

Une promenade entre physique et géométrie : des
anomalies au théorème de l'indice

Licence math physique

ENS Paris

12 septembre 2011

Pierrick Bousseau

Encadrants : Adel Bilal et Damien Gayet

Table des matières

1	Introduction	3
2	Notions préliminaires de physique et de géométrie	7
2.1	Géométrie des champs classiques	7
2.1.1	La notion de fibré	7
2.1.2	Connexion et courbure	10
2.1.3	Le cas particulier de la géométrie riemannienne	11
2.1.4	Applications à la construction de théories physiques	13
2.1.5	Fermions et spineurs	15
2.2	Physique quantique et théorie des champs	21
2.2.1	Principes de la physique quantique	21
2.2.2	Le problème de la quantification d'une théorie de champs classiques	22
2.2.3	L'approche perturbative	23
3	Anomalie chirale et théorème de l'indice pour les opérateurs de Dirac	28
3.1	Physique de l'anomalie chirale	28
3.1.1	Définition	28
3.1.2	Calcul perturbatif	30
3.1.3	Rôle des symétries globales et de leurs brisures dans la théorie de l'interaction forte	31
3.2	Vers le théorème de l'indice	34
3.2.1	Classes caractéristiques	34
3.2.2	Analyse fonctionnelle sur les variétés	37
3.2.3	Développement asymptotique du noyau de la chaleur	40
3.2.4	Problème de l'indice	45
3.3	Théorème de l'indice	48
3.3.1	Objectif et réduction à un problème local	48
3.3.2	Un changement d'échelle bien choisi	50
3.3.3	Oscillateur harmonique	52
3.4	Méthode de Fujikawa : application du théorème de l'indice au calcul de l'anomalie chirale	53
4	Anomalie de jauge	56
4.1	Quantification des théories de jauge et anomalies	56
4.1.1	Quantification : de la symétrie de jauge à la symétrie BRST	56
4.1.2	Anomalie de jauge	60
4.1.3	Conditions de Wess-Zumino	60
4.2	Calcul de l'anomalie de jauge	61
4.2.1	Théorie des formes de Chern-Simons	61
4.2.2	Résolution des contraintes de Wess-Zumino	65
4.2.3	Cas du Modèle Standard	66
4.3	La construction d'Alvarez-Gaumé et Ginsparg	69
4.3.1	But de la construction	69
4.3.2	Développement adiabatique	70
4.3.3	Anomalie : un problème de géométrie classique ou quantique? Un saut dans l'infini	75

1 Introduction

L'existence de liens entre la physique et la géométrie est un fait évident : la géométrie est sans doute née d'une volonté de description de l'« espace physique » et les développements ultérieurs de la notion d'« espace géométrique » se sont avérés particulièrement adaptés pour répondre à ce problème comme le montre l'exemple de la relativité générale.

Il est peut-être plus surprenant de voir le caractère géométrique des théories de champs (classiques) qui sous-tendent la description des particules élémentaires. En effet, il ne s'agit plus de décrire la géométrie de l'« espace-temps » mais les propriétés géométriques permettant de rendre compte des symétries de variables dynamiques « internes » aux champs. Le cadre mathématique des théories de champs classiques est maintenant bien compris : les champs sont décrits comme des sections de fibrés définis sur l'espace-temps.

Cependant, une théorie des champs classiques ne peut pas décrire de manière satisfaisante le monde des particules élémentaires car elle ne prend pas en compte les impératifs de la physique quantique : la théorie adaptée est la théorie quantique des champs obtenue par quantification de la théorie classique. La quantification est un problème difficile : on ne sait la mener à bien essentiellement que de manière perturbative. Mais même à ce niveau, des phénomènes nouveaux, typiquement quantiques, apparaissent. Par exemple, une symétrie de la théorie classique peut ne plus être présente au niveau quantique : on parle alors d'anomalie. Le présent mémoire est consacré à l'étude des liens existant entre ce phénomène et l'existence d'obstructions de nature topologique. Plus précisément, on présentera :

1. l'anomalie chirale, qui correspond à une brisure de la symétrie chirale dans une théorie décrivant un spineur (de Dirac) de masse nulle couplé à un champ de jauge. L'obstruction associée est une dissymétrie existante entre les modes nuls de chiralité différente de l'opérateur de Dirac, ce qu'on appelle l'indice de cet opérateur, et la nature topologique de cette quantité a priori issue de l'analyse est exprimée par le théorème de l'indice d'Atiyah-Singer. Il est tout à fait remarquable que ce théorème aux origines assez éloignées de la physique (il a été découvert suite au développement de généralisations du théorème de Riemann-Roch en géométrie algébrique) intervienne en théorie quantique des champs. L'existence de l'anomalie chirale ne pose pas de problème de cohérence à la théorie : au contraire, elle prédit l'existence de phénomènes qui auraient été interdits par la symétrie classique et elle a permis de comprendre par exemple la période de désintégration du pion neutre.
2. l'anomalie de jauge, qui correspond à une brisure de la symétrie de jauge dans une théorie décrivant un fermion chirale (de Weyl) couplé à un champ de jauge. Contrairement au cas de l'anomalie chirale, la présence d'une anomalie de jauge dans une théorie se traduit par une incohérence physique. L'intérêt physique d'étudier ce phénomène est le caractère contraignant de l'absence d'anomalie : par exemple, la condition d'absence d'anomalie de jauge dans le Modèle Standard, qui est bien vérifiée, est la seule relation non-triviale prédite théoriquement reliant les nombres quantiques des quarks et des leptons et les contraintes liées à l'absence d'anomalie de jauge ont joué un rôle important dans le développement de la théorie des cordes (pour l'annulation des anomalies dans les théories des champs limites à basse énergie des différentes théories des supercordes, sujet qui n'est pas traité dans ce qui suit, voir la fin de [6]). L'anomalie de jauge correspond au problème de la définition d'un déterminant invariant de jauge d'un opérateur de Dirac. Mathématiquement, il s'agit d'étudier les points d'annulation du déterminant d'une famille à deux paramètres d'opérateurs de Dirac et de manière surprenante, on verra que ce problème est relié à l'indice d'un opérateur de Dirac vivant sur un espace possédant deux dimensions de plus que l'espace-temps de départ.

Notations

- On se place dans un système d'unités où $\hbar = c = 1$. Dans certaines formules, on laissera apparent \hbar pour mettre en évidence la nature quantique de l'expression.
- On utilisera systématiquement la convention d'Einstein de sommation sur les indices répétés.
- On écrira ∂_i pour $\partial/\partial x^i$.
- $[\cdot, \cdot]$ désignera le commutateur : $[A, B] = AB - BA$, $\{\cdot, \cdot\}$ désignera l'anticommutateur $\{A, B\} = AB + BA$.
- $\epsilon_{\mu_1, \dots, \mu_n}$ désignera le tenseur totalement antisymétrique de rang n .

Conventions

Les conventions utilisées en mathématique et en physique sont souvent différentes et diffèrent même souvent parmi les auteurs d'une même discipline : elles concernent le choix de la signature lorentzienne ($(- + \dots +)$ ou $(+ - \dots -)$), l'écriture de l'algèbre de Clifford ($\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu, \nu}$ ou $\{\gamma_\mu, \gamma_\nu\} = -2g_{\mu, \nu}$) ou le choix des «générateurs» des algèbres de Lie (hermitiens ou antihermitiens). Pour garder une certaine similitude avec la littérature, on changera de conventions au cours du texte. On indiquera au début de chaque partie les conventions adoptées.

Terminologie

Ce qu'on appelle anomalie chirale est parfois désigné dans la littérature sous le nom d'anomalie abélienne ou d'anomalie axiale.

Ce qu'on appelle anomalie de jauge est parfois désigné dans la littérature sous le noms d'anomalie non-abélienne ou encore d'anomalie chirale.

Sources

Beaucoup de démonstrations mathématiques sont tirées de [12], alors que de nombreux développements physiques sont issus de [6]. Ces références sont les deux sources d'inspiration directe de ce mémoire. Les liens entre anomalies et géométries différentielle sont aussi évoqués dans [10] et [11]. Pour la théorie des fibrés et des connexions, voir [9]. Pour la théorie quantique des champs, mis à part les nombreux ouvrages de physique sur le sujet, une introduction se trouve au début de [8]. Des références ponctuelles sont signalées dans le texte lorsqu'elles apparaissent.

Remerciements

Je tiens à remercier M.Bilal de m'avoir proposé de travailler sur les anomalies, d'avoir encadré ce travail et M.Gayet d'avoir accepté de répondre à mes questions. Je remercie également M.Bilal et M.Gayet d'avoir relu ce mémoire, permettant ainsi de nombreuses corrections et améliorations.

Plan détaillé

Dans 2.1.1, on montre comment la notion de champ classique muni de symétries conduit aux notions de fibré principal et fibré associé. La symétrie de jauge présente en électromagnétisme sert de justification à l'introduction des notions de connexion et de courbure dans 2.1.2. Dans 2.1.3, on montre que ce formalisme admet comme cas particulier la géométrie riemannienne, ce qui nous donne les outils géométriques permettant de décrire les propriétés d'un espace-temps courbé. Ceci étant combiné avec la théorie des fibrés définis sur cet espace-temps, on peut définir dans 2.1.4 ce qu'est une théorie de jauge d'action de Yang-Mills décrivant des champs classiques. Les problèmes spécifiques posés par la description des fermions conduisent à introduire la théorie des algèbre de Clifford, les spineurs et la notion d'opérateur de Dirac dans 2.1.5. La section 2.1 est donc consacrée à exposer des notions de géométrie différentielle avec pour motivation physique l'écriture de théories décrivant des champs classiques.

Le problème de l'incorporation de la mécanique quantique dans ce cadre est l'objet de la section 2.2. Dans 2.2.1, on rappelle brièvement les fondements de la physique quantique. Dans 2.2.2, on donne un survol de la quantification d'une théorie de champs libre et dans 2.2.3, un aperçu de la formulation perturbative de la théorie quantique des champs : intégration fonctionnelle, graphes de Feynman et renormalisation.

La section 3.1 a pour objet la présentation de l'anomalie chirale. Dans 3.1.1, on essaye de donner l'intuition du fait que si un spineur de Dirac est couplé à un champ de jauge alors une dissymétrie entre les spineurs de chiralités positive et négative va apparaître, la symétrie chirale de la théorie classique va disparaître au niveau quantique. On vérifie ce résultat par un calcul perturbatif à l'ordre d'une boucle dans 3.1.2 : l'anomalie chirale apparaît comme une conséquence de l'impossibilité de construire une procédure de régularisation préservant l'intégralité des symétries de la théorie classique. Dans 3.1.3, on montre que la brisure de symétries globales par des effets quantiques n'est pas qu'une spéculation théorique : elle joue un rôle important dans la compréhension de la chromodynamique quantique, la théorie quantique des champs décrivant l'interaction forte.

La forme des expressions issues du calcul de l'anomalie chirale au niveau perturbatif sert de justification à l'introduction de la théorie des classes caractéristiques dans 3.2.1. À partir de 3.2.2 et jusqu'à 3.3.3, l'exposé est essentiellement mathématique. Dans 3.2.2, grâce aux outils standard de l'analyse fonctionnelle que constituent les espaces de Sobolev, on démontre un théorème spectral pour les opérateurs de Dirac. Dans 3.2.3, on introduit «l'équation de la chaleur associée à un opérateur de Dirac \mathcal{D} » dont la résolution revient à considérer $e^{-t\mathcal{D}^2}$. On montre que cet opérateur est à noyau et que ce noyau, «le noyau de la chaleur», admet un développement asymptotique. Ce résultat est en lui-même un résultat central de la géométrie spectrale. Dans 3.2.4, on introduit la notion d'indice d'un opérateur de Dirac et on montre qu'il peut être exprimé comme la supertrace de $e^{-t\mathcal{D}^2}$. Cette section se finit par deux exemples explicites sur lesquels on peut constater un lien entre indice et classes caractéristiques.

La section 3.3 a pour but d'exprimer l'indice d'un opérateur de Dirac en fonction de classes caractéristiques, ce résultat s'appelant le théorème de l'indice. Dans 3.3.1, on montre grâce à l'existence du noyau de la chaleur que la question est purement locale. Dans 3.3.2, on fait un changement d'échelle permettant d'isoler le terme intéressant dans le développement asymptotique du noyau de la chaleur. On se ramène ainsi à un problème simple qu'on résout par un calcul explicite dans 3.3.3.

Dans 3.4, on revient à la physique. On interprète l'anomalie chirale comme une variation de la mesure fermionique intervenant dans l'intégrale fonctionnelle définissant la théorie quantique des champs. La nécessité d'une régularisation invariante de jauge conduit à introduire un facteur $e^{-t\mathcal{D}^2}$ et l'anomalie est alors reliée à la fonctionnelle locale intervenant dans le théorème de

l'indice.

À partir de 4.1, on oublie l'anomalie chirale et on s'intéresse à l'anomalie de jauge, i.e. à la possibilité d'une brisure de la symétrie de jauge au niveau quantique dans une théorie contenant des fermions de Weyl. Dans 4.1.1, on explique comment quantifier une théorie de jauge à l'aide de la méthode de Faddeev-Popov et le rôle de la symétrie $BRST$. On en déduit dans 4.1.2 qu'une absence de symétrie de jauge au niveau quantique est fatale pour la cohérence de la théorie. Dans 4.1.3, on montre que la forme de l'anomalie de jauge est fortement contrainte par ce qui s'appelle la condition de Wess-Zumino.

Dans 4.2.1, on présente la théorie des formes de Chern-Simons qui est à l'anomalie de jauge ce que la théorie des classes caractéristiques est à l'anomalie chirale : le bon cadre mathématique pour exprimer le résultat du calcul de la quantité physique. L'interprétation de la condition de Wess-Zumino en termes de cohomologie $BRST$ combinée aux équations de la théorie des formes de Chern-Simons permet de déterminer la forme générale de l'anomalie de jauge dans 4.2.2. Dans 4.2.3, on met en évidence l'intérêt «concret» de l'étude de l'anomalie de jauge : la condition d'absence d'anomalie de jauge a un caractère contraignant pour construire une théorie physiquement cohérente et on vérifie qu'elle est bien vérifiée dans le Modèle Standard.

La section 4.3 est consacrée à une construction géométrique permettant de redériver de manière plus précise l'anomalie de jauge et d'explicitier le lien avec le théorème de l'indice en deux dimensions de plus. Dans 4.3.1, on interprète l'anomalie de jauge comme un problème de définition du déterminant d'un opérateur qui n'est pas un endomorphisme. Dans 4.3.2, on présente la construction géométrique et le lien avec le théorème de l'indice est fait via un développement perturbatif. Dans 4.3.3, on précise le lien entre anomalie et géométrie en dimension infinie d'une part et entre anomalie et topologie d'autre part.

2 Notions préliminaires de physique et de géométrie

2.1 Géométrie des champs classiques

On commence par montrer comment la théorie des fibrés et des connexions apparaît naturellement lorsqu'on cherche à décrire mathématiquement un champ classique obéissant à des symétries.

2.1.1 La notion de fibré

Tout d'abord, qu'est-ce qu'un champ ? Un champ est une fonction définie sur l'espace \mathbb{R}^3 . Mais une fonction à valeurs dans quoi ? Prenons des exemples : un champ de température est une fonction scalaire, i.e. à valeurs dans \mathbb{R} , un champ électrique est à valeurs dans \mathbb{R}^3 ... On aurait donc envie de considérer des fonctions de \mathbb{R}^3 dans un \mathbb{R}^N . Néanmoins, plusieurs remarques montrent qu'il faut considérer des fonctions plus générales :

1. Le champ magnétique, en tant que « champ de pseudo-vecteurs » n'est une fonction bien définie de \mathbb{R}^3 dans \mathbb{R}^3 que si une orientation a été choisie sur \mathbb{R}^3 .
2. Il existe des champs de « tenseurs » comme par exemple le tenseur champ électromagnétique.
3. On a envie de considérer des « espaces » plus généraux que \mathbb{R}^3 : la relativité générale nous enseigne que la physique doit se faire sur un « espace-temps » qui s'avère être une « variété de dimension 4 ».
4. Même sur \mathbb{R}^3 , on aimerait penser à un champ de vecteurs comme à la donnée d'un vecteur tangent en chaque point : les vecteurs correspondants à différents points devraient vivre dans des espaces différents, dans « les espaces tangents à chacun des points ».

Pour toutes ces raisons, on commence par introduire la notion de variété. On appelle variété topologique de dimension n un espace topologique M séparé dénombrable à l'infini dont tout point possède un voisinage homéomorphe à \mathbb{R}^n . On appelle cartes ces homéomorphismes et atlas un ensemble de cartes dont les domaines de définition recouvrent M . Deux cartes sont dites compatibles si la fonction de transition associée est un difféomorphisme de classe C^∞ . Étant donné un atlas de cartes compatibles, on peut considérer l'atlas maximal pour l'inclusion des cartes compatibles avec cet atlas. La donnée d'un atlas maximal sur une variété topologique définit une structure de variété différentielle : on dira variété par la suite. La structure de variété est faite de telle façon que les notions d'applications de classes C^∞ entre variétés, de difféomorphismes, d'espace tangent en un point, de formes différentielles ... aient un sens. On renvoie à un cours de géométrie différentielle élémentaire pour plus de détails.

L'ensemble des espaces tangents à une variété M de dimension n peut être muni de façon naturelle d'une structure de variété de dimension $2n$. Cette variété notée TM a une structure très particulière : localement, elle s'identifie à un produit d'un ouvert de M par \mathbb{R}^n . Cet exemple motive la notion générale suivante :

Définition 1. Soit E , M et F trois variétés. On dit que E est un fibré d'espace total E , de base M et de fibre F s'il existe $\pi: E \rightarrow M$ de classe C^∞ , appelée projection de E sur M telle que pour tout $x \in M$, $\pi^{-1}(x)$ est difféomorphe à F et au-dessus de tout point $x \in M$, E s'écrit localement comme le produit de F par un ouvert de M . E est dit trivial s'il s'écrit globalement comme un produit, i.e. si E est difféomorphe à $M \times F$. Une section de E est une application s de classe C^∞ de M dans E vérifiant $\pi \circ s = id_M$.

Il faut penser à un fibré en termes de verticalité : on a placé une copie de F au-dessus de chaque point de M . Dans le cas de TM , appelé le fibré tangent, on a $F = \mathbb{R}^n$. F est donc mieux qu'une

variété, c'est un espace vectoriel, on dit alors que TM est un fibré vectoriel. Une section de TM n'est rien d'autre qu'un champ de vecteurs sur M . On peut généraliser aux fibrés vectoriels des notions définies pour les espaces vectoriels : on peut faire le produit tensoriel de fibrés vectoriels, si E est un fibré vectoriel, on peut construire le fibré $End(E)$ des endomorphismes de E ...

