\psi(x), \partial\psi(x))\right)$$

et que l'on va ainsi pouvoir traiter, dans un unique formalisme, des systèmes ayant un nombre variable de particules, et donc de rendre compte des *processus* fondamentaux des créations et annihilations de particules au cours des interactions[4].

4.2.3 Brisure spontanée de symétrie, particules sans-masses et particules massives

Considérons, pour un champs classique scalaire complexe, un Lagrangien de la forme

$$\mathcal{L} = \partial^\mu \varphi^* \partial_\mu \varphi - \mu^2 |\varphi|^2 + \lambda |\varphi|^4$$

Si $\mu^2 > 0$, le seul champ invariant par transformation de Lorentz qui minimise ce Lagrangien est le champ nul. Par contre, si $\mu^2 < 0$, on trouve un ensemble de solutions de la forme

$$\varphi = \sqrt{\frac{-\mu^2}{2\lambda}} e^{i\theta}$$

Cela correspond à une brisure spontanée de symétrie dans le sens où l'une de ces solutions doit être choisi par le champs. Prenant l'un de ces états comme état de référence, $\varphi_0 = v/\sqrt{2}$, on fait le changement de variable

$$\varphi = \frac{\sqrt{2}}{2} (v + \sigma + i\eta)$$

et le Lagrangien prend alors la forme

$$\mathcal{L} = \frac{1}{2} \partial^\mu \sigma \partial_\mu \sigma - \lambda v^2 \sigma^2 + \frac{1}{2} \partial^\mu \eta \partial_\mu \eta + O(\text{champs}^3)$$

En interprétant les termes de plus grand degré comme des termes d'interaction, on se retrouve face au Lagrangien de deux champs σ et η de Klein-Gordon, de masses respectives $v\sqrt{2\lambda}$ et 0. Plus généralement, on peut montrer que chaque brisure de symétrie peut être traitée comme la perturbation d'une interaction impliquant un boson de masse nulle, le boson de Goldstone. Il s'agit à l'heure actuelle du seul mécanisme permettant d'attribuer des masses aux particules fondamentales : l'introduction de termes massifs directement dans le Lagrangien conduit en effet à des problèmes de renormalisation, alors qu'un Lagrangien invariant par transformation de jauge mène nécessairement à une théorie renormalisable [11].

4.3 Le modèle standard : équations des champs de jauge

4.3.1 Dynamique libre des fermions et des hadrons

Les dynamiques libres des fermions Ψ et des hadrons Φ sont données par les équations précédemment évoquées de Dirac et de Klein-Gordon, qui sont les équations d'Euler-Lagrange pour les Lagrangiens :

$$\mathcal{L}_{dyn,f}(\Psi) = \frac{i}{2} (\bar{\Psi} \gamma^\mu \partial_\mu \Psi - \partial_\mu \bar{\Psi} \gamma^\mu \Psi) - m \bar{\Psi} \Psi$$

et

$$\mathcal{L}_{dyn,h}(\Phi) = \partial_\mu \tilde{\Phi} \partial_\mu \Phi - m \tilde{\Phi} \Phi$$

Ces deux équations sont insatisfaisantes parce qu'elles font intervenir a priori des termes de masse extrinsèques, problème qui sera résolu par le mécanisme de Higgs dans le modèle unifié (c.f. interactions électrofaibles). Mais surtout, elles ne sont pas géométriques puisque, excepté dans le cas scalaire, il n'y a pas de dérivée canonique des sections comme implicitement supposé dans les notions $\partial_\mu \Psi$ et $\partial_\mu \Phi$.

4.3.2 Dynamique libre des bosons : équation de Yang-Mills

La dynamique libre des bosons découle alors de la recherche d'une dérivation "minimale" des sections. Les particules étant modélisées comme des sections de fibré linéaires, il est naturel de remonter à la source des symétries existant sur ces fibrés et donc de s'intéresser à la courbure du fibré principal au dessus de V dont le groupe de structure est le groupe des symétries. Étant donné une variété Lorentzienne (V, g) et un fibré principal sur V , on définit l'action de Yang-Mills de la connexion D :

$$\mathcal{S}_{YM}(D) = \int_M \mathcal{L}_{YM}(D) v = \int_V \|F\|^2 v$$

où F est la forme de courbure associée à D , v est la forme de volume associée à g , et $\|F\|$ est la norme induite sur les formes par la norme sur \mathfrak{g} dérivée du produit scalaire : $(A|B) = -tr(AB)$.

Les équations de Yang-Mills sont les équation d'Euler-Lagrange pour cette action. On trouve :

$$D^* F = 0 = D * F$$

où $D^* = - * D *$ et $* : T^{(0,k)} \rightarrow T^{(0,n-k)}$ est l'opérateur de Hodge tel que $\theta \wedge * \phi = (\theta | \phi) v$ sur chaque fibre. En dimension 4, $* : T^{(0,2)} \rightarrow T^{(0,2)}$ est involutif, et donc les 2-formes se décomposent en deux parties dites auto-duales et anti-auto-duales sur les 2 sous-espaces propres de $* : T^{(0,2)} = T_+^{(0,2)} \oplus T_-^{(0,2)}$. On a alors

$$\|F\|^2 = \|F_+\|^2 + \|F_-\|^2 \quad \text{et} \quad \text{tr}(F \wedge F) = -\|F_+\|^2 + \|F_-\|^2$$

ce qui permet de borner l'action de Yang-Mills en fonction de la classe caractéristique du fibré (c.f. annexe) :

$$\mathcal{S}_{YM}(D) = \int_M (\|F_+\|^2 + \|F_-\|^2) v \geq \left| \int_M (\|F_+\|^2 - \|F_-\|^2) v \right| = \int \text{tr}(F \wedge F) v$$

avec égalité si et seulement si F est auto-dual ou anti-auto-dual. Si cette borne est atteinte, les équations de Yang-Mills se réduisent donc à l'équation

$$F = \pm * F$$

4.3.3 Interactions des fermions et des mesons avec les bosons de jauge

On peut ainsi retrouver un sens géométrique pour les Lagrangiens $\mathcal{L}_{dyn,f}$ et $\mathcal{L}_{dyn,h}$ lorsqu'on considère l'interaction des fermions et mesons avec les forces fondamentales, ces dernières déterminant la connexion absente de la dynamique libre. Étant donné un spineur S sur V , un espace de multiplicité E et une représentation ρ dans le groupe unitaire de E du groupe de structure G du fibré principal précédent, les fermions et les mesons sont des sections des fibrés $S \otimes E$ et $V \otimes E$, sur lesquels on définit les connexions suivantes, images de la connexion de Yang-Mills sur le fibré principal :

$$D\Phi = d\Phi + SA\Phi$$

pour les mesons, où A est le potentiel de Yang-Mills (c.f. section "synthèse"), $S = \rho'(1) : \mathfrak{g} \rightarrow \mathfrak{u}$ envoie l'algèbre de Lie de G dans celle du groupe unitaire de E , et

$$\mathcal{D}\Psi = \gamma^\mu D_\mu \Psi$$

pour les fermions.

Le Lagrangien correspondant se déduit directement des précédents en remplaçant les différentielles ∂_μ par D et \mathcal{D} :

$$\mathcal{L} = \frac{1}{4} F_{\lambda\mu}^a F_a^{\lambda\mu} + \frac{i}{2} (\bar{\Psi} \mathcal{D}\Psi - \mathcal{D}\bar{\Psi} \Psi) + D\tilde{\Phi}D\Phi - m_2 \tilde{\Phi}\Phi - m_1 \bar{\Psi}\Psi$$

4.3.4 Interactions entre fermions et mesons

Cette interaction est prise en compte par l'ajout au Lagrangien \mathcal{L} d'un terme d'interaction \mathcal{L}_{int} . Selon les particules et les modèles considérés, ce terme peut prendre plusieurs formes.

Modèle de Yukawa, couplage minimal

Le modèle de Yukawa des interactions nucléaires cherche à rendre compte des interactions faibles, i.e. des interactions entre les particules W^+ , W^- , les doublets de leptons (ν_e, e) , (ν_μ, μ) et (ν_τ, τ) et les trois pions Π_0, Π_+, Π_- . Le groupe de structure du fibré est $SU(2)$, le champ de spineurs est un doublets $\Psi = (\Psi_p, \Psi_n)$ et le champs de scalaires un triplet $\Phi = (\Phi_0, \Phi_1, \Phi_2)$. Le couplage est

$$\mathcal{L}_{int} = \bar{\Psi} C \Psi \Phi + \tilde{\Phi} \bar{\Psi} \tilde{C} \Psi + \mu (\tilde{\Phi} \Phi)^2$$

où $C = k\gamma_5\tau$ est une carte trinéaire de $\mathbb{C}^l \times \mathbb{C}^l \times \mathbb{C}^m$ dans S_1^1 . Dans ce cas, les équations d'Euler sont les équations 4.2 avec

$$H = i(C\Phi + \tilde{\Phi}\tilde{C})\Psi \quad \text{et} \quad K = -\bar{\Psi}\tilde{C}\Psi - \mu(\tilde{\Phi}\Phi)\Phi$$