Notation : Pour E un fibré sur M et $m \in M$, on notera toujours E_m la fibre de E au-dessus de m .

En général, les champs de la physique seront des sections de fibrés vectoriels sur l'espace-temps, une variété M de dimension 4. Dans toute la suite, M désignera une variété de dimension n . Pour comprendre l'origine de ces fibrés, on va se poser la question suivante : qu'est-ce qu'un vecteur dans un espace vectoriel E de dimension n ? On pourrait avoir envie de répondre : c'est la donnée de n nombres, les coordonnées du vecteur. Mais les coordonnées par rapport à quelle base? Un vecteur de E est un élément de E , c'est un objet géométrique défini indépendamment de tout choix de base. On pourrait répliquer que de toute façon E s'identifie à \mathbb{R}^n . Certes, mais cette identification revient à choisir une base, elle n'a rien de canonique, deux choix de bases différents donnent deux systèmes de coordonnées différents. Cependant, on peut essayer de raisonner à l'envers : si on ne veut privilégier aucune base sur une autre et si on veut continuer à parler de coordonnées, on peut voir un vecteur comme l'ensemble de ses coordonnées dans toutes les bases possibles. À un espace vectoriel E , on est donc amené à associer l'ensemble de ses bases B . Si on choisit une base, i.e. un élément g de B , alors tout autre élément de B s'obtient à partir de g par action d'un élément de $GL_n(\mathbb{R})$ d'où une identification de B avec $GL_n(\mathbb{R})$, qui, attention, dépend du choix de g . Cette construction, le fait d'associer à un espace vectoriel l'ensemble de ses bases, peut par exemple se faire pour chacun des espaces tangents à M . On obtient ainsi un fibré sur M , le fibré des repères, qui n'est pas un fibré vectoriel puisque la fibre type est une variété homéomorphe à $GL_n(\mathbb{R})$. Le fibré des repères est l'exemple type d'une classe particulière de fibrés : les fibrés principaux. Pour en donner la définition, on rappelle que :

Définition 2. *On appelle groupe de Lie une variété G possédant une structure de groupe compatible avec la structure de variété, i.e. telle que le produit et le passage à l'inverse soient des applications de classe C^∞ respectivement de $G \times G$ et G à valeurs dans G . On appelle algèbre de Lie de G et on note $Lie(G)$ son espace tangent en l'élément neutre, identifié grâce à la structure de groupe à l'espace des champs de vecteurs invariants à gauche sur G et muni grâce à cette identification du crochet hérité du crochet des champs de vecteurs.*

$GL_n(\mathbb{R})$ est un groupe de Lie. La généralisation du fibré des repères est la notion suivante :

Définition 3. *Soit G un groupe de Lie. On appelle fibré principal sur M de groupe G un fibré sur M de fibre homéomorphe à G , muni d'une action à droite de classe C^∞ de G préservant les fibres et transitive sur chacune des fibres.*

Dans la définition, on a indiqué que la fibre était seulement homéomorphe à G pour bien indiquer que les fibres de G ne sont pas munies d'une structure de groupe. En revanche, le choix d'un élément particulier d'une fibre permet grâce à l'action de G de munir la fibre d'une structure de groupe de Lie l'identifiant avec G . Le fibré des repères de M est un fibré principal sur M de groupe $GL_n(\mathbb{R})$.

On a construit le fibré des repères à partir du fibré tangent : on peut inverser la construction. On peut considérer le fibré produit du fibré des repères avec \mathbb{R}^n , auquel il faut penser comme l'ensemble des couples « (base, coordonnées dans cette base) », quotienté par la relation : on identifie (b_1, v_1) et (b_2, v_2) si lorsqu'on change de base pour passer de b_1 à b_2 , v_1 se transforme en v_2 . Le fibré résultant est le fibré tangent. Cette construction peut se généraliser et permet d'associer à tout fibré principal un fibré vectoriel, à condition qu'on ait fixé une représentation linéaire de dimension finie de G , analogue de la représentation standard de $GL_n(\mathbb{R})$ sur \mathbb{R}^n .

Définition 4. Soit P un fibré principal de groupe G sur M . Soit $\rho: G \rightarrow GL_N(\mathbb{R})$ une représentation linéaire de G sur un espace vectoriel de dimension N (il s'agit d'une action à gauche de G). Alors on appelle fibré associé à G et à ρ le fibré vectoriel $P \times \mathbb{R}^N$ quotienté par la relation d'équivalence : (b_1, v_1) et (b_2, v_2) sont identifiés si et seulement s'il existe $g \in G$ tel que $b_1.g = b_2$ et $\rho(g^{-1}).v_1 = v_2$.

La notion de fibré associé à un fibré principal est la construction géométrique permettant de décrire un champ se transformant d'une certaine façon sous certaines symétries. Le cas le plus simple est celui des symétries spatio-temporelles : la relativité restreinte impose l'invariance des lois physiques par le groupe de Lorentz propre, i.e. de la composante connexe de $SO(3, 1)$. Étant donné l'espace de Minkowski (ou une variété lorentzienne de dimension 4) M , on a un fibré principal de fibre le groupe de Lorentz propre comme sous-fibré du fibré des repères. Un champ invariant relativiste sera donc caractérisé par une représentation linéaire de dimension finie du groupe de Lorentz et apparaîtra comme une section du fibré vectoriel associé à cette représentation. On parlera plus loin des problèmes spécifiques soulevés par le cas des fermions.

Un champ peut posséder des symétries autres que spatio-temporelles, on parle de symétries internes et il s'avère que toutes les interactions fondamentales hormis la gravitation peuvent être comprises à partir de symétries dites de jauge. Pour préciser, prenons l'exemple de l'interaction électromagnétique qui au niveau classique est décrite par les équations de Maxwell. Cette théorie décrit un tenseur antisymétrique à deux composantes $F_{\mu,\nu}$, le tenseur de champ électromagnétique, qui apparaît comme la dérivée d'un champ vectoriel : le potentiel vecteur A_μ , $F_{\mu,\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Ce dernier n'est pas défini de façon unique mais «à une transformation de jauge» près : si Λ est une fonction scalaire, alors $A_\mu + \partial_\mu \Lambda$ est encore un potentiel vecteur, on parle «d'invariance de jauge». Cette remarque d'apparence anodine est en fait essentielle pour comprendre le contenu géométrique de l'électromagnétisme. Supposons qu'on se place en mécanique quantique traditionnelle et qu'on essaye de décrire une particule de masse m électriquement chargée de charge q par une fonction d'onde $\psi(x)$. En présence d'un potentiel vecteur extérieur $A(x)$, $\psi(x)$ est solution de l'équation de Schrödinger $i \frac{d\psi}{dt} = \frac{1}{2m} (-i \frac{d}{dx} - qA(x))^2 \psi(x)$. L'introduction du terme $-qA$ peut se justifier à l'aide des règles de correspondance et du fait que le hamiltonien d'une particule classique dans un potentiel vecteur extérieur $A(x)$ s'écrit $\frac{1}{2m} (p - qA)^2$. Mais on peut aussi l'interpréter comme une conséquence d'une symétrie. Soit α une fonction C^∞ , $x \rightarrow \alpha(x)$ et considérons la transformation :

$$\psi(x) \rightarrow \psi'(x) = e^{i\alpha(x)q} \psi(x)$$

$$A(x) \rightarrow A'(x) = A(x) + \frac{i}{q} e^{i\alpha(x)q} \frac{d}{dx} (e^{-i\alpha(x)q}) = A(x) + \frac{d}{dx} \alpha(x)$$

Par invariance de jauge, cette transformation, dite de jauge, ne change pas le contenu physique du champ A . Une équation décrivant ψ doit donc garder la même forme sous cette transformation. Ceci montre que d/dx n'est pas un opérateur différentiel adapté à la situation car il n'est pas invariant, alors qu'un calcul facile montre $d/dx - iqA$ est invariant. On appelle cet opérateur une dérivation covariante et l'exigence d'une théorie invariante sous transformation de jauge impose donc que la dérivation covariante soit l'opérateur naturel de différentiation devant intervenir dans l'écriture des équations. La transformation qui est au coeur de ce qui précède est dite locale car elle dépend du point d'espace-temps considéré et de groupe $U(1)$ car ψ se transforme suivant une représentation du groupe $U(1)$. Remarquons que la charge électrique s'interprète comme la quantité déterminant la représentation.

La section suivante donne les outils géométriques qui permettront d'explicitier et de généraliser cette situation.

2.1.2 Connexion et courbure

Définition 5. Soit E un fibré vectoriel sur M . Une connexion sur E est une application linéaire $\nabla : C^\infty(M, TM) \otimes C^\infty(E) \rightarrow C^\infty(E)$ associant à un champ de vecteurs X et une section s de E une nouvelle section $\nabla_X s$ de E telle que pour tout $f \in C^\infty(M)$:

1. $\nabla_{fX} s = f \nabla_X s$
2. $\nabla_X (fs) = f \nabla_X s + (X(f))s$

Si γ est une courbe C^∞ sur M , l'équation différentielle $\nabla_{\dot{\gamma}} s = 0$ est une équation différentielle linéaire du premier ordre sur les sections s de E le long de γ , qui, du moins localement, a une unique solution pour une valeur initiale donnée : on dit que la solution s est obtenue par transport parallèle le long de γ .

Une connexion peut s'écrire en coordonnées locales : soit s_α une base locale de sections de E , soit x^i des coordonnées locales sur M , alors on peut écrire : $\nabla_i s_\alpha = \Gamma_{i,\alpha}^\beta s_\beta$. La connexion est totalement déterminée localement par la donnée des coefficients $\Gamma_{i,\alpha}^\beta$.

Si le fibré E est muni d'une métrique (euclidienne ou hermitienne), on dit que ∇ est compatible avec la métrique si pour tout X champ de vecteurs sur M et s_1, s_2 sections C^∞ de E , on a : $X \cdot (s_1, s_2) = (\nabla_X s_1, s_2) + (s_1, \nabla_X s_2)$

Courbure : La courbure K d'une connexion sur E est définie par :

$$K(X, Y)s = \nabla_X \nabla_Y s - \nabla_Y \nabla_X s - \nabla_{[X, Y]} s.$$

On vérifie facilement que K est un opérateur tensoriel. On peut donc écrire :

$$K(\partial_j, \partial_k)s_\alpha = K_{\alpha, j, k}^\beta s_\beta$$

où :

$$K_{\alpha, j, k}^\beta = \partial_j \Gamma_{k, \alpha}^\beta - \partial_k \Gamma_{j, \alpha}^\beta + \Gamma_{k, \alpha}^\gamma \Gamma_{j, \gamma}^\beta - \Gamma_{j, \alpha}^\gamma \Gamma_{k, \gamma}^\beta.$$

Cette expression étant antisymétrique en j et k , K peut être identifié avec une 2-forme à valeurs dans $End(E)$:

$$F_\alpha^\beta = K_{\alpha, j, k}^\beta dx^j \wedge dx^k.$$

Repère synchrone : Soit x^i des coordonnées locales sur M et s_α un repère local de E . s_α est dit synchrone par rapport à x^i si chaque section s_α est parallèle le long des lignes radiales issues de l'origine. Les repères locaux synchrones existent grâce au transport parallèle. Si E est muni d'une métrique compatible avec la connexion, on peut de plus choisir des repères locaux synchrones orthonormés.

Proposition 1. À l'origine d'un repère synchrone, on a : $\Gamma_{j, \alpha}^\beta = -\frac{1}{2} K_{\alpha, j, k}^\beta x^k + O(|x|^2)$

Soit P un fibré principal sur M de groupe G et soit $\pi : P \rightarrow M$ la projection canonique associée. Alors $Ker(T_\pi)$ est un sous-fibré du fibré tangent TP de P appelé sous-fibré des «vecteurs tangents verticaux» : il s'agit de l'ensemble des vecteurs tangents au fibré P le long des fibres. Un choix G -équivariant d'un supplémentaire de $Ker(T_\pi)$ dans TP s'appelle une connexion sur P . Autrement dit, il s'agit d'une application G -équivariante $h : \pi^* TM \rightarrow TP$ telle que $T_\pi \circ h = id$. L'image de h est alors un sous-fibré supplémentaire de $Ker(T_\pi)$ dans TP , appelé le fibré des vecteurs horizontaux. On obtient une définition équivalente en remarquant que se donner un fibré de vecteurs horizontaux revient à se donner une 1-forme ω sur P à valeurs dans $Lie(G)$ vérifiant :

1. ω est G -équivariante : $\omega(\xi.g) = Ad(g^{-1})\omega(\xi)$, où Ad est la représentation de $Lie(G)$ sur elle-même donnée par conjugaison.

2. Pour $u \in \text{Lie}(G)$, soit X_u le champ de vecteurs sur P induit par u . Alors $\omega(X_u) = u$.

Si on se donne une telle 1-forme, on obtient un champ de vecteurs horizontaux : $\text{Ker}\omega$. Réciproquement, un champ de vecteurs horizontaux au sens précédent donne naturellement une 1-forme vérifiant 1) et 2). Une 1-forme à valeurs dans $\text{Lie}(G)$ vérifiant 1) et 2) s'appelle une 1-forme de connexion ou simplement une connexion.

Soit P un fibré principal sur M de groupe G et muni d'une connexion. Soit $\gamma : [0, 1] \rightarrow M$ un chemin tracé sur M et un point $b \in P_{\gamma(0)}$, il existe un unique chemin $\tilde{\gamma} : [0, 1] \rightarrow P$ partant de b , tel que $\pi \circ \tilde{\gamma} = \gamma$ et tel que pour tout $t \in [0, 1]$, $\tilde{\gamma}'(t)$ soit horizontal : $\tilde{\gamma}$ s'appelle le relèvement horizontal de γ .

Soit ρ une représentation linéaire de dimension finie de G et soit E le fibré vectoriel associé à E et à ρ . On va voir qu'étant donnée une connexion sur P , on peut construire naturellement une connexion sur E , dite connexion associée. Soit $\gamma : [0, 1] \rightarrow M$ une courbe tracée sur M et soit $e \in E_{\gamma(0)}$. Par construction du fibré associé, on peut écrire $e = (b, v)$ avec $b \in G$ et v élément de l'espace de ρ . On définit alors le vecteur transporté parallèle de e suivant γ comme étant $f \in E_{\gamma(1)}$ défini par $(\tilde{\gamma}(1), v)$ où $\tilde{\gamma}$ est le relèvement horizontal de γ partant de e défini par la connexion de P . On vient donc de définir une notion de transport parallèle sur E et c'est suffisant pour définir une connexion : si ξ est un vecteur tangent M en un point $m \in M$ et si w est une section de E , on choisit γ une courbe sur M telle que $\gamma(0) = m$ et $\gamma'(0) = \xi$, soit $\theta_t : E_{\gamma(t)} \rightarrow E_{\gamma(0)}$ l'isomorphisme donné par le transport parallèle, on définit alors : $\nabla_\xi w = \frac{d}{dt}|_{t=0}(\theta_t w_{\gamma(t)})$. On vérifie facilement qu'on définit ainsi une connexion ∇ sur E indépendante des choix intermédiaires faits au cours de la construction.

Courbure : Soit ω la 1-forme à valeurs dans $\text{Lie}(G)$ d'une connexion sur le fibré principal P . On définit la courbure Ω de ω comme étant la 2-forme sur P à valeurs dans $\text{Lie}(G)$ définie par : $\Omega = d\omega + \frac{1}{2}[\omega, \omega] = d\omega + \omega \wedge \omega$.

On vient de définir la courbure d'une connexion sur P . Soit ρ une représentation de G sur un espace vectoriel F et E le fibré vectoriel associé, on a vu qu'on pouvait définir une connexion associée sur E qui possède donc une courbure K . Comment la courbure de la connexion associée sur E est-elle reliée à la courbure Ω de la connexion sur P ?

Proposition 2. Soit $\rho_* : \text{Lie}(G) \rightarrow \text{End}(F)$ le morphisme d'algèbre de Lie induit par ρ . Rappelons de plus que, si $m \in M$, un choix d'un élément $p \in P_m$ permet d'identifier P_m à G , $T_p P$ à $\text{Lie}(G)$, E_m à F d'où une identification $\sigma_p : \text{End}(F) \rightarrow \text{End}(E_m)$. Soit ξ_1 et ξ_2 dans $T_m M$ et η_1, η_2 des relèvements à $T_p P$, alors : $K(\xi_1, \xi_2) = 2\sigma_p \rho_* \Omega(\eta_1, \eta_2)$.

On ne donne pas la démonstration, il suffit de jouer avec les définitions (voir par exemple [12] ou [9]).

2.1.3 Le cas particulier de la géométrie riemannienne

Une variété riemannienne (M, g) est une variété M muni d'un champ C^∞ de produit scalaire sur les espaces tangents : en chaque $m \in M$, $g(m)$ est un produit scalaire sur $T_m M$, g s'appelle une métrique sur M . On notera souvent $\langle X, Y \rangle$ pour $g(X, Y)$. La considération des repères orthonormés (resp. et directs) des espaces tangents donne naissance à un sous-fibré du fibré des repères qui s'avère être un fibré principal de groupe $O_n(\mathbb{R})$ (resp. $SO_n(\mathbb{R})$). Le résultat fondateur de la géométrie riemannienne est le théorème suivant :

Théorème 1. Une variété riemannienne possède une unique connexion, dite de Levi-Civita, ∇ sur TM vérifiant :

1. ∇ est compatible avec la métrique : si X, Y et Z sont des champs de vecteurs sur M , alors : $X \cdot \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle$. En particulier, le transport parallèle associé à ∇ préserve la métrique.

2. ∇ est sans torsion : si X et Y sont des champs de vecteurs sur M , alors : $\nabla_X Y - \nabla_Y X - [X, Y] = 0$. Autrement dit, si on regarde en coordonnées locales, les coefficients de la connexion (qu'on appelle les coefficients de Christoffel pour la connexion de Levi-Civita) vérifient une propriété de symétrie : $\Gamma_{i,j}^k = \Gamma_{j,i}^k$

Démonstration. On cherche la forme nécessaire de ∇ . On écrit la condition de compatibilité avec la métrique pour les trois permutations cycliques de X, Y, Z . On obtient trois formules : on ajoute les deux premières, on soustrait la troisième, on utilise l'absence de torsion pour simplifier et on obtient finalement une formule qui exprime $\langle \nabla_X Y, Z \rangle$ en fonction de X, Y, Z . On en déduit l'unicité et on obtient l'existence en vérifiant que cette formule définit un ∇ vérifiant les propriétés désirées. En coordonnées locales, les symboles de Christoffel sont déterminés par la métrique suivant : $\Gamma_{i,j}^k = \frac{1}{2} g^{k,l} (\partial_i g_{j,l} + \partial_j g_{i,l} - \partial_l g_{i,j})$ \square

Dans la théorie générale des fibrés, il est fondamental de distinguer les variables sur la variété de base des variables sur le fibré mais dans le cas d'une connexion sur TM cette distinction disparaît, ce qui rend la géométrie riemannienne à la fois plus simple que la théorie générale mais en même temps plus compliquée du fait de nombreuses identifications (en plus de l'identification des indices de base et de fibre, la présence de la métrique permet d'identifier TM à T^*M et donc de «monter et descendre les indices» comme on veut).

La courbure de la connexion de Levi-Civita s'appelle le tenseur de Riemann :

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z$$

En coordonnées locales : $R(\partial_j, \partial_k)\partial_l = R_{l,j,k}^i \partial_i$ où :

$$R_{l,j,k}^i = (\partial_j \Gamma_{k,l}^i - \partial_k \Gamma_{j,l}^i + \Gamma_{k,l}^m \Gamma_{j,m}^i - \Gamma_{j,l}^m \Gamma_{k,m}^i) \partial_i$$

Symétries du tenseur de courbure : Soit $R_{l,i,j,k} = g_{l,m} R_{i,j,k}^m$. On peut alors montrer, par exemple en regardant les formules en coordonnées, que :

1. $R_{j,i,k,l} = -R_{i,j,k,l}$, $R_{i,j,l,k} = -R_{i,j,k,l}$
2. $R_{k,l,i,j} = R_{i,j,k,l}$
3. $R_{i,j,k,l} + R_{i,k,l,j} + R_{i,l,j,k} = 0$

Définition 6. Le tenseur de Ricci est défini en coordonnées locales par : $Ric_{a,b} = R_{a,i,b}^i$. C'est un tenseur symétrique : $Ric_{a,b} = Ric_{b,a}$. La courbure scalaire est la trace par rapport à la métrique du tenseur de Ricci : $R = g^{a,b} Ric_{a,b}$.

Coordonnées géodésiques : On appelle géodésique une courbe γ sur M vérifiant $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$ autrement dit telle que le champ le long de γ des vecteurs tangents soit parallèle. C'est une équation différentielle linéaire du second ordre qui possède donc, au moins localement, une unique solution pour des conditions initiales données. Si $m \in M$, l'application exponentielle $exp_m : U \rightarrow M$, définie sur un voisinage de 0 dans $T_m M$, envoie un vecteur $v \in T_m M$ sur la valeur $\gamma(1)$ de l'unique géodésique γ vérifiant $\gamma(0) = m$, $\dot{\gamma}(0) = v$. Par le théorème d'inversion locale, exp_m est un difféomorphisme d'un voisinage de 0 dans $T_m M$ sur un voisinage de m dans M . Le choix d'une base orthonormée de $T_m M$ donne alors un système de coordonnées, appelé système de coordonnées géodésiques, sur un voisinage de m . Un tel système de coordonnées sera très pratique pour les calculs explicites grâce au résultat suivant :

Proposition 3. Dans un système de coordonnées géodésiques, on a $\Gamma_{i,j}^k(0) = 0$, $g_{i,j}(0) = \delta_{i,j}$ et plus précisément

$$g_{i,j}(x) = \delta_{i,j} + \frac{1}{3} x^p x^q R_{i,p,q,j}(0) + O(|x|^3)$$

Pour la preuve, voir [12] ou [5] ou n'importe quel ouvrage de géométrie riemannienne.
Quelques remarques concernant les formes différentielles :

1. La courbure peut se voir comme une 2-forme à valeurs dans $End(TM)$:

$$F_j^i = R_{j,k,l}^i dx^k \wedge dx^l.$$

2. Sur une variété riemannienne orientée, on a une forme volume naturelle :

$$vol = \sqrt{\det(g)} dx^1 \wedge \dots \wedge dx^n.$$

3. La donnée d'une métrique permet de construire un opérateur $*$, l'étoile de Hodge, envoyant les k -formes sur les $(n-k)$ -formes. Localement, si (e_1, \dots, e_n) est un champ local de bases orthonormées et si (e'_1, \dots, e'_n) sont les 1-formes associées par la dualité définie par la métrique, alors : $*(e_1 \wedge \dots \wedge e_k) = e'_{k+1} \wedge \dots \wedge e'_n$.

2.1.4 Applications à la construction de théories physiques

Le formalisme mathématique développé dans les sections précédentes permet la construction de théorie des champs. Pour ce faire, il faut se donner des champs, une fonctionnelle en ces champs qu'on appelle l'action, qui est une intégrale sur l'espace-temps M d'une fonctionnelle locale appelée densité lagrangienne. À partir de l'action, on obtient par un principe variationnel des équations de type Euler-Lagrange qui sont des équations aux dérivées partielles en les champs et qui constituent les équations de la théorie.

On obtient par exemple de cette façon la relativité générale : on se donne M une variété de dimension 4, «l'espace-temps». Le champ de la théorie est une métrique g de signature $(-+++)$ sur M , autrement dit la donnée d'une forme quadratique de signature $(-+++)$ sur chaque espace tangent. La géométrie riemannienne, rapidement présentée précédemment, considère des métriques de signature $(++++)$ mais les notions essentielles (connexion de Levi-Civita, géodésique, courbure...) n'utilisent que le caractère non-dégénéré de la métrique et ont donc encore un sens en signature $(-+++)$. g s'appelle le champ gravitationnel et on prend pour action de ce champ

$$S(g) = \int_M R(g) vol$$

où $R(g)$ est la courbure scalaire de g . Si on inclut en plus dans l'action un terme décrivant les champs de matière vivant sur M , l'équation d'Euler-Lagrange associée est l'équation d'Einstein :

$$Ric_{i,j} - \frac{1}{2} g_{i,j} R = Cste T_{i,j}$$

où $T_{i,j}$ est le tenseur d'énergie-impulsion associé aux champs de matière et où $Cste$ est une constante ($\frac{8\pi G}{c^4}$ dans les unités du système international et avec les conventions usuelles).

On construit également de cette manière les théories des champs classiques servant de base à la description des interactions et des particules élémentaires. On se donne (M, g) une variété lorentzienne de dimension 4, i.e. g est une métrique de signature $(-+++)$. Ce choix de signature est nécessaire si on veut interpréter (M, g) comme un «espace-temps». Néanmoins, pour des raisons de commodité mathématique, on travaillera dans la suite de ce mémoire souvent avec une métrique riemannienne, i.e. de signature $(++++)$.

Les données géométriques pour formulées une théorie de jauge sont un groupe de Lie G compact, produit direct de facteurs $U(1)$ et de facteurs compacts simples, et un fibré principal P de groupe G sur M . Le champ de la théorie, le champ de jauge, est une connexion A sur P .

Localement, A est une 1-forme à valeurs dans $Lie(G)$: $A = A_\mu dx^\mu$. On appelle transformation de jauge un automorphisme vertical du fibré P , i.e. une application de P dans P préservant les fibres G -équivariante. Localement, elle est donnée par une fonction à valeurs dans G : $x \rightarrow g(x)$. On vérifie facilement que sous une transformation de jauge, la connexion se transforme selon :

$$A_\mu(x) \rightarrow g(x)A_\mu(x)g(x)^{-1} + g(x)\partial_\mu(g(x)^{-1}).$$

Soit F la courbure de la connexion A : localement, F est une 2-forme à valeurs dans $Lie(G)$:

$$F = F_{\mu,\nu}dx^\mu \wedge dx^\nu = \frac{1}{2}F_{\mu,\nu}dx^\mu dx^\nu,$$

avec :

$$F_{\mu,\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu].$$

On vérifie que sous une transformation de jauge, F se transforme suivant

$$F(x) \rightarrow g(x)F(x)g(x)^{-1}.$$

Pour définir une théorie de jauge, on exige que l'action de la théorie soit invariante sous les transformations de jauge. On appelle action de Yang-Mills :

$$S[A] = \int_M tr(F \wedge *F) = \int_M tr(F_{\mu,\nu}F^{\mu,\nu})vol,$$

où tr est sur chaque facteur $U(1)$ ou simple composant G un multiple de la trace usuelle, la constante de proportionnalité étant reliée à ce qu'on appelle la constante de couplage du facteur. $S[A]$ est clairement invariante de jauge. L'équation d'Euler-Lagrange associée est l'équation de Yang-Mills $d_A * F = 0$ où $d_A = d + [A, \cdot]$ est la différentielle extérieure covariante agissant sur les formes à valeurs dans $Lie(G)$. On retrouve la théorie de l'électromagnétisme en prenant $G = U(1)$, P un fibré principal de groupe $U(1)$ et les équations de Maxwell dans le vide sont $dF = 0$ et $d*F = 0$. Un point essentiel est que $U(1)$ étant abélien, on a simplement $F = dA$, ce qui a pour conséquence de fait que les équations de Maxwell dans le vide sont linéaires. Dans le cas d'une théorie de jauge non-abélienne, i.e. avec G non-abélien, $F = dA + A \wedge A$, l'action contient des termes cubiques et quartiques en A , les équations de Yang-Mills sont alors non-linéaires.

Physiquement, un champ de jauge sert à décrire un champ véhiculant une interaction entre champs de matière. Les équations de Yang-Mills ci-dessus décrivent la propagation des champs de jauge dans le vide. Dans ce cadre, un champ de matière est une section d'un fibré vectoriel associé à P et à une représentation de G (et aussi associé au fibré des repères et à une représentation du groupe de Lorentz pour être compatible avec la relativité restreinte), la représentation de G nous disant comment le champ de matière se transforme sous une transformation de jauge. La construction d'une action invariante de jauge pour un champ de matière impose d'utiliser des dérivations covariantes : la présence de A dans les dérivées covariantes exprime le couplage entre le champ de jauge et le champ de matière.

Remarque : Les outils géométriques permettant de formuler la relativité générale et les théories de jauge sont les mêmes : ce sont ceux de la théorie des fibrés et des connexions. Cependant, ces théories sont très différentes sur le plan physique : la variable dynamique intervenant dans l'action d'une théorie de jauge est la connexion A , alors qu'en relativité générale, ce ne sont pas les symboles de Christoffel (qui sont les coefficients de la géométrie de Levi-Civita) qui interviennent mais un objet géométrique plus fondamental, à savoir la métrique. Une théorie de jauge vit sur une variété riemannienne (ou lorentzienne) (M, g) avec g fixée alors qu'en relativité générale la métrique g est une inconnue solution des équations d'Einstein.

Remarque : Dans la littérature physique, on utilise d'autres conventions. Soit A^{ph} et F^{ph} les grandeurs appelées A et F en physique. Alors : $A = -iA^{ph}$ et $F = -iF^{ph}$ avec A et F définis comme précédemment.

2.1.5 Fermions et spineurs

On est arrivé plus haut à la conclusion qu'un champ classique était une section d'un fibré vectoriel associé à une représentation du groupe de Lorentz. En fait, ce n'est pas toujours le cas : il existe des particules, les fermions, qui se transforment lors d'une transformation de Lorentz infinitésimale sous une représentation d'algèbre de Lie qui ne descend pas au groupe du fait d'une obstruction topologique : sous une rotation spatiale d'angle 2π , un champ fermionique change de signe, il ne vit donc pas dans une représentation du groupe de Lorentz. Si on se restreint pour simplifier au groupe des rotations spatiales $SO(3)$, le problème vient de la non simple connexité de $SO(3)$: le groupe fondamental est isomorphe à $\mathbb{Z}/2$ et $SO(3)$ possède un revêtement universel simplement connexe à deux feuillets qui est un groupe de Lie et qu'on appelle $Spin(3)$ (qui s'identifie en fait à $SU(2)$ du fait d'un isomorphisme exceptionnel de basse dimension). $SO(3)$ et $Spin(3)$ ont la même algèbre de Lie $Lie(SO(3)) = Lie(Spin(3))$ dont les représentations (de dimension finie) sont paramétrées par un $j \in \mathbb{N} + \frac{1}{2}\mathbb{N}$. Une représentation de $Lie(SO(3))$ associée à $j \in \mathbb{N} + \frac{1}{2}\mathbb{N}$ descend à $SO(3)$ seulement si j est entier (et descend toujours à $Spin(3)$ par simple connexité). Un champ qui se transforme sous la représentation j est dit de spin j : les bosons sont les particules de spin entier alors que les fermions sont les particules de spin demi-entier.

Le cadre algébrique adéquat pour décrire les fermions est celui des algèbres de Clifford, en particulier, il permet de donner une réalisation explicite des groupes $Spin$.

Définition 7. Soit V un espace vectoriel réel et Q une forme quadratique sur V . On appelle algèbre de Clifford associée à Q et on note $Cl(V, Q)$ l'algèbre réelle engendrée par V satisfaisant les relations : $v.w + w.v = -2Q(v, w)$. Autrement dit $Cl(V, Q)$ est le quotient de l'algèbre tensorielle de V par l'idéal bilatère engendré par les éléments de la forme $v \otimes w + w \otimes v + 2Q(v, w)$ avec $v, w \in V$. Lorsque Q est fixée, on notera $Cl(V)$ pour $Cl(V, Q)$.

Si M est une variété riemannienne alors TM est un fibré dont les fibres sont des espaces munis de produits scalaires, on peut donc former un fibré en algèbres de Clifford $Cl(TM) \otimes_{\mathbb{R}} \mathbb{C}$. On appelle fibré en modules de Clifford un fibré vectoriel E tel que chaque fibre E_m , $m \in M$, soit un module sur $Cl(T_m M) \otimes \mathbb{C}$.

Définition 8. Soit E un fibré en modules de Clifford sur une variété riemannienne M . On dit que E est un fibré de Clifford s'il est muni d'une métrique hermitienne et d'une connexion compatible telles que :

1. L'action de Clifford d'un vecteur $\xi \in T_m M$ sur E_m est antiautoadjointe :

$$\langle \xi s_1, s_2 \rangle + \langle s_1, \xi s_2 \rangle = 0$$

2. La connexion sur E est compatible avec la connexion de Levi-Civita sur M au sens où pour tout champs de vecteurs X, Y et $s \in C^\infty(M, E)$:

$$\nabla_X(Ys) = (\nabla_X Y)s + Y(\nabla_X s)$$

Remarquons que si E est un fibré de Clifford et V est un fibré vectoriel complexe muni d'une métrique hermitienne et d'une connexion compatible, alors, en faisant agir trivialement $Cl(TM \otimes \mathbb{C})$ sur V , on munit $E \otimes V$ d'une structure de fibré de Clifford.

Définition 9. L'opérateur de Dirac \mathcal{D} d'un fibré de Clifford E est l'opérateur différentiel du premier ordre sur $C^\infty(M, E)$ défini par la composition suivante :

$$C^\infty(M, E) \longrightarrow C^\infty(M, T^*M \otimes E) \longrightarrow C^\infty(M, TM \otimes E) \longrightarrow C^\infty(M, E)$$

où la première flèche est donnée par la connexion, la deuxième par la métrique et la troisième par l'action de Clifford. En termes d'une base orthonormée e_α de sections de TM , on a :

$$\mathcal{D}s = \sum_{\alpha} e_{\alpha} \nabla_{\alpha} s$$

Pour calculer \mathcal{D}^2 , on choisit un repère local synchrone en un point $m \in M$. Alors $\nabla_{\alpha} e_{\beta}(m) = 0$ et $[e_{\alpha}, e_{\beta}](m) = 0$. Ainsi, en m :

$$\mathcal{D}^2 s = \sum_{\alpha, \beta} e_{\beta} \nabla_{\beta} (e_{\alpha} \nabla_{\alpha} s) = \sum_{\alpha, \beta} e_{\beta} e_{\alpha} \nabla_{\beta} \nabla_{\alpha} s = - \sum_{\alpha} \nabla_{\alpha}^2 s + \sum_{\beta < \alpha} e_{\beta} e_{\alpha} (\nabla_{\beta} \nabla_{\alpha} - \nabla_{\alpha} \nabla_{\beta}) s$$

Soit $\mathcal{K} = \sum_{\beta < \alpha} e_{\beta} e_{\alpha} K_{\beta, \alpha}$, alors on a montré la formule de Weitzenböck :

$$\mathcal{D}^2 s = \nabla^* \nabla s + \mathcal{K} s$$

Il faut expliquer le notation ∇^* . ∇ est un opérateur différentiel de $C^{\infty}(M, E)$ dans $C^{\infty}(M, T^*M \otimes E)$. Ces fibrés étant munis de métriques, on obtient par intégration sur M un produit scalaire naturel sur les espaces de sections. Par rapport à ces produits scalaires, ∇ a un adjoint formel ∇^* et $\nabla^* \nabla$ est bien un opérateur différentiel de $C^{\infty}(M, E)$ dans lui-même.

Lemme 1. En coordonnées locales, ∇^* est donné par :

$$\nabla^*(dx^j \otimes s_j) = -g^{j,k} (\nabla_j s_k - \Gamma_{j,k}^i s_i)$$

En particulier, à l'origine d'un repère orthonormé synchrone :

$$\nabla^* \left(\sum_{\alpha} e_{\alpha} \otimes s_{\alpha} \right) = - \sum_{\alpha} \nabla_{\alpha} s_{\alpha}$$

ce qui montre que $\nabla^* \nabla$ est bien le terme intervenant dans la formule de Weitzenböck.

Remarquons que $\nabla^* \nabla$ est un opérateur positif, i.e. $\langle \nabla^* \nabla s, s \rangle = \|\nabla s\|^2 \geq 0$, d'où on déduit :

Théorème 2. Si l'opérateur \mathcal{K} est défini positif (en tant qu'endomorphisme autoadjoint sur E) alors l'équation $\mathcal{D}^2 s = 0$ ne possède pas de solution autre que $s = 0$.

Voici une propriété essentielle de \mathcal{D} :

Proposition 4. L'opérateur \mathcal{D} est autoadjoint, i.e. si s_1 et s_2 sont des sections C^{∞} de E , alors : $\langle \mathcal{D} s_1, s_2 \rangle = \langle s_1, \mathcal{D} s_2 \rangle$.