Modèle de Glashow-Weinberg-Salam, interaction électrofaibles

Le modèle de Weinberg-Salam est une théorie unifiant les interactions faibles et électromagnétiques. Le groupe de structure du fibré est $SU(2) \times U(1)$, le champ de spineurs est un triplet $\Psi = (L_1, L_2, R)$ et le champs de scalaires est un doublet $\Phi = (\Phi_1, \Phi_2)$. Le couplage est ($L \equiv (L_1, L_2)$):

$$\mathcal{L}_{int} = (\bar{L}\Phi)R + \bar{R}(\tilde{\Phi}L) + \mu(\tilde{\Phi}\Phi)^2$$

Ce modèle conduit à considérer a priori trois bosons de jauge pour $SU(2)$ (les bosons W_+ , W_- et W^0), et un pour $SU(1)$ (le boson B). En fait, W_0 et B ont la même constante d'échelle et donnent lieu, dans le cadre d'une brisure de symétrie sur le multiplet Φ en l'absence de fermions ($\Psi = 0$), à considérer qu'ils sont en fait des états mixtes de deux particules. On pose

$$\Phi(x) = \begin{pmatrix} \eta_1(x) + i\eta_2(x) \\ \frac{1}{\sqrt{2}}(v + \sigma(x) + i\eta_3(x)) \end{pmatrix}$$

ou $v/\sqrt{2}$ est le potentiel minimal (c.f. section "brisure de symétrie"), et on trouve qu'il existe un choix de jauge qui fait disparaître les champs scalaires η_i pour finalement arriver à

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}\partial^\mu\sigma\partial_\mu\sigma - \frac{1}{2}m_H^2\sigma^2 \\ & - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ & - \frac{1}{4}\tilde{F}_{W\mu\nu}F_W^{\mu\nu} + \frac{1}{2}m_W^2\tilde{W}_\mu W^\mu \\ & - \frac{1}{4}F_{Z\mu\nu}F_Z^{\mu\nu} + \frac{1}{2}m_Z^2Z_\mu Z^\mu \\ & + O(champs^3) \end{aligned}$$

dans la base de $\mathfrak{su}(2) \oplus \mathfrak{su}(1)$ définie à partir de la base canonique par la "rotation de Weinberg" (où A est le potentiel de F , Z celui de F_Z):

$$\begin{pmatrix} Z \\ A \end{pmatrix} = \begin{pmatrix} c_w & s_w \\ -s_w & c_w \end{pmatrix} \begin{pmatrix} W^0 \\ B \end{pmatrix}$$

Ainsi on arrive grâce au boson de Higgs σ à un modèle comprenant trois bosons massifs, W_+ , W_- et Z , un boson sans masse γ . Pour que cette théorie soit validée, il faudrait donc déceler une nouvelle particule de masse m_H , ce qui est actuellement tenté au CERN, sur le LHC (Large Hadron Collider).

Chromodynamique, interactions fortes

La chromodynamique unifie les interactions électro-faibles avec les interactions fortes, médiées par le champ de jauge des gluons. Le groupe de structure du fibré est $SU(3)$, le champ de spineurs est un triplet (ou plus), il n'y a pas de champs scalaires ni de Lagrangien d'interactions entre les quarks.

4.4 Modèle standard : synthèse et couplage avec la gravitation

4.4.1 Synthèse, théorèmes d'existence

Notations

Les données d'une théorie de champ de jauge sont constituées par un espace-temps, i.e. une variété V de dimension $n + 1$ avec une métrique g de signature $(+, -, -, \dots, -)$, et une structure de spin σ sur (V, g) (c.f. annexe). C'est dans le rôle joué par cette métrique et cette structure spin dans les équations qui vont suivre que s'effectue un premier couplage du modèle standard avec la gravitation, mais il s'agit a priori d'un couplage exogène car présent dans les *données* du modèle. Mais la métrique est ici supposée *précéder* les interactions fondamentales, lorsque les équations d'Einstein sont précisément des équations sur l'interaction entre la gravité et les électromagnétique et de matière.

Une *connexion* de Yang-Mills est la 1-forme de connexion ω associée à une connexion D sur un fibré principal P , de base V et de groupe de structure G . Étant donnée une section s de P , qu'on nomme une jauge, ω est représentée par le *potentiel* de Yang-Mills $A = s^*\omega$. Le *champ* de Yang-Mills est la représentation de la 2-forme de courbure Ω :

$$F = s^*\Omega = dA + \frac{1}{2}[A, A]$$

Un champ de multiplet de scalaires Φ est la section d'un fibré linéaire associé par une représentation unitaire ρ de G . La dérivée covariante selon l'image de la connexion ω dans ce fibré est définie par

$$\hat{\nabla}\Phi = d\Phi + SA\Phi$$

où $S = \rho'(1) : \mathfrak{g} \rightarrow \mathfrak{u}$ transporte l'algèbre de Lie de G sur celle du group unitaire. On note $\tilde{\Phi}$ l'adjoint hermitien pour la métrique g .