Démonstration. Par le théorème de Stokes, il suffit de montrer que l'expression locale de la différence entre les deux membres est une divergence. Notons encore \langle, \rangle la version locale du produit scalaire (avant intégration sur M), alors :

$$\begin{aligned} \langle \mathcal{D} s_1, s_2 \rangle(x) - \langle s_1, \mathcal{D} s_2 \rangle(x) &= \sum_{\alpha} \langle e_{\alpha} \nabla_{\alpha} s_1, s_2 \rangle(x) - \langle s_1, e_{\alpha} \nabla_{\alpha} s_2 \rangle(x) \\ &= \sum_{\alpha} \langle \nabla_{\alpha} e_{\alpha} s_1, s_2 \rangle(x) + \langle e_{\alpha} s_1, \nabla_{\alpha} s_2 \rangle(x) = \sum_{\alpha} \nabla_{\alpha} (\langle e_{\alpha} s_1, s_2 \rangle)(x) \end{aligned}$$

□

Le fibré extérieur comme fibré de Clifford : Soit E le fibré $\Lambda TM \otimes \mathbb{C}$ où ΛTM est le fibré extérieur. En tant que fibré vectoriel, E est isomorphe à $Cl(TM) \otimes \mathbb{C}$ via l'isomorphisme qui envoie un élément $e_1 \wedge \dots \wedge e_k$ (e_1, \dots, e_k orthonormée) de l'algèbre extérieure sur l'élément $e_1 \dots e_k$ de l'algèbre de Clifford. En transportant l'action de $Cl(TM) \otimes \mathbb{C}$ sur lui-même par multiplication à gauche via cet isomorphisme on obtient une structure de fibré de Clifford sur $\Lambda TM \otimes \mathbb{C}$. On rappelle qu'on identifie TM et T^*M grâce à la métrique sur M . L'action de Clifford d'un covecteur e sur une k -forme ω s'écrit : $e.\omega = e \wedge \omega + i_e \omega$, où i_e désigne le produit intérieur par e . ΛTM est bien un fibré de Clifford, vérifions les points de la définition : 1) résulte du fait que la multiplication intérieure est au signe près l'adjoint du produit extérieur et 2) est une conséquence du fait que la connexion de Levi-Civita est compatible avec la métrique et avec le produit extérieur.

Calculons l'opérateur de Dirac de ce fibré de Clifford :

$$\mathcal{D}\omega = \sum_{\alpha} e_{\alpha} \nabla_{\alpha} \omega = \sum_{\alpha} e_{\alpha} \wedge \nabla_{\alpha} \omega + \sum_{\alpha} i_{e_{\alpha}} \nabla_{\alpha} \omega = d\omega + d^* \omega$$

Ainsi $\mathcal{D} = d + d^*$ et $\mathcal{D}^2 = dd^* + d^*d$ n'est autre que le laplacien usuel sur les formes différentielles.

La représentation spinorielle : Soit V un espace vectoriel réel de dimension paire $n = 2r$ muni d'un produit scalaire et soit C l'algèbre de Clifford $Cl(V) \otimes \mathbb{C}$. Soit e_1, \dots, e_n une base orthonormée de V . Soit G le groupe d'ordre 2^{n+1} constitué des éléments de la forme $\pm e_1^{i_1} \dots e_n^{i_n}$ où chacun des i_1, \dots, i_n est soit 0 soit 1. En particulier $-1 \in G$. Il est utile d'avoir introduit ce groupe car il est clair que les représentations de C en tant qu'algèbre, i.e. les modules de Clifford, s'identifient aux représentations du groupe fini G dans lesquelles -1 agit comme moins l'identité. Pour étudier les modules de Clifford possibles, il ne reste plus qu'à appliquer la théorie des représentations des groupes finis.

$-1 \in G$ est une involution centrale de G et agit donc comme l'identité ou moins l'identité sur chaque représentation irréductible de G . Soit Z le sous-groupe de G à deux éléments engendré par -1 , alors un petit calcul montre que Z est le centre de G . Le quotient de G par Z est un groupe abélien d'ordre 2^n , ce qui donne 2^n représentations de dimension 1 de G . Puisque G/Z est abélien, la classe de conjugaison d'un élément $g \in G$ est réduite à $\{g\}$ si g est central et à $\{g, -g\}$ sinon. G contient donc $2^n + 1$ classes de conjugaison. Mais d'après la théorie des représentations, le nombre de classes de conjugaison est égal aux nombres de représentations irréductibles de G : G possède donc une unique représentation irréductible sur laquelle -1 agit comme moins l'identité. On en déduit qu'il existe une unique représentation irréductible de C qu'on appelle la représentation spinorielle Δ et que C s'identifie à l'algèbre des endomorphismes de l'espace vectoriel Δ , d'où la dimension de Δ : $2^{n/2} = 2^r$. On peut donner une construction explicite de Δ . Soit J une structure complexe sur V , i.e. un endomorphisme de V tel que $J^2 = -1$, compatible avec la métrique, i.e. telle que $\langle Jx, Jy \rangle = \langle x, y \rangle$. Alors on a une décomposition $V \otimes \mathbb{C} = P \oplus Q$ où P et Q sont les sous-espaces propres de $J \otimes 1$ associés aux valeurs propres respectivement $+i$ et $-i$. On peut alors munir l'algèbre extérieure ΛP d'une structure de module sur C : si $x \in \Lambda P$ et $p + q \in V \otimes \mathbb{C}$ avec $p \in P$ et $q \in Q$, on pose : $p.x = \sqrt{2}p \wedge x$, $q.x = \sqrt{2}i_q(x)$, ce qui s'étend en une action de C car p^2 et q^2 agissent comme zéro et $pq + qp$ agit comme $-2 \langle p, q \rangle$. ΛP est alors une représentation de C de dimension $2^{n/2} = 2^r$ qui est donc isomorphe à la représentation spinorielle Δ .

On appelle matrice de Dirac γ_i la représentation de e_i sur la représentation spinorielle.

Remarque : Avec la définition de l'algèbre de Clifford adoptée précédemment, on a

$$\{\gamma_i, \gamma_j\} = \gamma_i \gamma_j + \gamma_j \gamma_i = -2g_{i,j}$$

où $g_{i,j} = Q(e_i, e_j)$. Cette convention est pratique d'un point de vue mathématique car c'est elle qui assure que \mathcal{D} est autoadjoint (Proposition 4 ci-dessus) et que \mathcal{D}^2 est un opérateur

positif lorsque toutes les courbures sont nulles. Cependant, les matrices de Dirac sont définies en physique par la condition $\{\gamma_i, \gamma_j\} = 2g_{i,j}$. C'est alors $i\mathcal{D}$ qui est autoadjoint et $(i\mathcal{D})^2$ qui est positif lorsque toutes les courbures sont nulles. Dans les parties mathématiques, on utilisera la première convention et on utilisera la deuxième dans les parties physiques pour garder des formules proches de celles de la littérature.

Définition 10. Une algèbre A est appelée une superalgèbre si elle est $\mathbb{Z}/2\mathbb{Z}$ graduée, i.e. s'il existe une décomposition $A = A_0 \oplus A_1$ telle que $A_0.A_0 \subset A_0$, $A_1.A_1 \subset A_0$, $A_0.A_1 \subset A_1$ et $A_1.A_0 \subset A_1$. A_0 et A_1 s'appellent respectivement les parties paires et impaires de la superalgèbre A . Si $x \in A_i$ pour $i = 0$ ou 1 , on dit que x est homogène de degré $\deg(x) = i$.

Proposition 5. Soit V un espace euclidien réel, alors l'algèbre de Clifford de V est munie d'une unique structure de superalgèbre telle que les éléments de V soient impaires.

Démonstration. Soit e_i une base de V , on appelle partie paire (resp. impaire) de $Cl(V)$ l'espace vectoriel engendré par les produits d'un nombre pair (resp. impair) de e_i . Il résulte de la structure d'algèbre de $Cl(V)$ que ces sous-espaces sont indépendants de la base de V choisie. \square

Si A est une superalgèbre, on peut se restreindre à définir une application linéaire sur A sur les éléments homogène puis l'étendre à tout A par linéarité. On définit ainsi le supercommutateur : $[x, y]_S = xy - (-1)^{\deg(x)\deg(y)}yx$ sur les éléments homogènes et le supercentre : $Z_S(A) = \{x \in A, [x, y]_S = 0, \forall y \in A\}$.

Lemme 2. $Z_S(Cl(V))$ est réduit au corps des scalaires.

Démonstration. Soit e_1, \dots, e_k une base orthonormée de V . Soit $x \in Z_S(Cl(V))$, écrivons : $x = a + e_1b$ où a et b ne contiennent pas e_1 lorsqu'ils sont exprimés en fonction des e_i . On peut supposer x homogène et donc $\deg(x) = \deg(a) = \deg(b) + 1$. Par calcul, on montre que $xe_1 = (-1)^{\deg(x)}(e_1a + b)$ et $e_1x = e_1a - b$. Le fait que x soit dans le supercentre de A implique donc que $b = 0$, i.e. x ne contient pas e_1 . Le même argument montre que x ne contient aucun des e_i et donc que x est un scalaire. \square

On note $Cl(k)$ l'algèbre de Clifford de \mathbb{R}^k muni de son produit scalaire canonique.

Définition 11. On définit $*$: $Cl(k) \rightarrow Cl(k)$ l'opération de renversement des facteurs : si $x \in Cl(k)$ est un produit $v_1 \dots v_j$ de vecteurs de \mathbb{R}^k , alors $*x = v_j \dots v_1$.

Définition 12. On appelle $Pin(k)$ le sous-groupe du groupe des éléments inversibles de $Cl(k)$ formé des éléments x tels que :

1. x est homogène
2. $xx^* = x^*x = (-1)^{\deg(x)}$
3. pour tout $v \in \mathbb{R}^k$ on a $vxv^* \in \mathbb{R}^k$.

Le groupe $Spin(k)$ est le sous-groupe de $Pin(k)$ formé des éléments pairs.

Grâce à 3), si $x \in Pin(k)$ on définit une transformation linéaire ρ de \mathbb{R}^k en posant $\rho(x) = vxv^*$ et on vérifie facilement que $\rho : Pin(k) \rightarrow GL(k)$ est un homomorphisme de groupes.

Proposition 6. L'image de ρ est incluse dans $SO(k)$ et on a une suite exacte :

$$0 \rightarrow \mathbb{Z}/2 \rightarrow Spin(k) \rightarrow SO(k) \rightarrow 0$$

Démonstration. Si $v \in \mathbb{R}^k$, on a $|v|^2 = -v^2$. En utilisant de plus le point 2) de la définition de $Pin(k)$, on voit que si $x \in Pin(k)$, alors $\rho(x)$ préserve la norme $|\cdot|$ de \mathbb{R}^k , i.e. $\rho(Pin(k)) \subset O(k)$.

Si x est dans le noyau de $\rho: Pin(k) \rightarrow O(k)$, alors x est dans le supercentre de $Cl(V)$. Par le lemme, x est donc un scalaire et comme de plus $x^2 = 1$ on a $x = \pm 1$.

Si $w \in \mathbb{R}^k$, $|w| = 1$, alors $w \in Pin(k)$ et $\rho(w)$ est la réflexion par rapport à l'hyperplan orthogonal à w (il suffit de poser $e_1 = w$, de compléter en e_2, \dots, e_k base orthonormée de \mathbb{R}^k et de calculer l'action de $w = e_1$ sur cette base).

$\rho(Spin(k)) = SO(k)$: on utilise le fait que toute rotation est produit d'un nombre pair de réflexions. \square

ρ définit un revêtement à deux feuillet de $SO(k)$ par $Spin(k)$, l'algèbre de Lie de $Spin(k)$ s'identifie donc à celle de $SO(k)$.

Lemme 3. *Lie(Spin(k)) s'identifie au sous-espace vectoriel de $Cl(k)$ engendré par les produits $e_i e_j$, $i \neq j$. Elle s'identifie à $Lie(SO(k))$, l'espace des matrices $k \times k$ antisymétriques, via l'application qui associe à une matrice antisymétrique $(a_{i,j})_{i,j}$ l'élément $-1/4 a_{i,j} e_i e_j$ de $Cl(k)$.*

Démonstration. Puisque $(e_i e_j)^2 = -1$, on a $\exp(te_i e_j) = (\cos(t)) + e_i e_j (\sin(t)) \in Spin(k)$. Ainsi, tous les $e_i e_j$ sont dans $Lie(Spin(k))$ et en fait l'engendrent pour une raison de dimension : $\dim(Lie(Spin(k))) = k(k-1)/2$.

Si $u \in Lie(Spin(k))$ alors u agit sur $Cl(k)$ par crochet : $[u, v] = uv - vu$ et on prend pour $(a_{i,j})_{i,j}$ la matrice de cette application dans la base $(e_i)_i$ (pour les vérifications, prendre $u = e_1 e_2$ et faire les calculs sur la base des e_i). \square

Définition 13. *On appelle structure spin sur M la donnée d'un fibré principal $Spin(M)$ de groupe $Spin(n)$ sur M telle que TM soit isomorphe au fibré vectoriel associé à la représentation standard de $Spin(n)$ sur \mathbb{R}^n .*

Dans ce cas, $Spin(M)$ est un double revêtement du fibré des repères orthonormés orientés $SO(M)$ de M . Il faut faire attention : une structure spin n'existe pas forcément sur une variété M donnée, il existe une obstruction topologique. C'est relativement compréhensible : se donner une structure spin, c'est se donner un relèvement (continu) de $SO(n)$ en $Spin(n)$ en chaque point de M : puisqu'il y a en chaque point deux choix possibles, il se peut qu'il ne soit pas possible de faire un ensemble de choix cohérents les uns avec les autres. L'exemple le plus simple de variété compacte orientable n'admettant pas de structure spin est le plan projectif complexe $\mathbb{C}P^2$.

Définition 14. *Soit M une variété spin de dimension paire. On appelle fibré spinoriel (ou fibré des spineurs) S le fibré vectoriel associé au fibré principal $Spin(M)$ et à la représentation de $Spin(n)$ sur les spineurs.*

Remarquons que par construction $End(S) = Cl(TM) \otimes \mathbb{C}$.

Définition 15. *On appelle connexion spin sur $Spin(M)$ la connexion obtenue par relèvement de la connexion sur $SO(M)$ induite par la connexion de Levi-Civita sur TM . La connexion spin sur S est la connexion associée à la connexion spin sur $Spin(M)$.*

Proposition 7. *Le fibré spinoriel S (muni de sa métrique hermitienne et de la connexion spin) sur une variété spin M est un fibré de Clifford. De plus, l'opérateur \mathcal{K} apparaissant dans la formule de Weitzenbock de son opérateur de Dirac est l'opérateur de multiplication par $1/4$ de la courbure scalaire : soit \mathcal{D} l'opérateur de Dirac de S , ∇ la connexion spin, R la courbure scalaire, alors :*

$$\mathcal{D}^2 = \nabla^* \nabla + \frac{1}{4} R$$

Ce cas particulier de la formule de Weitzenbock s'appelle la formule de Lichnerowicz.

Démonstration. Vérification des points de la définition d'un fibré de Clifford.

La représentation spin est unitaire, la connexion spin et la métrique hermitienne sur le fibré spinoriel sont donc compatibles.

1. soit $\xi \in T_m M$, alors ξ agit par isométrie sur S_m d'où :

$$\langle \xi s_1, s_2 \rangle = \langle \xi \cdot \xi s_1, \xi s_2 \rangle = - \langle s_1, \xi s_2 \rangle$$

2. compatibilité entre la connexion spin et la connexion de Levi-Civita : conséquence de la construction de la connexion spin sur $Spin(M)$ par relèvement de la connexion de Levi-Civita.

Calcul pour la formule de Lichnerowicz : soit $(e_i)_i$ une base orthonormée de $T_m M$. L'expression de \mathcal{K} fait intervenir la courbure de la connexion spin sur S qui peut donc s'exprimer en fonction de la connexion spin sur $Spin(M)$, i.e. on peut voir $K_{i,j}$ comme une 2-forme à valeurs dans $Lie(Spin(n))$. D'après le lemme reliant $Lie(Spin(n))$ et $Lie(SO(n))$, on a donc : $K_{i,j} = -\frac{1}{4} R_{k,l,i,j} e_k e_l$ où R est le tenseur de courbure de Riemann. Ainsi :

$$\mathcal{K} = \frac{1}{2} e_i e_j K_{i,j} = -\frac{1}{8} e_i e_j e_k e_l R_{k,l,i,j} = \frac{1}{8} e_i e_j e_l R_{k,l,i,j} e_k$$

Or : $e_i e_j e_l R_{k,l,i,j} = -2e_i Ric_{i,k}$ d'où : $\mathcal{K} = -\frac{1}{4} e_i e_k Ric_{i,k} = \frac{1}{4} Ric_{i,i} = \frac{R}{4}$. \square

Corollaire 1. *Soit M une variété spin compacte de courbure scalaire positive en tout point et strictement positive en au moins un point, alors $Ker \mathcal{D} = 0$, i.e. M ne possède pas de spineurs harmoniques (en particulier $ind \mathcal{D} = 0$).*

Ce résultat obtenu par Lichnerowicz en 1963, combiné au théorème de l'indice (voir plus loin), a donné le premier exemple d'une obstruction topologique (à savoir M spin + $\hat{A}(M) \neq 0$) à l'existence d'une métrique de courbure scalaire strictement positive sur une variété, ce qui constitue un résultat non-trivial : par exemple, toute variété compacte M admet une métrique de courbure scalaire strictement négative (en fait, on peut même imposer à la courbure de Ricci d'être strictement négative).

Revenons à la physique : on se place en signature lorentzienne $(- + \dots +)$ et on adopte la convention «physique» pour la définition des matrices de Dirac et donc de l'opérateur de Dirac : $\{\gamma_i, \gamma_j\} = 2g_{i,j}$ (Voir la remarque qui précède 2.1.5, Définition 10). On appelle spineur de Dirac une section d'un fibré de la forme $S \otimes E$. Soit $\gamma_1, \dots, \gamma_n$ les éléments d'une base locale du fibré tangent, identifiés comme ci-dessus à des éléments du fibré en algèbres de Clifford. Soit $\gamma_5 = i^{\frac{n}{2}-1} \nu$ où $\nu = \gamma_1 \dots \gamma_n$. Attention, il y a ici un risque de confusion dans les notations : ce γ_5 n'est pas γ_i avec $i = 5$, dans la suite, quand on écrira explicitement γ_5 , il s'agira de $i^{\frac{n}{2}-1} \nu$. On vérifie facilement que ν et γ_5 sont définis indépendamment du choix de la base (γ_i) , qu'ils anticommulent avec les e_i et que γ_5 a été construit à partir de ν de façon à être une involution : $\gamma_5^2 = 1$. γ_5 s'appelle l'opérateur de chiralité. On peut alors décomposer S en une somme directe : $S = S^+ \oplus S^-$ correspondant aux espaces propres de γ_5 . Une section v de S^+ (resp. S^-) vérifie : $\gamma_5 v = v$ (resp. $-v$). v s'appelle un spineur de chiralité positive (resp. négative). Un spineur de chiralité donnée s'appelle un spineur de Weyl. Un spineur de Dirac peut être considéré comme la donnée de deux spineurs de Weyl, un de chiralité positive et un autre de chiralité négative.

Physiquement, au niveau d'une théorie de champs classiques, on décrit un fermion de spin 1/2 par un spineur. Ce qui précède montre qu'on a plusieurs choix possibles : spineurs de Dirac,

de Weyl ... (il existe d'autres possibilités suivant les valeurs de n). Par exemple, un électron dans notre espace-temps, $n = 4$, est décrit par un spineur de Dirac. La fibre de S en chaque point est alors de dimension $2^{4/2} = 4$: localement, un spineur de Dirac est une fonction à quatre composantes complexes. Dans les modèles les plus simples, un neutrino est décrit par un spineur de Weyl : localement, on considère une fonction à deux composantes complexes.

On vient de voir les objets permettant de décrire des fermions de spin 1/2 dans une théorie de champs classiques. Passons à la construction d'une action. Soit ψ un spineur de Dirac. On note $\bar{\psi} = \psi^\dagger i\gamma_0$ où \dagger est la conjugaison hermitienne. $\bar{\psi}$ a été définie de telle façon que $\bar{\psi}\psi$ soit une quantité scalaire bien définie. La densité lagrangienne décrivant un spineur de Dirac de masse nulle est alors :

$$\mathcal{L}(\psi, \bar{\psi}) = -\bar{\psi}\mathcal{D}\psi$$

L'équation du mouvement associée est l'équation de Dirac $\mathcal{D}\psi = 0$. Pour un spineur de Dirac massif, il suffit de rajouter un terme proportionnel à $\bar{\psi}\psi$ dans la densité lagrangienne.

2.2 Physique quantique et théorie des champs

Qu'est-ce qu'un champ quantique ?

Dans la partie précédente, on a montré comment construire une théorie de champs classiques sur une variété. On peut donc se demander s'il existe une version quantique de cette théorie.

2.2.1 Principes de la physique quantique

Commençons par rappeler ce qu'est une théorie quantique. Une description quantique d'un phénomène physique correspond à la donnée d'un espace de Hilbert complexe \mathcal{H} , l'espace des états du système étudié, et d'opérateurs (en général non bornés) autoadjoints agissant sur \mathcal{H} , qui décrivent les quantités physiques observables du système. La physique quantique prédit des résultats probabilistes : étant donné un état $|\psi\rangle$ normalisé et un opérateur A , une mesure sur $|\psi\rangle$ de la quantité physique associée à A donnera pour résultat une valeur propre (réelle) de A , la probabilité d'obtenir λ étant le carré de la norme de la projection orthogonale de $|\psi\rangle$ sur l'espace propre associé à λ . Immédiatement après la mesure, le système est dans l'état projeté orthogonal sur l'espace propre associé au résultat de la mesure. Il y a (au moins) deux façons d'envisager la dynamique en mécanique quantique :

1. Le point de vue de Schrödinger : les observables sont indépendantes du temps, les états évoluent au cours du temps, cette évolution étant caractérisée par la donnée d'un opérateur autoadjoint H appelé hamiltonien et l'équation de Schrödinger : $i\hbar \frac{d|\psi(t)\rangle}{dt} = H|\psi(t)\rangle$ où $\hbar = 1,054.10^{-34} J.s$ est la constante de Planck réduite.
2. Le point de vue de Heisenberg : les états sont indépendants du temps, les observables évoluent au cours du temps, cette évolution étant caractérisée par la donnée d'un opérateur auto-adjoint H appelé hamiltonien et l'équation de Heisenberg : $i\hbar \frac{dA}{dt} = [A, H]$

Ces deux points de vue sont entièrement équivalents.

Il est important de souligner que ce qu'on appelle physique quantique est le cadre abstrait formulé ci-dessus et non pas une de ses réalisations particulières. Étant donné un phénomène physique réel, donner une description quantique revient (en théorie) à construire un espace de Hilbert, des observables et un hamiltonien décrivant la réalité expérimentale. Une manière de réaliser cette construction est par exemple de partir d'un modèle classique et d'essayer de le quantifier.

2.2.2 Le problème de la quantification d'une théorie de champs classiques

Une première tentative de description quantique d'une particule élémentaire consiste à prendre pour espace de Hilbert l'espace des fonctions définies sur \mathbb{R}^3 à valeurs complexes de carré intégrable, de décrire les observables de position et d'impulsion respectivement par des opérateurs de multiplication et de dérivation et de prendre pour hamiltonien une copie d'un hamiltonien classique en ayant remplacé les variables de position et d'impulsion classiques par leurs analogues opératoriels quantiques. L'idée est ainsi de décrire par exemple un électron par une fonction d'onde donnant l'amplitude de probabilité d'observer un électron en un point fixé. Dans ce cadre, la quantification consiste à penser la matière en termes ondulatoires. Néanmoins, décrire une particule par une onde, i.e. par un champ classique n'est pas suffisant : la relativité restreinte impose la non-conservation du nombre de particules, les équations de champs classiques (Klein-Gordon, Dirac...) sont difficiles à interpréter physiquement dans ce cadre ... Il s'agit donc de savoir comment quantifier une théorie de champs classiques. Remarquons que ce problème se pose également naturellement lorsqu'on s'intéresse aux propriétés quantiques du champ électromagnétique : le cadre de la théorie quantique des champs permet un traitement unifié des particules de matières et des champs véhiculant les interactions autres que la gravitation.

Commençons par donner un cadre axiomatique. On part d'une théorie de champs classiques : on a un espace-temps M (une variété lorentzienne de dimension n , par exemple l'espace de Minkowski) et un fibré vectoriel E sur M dont les sections sont les différentes configurations du champ classique. Une théorie quantique des champs obtenue par quantification de cette théorie classique est la donnée d'un espace de Hilbert \mathcal{H} , d'une «application de quantification» qui à toute section ϕ de E associe un champ d'opérateurs Φ (en général non bornés) agissant sur \mathcal{H} et d'un vecteur de \mathcal{H} , appelé «vide» et notée $|0\rangle$, ces données étant soumises à des conditions de cohérence physique : covariance sous le groupe de Poincaré, causalité, unitarité ... (il s'agit des axiomes de Wightman qu'on ne développera pas ici). Un champ quantique Φ est donc un champ d'opérateurs : pour tout $x \in M$, $\Phi(x)$ est un opérateur agissant sur \mathcal{H} . On demande que \mathcal{H} soit l'adhérence de l'espace vectoriel engendré par les éléments de la forme $\Phi(x_1)\dots\Phi(x_N)|0\rangle$.

Les grandeurs d'intérêt physique sont les fonctions $(x_1, \dots, x_N) \rightarrow \langle 0|T\Phi(x_1)\dots\Phi(x_N)|0\rangle$, appelées amplitudes de probabilité, ou encore, par analogie avec la physique statistique, fonctions de corrélation. Le T signifie qu'on prend le T -produit (ou produit chronologique) : $T\Phi(x_1)\dots\Phi(x_N) = \Phi(y_1)\dots\Phi(y_N)$ où (y_1, \dots, y_N) est une permutation de (x_1, \dots, x_n) qui vérifie au niveau des composantes temporelles $y_1^0 > \dots > y_N^0$. Les fonctions de corrélation sont les quantités de base que l'on cherche à calculer en théorie quantique des champs : du point de vue pratique, toutes les grandeurs physiques mesurables expérimentalement (sections efficaces, périodes de désintégration ...) sont calculables à partir de ces fonctions de corrélation et d'un point de vue théorique, on peut montrer que la connaissance des fonctions de corrélation est équivalente à la donnée de l'espace de Hilbert \mathcal{H} et de l'application de quantification (plus précisément, si pour tout champ classique ϕ , on se donne des fonctions $f_{\phi,N}$ satisfaisant certaines conditions, alors on peut construire un espace de Hilbert \mathcal{H} , un vide $|0\rangle$ et une application de quantification telle que $\langle 0|T\Phi(x_1)\dots\Phi(x_N)|0\rangle = f_{\phi,N}$).

Tout le problème est de savoir comment associer une théorie quantique des champs à une théorie de champs classiques, autrement dit comment construire des données (espace de Hilbert, vide, application de quantification) satisfaisant aux conditions précédentes. La réponse est connue pour une théorie de champs classiques définies sur \mathbb{R}^n décrivant un champ de matière libre, i.e. soumis à aucune interaction. Dans ce cas, le point clé est que les équations de la théorie classique sont linéaires (par exemple, équation de Klein-Gordon pour un champ de spin 0 ou équation de Dirac pour un champ de spin 1/2) : on peut donc les résoudre par transformation de Fourier et on obtient une description du champ classique sous la forme d'une somme infinie d'os-

cillateurs harmoniques. La quantification d'un oscillateur harmonique en mécanique quantique non-relativiste étant bien connue, on se doute qu'on va s'en sortir. Dans le cas le plus simple d'un champ scalaire réel, on définit \mathcal{H}_1 l'espace de Hilbert engendré par les modes propres de fréquence positive de l'équation de Klein-Gordon, on pose $\mathcal{H} = \bigoplus_{k=0}^{\infty} \mathcal{H}_1^{\otimes k}$ et on construit explicitement des opérateurs $a(\vec{p})$ et $a^\dagger(\vec{p})$, dits respectivement d'annihilation et de création, agissant sur \mathcal{H} et satisfaisant les relations de commutation caractéristiques d'un oscillateur harmonique. Physiquement, $a^\dagger(\vec{p}_1)\dots a^\dagger(\vec{p}_N)|0\rangle$ décrit un état constitué de N particules d'impulsions respectives $\vec{p}_1, \dots, \vec{p}_N$. L'application de quantification est construite en associant $a(\vec{p})$ et $a^\dagger(\vec{p})$ aux ondes planes d'impulsion p solutions de Klein-Gordon d'énergie respectivement négative et positive et en la prolongeant par linéarité à l'ensemble des solutions de Klein-Gordon. Le cas d'un champ spinoriel est analogue, à ceci près qu'il faut imposer des conditions d'anticommutation et non de commutation aux opérateurs de création et d'annihilation.

On vient de décrire très brièvement comment quantifier une théorie de champs de matière libre, la méthode suivie est dite de quantification canonique. On pourrait se demander si on peut quantifier de la même manière une théorie comprenant uniquement un champ de jauge : en effet, sans autre champ, on a l'impression d'avoir affaire à un champ classique libre et la situation devrait donc être aussi simple que ci-dessus. Néanmoins, il y a deux niveaux de difficulté. Tout d'abord, il faut prendre en compte l'invariance de jauge pour ne pas compter plusieurs fois les mêmes états physiques et se retrouver avec des incohérences : pour le champ électromagnétique, i.e. une théorie de groupe $U(1)$, abélien, les équations classiques sont linéaires (équations de Maxwell) et on peut réussir à le quantifier canoniquement à condition de bien prendre en compte l'invariance de jauge mais ce n'est pas si facile. D'autre part, les théories de jauge non-abéliennes ne sont pas libres au sens où les équations classiques seraient linéaires : elles ne le sont pas (équations de Yang-Mills), le champ de jauge interagit avec lui-même du fait de la présence de termes cubiques et quartiques dans l'action.

2.2.3 L'approche perturbative

On est donc conduit à la question suivante : comment quantifier les théories avec interactions (qui en fait sont les seules d'intérêt physique car tout phénomène physique est le résultat d'une interaction) ? Dans ce cas, les équations classiques sont non-linéaires, en général, on ne sait pas les résoudre et il n'y a pas d'espoir pour la quantification canonique. Trouver une manière de faire mathématiquement rigoureuse est le but de la théorie constructive des champs. Pendant longtemps, aucun exemple non trivial (i.e. avec interaction) de théorie quantique des champs entièrement bien définie n'était connu et ce n'est qu'à la fin des années 1970, début 1980, que les premiers exemples furent construits, essentiellement en basse dimension (exemple : modèle de Gross-Neveu massif à deux dimensions, les théories de champs conformes à deux dimensions ...) mais le cas de plus grand intérêt physique, celui des théories de jauge non-abéliennes à quatre dimensions reste largement ouvert (c'est en gros un des « problèmes du millénaire » proposés par l'institut Clay).

Quoiqu'il en soit, les physiciens ont trouvé depuis longtemps (fin des années 1940 pour l'électrodynamique) une manière d'exploiter physiquement, i.e. pour faire des prévisions expérimentales, les idées de la théorie quantique des champs : il est possible de définir ces théories de manière perturbative, i.e. par développement en série en la constante de couplage caractéristique des interactions autour d'une théorie libre. La façon la plus commode de présenter ce formalisme consiste à introduire des « intégrales fonctionnelles ». Pour simplifier, considérons une théorie classique comprenant un unique champ de manière ϕ et décrite par une densité lagrangienne $\mathcal{L}(\phi) = \mathcal{L}_0(\phi) + \mathcal{L}_{int}(\phi)$ où $\mathcal{L}_0(\phi)$ est la densité lagrangienne d'une théorie libre, i.e. une fonction quadratique en ϕ et $\partial\phi$ alors que $\mathcal{L}_{int}(\phi)$ comprend les termes d'interaction. On note

$S(\phi) = S_0(\phi) + S_{int}(\phi) = \int d^n x \mathcal{L}(\phi)$ l'action correspondante. Dans la théorie classique, on obtient les équations du mouvement à partir d'un principe variationnel, par minimisation de S sur tous les champs ϕ possibles. Ainsi, si \mathcal{O} est une fonction de ϕ , la valeur observée expérimentalement sera $\mathcal{O}(\phi_{ph})$ avec $S(\phi_{ph}) = \inf_{\phi} S(\phi)$. Feynman a compris l'analogie quantique de ce principe : plutôt que de prendre la valeur sur le champ minimisant l'action, il faut sommer sur tous les champs possibles en affectant un champ ϕ d'un poids $e^{i\frac{S(\phi)}{\hbar}}$: la valeur moyenne de \mathcal{O} sera : $\langle \mathcal{O} \rangle = \mathcal{N} \int e^{i\frac{S(\phi)}{\hbar}} \mathcal{O}(\phi) D\phi$ où l'intégration se fait sur tous les champs classiques. On parle d'intégrale fonctionnelle car il faut intégrer sur un espace de fonctions ce qui malheureusement n'est pas bien défini mathématiquement en général. Une théorie quantique des champs étant définie par ses fonctions de corrélation, on sait formellement en définir une en posant :

$$\langle 0|T\Phi(x_1)\dots\Phi(x_N)|0\rangle = \frac{\int e^{i\frac{S(\phi)}{\hbar}} \phi(x_1)\dots\phi(x_N) D\phi}{\int e^{i\frac{S(\phi)}{\hbar}} D\phi}$$

On peut considérer que le but de la théorie constructive des champs est de donner un sens à ces intégrales. Bien sûr, ces expressions formelles ne sortent pas de nulle part : dans le cas d'une théorie libre, on a à calculer une intégrale gaussienne en ϕ , ce qu'on peut faire formellement en utilisant les formules donnant une intégrale gaussienne sur \mathbb{R}^n et on trouve le même résultat que par quantification canonique. En fait, ceci est bien défini mathématiquement car on sait construire des mesures gaussiennes sur des espaces de fonctions : pour s'en convaincre, il suffit de se rappeler que construire un mouvement brownien est équivalent à construire la mesure de Wiener. En fait on ne sait faire cette construction mathématique rigoureuse que dans le cas d'une «vraie» gaussienne $e^{-\frac{S(\phi)}{\hbar}}$ avec $S(\phi) > 0$ et non pas dans le cas d'une intégrale «oscillante» comme $e^{i\frac{S(\phi)}{\hbar}}$. On rencontre une vraie gaussienne lorsqu'on se place en signature euclidienne : si on note S_E l'action euclidienne, alors les fonctions de corrélation euclidiennes s'écrivent :

$$\langle 0|\Phi(x_1)\dots\Phi(x_N)|0\rangle = \frac{\int e^{-\frac{S(\phi)}{\hbar}} \phi(x_1)\dots\phi(x_N) D\phi}{\int e^{-\frac{S(\phi)}{\hbar}} D\phi}$$

Pour la fin de cette section, on reste en signature lorentzienne.

L'idée de l'approche perturbative est de poser $d\mu = e^{i\frac{S_0(\phi)}{\hbar}} D\phi$, mesure gaussienne par rapport à laquelle on sait intégrer formellement (au moins pour des polynômes) et on développe l'exponentielle restante en série :

$$\langle 0|T\Phi(x_1)\dots\Phi(x_N)|0\rangle = \frac{\sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{i}{\hbar}\right)^k \int \phi(x_1)\dots\phi(x_N) S_{int}(\phi)^k d\mu}{\sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{i}{\hbar}\right)^k \int S_{int}(\phi)^k d\mu}$$

Chaque terme se calcule formellement par intégration par partie : si Q est une forme quadratique non-dégénérée sur un espace vectoriel de dimension finie, on définit L une forme linéaire par la condition $\partial_{Q^{-1}(L)}(Q/2) = L$ où $\partial_{Q^{-1}(L)}$ signifie dérivée partielle évaluée en $Q^{-1}(L)$ et alors la formule d'intégration par partie s'écrit, si P est un polynôme :

$$\int P(X) L(X) e^{-\frac{Q(X)}{2}} DX = - \int P(X) \partial_{Q^{-1}(L)} (e^{-\frac{Q(X)}{2}}) DX = \int (\partial_{Q^{-1}(L)} P(X)) e^{-\frac{Q(X)}{2}} DX$$

Dans l'application formelle qu'on fait de cette formule, Q est une forme quadratique contenant des opérateurs différentiels, les quantités du type Q^{-1} doivent alors s'interpréter comme des fonctions de Green (ou propagateurs).

L'application répétée d'intégrations par parties donne une série de termes qui peuvent s'interpréter graphiquement à l'aide des diagrammes de Feynman. On renvoie à un ouvrage de théorie

quantique des champs pour ces aspects plus techniques (théorème de Wick ...). Un graphe de Feynman est formé d'arêtes, on a un type d'arête pour chaque type de champs, de sommets externes, on a autant de sommets externes que de champs dans la fonction de corrélation qu'on cherche à calculer, et de sommets internes, on a un type de sommet pour chaque terme d'interaction dans le lagrangien. La conclusion de cette démarche est l'écriture :

$$\langle 0|T\Phi(x_1)\dots\Phi(x_N)|0\rangle = \sum_G \frac{1}{\sigma(G)} \int \frac{dp_1}{(2\pi)^n} \dots \frac{dp_N}{(2\pi)^n} V(G)(p_1, \dots, p_N) e^{i(x_1 p_1 + \dots + x_N p_N)}$$

où on somme sur tous les graphes de Feynman G ayant N lignes externes de champs ϕ , $\sigma(G)$ est un facteur de symétrie de nature combinatoire et $V(G)(p_1, \dots, p_N)$ est la contribution du graphe G avec p_1, \dots, p_N pour impulsions entrantes. Un graphe avec N sommets internes décrit une situation avec N interactions, i.e. est d'ordre N au niveau perturbatif. La mise en place de la théorie des perturbations pour calculer $\langle 0|T\Phi(x_1)\dots\Phi(x_N)|0\rangle$ consiste donc à calculer la contribution des graphes de Feynman dans un ordre croissant de complexité.