Un champ de multiplets de spineurs Ψ est la section d'un fibré linéaire associé, avec comme fibre $\mathbb{C}^k \times \mathbb{C}^l$, où le premier facteur est l'espace de représentation fondamental du groupe $Spin(n + 1)$ (c.f. annexe) et le second est l'espace d'une représentation unitaire r de G . La dérivée covariante, selon l'image des connexions ω et σ dans ce fibré, est définie par

$$\hat{\nabla}\Psi = d\Psi + (\sigma, r'(1)A)\Psi$$

On définit l'adjoint de Dirac de Ψ par $\bar{\Psi} = \tilde{\Psi}\gamma_0$ (notations de la section "équation de Dirac"), et l'opérateur de Dirac, défini sur Ψ par (c.f. annexe pour une définition plus détaillée)

$$\hat{\nabla}\Psi = \gamma^\alpha \hat{\nabla}_\alpha \Psi$$

Équations

La dynamique d'une théorie de jauge est déterminée, étant donné un espace temps, une structure de spin, et un fibré principal, par la minimisation d'un Lagrangien de la forme

$$\mathcal{L}(D, \Phi, \Psi) = \frac{1}{4}F_{\lambda\mu}^\alpha F_a^{\lambda\mu} + \frac{1}{2}i(\bar{\Psi}\hat{\nabla}\Psi - \hat{\nabla}\bar{\Psi}\Psi) + \hat{\nabla}_\mu \tilde{\Phi} \hat{\nabla}_\mu \Phi + \mathcal{L}_{int}(\Phi, \Psi)$$

où le premier terme caractérise la courbure de la connexion, le second terme la dynamique libre des champs, et le dernier terme leurs interactions.

Les équations d'Euler prennent la forme

$$\hat{\nabla}\Psi = H \quad \text{et} \quad \hat{\square}\Phi = K \quad (4.2)$$

où l'opérateur onde $\hat{\square}$ est défini par

$$\hat{\square}\Phi = \hat{\nabla}^\lambda \hat{\nabla}_\lambda \Phi = g^{\lambda\mu} \hat{\nabla}_\lambda \hat{\nabla}_\mu \Phi$$

H et K sont déterminés par le Lagrangien d'interaction \mathcal{L}_{int} , et sont nuls en l'absence d'interaction.

Théorèmes d'existence

On peut citer deux théorèmes d'existence généraux sur ces équations [2] :

Théorème 5. Existence locale et d'unicité. *Supposons donnés, sur $S_0 = \{t = 0\} \subset V$, les valeurs de Φ , Ψ , A et $n \cdot \nabla\Phi|_{S_0}$, satisfaisant $\hat{d}iv(F \cdot n) = \hat{\nabla}_\lambda F \cdot n$. Alors il existe des solutions définies sur un voisinage de S_0 dont les données précédentes soient les restrictions à S_0 . De plus, pour des données de Cauchy équivalentes à transformation de jauge près, les solutions restent équivalentes à transformation de jauge près.*

Théorème 6. Existence globale *Les mêmes données pour un système invariant conforme donne lieu à une solution globale si les données sont assez régulières.*

4.4.2 Couplage avec la gravitation

On ne peut pas agglomérer les équations d'Einstein et les interactions fondamentales par la sommation directe de leurs Lagrangiens respectifs. Il faut tout d'abord découpler les deux problèmes en considérant la construction auxiliaire d'une structure et d'une connexion spin ω , dans laquelle est a priori instanciée l'équation de Dirac, puis se poser la question de son insertion dans le fibré tangent de l'espace-temps via le choix d'un vilbein (e_a). C'est sur la connexion spin et le choix du vilbein que porte alors le Lagrangien :

$$\mathcal{L} = -\frac{1}{2}e\mathcal{R} + \frac{1}{2}e\epsilon^{klmn} (\bar{\psi}_k \gamma_l D_m \psi_n - \psi_k \gamma_l D_m \bar{\psi}_n)$$

avec

$$\mathcal{R} = e_a e_b^m (\partial_n \omega_m^{ab} - \partial_m \omega_n^{ab} + \omega_m^{ac} \omega_{nc}^b - \omega_n^{ac} \omega_{mc}^b)$$

Cette construction est légitimée par le fait qu'elle redonne les équations d'Einstein lorsque les variables concernant les interactions fondamentales sont fixées, et les équations du modèle standard lorsque les variables concernant la gravitation sont fixées. Elle n'est toujours pas testée à l'heure actuelle.