Il est donc naturel de se demander : la combinatoire des graphes de Feynman peut-elle se simplifier ? On va voir que oui et que les objets mis en jeu sont d'une grande importance physique. Tout d'abord, on peut rassembler toutes les fonctions de corrélation en une unique fonction génératrice :

$$Z[J] = \mathcal{N} \int e^{iS(\phi) + \langle J, \phi \rangle} D\phi = \sum_{N=0}^{+\infty} \frac{i^N}{N!} \int J(x_1) \dots J(x_N) \langle 0|T\Phi(x_1) \dots \Phi(x_N)|0\rangle dx_1 \dots dx_N$$

où on a introduit un terme de source : J vit dans l'espace dual des champs classiques, $\langle J, \phi \rangle = \int J(x)\phi(x)dx$. On retrouve les fonctions de corrélations à N points par N dérivations fonctionnelles de $Z[J]$ par rapport à J , évaluées en $J = 0$. La contribution d'un graphe de Feynman disconnexe s'obtient en faisant le produit des contributions de ses composantes connexes. En tenant compte de l'invariance du graphe sous le groupe des permutations de ses composantes connexes, on obtient donc

$$Z[J] = \sum_{N=0}^{\infty} \frac{1}{N!} (iW[J])^N = e^{iW[J]}$$

avec

$$iW[J] = \log Z[J] = \sum_{N=0}^{+\infty} \frac{i^N}{N!} \int J(x_1) \dots J(x_N) \langle 0|T\Phi(x_1) \dots \Phi(x_N)|0\rangle_c dx_1 \dots dx_N$$

où $T\langle 0|\Phi(x_1)\dots\Phi(x_N)|0\rangle_c$ désigne une fonction de corrélation calculée en prenant en compte uniquement les graphes de Feynman connexes. $iW[J]$ est donc la fonction génératrice de ces fonctions de corrélation connexes. Pour aller plus loin dans la simplification de la combinatoire des graphes de Feynman, on introduit la notion de graphe « irréductible à une particule », 1PI pour simplifier :

Définition 16. *Un graphe de Feynman connexe est dit 1PI s'il n'est pas un arbre et s'il ne peut être rendu disconnexe en coupant une seule arête.*

Pour faire apparaître cette notion, on introduit l'action effective quantique Γ qui n'est autre que la transformée de Legendre de $W[J]$, i.e. $\Gamma(\phi) = W[J(\phi)] - \langle \phi, J(\phi) \rangle$ où J comme fonction de ϕ est défini par la condition $\frac{\delta W}{\delta J}[J(\phi)] = \phi = \langle 0|\Phi(x)|0\rangle_J$ où Φ est l'opérateur quantique associé au champ classique ϕ . On peut reformuler le lien entre ϕ et $J(\phi)$ de la façon suivante : ϕ est l'amplitude vide-vide du champ quantique Φ en présence du courant classique $J(\phi)$. On peut

montrer que Γ est la fonction génératrice des fonctions de corrélations calculées en ne tenant compte que des graphes PI :

$$i\Gamma[\phi] = \sum_{N=0}^{+\infty} \frac{i^N}{N!} \int J(x_1) \dots J(x_N) \langle 0 | T \Phi(x_1) \dots \Phi(x_N) | 0 \rangle_{1PI} dx_1 \dots dx_N$$

Puisque n'importe quel graphe peut se décomposer en graphe $1PI$, on voit que la théorie quantique des champs est équivalente à une théorie des champs classiques d'action Γ , d'où le nom d'action quantique effective (ceci peut également se voir de manière non-perturbative : on a $\frac{\delta \Gamma[\phi]}{\delta \phi(y)} = -J(\phi)(y)$: les valeurs possibles des champs ϕ en l'absence du courant J , i.e. les valeurs moyennes $\langle 0 | \Phi | 0 \rangle$, sont donc données par les points stationnaires de Γ). Le problème de base d'une théorie quantique des champs est de déterminer la fonctionnelle non-linéaire en les champs classiques formant l'action quantique effective. De manière perturbative, cela se fait par évaluation des graphes $1PI$.

Renormalisation : Dans ce qui précède, on a passé sous silence une difficulté qui est essentielle pour avoir un aperçu fidèle de ce qu'est une théorie quantique des champs. On a parlé de la contribution $V(G)(p_1, \dots, p_N)$ d'un graphe de Feynman G , qui s'obtient après application répétée d'intégrations par parties formelles. On a déjà noté que ceci faisait apparaître des fonctions de Green d'opérateurs différentiels qui dans l'espace des impulsions sont des fractions rationnelles. Un graphe de Feynman peut contenir des boucles et pour calculer $V(G)(p_1, \dots, p_N)$, il faut intégrer sur toutes les impulsions possibles pouvant circuler dans cette boucle. Autrement dit, $V(G)(p_1, \dots, p_N)$ apparaît comme une intégrale multiple, une intégrale par boucle, d'une fraction rationnelle dont les coefficients sont des fonctions de p_1, \dots, p_N . Le problème est que cette intégrale est en général divergente.

Ce problème des divergences a été observé dès les premières tentatives de quantifier le champ électromagnétique dans les années 30. Considérons par exemple un électron dans un champ électromagnétique : l'électron est couplé à toutes les fluctuations quantiques du champ électromagnétique, celles-ci pouvant être de fréquence arbitrairement élevée, d'où des divergences dites ultraviolettes. Une observation naïve de ces divergences conduirait à penser qu'un électron possède une masse infinie, que le champ électromagnétique contient une énergie infinie... tout ceci étant physiquement inacceptable.

En fait, il n'y a pas d'incohérence physique. En effet, les divergences précédentes se produisent si les paramètres intervenant dans le lagrangien classique sont finis. Or ces grandeurs qui peuvent être des masses, des charges n'ont pas de signification physique. Considérons par exemple le cas de l'électrodynamique. Dans le lagrangien de la théorie classique apparaît une masse qu'on aurait envie d'appeler la masse de l'électron. Mais la véritable interprétation de cette masse est d'être la masse d'un électron «nu», i.e. non-couplé au champ électromagnétique. Ce qu'on observe expérimentalement comme étant la masse de l'électron est une masse effective prenant en compte l'interaction avec le champ électromagnétique (techniquement, c'est la grandeur qui apparaît dans la théorie comme le pôle du propagateur quantique effectif de l'électron). La masse «nue» n'a pas de réalité physique, la masse «réelle» en a une. Si on introduit un paramètre de cut-off Λ , i.e. si on n'intègre que sur des fluctuations d'énergie plus petite que Λ de façon à avoir en permanence des intégrales convergentes, alors il n'y a pas d'empêchement à faire dépendre les paramètres «nus» de Λ de façon à ce que les paramètres réels soient égaux aux valeurs observées expérimentalement quand $\Lambda \rightarrow \infty$. Dans cette limite, les paramètres «nus» divergent mais comme on vient de l'expliquer, ceci ne pose pas de problème physique.

S'il faut ajuster toutes les grandeurs physiques en fonction de leurs valeurs expérimentales, alors la théorie n'a plus de pouvoir prédictif et est donc sans intérêt. On dit qu'une théorie est

renormalisable s'il suffit de fixer expérimentalement un nombre fini de paramètres. C'est par exemple le cas de l'électrodynamique quantique : il faut fixer la charge électrique de l'électron, la masse de l'électron ... à partir de l'expérience mais on peut prévoir le moment magnétique anormal de l'électron, le déplacement de Lamb ... La théorie étant définie de manière perturbative, le processus d'ajustement des paramètres «nus», ce qu'on appelle la renormalisation, se refait à chaque ordre perturbatif. Pour qu'une théorie quantique des champs est un sens, il faut qu'elle soit renormalisable : on peut montrer que c'est une contrainte très forte sur les termes d'interaction pouvant être mis dans la densité lagrangienne classique.

Techniquement, la renormalisation peut se faire de multiples façons. L'introduction d'un cut-off Λ n'est pas forcément la meilleure manière de faire car ceci peut briser des symétries de la théorie (par exemple une symétrie de jauge). Une méthode très efficace est la régularisation dimensionnelle : les intégrales divergentes sont des intégrales sur l'espace-temps de dimension n , qui deviennent en générales convergentes en des dimensions d d'espace-temps différentes. Les formules obtenues pour ces dimensions s'étendent formellement pour définir des fonctions $I(d)$ avec $d \in \mathbb{R}$ avec $I(d)$ qui diverge quand $d \rightarrow n$. On peut montrer que I s'étend en une fonction méromorphe $I(z)$ sur \mathbb{C} avec un pôle en $z = n$, on définit alors l'intégrale régularisée en dimension n comme étant la partie finie de I en $z = n$. Mais il existe quelques cas pour lesquels il est difficile de mettre en place la régularisation dimensionnelle et ce sont ces cas qu'on rencontrera par la suite : il n'est pas a priori facile de donner un équivalent en dimension réelle non-entière à une matrice γ_5 . On verra quand on en aura besoin comment se sortir de cette situation.

3 Anomalie chirale et théorème de l'indice pour les opérateurs de Dirac

3.1 Physique de l'anomalie chirale

3.1.1 Définition

On se place dans le cadre d'une théorie quantique des champs obtenue par quantification d'une théorie des champs classique. De façon très générale, on dit que cette théorie possède une anomalie si une des symétries de la théorie classique est absente de la théorie quantique, autrement dit si cette symétrie est brisée par des effets purement quantiques.

La notion d'anomalie est à bien distinguer de celle de symétrie spontanément brisée dont on reparlera un peu plus loin et qui est aussi très importante en physique des particules. On dit qu'une symétrie de l'action de la théorie est spontanément brisée si l'état du vide de cette théorie ne possède pas cette symétrie. Le même phénomène peut se produire dans une théorie purement classique : une solution classique peut ne pas posséder une symétrie de l'action (c'est un phénomène courant en physique statistique lorsqu'on étudie les transitions de phase).

Revenons aux anomalies. La première étudiée, qu'on appellera « anomalie chirale », concerne le couplage d'un fermion de Dirac sans masse sur un espace-temps de dimension paire à un champ de jauge. Plus précisément, soit (M, g) une variété lorentzienne de dimension paire $n = 2r$, on se place en signature $(- + \dots +)$ et on adopte la convention « physique » pour la définition des matrices de Dirac et donc de l'opérateur de Dirac : $\{\gamma_i, \gamma_j\} = 2g_{i,j}$ (Voir la remarque qui précède 2.1.5, Définition 10). Le fibré spinoriel se décompose alors en : $S = S^+ \oplus S^-$. On rappelle qu'un spineur de Dirac ψ est une section de S et peut donc se décomposer en deux spineurs de Weyl ψ_+ et ψ_- dits de chiralité positive et négative vérifiant respectivement $\gamma_5 \psi_+ = \psi_+$ et $\gamma_5 \psi_- = -\psi_-$. Soit G un groupe de jauge, on couple les spineurs de Dirac à G via une représentation de G . La densité lagrangienne fermionique de la théorie classique considérée est alors $\mathcal{L} = -\psi \not{D} \psi$ où \not{D} est l'opérateur de Dirac associé aux données ci-dessus. Cette densité lagrangienne possède une symétrie globale de groupe $U(1)$ donnée par $\psi(x) \rightarrow e^{i\alpha \gamma_5} \psi(x)$ qui échange composantes de chiralité positive et négative. On parle donc de symétrie chirale. On va voir que cette symétrie n'est pas conservée au niveau quantique : on parle d'anomalie chirale. Par le théorème de Noether, la symétrie chirale est associée à la conservation d'un courant $J_5^\mu = i\bar{\psi} \gamma^\mu \gamma_5 \psi$ qui s'écrit $\partial_\mu J^\mu = 0$. L'anomalie chirale se traduit donc par la non-conservation du courant J^μ , ce qui donne un moyen de quantifier l'anomalie par la valeur de $\partial_\mu J^\mu$.

On va commencer par donner un exemple simple, à deux dimensions, pour essayer de comprendre comment l'anomalie chirale peut apparaître. Cet exemple est tiré de [2] et les formules suivantes sont écrites en signature $(+-)$ et non pas $(-+)$. On considère un espace-temps à deux dimensions, muni de coordonnées (x^0, x^1) , la dimension spatiale étant compactifiée en un cercle de longueur L pour simplifier. Soit $\psi = \begin{pmatrix} u_+ \\ u_- \end{pmatrix}$ un champ fermionique constitué de deux spineurs de Weyl. On commence par considérer une théorie de champs libres. On choisit comme représentation des matrices de Dirac : $\gamma_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\gamma_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ et $\gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. L'équation de Dirac s'écrit : $(\partial_0 - \partial_1)u_+ = 0$, $(\partial_0 + \partial_1)u_- = 0$ dont la solution générale est $u_+ = u_+(x^0 + x^1)$, $u_- = u_-(x^0 - x^1)$. Du point de vue des champs classiques, les spineurs de Weyl sont donc deux paquets d'onde se déplaçant suivant la direction spatiale, suivant les x^1 décroissants pour u_+ et croissants pour u_- . Les particules libres de moment bien défini $p^\mu = (E, p)$ sont alors :

$$u_\pm^E(x^0 \pm x^1) = \frac{1}{\sqrt{L}} e^{-iE(x^0 \pm x^1)}, p = \mp E$$

Remarquons qu'on obtient des solutions d'énergies positives et négatives : pour u_+ , les solutions d'énergie positive (resp. négative) sont celles de moment négatif (resp. positif) et inversement pour u_- . D'autre part, le caractère compact de la dimension spatiale impose une quantification des moments : $p = 2\pi n/L$ où $n \in \mathbb{Z}$. On a donc déterminé le spectre de la théorie. Passons à la quantification : à une solution d'énergie positive, on associe un opérateur d'annihilation et à une solution d'énergie négative, on associe un opérateur de création de l'antiparticule correspondante. Le vide de la théorie quantique, noté $|0, \pm\rangle$, peut s'interpréter comme l'ensemble des états classiques d'énergie négative (la fameuse « mer de Dirac »). Les champs généraux s'écrivent

$$u_{\pm} = \sum_{E>0} [a_{\pm}(E)v_{\pm}^E(x) + b_{\pm}^{\dagger}(E)v_{\pm}^{E*}(x)]$$

où $a_{\pm}(E)$ annihile une particule d'énergie E positive, de moment $\mp E$, $b_{\pm}^{\dagger}(E)$ crée une antiparticule d'énergie E positive, de moment $\mp E$. Ces opérateurs vérifient les relations d'anticommuation habituelles des champs fermioniques. La densité lagrangienne de la théorie, $\mathcal{L} = iu_+^{\dagger}(\partial_0 + \partial_1)u_+ + iu_-^{\dagger}(\partial_0 - \partial_1)u_-$ possède les symétries suivantes :

1. Une symétrie vectorielle $U(1)_V : u_{\pm} \rightarrow e^{i\alpha}u_{\pm}$, de courant de Noether associé

$$J_V^{\mu} = \bar{\psi}\gamma^{\mu}\psi = \begin{pmatrix} u_+^{\dagger}u_+ + u_-^{\dagger}u_- \\ -u_+^{\dagger}u_+ + u_-^{\dagger}u_- \end{pmatrix}$$

de charge

$$Q_V = \int_0^L (u_+^{\dagger}u_+ + u_-^{\dagger}u_-)dx^1 = \sum_{E>0} [a_+^{\dagger}(E)a_+(E) - b_+^{\dagger}(E)b_+(E) + a_-^{\dagger}(E)a_-(E) - b_-^{\dagger}(E)b_-(E)]$$

Q_V compte algébriquement, i.e. avec un signe positif les particules et avec un signe négatif les antiparticules, le nombre d'états de chiralité positive plus le nombre d'états de chiralité négative.

2. Une symétrie axiale $U(1)_A : u_{\pm} \rightarrow e^{\pm i\alpha}u_{\pm}$, de courant de Noether associé

$$J_A^{\mu} = \bar{\psi}\gamma^{\mu}\gamma_5\psi = \begin{pmatrix} u_+^{\dagger}u_+ - u_-^{\dagger}u_- \\ -u_+^{\dagger}u_+ - u_-^{\dagger}u_- \end{pmatrix}$$

de charge

$$Q_A = \int_0^L (u_+^{\dagger}u_+ - u_-^{\dagger}u_-)dx^1 = \sum_{E>0} [a_+^{\dagger}(E)a_+(E) - b_+^{\dagger}(E)b_+(E) - a_-^{\dagger}(E)a_-(E) + b_-^{\dagger}(E)b_-(E)]$$

Q_A compte algébriquement, i.e. avec un signe positif les particules et avec un signe négatif les antiparticules, le nombre d'états de chiralité positive moins le nombre d'états de chiralité négative.

Dans une théorie libre, les nombres d'occupations des états étant fixés, les charges Q_V et Q_A sont conservées. Les choses vont devenir intéressantes en présence d'interaction. On considère donc un couplage à un champ de jauge $U(1)$. En deux dimensions d'espace-temps, la courbure de jauge $U(1)$ ne possède qu'une seule composante indépendante qui peut s'interpréter comme un champ électrique \mathcal{E} . On fait un choix de jauge en posant $A_0 = 0$. Pour décrire la dynamique des champs fermioniques, il faut remplacer p dans les expressions précédentes par $p - qA_1$.

Physiquement, \mathcal{E} crée à partir du vide un certain nombre de paires particule-antiparticule, particule et antiparticule étant de chiralité opposée par conservation du moment cinétique. Ainsi,

Q_V est conservé mais pas Q_A , d'où une non conservation du courant axial qu'on attend de la forme

$$\partial_\mu J_A^\mu \sim e\hbar\mathcal{E}$$

En fait, on verra plus loin que

$$\partial_\mu J_A^\mu = -\frac{i}{2\pi}\epsilon\hbar\epsilon^{\nu\sigma}F_{\nu\sigma}$$

où $\epsilon^{01} = -\epsilon^{10} = 1$.