Bibliographie

- [1] Y. Choquet-Bruhat. Théorème global d'unicité pour les solutions des équations d'einstein. *Bulletin de la SMF*, 1968.
- [2] Y. Choquet-Bruhat and D. Christodoulou. Existence of global solutions of the yang-mills, higgs and spinor field equations in $3 + 1$ dimensions. *Bulletin de la SMF*, 1981.
- [3] A. Einstein. Considerations concerning the fundamentals of theoretical physics. *Science*, 1940.
- [4] R. P. Feynman. *QED*. Princeton University Press, 1985.
- [5] D. S. Freed and K. K. Uhlenbeck. *Instantons and Four-Manifolds*. Springer-Verlag, 1984.
- [6] M. B. Green, J. H. Schwarz, and E. Witten. *Superstring theory*, volume 2. Cambridge Monographs on Mathematical Physics, 1987.
- [7] L. Landau and E. M. Lifshitz. *Mechanics, Course of Theoretical Physics, Volume I*. Pergamon Press, 1976.
- [8] H. B. Lawson and M.-L. Michelsohn. *Spin Geometry*. Princeton University Press, 1989.
- [9] R. Penrose. *Techniques of Differential Topology in Relativity*. CBMS-NSF Regional Conference Series in Applied Mathematics, 1972.
- [10] R. K. Sachs and H. Wu. *General Relativity for Mathematicians*. Springer-Verlag, 1977.
- [11] G. 't Hooft. Renormalizable lagrangians for massive yang-mills fields. *Nuclear Physics B*, 35 :167–88, 1971.

Chapitre 5

Annexes

5.1 Géométrie différentielle

5.1.1 Rappels et notations

Dualité métrique La donnée d'une métrique g sur une variété M définit une notion de dualité à l'intérieur de l'algèbre tensorielle $T(M)$. En particulier, à $\omega \in T^{(0,1)}(M)$ on associe naturellement $\hat{\omega} \in T^{(1,0)}(M)$ vérifiant $\omega = g(\hat{\omega}, \cdot)$, à $F = \omega_1 \otimes \omega_2 \in T^{(0,2)}(M)$ on associe $\tilde{F} = \hat{\omega}_1 \otimes \omega_2 \in T^{(1,1)}(M)$ et $\hat{F} = \hat{\omega}_1 \otimes \hat{\omega}_2 \in T^{(2,0)}(M)$.

Opérateur divergence Si n est la dimension de M , $T^{(0,n)}(M)$ est de dimension 1 et l'on appelle forme de volume l'élément $\Omega = \omega_1 \otimes \dots \otimes \omega_n$ où $(\hat{\omega}_i)$ est une base g -orthonormée de $T^{(1,0)}(M)$. Pour $X \in T^{(1,0)}(M)$, $d(i(X)\Omega) \in T^{(0,n)}(M)$ et l'on peut alors définir l'opérateur divergence $div : T^{(1,0)}(M) \rightarrow \mathcal{F}(M, \mathbb{R})$ qui vérifie $d(i(X)\Omega) = div(X)\Omega$. On peut étendre cet opérateur sur les tenseurs antisymétriques $C \in T^{(p,0)}(M)$ de manière à ce que $div : T^{(p,0)}(M) \rightarrow T^{(p-1,0)}(M)$ et $d(i(C)\Omega) = i(div(C))\Omega$.

Courbures Étant donné la connexion de Levi-Civita D associée à une métrique g , on définit l'opérateur de courbure $R_{XY} = D_X D_Y - D_Y D_X - D_{[X,Y]}$, le tenseur de courbure $R : (\omega, Z, X, Y) \mapsto \omega(R_{XY}Z)$, le tenseur de Ricci $Ric : (X, Y) \mapsto R(\omega^i, X, X_i, Y)$ (somme implicite, à partir d'une base g -orthonormale et sa duale), la courbure scalaire $S = Ric(X_i, X_i)$.

5.1.2 Variétés Lorentziennes

Orientation du temps Pour définir cette notion d'orientation du temps on commence par considérer le cas d'un espace de Lorentz (V, g) : un sous-espace $W \subset V$ a comme genre : spatial ssi g est définie positive sur W , lumière ssi g est positive semi-définie mais pas définie, temporel dans les autres cas. Un vecteur est causal ssi il n'est pas spatial. Dans le cas d'une variété Lorentzienne, ces notions s'appliquent naturellement sur le fibré tangent. On montre alors que l'ensemble des points temporels de TM est une variété à une ou deux composantes connexes. Lorsqu'il y en a deux, on dit que M est orientable et une orientation est le choix d'une des composantes comme étant le futur.

5.1.3 Classes caractéristiques

En se restreignant aux fibrés de groupe de structure G matriciel, un point de vue élémentaire sur la classification est le suivant. Étant donnée des sections locales $s_i : \mathcal{U}_i \rightarrow G$ d'un fibré P , on note $F_i = s_i^* \Omega$ les trivialisations locales de la 2-forme de courbure, et on considère les $2p$ -formes

$$P(F_i) : (X_1, Y_1, \dots, X_p, Y_p) \in (T\mathcal{U}_i)^{2p} \mapsto P(F_i(X_1, Y_1), \dots, F_i(X_p, Y_p)) \in \mathfrak{g}$$

P étant un polynôme homogène de degré p . Si ce polynôme est G -invariant, i.e.

$$P(Ad_g X_1, \dots, Ad_g X_p) = P(X_1, \dots, X_p)$$

alors on peut recoller les formes précédentes, puisque $P(F_i) = P(Ad_{g_{ij}} F_j) = P(F_j)$, pour définir un objet global sur P .

Considérant deux connexions ω et ω' , on définit la famille de connexions $\omega_t = \omega + t(\omega' - \omega)$, Du fait de l'identité de Bianchi, on montre que $dP(F) = 0$, et

$$dP(A' - A, F_t, \dots, F_t) = \frac{1}{p} \frac{dP(F_t)}{dt}$$

donc

$$P(F') - P(F) = d \int pP(A' - A, F_t, \dots, F_t) dt$$

Ceci implique que $P(F)$ est fermée et que sa classe de cohomologie ne dépend pas de la connexion.

Les polynômes invariants utilisés couramment sont les monômes issus de $\det(1 + \frac{iX}{2\pi})$ pour $G = GL_n(\mathbb{C})$ (classes de Chern), ou $\det(1 + \frac{X}{2\pi})$ pour $G = O_n(\mathbb{R})$ (classes de Pontrjagin). On trouve entre autre : $1, Tr(F), Tr F \wedge Tr F - Tr(F \wedge F), \dots, \det(F)$.

5.2 Spineurs

5.2.1 Algèbre de Clifford

L'algèbre de Clifford $Cl(V, q)$ est le quotient de l'algèbre tensorielle

$$T(V) = \sum_{r=0}^{\infty} \otimes^r V$$

par l'idéal bilatère engendré par les éléments de la forme $v \otimes v + q(v)$. On note π_q l'injection canonique. V s'immerge naturellement dans $Cl(V, q)$, i.e. $\pi_q|_V$ est injective. On a $q(v)1 = -v \cdot v$ et $v \cdot w + w \cdot v = -2q(v, w)$. Il s'agit du modèle canonique d'algèbre associative vérifiant la précédente relation : toute application linéaire f de V dans une algèbre associative (\mathcal{A}, \cdot) telle que

$$f(v) \cdot f(v) = -q(v)1$$

s'étend en un homomorphisme d'algèbre $\tilde{f} : Cl(V, q) \rightarrow \mathcal{A}$.

Il existe un unique antiautomorphisme t égal à l'identité sur V et une involution canonique α égale à moins l'identité sur V . La décomposition spectrale de α montre que $Cl(V, q)$ est une algèbre graduée :

$$Cl(V, q) = Cl^0(V, q) \oplus Cl^1(V, q)$$

Quelque soit la forme quadratique q , l'algèbre graduée associée à $Cl(V, q)$:

$$\mathcal{G} = \bigoplus_{r \geq 0} (\mathcal{F}_r / \mathcal{F}_{r-1}) \quad , \quad \mathcal{F}_r = \pi_q \left(\sum_{s \leq r} \otimes^s V \right)$$

est isomorphe à l'algèbre extérieure $\Lambda^* V$. Il y a un isomorphisme *d'espace vectoriel* $\Lambda^* V \rightarrow Cl(V, q)$ compatible avec la filtration des (\mathcal{F}_i) .