On revient à la situation générale, en dimension $n = 2r$ paire et on retourne en signature $(-+\dots+)$. Comme précédemment, on a un courant vectoriel $J^\mu = i\bar{\psi}\gamma^\mu\psi$ et un courant axial $J_5^\mu = i\bar{\psi}\gamma^\mu\gamma_5\psi$. On va essayer de mettre en évidence l'existence d'une anomalie au niveau perturbatif, i.e. par calcul de graphes de Feynman. Une anomalie étant un phénomène quantique, elle ne peut pas être présente au niveau des graphes sans boucles (arbres). Les graphes les plus simples à étudier sont donc ceux à une boucle. On considère un diagramme avec N courants externes : un courant axial et $(N-1)$ courants vectoriels, et une boucle fermionique. Si la symétrie chirale était présente au niveau quantique, les contributions des deux orientations possibles de la boucle fermionique devraient se compenser et la contribution totale du graphe devrait être nulle. Si on arrive à trouver un diagramme de ce type dont la contribution n'est pas nulle, on aura la preuve qu'il existe une anomalie, plus précisément si la contraction de la contribution et de l'impulsion contenue dans le courant axial est non-nulle, on aura une contribution non-triviale à la non-conservation du courant axial. Cherchons le plus petit N donnant une situation non-triviale. Lors du calcul de la contribution d'un graphe du type précédent avec N courants externes, il faut prendre une trace sur un produit de matrices de Dirac contenant une matrice γ_5 du fait de la présence du courant axial : on s'attend donc à un résultat proportionnel à $\epsilon_{\mu_1, \dots, \mu_n}$ avec $N-1$ indices devant être les indices des courants vectoriels (on contracte sur l'indice du courant axial) et les $n - (N-1)$ autres indices devant être contactés avec les impulsions des courants. Ces impulsions représentent $N-1$ quantités indépendantes (et non pas N car on a la loi de conservation de l'impulsion qui impose une relation non-triviale). Il faut donc nécessairement : $N-1 \geq n - (N-1)$ d'où : $N \geq n/2 + 1$. Le premier diagramme de Feynman à calculer est donc celui avec $n/2 + 1$ courants externes. Pour $n = 4$, il s'agit d'un diagramme avec trois courants externes : la boucle fermionique est un triangle, on parle donc de diagramme triangulaire. Le calcul de ce diagramme est l'objet de la section suivante.

3.1.2 Calcul perturbatif

On considère une symétrie globale de groupe $G : \psi(x) \longrightarrow e^{i\epsilon^\alpha t_\alpha \gamma_5} \psi(x)$. La discussion de la section précédente correspond au cas $G = U(1)$.

Il s'agit de calculer la contribution d'un diagramme triangulaire entre un courant axial $j_{5,\alpha}^\mu = i\bar{\psi}\gamma^\mu\gamma_5 t_\alpha \psi$ d'impulsion $-p-q$ et deux courants vectoriels j_ν^β et j_γ^ρ d'impulsions respectives p et q . On appelle $\Gamma_{5,\alpha,\beta,\gamma}^{\mu,\nu,\rho}(-p-q, p, q)$ cette contribution. Pour calculer $i\Gamma_{5,\alpha,\beta,\gamma}^{\mu,\nu,\rho}$ il y a deux diagrammes à considérer correspondant aux deux façons de contracter les champs fermioniques contenus dans les courants. On rappelle les règles de Feynman : les sommets donnent respectivement $-\gamma^\mu\gamma_5 t_\alpha$, $-\gamma^\nu t_\beta$ et $-\gamma^\rho t_\gamma$, la boucle fermionique donne un signe moins et le propagateur d'un fermion d'impulsion k et de masse m est $-i/(i\not{k} + m - i\epsilon) = (-i)(-i\not{k} + m)/(k^2 + m^2 - i\epsilon)$. Ainsi :

$$\begin{aligned} & i\Gamma_{5,\alpha,\beta,\gamma}^{\mu,\nu,\rho}(-p-q, p, q) \\ = & -\int \frac{d^4k}{(2\pi)^4} \text{tr}_D \left\{ (-\gamma^\mu\gamma_5)(-i) \frac{-i(\not{k} + \not{p})}{(k+p)^2 - i\epsilon} (-\gamma^\nu)(-i) \frac{-i\not{k}}{k^2 - i\epsilon} (-\gamma^\rho)(-i) \frac{-i(\not{k} - \not{q})}{(k-p)^2 - i\epsilon} \right\} \text{tr}_R(t_\alpha t_\beta t_\gamma) \end{aligned}$$

$$+(p \longleftrightarrow q, \nu \longleftrightarrow \rho, \beta \longleftrightarrow \gamma)$$

L'intégrale est divergente : pour régulariser les choses, on va utiliser la régularisation de Pauli-Villars qui consiste à ajouter pour chaque fermion un autre de masse M grande et de statistique opposée (i.e. sans le signe moins accompagnant la boucle fermionique). On obtient ainsi une intégrale convergente qu'on va transformer jusqu'à ce que la limite $M \rightarrow +\infty$ ait un sens.

$$[\Gamma_{5,\alpha,\beta,\gamma}^{\mu,\nu,\rho}(-p-q, p, q)]_M = -i \int \frac{d^4k}{(2\pi)^4} (I_0^{\mu,\nu,\rho}(k, p, q) - I_M^{\mu,\nu,\rho}(k, p, q)) \text{tr}_R(t_\alpha t_\beta t_\gamma) + (p \longleftrightarrow q, \nu \longleftrightarrow \rho, \beta \longleftrightarrow \gamma)$$

où :

$$I_M^{\mu,\nu,\rho}(k, p, q) = \text{tr}_D \left\{ \gamma_5 \gamma^\mu \frac{\not{k} + \not{p} + iM}{(k+p)^2 + M^2 - i\epsilon} \gamma^\nu \frac{\not{k} + iM}{k^2 + M^2 - i\epsilon} \gamma^\rho \frac{\not{k} - \not{q} + iM}{(k-q)^2 + M^2 - i\epsilon} \right\}$$

Pour calculer $(p+q)_\mu I_M^{\mu,\nu,\rho}(k, p, q)$, on utilise : $\gamma_5(\not{p}+q) = \gamma_5(\not{p}+\not{k}-iM) + (\not{k}-q-iM) + 2iM\gamma_5$ de façon à écrire $(p+q)_\mu I_M^{\mu,\nu,\rho}(k, p, q)$ comme la somme de trois termes. Dans deux de ces termes, le calcul de la trace de Dirac montre qu'on obtiendra 0 après intégration. Ainsi :

$$-i(p+q)_\mu [\Gamma_{5,\alpha,\beta,\gamma}^{\mu,\nu,\rho}(-p-q, p, q)]_M = 8iM^2 \epsilon^{\nu,\rho,\lambda,\sigma} p_\lambda q_\sigma I(p, q, M) \text{tr}_R(t_\alpha t_\beta t_\gamma) + (p \longleftrightarrow q, \nu \longleftrightarrow \rho, \beta \longleftrightarrow \gamma)$$

où :

$$I(p, q, M) = \int \frac{d^4k}{(2\pi)^4} \frac{1}{[(k+p)^2 + M^2 - i\epsilon][k^2 + M^2 - i\epsilon][(k-q)^2 + M^2 - i\epsilon]}$$

Dans la limite $M \rightarrow +\infty$, on a, en posant $k = Ml$:

$$I(p, q, M) \sim 1/M^2 \int \frac{d^4l}{(2\pi)^4} \frac{1}{[l^2 + 1 - i\epsilon]^3} = \frac{i}{32\pi^2 M^2}$$

On obtient ainsi :

$$\begin{aligned} -i(p+q)_\mu \Gamma_{5,\alpha,\beta,\gamma}^{\mu,\nu,\rho}(-p-q, p, q) &= -\frac{1}{4\pi^2} \epsilon^{\nu,\rho,\lambda,\sigma} p_\lambda q_\sigma \text{tr}_R(t_\alpha t_\beta t_\gamma) + (p \longleftrightarrow q, \nu \longleftrightarrow \rho, \beta \longleftrightarrow \gamma) \\ &= -\frac{1}{2\pi^2} \epsilon^{\nu,\rho,\lambda,\sigma} p_\lambda q_\sigma \text{tr}_R(t_\alpha \{t_\beta, t_\gamma\}) \end{aligned}$$

On peut montrer que cette expression est cohérente avec :

$$\langle \partial_\mu J_{5,\alpha}^\mu \rangle_A = -\frac{1}{16\pi^2} \epsilon^{\mu,\nu,\rho,\sigma} \text{tr}(t_\alpha F_{\mu,\nu} F_{\rho,\sigma})$$

où $\langle . \rangle_A$ signifie qu'on prend la valeur moyenne en présence du champ classique A . Cette formule donne donc la contribution à l'anomalie chirale du diagramme triangulaire. On peut montrer par une étude de diagrammes de Feynman qu'il n'y a pas d'autre contribution perturbative. En fait, on verra plus loin que cette formule est même valable au niveau non-perturbatif.

3.1.3 Rôle des symétries globales et de leurs brisures dans la théorie de l'interaction forte

L'interaction forte est l'interaction assurant la cohésion des noyaux atomiques constitués de protons et de neutrons. Les protons et les neutrons sont formés de quarks et on décrit l'interaction forte par une théorie de jauge non-abélienne, de groupe $SU(3)$, appelée chromodynamique quantique (QCD). L'étude de cette théorie est très compliquée du fait de la valeur importante de la

constante de couplage et donc de la difficulté à obtenir des résultats précis avec les méthodes perturbatives. Dans ce cadre, l'étude des symétries présentes au niveau quantique, même si certaines ne sont qu'approchées du fait de l'utilisation d'un modèle simplifié, est d'une grande importance pour comprendre, au moins qualitativement, certaines propriétés générales de la QCD.

Plaçons nous dans un cadre un tout petit peu plus général : on considère une théorie de groupe de jauge $SU(N_c)$ décrivant N_f champs de quarks, fermions de spin 1/2. L'indice de champ $f = 1, \dots, N_f$ s'appelle la saveur : Q^f . Chacun de ces champs se transforme suivant la représentation standard de $SU(N_c)$ d'où pour chaque $i = 1, \dots, N_c$ une composante $Q^{f,i}$. L'indice i s'appelle la couleur. $SU(N_c)$ étant un groupe de jauge de dimension $N_c^2 - 1$, on a $N_c^2 - 1$ bosons de jauge : les gluons, A_μ^a , $a = 1, \dots, N_c^2 - 1$.

Dans le monde réel $N_c = 3$, i.e. le groupe de jauge est $SU(3)$: on a trois couleurs et huit champs gluoniques et $N_f = 6$: on a six saveurs de quarks, u (up), d (down), s (strange), c (charm), b (bottom), t (top) dans l'ordre croissant de masses. Un proton est constitué de deux quarks u et d'un quark d alors qu'un neutron est constitué de deux quarks d et d'un quark u .

On se place dans la limite $m_f = 0$. La densité lagrangienne (classique) de la théorie est alors, en signature $(+ - - -)$ (et non pas $(- + + +)$) et en séparant les composantes chirales Q_R et Q_L de Q :

$$\mathcal{L}_{QCD} = \frac{1}{4} F_{\mu,\nu}^a F^{a,\mu,\nu} + \sum_{f=1}^{N_f} [i\bar{Q}_L^f \not{D} Q_L^f + i\bar{Q}_R^f \not{D} Q_R^f]$$

Examinons les symétries de la densité lagrangienne : on a bien entendu par construction la symétrie de jauge sous $SU(N_c)$ mais aussi une symétrie globale sous $U(N_f)_L \times U(N_f)_R$ où $U(N_f)_L$ agit par transformation unitaire globale sur le vecteur $(Q_L^f)_{f=1,\dots,N_f}$ alors que $U(N_f)_R$ agit par transformation unitaire globale sur le vecteur $(Q_R^f)_{f=1,\dots,N_f}$. En utilisant la décomposition $U(N) = SU(N) \times U(1)$, on peut écrire le groupe de cette symétrie globale sous la forme $SU(N_f)_L \times U(1)_L \times SU(N_f)_R \times U(1)_R$. Le sous-groupe abélien $U(1)_L \times U(1)_R$ peut se décomposer d'une autre façon : $U(1)_V \times U(1)_A$ où $U(1)_V$ agit de manière «vectorielle» :

$$Q_L^f \longrightarrow e^{i\alpha} Q_L^f, Q_R^f \longrightarrow e^{i\alpha} Q_R^f$$

alors que $U(1)_A$ agit de manière «axiale» :

$$Q_L^f \longrightarrow e^{i\alpha} Q_L^f, Q_R^f \longrightarrow e^{-i\alpha} Q_R^f$$

Les courants conservés associés à ces deux symétries sont respectivement :

$$J_V^\mu = \sum_{f=1}^{N_f} \bar{Q}^f \gamma^\mu Q^f, J_A^\mu = \sum_{f=1}^{N_f} \bar{Q}^f \gamma^\mu \gamma_5 Q^f$$

De même, on peut décomposer $SU(N_f)_L \times SU(N_f)_R$ sous la forme $SU(N_f)_V \times SU(N_f)_A$ où $SU(N_f)_V$ agit de manière «vectorielle» :

$$Q_L^f \longrightarrow \sum_g (U_V)_{f,g} Q_L^g, Q_R^f \longrightarrow \sum_g (U_V)_{f,g} Q_R^g$$

alors que $SU(N_f)_A$ agit de manière «axiale» :

$$Q_L^f \longrightarrow \sum_g (U_A)_{f,g} Q_L^g, Q_R^f \longrightarrow \sum_g (U_A^{-1})_{f,g} Q_R^g$$

Les courants conservés associés à ces symétries sont respectivement :

$$J_V^{I,\mu} = \sum_{f,g}^{N_f} \bar{Q}^f \gamma^\mu (T^I)_{f,g} Q^g, \quad J_A^{I,\mu} = \sum_{f,g}^{N_f} \bar{Q}^f \gamma^\mu \gamma_5 (T^I)_{f,g} Q^g$$

où les T^I sont les générateurs de l'algèbre de Lie de $SU(N_f)$.

On vient de réécrire le groupe de la symétrie globale de la densité lagrangienne sous la forme : $SU(N_f)_V \times SU(N_f)_A \times U(1)_V \times U(1)_A$. Il s'agit d'une symétrie de la théorie classique. La question qui se pose est : que reste-t-il de cette symétrie au niveau quantique ?

Le couplage de J_A^μ au champ gluonique se traduit par une anomalie :

$$\partial_\mu J_A^\mu = -\frac{g^2 N_f}{32\pi^2} \epsilon^{\mu,\nu,\rho,\sigma} F_{\mu,\nu}^a F_{\rho,\sigma}^a$$

Le facteur $g^2/2$ provient du fait que les générateurs T_I de l'algèbre de Lie de $SU(N_f)$ sont normalisés de telle façon que $\text{tr}(T^I T^J) = \frac{1}{2} g^2 \delta^{i,j}$. Le couplage de $J_A^{I,\mu}$ avec le champ gluonique donne une anomalie proportionnelle à $\text{tr} T^I$ qui est donc nulle. Mais attention, dans la densité lagrangienne précédente, on a oublié que les quarks étaient chargés et donc couplés au champ électromagnétique. On appelle q_f la charge du quark de saveur f et on note $F_{\mu,\nu}$ le champ électromagnétique. Le couplage de $J_A^{I,\mu}$ avec le champ photonique se traduit par une anomalie :

$$\partial_\mu J_A^{I,\mu} = -\frac{N_c}{16\pi^2} \left[\sum_{f=1}^{N_f} (T^I)_{f,f} q_f^2 \right] \epsilon^{\mu,\nu,\sigma,\lambda} F_{\mu,\nu} F_{\sigma,\lambda}$$

Ainsi, les symétries $U(1)_A$ et $SU(N_f)_A$ ne sont pas présentes au niveau quantique du fait d'une anomalie chirale. Mais en fait, il y a aussi d'autres raisons à ces absences :

1. Si l'anomalie était seule à l'origine de l'absence de $U(1)_A$, ceci n'expliquerait pas qu'expérimentalement on n'observe pas de multiplet $U(1)_A$ dans le spectre des hadrons. Ce problème, dit problème $U(1)$, est longtemps resté un obstacle pour le développement de la chromodynamique quantique jusqu'à ce qu'il soit résolu par 't Hooft en 1986 : la symétrie $U(1)_A$ est explicitement brisée du fait de l'existence des instantons.
2. $SU(N_f)_A$ est spontanément brisée à basse énergie : $\langle 0 | \bar{Q}^f Q^f | 0 \rangle \neq 0$.

Ainsi, à basse énergie, le groupe des symétries globales présentes dans la théorie quantique est réduit à $SU(N_f)_V \times U(1)_V$ et la brisure spontanée de $SU(N_f)_A$ a pour conséquence la présence de $N_f^2 - 1$ bosons de Goldstone (il s'agit d'un résultat général : une brisure spontanée d'une symétrie à N paramètres à basse énergie en une symétrie à m paramètres s'accompagne de la formation de $N - m$ bosons de masse nulle appelés bosons de Goldstone).

Dans le QCD réelle, à basse énergie, on peut se restreindre à $N_f = 2$ et ne considérer que les quarks u et d . Le groupe de symétrie globale est alors $SU(2)_V \times U(1)_V$ et on a trois bosons de Goldstone associés aux valeurs non-nulles de $\langle \bar{u}u \rangle$, $\langle \bar{v}v \rangle$ et $\langle \bar{u}d - \bar{d}u \rangle$ qu'on croit reconnaître en les pions π^0 et π^\pm . En fait, les quarks ne sont pas de masse nulle et les symétries globales précédentes ne sont qu'approchées : les pions, vus comme boson de Goldstone, ne sont donc pas de masse nulle bien que leurs masses réelles soient très petites par rapport à celles des autres hadrons (de l'ordre de la centaine de MeV).

On a $q_u = \frac{2}{3}e$, $q_d = -\frac{1}{3}e$ (où e est la charge élémentaire) et $T^K = \frac{1}{2}\sigma^K$ (où les σ^K sont les matrices de Pauli), d'où le calcul explicite des expressions intervenant dans l'anomalie de $SU(2)_A$ liée au couplage avec le champ électromagnétique : $\sum_{f=u,d} (T^1)_{f,f} q_f^1 = \sum_{f=u,d} (T^2)_{f,f} q_f^2 = 0$ (car

σ^1 et σ^2 ont une diagonale nulle) alors que : $\sum_{f=u,d}(T^3)_{f,f}q_f^2 = \frac{e^2}{6}$. L'anomalie est donc due au courant $J_A^{3,\mu}$. C'est un résultat physiquement important car $|\pi^0\rangle = \frac{1}{\sqrt{2}}(|\bar{u}\rangle|u\rangle - |\bar{d}\rangle|d\rangle)$ étant le boson de Goldstone associé à la non-conservation de $J_A^{3,\mu}$, on a : $|\pi^0\rangle \sim \partial_\mu J_A^{3,\mu}|0\rangle$. Le diagramme triangulaire donnant naissance à l'anomalie peut donc s'interpréter comme la contribution à l'ordre d'une boucle de la réaction $\pi^0 \rightarrow 2\gamma$ qui est l'embranchement principal de désintégration du pion neutre. Historiquement, c'est l'étude de cette désintégration qui a conduit à l'étude de l'anomalie chirale : on n'arrivait pas à expliquer le temps de vie extrêmement court du π^0 car on pensait qu'un couplage chirale au champ électromagnétique était interdit jusqu'à ce qu'on se rende compte que ce n'était pas le cas.