Si $V = V_1 \oplus^\perp V_2$, il y a un isomorphisme d'algèbre de Clifford

$$Cl(V, q) \rightarrow Cl(V_1, q_1) \hat{\otimes} Cl(V_2, q_2)$$

où $\hat{\otimes}$ est le produit tensoriel \mathbb{Z}_2 gradué dans lequel le produit est, pour b et a' de pur degré :

$$(a \otimes b) \cdot (a' \otimes b') = (-1)^{\deg(b)\deg(a)} (aa') \otimes (bb')$$

5.2.2 Groupes Pin et Spin

Le groupe multiplicatif des unités :

$$Cl^\times(V, q) = \{ \varphi \in Cl(V, q) : \exists \varphi^{-1} \text{ avec } \varphi^{-1} \varphi = \varphi \varphi^{-1} = 1 \}$$

contient tous les éléments de V tels que $q(v) \neq 0$, c'est un groupe de Lie de dimension 2^n . Il se représente sur le groupe des automorphismes de $Cl(V, q)$ par la représentation adjointe

$$Ad_\varphi(x) = \varphi x \varphi^{-1}$$

On a la relation, pour $q(v) \neq 0$

$$-Ad_v(w) = w - 2 \frac{q(v, w)}{q(v)} v$$

qui attire l'attention sur le sous-groupe $P(V, q)$ de $Cl^\times(V, q)$ généré par les éléments de V tels que $q(v) \neq 0$: la relation précédente montre qu'il se représente dans le groupe orthogonal de q par $\varphi \rightarrow Ad_\varphi$.

Le groupe $Pin(V, q)$ est le sous-groupe de $P(V, q)$ généré par les éléments de V tels que $q(v) = \pm 1$, le groupe $Spin(V, q)$ est défini par

$$Spin(V, q) = Pin(V, q) \cap Cl^0(V, q)$$

La représentation adjointe twisté est définie par

$$\widetilde{Ad}_\varphi(x) = \alpha(\varphi) y \varphi^{-1}$$

dont les éléments pairs du noyau commutent et les éléments impairs anti-commutent. On peut montrer par cette particularité que ce noyau est en fait exactement le groupe k^\times des éléments non nuls du corps k si q est non-dégénérée.

On définit la norme $N(\varphi) = \varphi \cdot t(\alpha(\varphi))$. Si $\tilde{P}(V, q)$ est l'ensemble des éléments pour lesquels V est stable par leur représentation adjointe, N définit un homomorphisme de $\tilde{P}(V, q)$ dans k^\times .

On peut montrer que $\widetilde{Ad}(Pin(V, q))$ et $\widetilde{Ad}(Spin(V, q))$ sont des sous-groupes normaux du groupe orthogonal de q , et le résultat essentiel suivant : il existe des séquences exactes

$$0 \rightarrow F \rightarrow Spin(V, q) \rightarrow SO(V, q) \rightarrow 1$$

$$0 \rightarrow F \rightarrow Pin(V, q) \rightarrow O(V, q) \rightarrow 1$$

où $F = \mathbb{Z}_2 = \{-1, 1\}$ si $\sqrt{-1} \notin k$, et $F = \mathbb{Z}_2 = \{\pm 1, \pm\sqrt{-1}\}$ sinon. On a $Spin(1) \simeq \mathbb{Z}/2\mathbb{Z}$, $Spin(2) \simeq U(1)$, $Spin(3) \simeq SU(2)$.

On note $Cl_{r,s}$ l'algèbre de Clifford $Cl(\mathbb{R}^{r+s}, q)$ où q est de signature (r, s) , $Cl_n = Cl_{n,0}$ et $Cl_n^* = Cl_{0,n}$. $Cl_{r,s}$ est générée en tant qu'algèbre par une base q-orthonormale (e_1, \dots, e_n) de \mathbb{R}^{r+s} , avec les relations $e_j e_k + e_k e_j = \pm 2\delta_{ij}$ selon que $i > r$ ou non. On voit donc que $Cl_1 \cong \mathbb{C}$ et $Cl_1^* \cong \mathbb{R} \oplus \mathbb{R}$. On note $Spin_{r,s}$ et $Pin_{r,s}$ les groupes spin et pin associés. On a immédiatement une décomposition

$$Cl_{r,s} = Cl_1 \hat{\otimes} \dots \hat{\otimes} Cl_1 \hat{\otimes} Cl_1^* \hat{\otimes} \dots \hat{\otimes} Cl_1^*$$

d'où on déduit que $dim Cl_{r,s} = 2^{r+s}$.

5.2.3 Représentation

Une représentation réelle de $Spin_n$ est la restriction d'une représentation irréductible de Cl_n dans $Hom_{\mathbb{R}}(S, S)$:

$$\Delta_n : Spin_n \rightarrow GL(S)$$

On a la propriété : quand $n = 3(mod 4)$, cette définition est indépendante de la représentation de Cl_n , pour $n \neq 0(mod 4)$, la représentation est soit irréductible soit la somme directe de deux représentations équivalentes irréductibles, dans tous les autres cas

$$\Delta_{4m} = \Delta_{4m}^+ \oplus \Delta_{4m}^-$$

où Δ_{4m}^+ et Δ_{4m}^- sont des représentations irréductibles non équivalentes de $Spin_{4m}$.

Dans le cas complexe, restriction d'une représentation de Cl_n dans $Hom_{\mathbb{C}}(S, S)$, les cas sont moins nombreux : si n est impair $\Delta_n^{\mathbb{C}}$ est indépendant de la représentation irréductible de Cl_n et est elle même irréductible, si n est pair on a une décomposition en la somme de deux représentation non équivalentes irréductibles.

5.2.4 Structures Spin, variétés spinorielles, spineurs, opérateur de Dirac

Le fibré $P_O(E)$ des repères orthonormés d'un fibré vectoriel $\pi : E \rightarrow X$ est le fibré constitué, au dessus de chaque point $x \in X$, de l'ensemble des bases orthonormales de $\pi^{-1}(x)$. Si X est orientable, on peut considérer le sous-fibré $P_{SO}(E)$. Une structure de spin sur E est un fibré principal de groupe de structure $Spin$ ainsi qu'un relèvement à 2 feuillet de $P_{SO}(E)$:

$$\xi : P_{Spin}(E) \rightarrow P_{SO}(E)$$

tel que $\xi(pg) = \xi(p)\xi_0(g)$ pour $p \in P_{Spin}(E)$ et $g \in Spin$, ξ_0 étant le relèvement universel de SO . L'existence d'une telle structure est une condition topologique, quantifiée par les classes de Stiefel-Whitney.

Une variété spinorielle est une variété riemannienne avec une structure de spin sur son fibré tangent.

Le fibré de Clifford d'un fibré vectoriel riemannien E est le fibré associé à $P_{SO}(E)$ par une représentation de SO_n dans les automorphismes de Cl_n (il a donc un fibré isomorphe à Cl_n). Il y a une isométrie canonique de fibré vectoriel :

$$\lambda : \Lambda^*(E) \rightarrow Cl(E)$$

telle que $\lambda(\Lambda^{pair} E) = Cl^0(E)$ et $\lambda(\Lambda^{impair} E) = Cl^1(E)$

Un spineur réel sur E est un fibré associé à une structure de spin sur E par une représentation μ de $Spin$ dans $SO(M)$ par la multiplication à gauche, où M est un module à gauche pour Cl_n . Une connexion riemannienne sur $P_{SO}(E)$ induit une connexion sur le fibré de Clifford, qui est une dérivation sur l'algèbre de ses sections :

$$\nabla(\varphi \cdot \psi) = (\nabla\varphi) \cdot \psi + \varphi \cdot (\nabla\psi)$$

qui préserve les sous-fibrés $Cl^0(E)$ et $Cl^1(E)$ et pour laquelle la forme de volume $\omega = e_1 \dots e_n$ satisfait $\nabla\omega = 0$.

En notant ω la 1-forme de connexion sur $P_{SO}(E)$ et ∇^S la dérivée covariante sur un spineur associé à E est donnée localement par

$$\nabla^S \sigma_\alpha = \frac{1}{2} \sum_{i < j} \tilde{\omega}_{ji} \otimes e_i e_j \cdot \sigma_\alpha$$

où $\mathcal{E} = (e_1, \dots, e_n)$ est une section de $P_{SO}(E)$ (un vilbein) qui se relève en une section $\tilde{\mathcal{E}}$ de $P_{Spin}(E)$ qui permet de relever la connexion en $\tilde{\omega} = \mathcal{E}^*(\omega)$.

Étant donnée un fibré S de modules à gauche sur $Cl(X)$ doté d'une structure riemannienne et d'une connexion compatible, on peut définir un opérateur canonique, l'opérateur de Dirac :

$$D\sigma = \sum_j e_j \cdot \nabla_{e_j} \sigma$$

Dans le cas où S est un spineur associé au fibré tangent de X , on appelle l'opérateur de Dirac opérateur d'Atiyah-Singer, et on le note \not{D} .

5.3 Notes physiques

Ces deux notes sont tirés des cours de Gilles Cohen-Tannoudji.

5.3.1 La cohérence quantique et les probabilités complexes

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.

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.

5.3.2 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 $x = (x, t)$ 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 DE et Dp vers zéro (si l'on veut limiter la région d'interaction à un point d'espace-temps), sans faire tendre DE et Dp 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 DE et Dp , les lois de conservation de l'énergie et de l'impulsion peuvent être violées, avec des erreurs DE et Dp reliées à DE et Dp 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 particulières 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 particulières 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.