3.2 Vers le théorème de l'indice

Dans la partie précédente, on a vu apparaître des expressions de la forme $tr(\epsilon^{\mu,\nu,\rho,\sigma}F_{\mu,\nu}F_{\rho,\sigma})$. Une quantité de cette forme possède des propriétés remarquables : c'est localement une divergence et son intégrale sur M va donc posséder des propriétés d'intégralité et d'invariance topologique. La théorie des classes caractéristiques donne un cadre général pour l'étude de ces quantités.

3.2.1 Classes caractéristiques

Définition 17. Une classe caractéristique c est une transformation naturelle qui à chaque fibré vectoriel E sur une variété M associe un élément $c(E)$ du groupe de cohomologie $H^*(M)$.

L'idée est de prendre un polynôme en la courbure d'une connexion sur E . Pour espérer avoir des quantités bien définies (invariantes de jauge), il faut se restreindre à des polynômes possédant certaines propriétés d'invariance.

Définition 18. Soit $M_m(\mathbb{C})$ l'anneau des matrices $m \times m$ sur \mathbb{C} . Un polynôme invariant sur $M_m(\mathbb{C})$ est une fonction polynomiale $P : M_m(\mathbb{C}) \rightarrow \mathbb{C}$ telle que pour tout $X, Y \in M_m(\mathbb{C})$, $P(XYX^{-1}) = P(Y)$. Une série formelle invariante est une série formelle sur $M_m(\mathbb{C})$, chacune des composantes homogènes étant un polynôme invariant.

Lemme 4. L'anneau des polynômes invariants sur $M_m(\mathbb{C})$ est engendré par les polynômes : $c_k(X) = (4\pi i)^{-k} tr(\Lambda^k X)$. On rappelle que $tr(\Lambda^k X)$ est la k -ième fonction symétrique élémentaire en les valeurs propres de X .

Démonstration. Soit P un polynôme invariant. Par densité des matrices diagonalisables dans $M_m(\mathbb{C})$, il suffit de déterminer P sur cet ensemble. Par invariance sous conjugaison de P , il suffit de déterminer P sur les matrices diagonales. On voit alors que P est une fonction polynomiale en les éléments diagonaux et puisque on peut permuter les éléments diagonaux par conjugaison, il s'agit d'une fonction polynomiale symétrique. Il suffit alors d'appliquer le résultat classique qui dit que l'anneau des fonctions polynomiales symétriques est engendré par les fonctions symétriques élémentaires. □

Soit E un fibré vectoriel complexe sur M avec une connexion ∇ de courbure F qui est une 2-forme sur E à valeurs dans $End(E)$. Si on choisit une trivialisations locale de E , on peut identifier F avec une matrice de 2-formes ordinaires. Si P est un polynôme invariant, on peut appliquer P à cette matrice et obtenir une forme différentielle paire $P(F)$. Du fait de l'invariance de P , la forme $P(F)$ est en fait indépendante du choix de trivialisations locale et est donc une forme bien définie. Puisque les 2-formes sont des éléments nilpotents de l'algèbre extérieure, toutes les séries formelles avec des 2-formes pour variables sont en fait des polynômes et donc sont bien définies. La construction précédente a donc un sens si P est une série formelle invariante.

Proposition 8. *Si P est un polynôme (ou une série formelle) invariante, alors la forme différentielle $P(F)$ est fermée et sa classe de cohomologie est indépendante du choix de la connexion ∇ sur E*

Démonstration. Ce résultat étant le coeur de la théorie des classes caractéristiques, on va en donner deux démonstrations : la première est la plus simple et la plus conceptuelle, la deuxième a l'avantage de donner une formule explicite pour la variation de la forme $P(F)$ en fonction de la connexion : on verra une application de cette formule pour la théorie des formes de Chern-Simons. (Le lecteur se convaincra que ces deux démonstrations sont fondamentalement les mêmes).

Preuve 1

1. $P(F)$ fermée

On peut supposer P homogène de degré j . Soit \tilde{P} le polynôme à k variables, forme polarisée de $P : \tilde{P}(X_1, \dots, X_j)$ est défini comme étant $1/j!$ fois le coefficient de $t_1 \dots t_j$ dans le développement de $P(t_1 X_1 + \dots + t_j X_j)$. Par construction, on a $P(X) = \tilde{P}(X, \dots, X)$. Soit $m \in M$. On veut montrer que $dP(F)(m) = 0$. Choisissons un repère synchrone centré sur m , alors $dF(m) = 0$. Mais alors : $dP(F)(m) = d\tilde{P}(F, \dots, F)(m) = j\tilde{P}(dF, F, \dots, F)(m) = 0$.

2. Indépendance de la classe de cohomologie vis-à-vis de ∇

Soit ∇_0 et ∇_1 deux connexions sur E . Soit $\nabla_\epsilon = \epsilon\nabla_1 + (1-\epsilon)\nabla_0$. On considère $M \times [0, 1]$ muni de sa projection π sur M . π^*E est alors un fibré sur $M \times [0, 1]$. Pour tout $\epsilon \in [0, 1]$, soit i_ϵ la section de $M \times [0, 1]$ donnée par : $i_\epsilon : x \rightarrow (x, \epsilon)$. On peut alors construire sur π^*E une connexion ∇ dont la restriction à $M \times \{\epsilon\}$ est ∇_ϵ , autrement dit, telle que $(i_\epsilon^*)(\nabla) = \nabla_\epsilon$. Alors : $P(F_{\nabla_\epsilon}) = (i_\epsilon^*)(P(F_\nabla))$. i_0 et i_1 étant homotopes, les applications i_0^* et i_1^* induites en cohomologie sont identiques.

Preuve 2

1. $P(F)$ fermée

Il suffit de prouver le résultat pour $P_j(F) = \text{tr}(F^j)$ puisque les fonctions symétriques élémentaires peuvent s'exprimer à partir de ces fonctions polynomiales. Or : $d(\text{tr}(F^j)) = j\text{tr}(dFF^{j-1})$ et d'après l'identité de Bianchi : $dF = FA - AF$, on a : $d(\text{tr}(F^j)) = j\text{tr}((FA - AF)F^{j-1}) = 0$ par cyclicité de la trace (puisque ici on a des matrices à valeurs dans les formes différentielles, on a : $\text{tr}(AB) = (-1)^{pq}\text{tr}(BA)$ si A est une matrice de p -formes et B est une matrice de q -formes).

2. Indépendance de la classe de cohomologie vis-à-vis de la connexion

On peut supposer P homogène de degré j . Soit A_0 et A_1 les 1-formes à valeurs dans $\text{Lie}(G)$ définissant ∇_0 et ∇_1 . Soit $A_t = A_0 + t(A_1 - A_0)$ et $F_t = dA_t + A_t^2$. Soit \tilde{P} la forme polarisée de P . On a :

$$\frac{\partial P(F_t)}{\partial t} = j\tilde{P}\left(\frac{\partial F_t}{\partial t}, F_t, \dots, F_t\right)$$

Or : $\partial F_t / \partial t = d(A_1 - A_0) + (A_1 - A_0)A_0 + A_0(A_1 - A_0) + 2t(A_1 - A_0)^2 = d(A_1 - A_0) + (A_1 - A_0)A_t + A_t(A_1 - A_0) = D_t(A_1 - A_0)$ où D_t est la dérivation covariante par rapport à la connexion A_t . D'après l'identité de Bianchi : $D_t F_t = 0$ d'où :

$$\frac{\partial P(F_t)}{\partial t} = jD_t \tilde{P}(A_1 - A_0, F_t, \dots, F_t)$$

Or $\tilde{P}(A_1 - A_0, F_t, \dots, F_t)$ est invariant de jauge puisque P est invariant de jauge, on peut donc remplacer D_t par d d'où par intégration :

$$P(F_1) - P(F_0) = dT$$

où :

$$T = j \int_0^1 \tilde{P}(A_1 - A_0, F_t, \dots, F_t) dt$$

□

On a mis en évidence des polynômes invariants particuliers : les fonctions symétriques élémentaires en les valeurs propres. Les classes de cohomologie associées aux polynômes $c_k(X)$ s'appellent les classes de Chern : $c_k \in H^{2k}(M)$, $k = 0, \dots, n$. On appelle classe de Chern totale $c = c_0 + \dots + c_n \in H^*(M)$. On retrouve la k -ième classe de Chern à partir de c en prenant la composante de degré $2k$. Par construction, toute classe caractéristique défini par un polynôme invariant est une fonction polynômiale en les classes de Chern.

Remarque On peut montrer que les classes de Chern vivent dans la cohomologie à coefficients entiers $H^*(M, \mathbb{Z})$, ce qui explique le choix de la normalisation.

Le caractère de Chern est une version multiplicative des classes de Chern :

$$S(X) = \text{tr}(e^{-\frac{X}{2i\pi}}) = \sum_j \frac{\text{Tr}((\frac{iX}{2\pi})^j)}{j!}$$

est une série formelle invariante sur $M_m(\mathbb{C})$. On peut donc définir $ch(E)$ comme étant la classe de cohomologie défini par $S(F)$ où F est la courbure d'une connexion choisie sur E . Le caractère de Chern vérifie : $ch(E \oplus E') = ch(E) + ch(E')$, $ch(E \otimes E') = ch(E)ch(E')$.

Jusqu'à présent, on a défini des classes caractéristiques de fibrés vectoriels complexes. Considérons maintenant le cas où E est un fibré vectoriel réel. Soit ∇ une connexion sur E préservant une métrique, F sa courbure, qui est alors une 2-forme à valeurs dans $Lie(O(m))$, on appelle classe de Pontjagin totale, notée $p(E)$, la classe de cohomologie associée à $\det(1 - \frac{1}{4\pi}F)$. Grâce à l'hypothèse sur ∇ , $p(E)$ ne contient que des composantes de degré multiple de 4 et on appelle k -ième classe de Pontjagin $p_k(E)$ la composante de degré $4k$ de $p(E)$. Autrement dit : $p_k(E) = (-1)^k c_{2k}(E \otimes \mathbb{C}) \in H^{4k}(M, \mathbb{R})$.

Si E est un fibré vectoriel réel orienté, on peut lui associer une classe caractéristique supplémentaire : la classe d'Euler (associée au pfaffien).

Revenons à des fibrés vectoriels complexes. Soit $f(z)$ une fonction holomorphe au voisinage de $z = 0$, on peut utiliser f pour construire des séries formelles invariantes, on pose : $\Sigma_f(X) = \text{tr}(f(X/(4i\pi)))$, «série additive associée à f » et $\Pi_f(X) = \det(f(X/(4i\pi)))$, « série multiplicative associée à f ». Si les valeurs propres de la matrice $X/(4i\pi)$ sont notées par (x_j) , alors : $\Sigma_f(X) = \sum_j f(x_j)$, $\Pi_f(X) = \prod_j f(x_j)$, qui sont bien des séries formelles symétriques en les x_j et qui sont donc exprimables au moyen des fonctions symétriques élémentaires. Mais les classes associées aux fonctions symétriques élémentaires ne sont autres que les classes de Chern, aussi écrit-on souvent les classes associées $\Sigma_f(E)$ et $\Pi_f(E)$ sous la forme $\sum_j f(x_j)$ et $\prod_j f(x_j)$ où les x_j sont définis formellement par : $c_1 = x_1 + \dots + x_m$, $c_2 = x_1x_2 + \dots + x_{m-1}x_{m-2}$, ..., $c_m = x_1 \dots x_m$.

Par exemple, la classe de Chern totale est la classe multiplicative associée à $f(z) = 1 + z$ et le caractère de Chern est la classe additive associée à $f(z) = e^z$.

Définition 19. La \hat{A} -classe est la classe multiplicative associée à la fonction analytique :

$$f(z) = \left(\frac{z/2}{sh(z/2)} \right)^{1/2}$$

La L -classe est la classe multiplicative associée à la fonction analytique :

$$f(z) = \left(\frac{z}{th(z)} \right)^{1/2}$$

Les \hat{A} - et L - classes sont toutes les deux associées à des séries formelles paires et s'expriment donc en fonction des classes de Chern paires. Ainsi, les \hat{A} et L classes de la complexification d'un fibré vectoriel réel peuvent être exprimées en fonction des classes de Pontrjagin.

Proposition 9. *Soit f paire et holomorphe au voisinage de 0. Alors pour un fibré vectoriel réel E , $\Pi_f(E \otimes \mathbb{C}) = \prod_j (f(y_j))^2$ où les y_j sont définis formellement par la condition que les classes de Pontrjagin de E soient les fonctions symétriques élémentaires en les y_j^2 .*

Démonstration. L'identité désirée est une égalité entre polynômes $Lie(O(m))$ -invariants. Toute matrice de $Lie(O(m))$ est semblable à une matrices diagonale par bloc où les blocs sont des matrices 2×2 de la forme $X = \begin{pmatrix} 0 & \lambda \\ -\lambda & 0 \end{pmatrix}$ qui a pour valeurs propres $\pm i\lambda$. Les deux membres de l'égalité étant multiplicatifs vis à vis des sommes directes, il suffit de la prouver pour cette matrice X . $c_1(X) = 0$, $c_2(X) = \frac{1}{(4i\pi)^2} (i\lambda)(-i\lambda) = -\frac{\lambda^2}{16\pi^2}$. Alors, $p_1(X) = \frac{\lambda^2}{16\pi^2}$, $y = \frac{\lambda}{4\pi}$. D'autre part, X étant semblable sur \mathbb{C} à $\begin{pmatrix} -i\lambda & 0 \\ 0 & i\lambda \end{pmatrix}$ on a $\Pi_f(X) = f(-\frac{\lambda}{4\pi})f(\frac{\lambda}{4\pi}) = f(y)^2$ car f est paire. \square

À partir de maintenant et dans les sections suivantes, on rompt le fil directeur de l'exposé : on oublie les anomalies, les classes caractéristiques et on s'intéresse de façon purement mathématique aux propriétés de l'opérateur de Dirac jusqu'à parvenir au théorème de l'indice. L'intérêt de ce travail et le lien avec ce qui précède n'apparaîtra qu'à la section 3.4.

De 3.2.2 à 3.3 inclus, on se place sur une variété riemannienne (i.e. en signature $(+\dots+)$) et on adopte la convention «mathématique» pour la définition des matrices de Dirac et donc de l'opérateur de Dirac : $\{\gamma_i, \gamma_j\} = -2g_{i,j}$ (Voir la remarque qui précède 2.1.5, Définition 10).

3.2.2 Analyse fonctionnelle sur les variétés

L'objectif de cette partie est d'obtenir un théorème spectral pour les opérateurs de Dirac. On commence par définir les espaces de Sobolev pour les fonctions définies sur le tore $\mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$.

Définition 20. *Soit $f : \mathbb{T}^n \rightarrow \mathbb{R}$ une fonction intégrable. La série de Fourier de f est la série formelle $\sum_{\nu \in \mathbb{Z}^n} a_\nu e^{i\nu x}$ où $a_\nu = \hat{f}_\nu = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} f(x) e^{-i\nu x} dx$.*

Définition 21. *Soit $k \in \mathbb{N}$. Sur $C^\infty(\mathbb{T}^n)$, on définit le k -ième produit scalaire de Sobolev par : $\langle f_1, f_2 \rangle_k = (2\pi)^n \sum_{\nu} \hat{f}_1(\nu) \overline{\hat{f}_2(\nu)} (1 + |\nu|^2)^k$. La k -ième norme de Sobolev $|\cdot|_k$ est la norme induite par ce produit scalaire. Le k -ième espace de Sobolev, noté W^k , est le complété de $C^\infty(\mathbb{T}^n)$ par rapport à $|\cdot|_k$.*

Soit M une variété compacte. Soit (U_j) un recouvrement de M par des cartes de coordonnées et ψ_j une partition de l'unité subordonnée à ce recouvrement. Soit ϕ_j un difféomorphisme de U_j sur $(0, 2\pi)^n \subset \mathbb{T}^n$.

Définition 22. *On définit la k -ième norme de Sobolev sur $C^\infty(M)$ par : $\|f\|_k = \sum_j \|(\psi_j f) \circ \phi_j^{-1}\|_k$ où dans le membre de droite $\|\cdot\|_k$ est la k -ième norme de Sobolev sur \mathbb{T}^n*

La norme dépend des différents choix faits dans la définition mais ce qui compte, c'est que si on modifie un de ces choix, on obtient une norme équivalente, et on a donc une topologie bien définie sur $C^\infty(M)$. On définit $W^k(M)$, le k -ième espace de Sobolev de M comme étant l'espace vectoriel topologique obtenu par complétion de $C^\infty(M)$ pour cette topologie. De même, si E est un fibré vectoriel sur E , on définit les espaces $W^k(M, E)$ de sections de E en prenant en plus des trivialisations locales pour définir $\|\cdot\|_k$.

Les deux propositions suivantes sont les deux résultats essentiels permettant d'appliquer la notion d'espace de Sobolev. Ces résultats étant classiques, on ne donnera pas les démonstrations (on se ramène au cas de \mathbb{T}^n , le théorème de plongement de Sobolev est alors une conséquence de l'inégalité de Cauchy-Schwarz et le théorème de Rellich est une conséquence du théorème d'Ascoli).

Proposition 10. *Théorème de plongement de Sobolev*

Pour tout entier $p > n/2$, l'espace $W^{k+p}(M, E)$ s'injecte continûment dans $C^k(M, E)$.

Proposition 11. *Théorème de Rellich*

Si $k_1 < k_2$, alors l'injection de W^{k_2} dans W^{k_1} est compacte.

On s'intéresse maintenant à un opérateur de Dirac \mathcal{D} sur un fibré de Clifford E sur la variété M . Les inégalités suivantes, dites estimées elliptiques, seront essentielles pour prouver des résultats de régularité.

Proposition 12. *Il existe une constante C telle que pour tout $s \in C^\infty(M, E)$, $\|s\|_1 \leq C(\|s\|_0 + \|\mathcal{D}s\|_0)$*

Plus généralement, pour tout entier $k > 0$, il existe une constante C_k , telle que pour tout $s \in C^\infty(M, E)$: $\|s\|_{k+1} \leq C_k(\|s\|_k + \|\mathcal{D}s\|_k)$

Démonstration. On raisonne par récurrence sur k .

$k = 0$. La formule de Weitzenbock s'écrit $\mathcal{D}^2 = \nabla^* \nabla + \mathcal{K}s$ d'où :

$$\|\mathcal{D}s\|_0^2 = \|\nabla s\|_0^2 + \langle \mathcal{K}s, s \rangle_{>0}$$

et donc $\|\nabla s\|_0 \leq C_1(\|s\|_0 + \|\mathcal{D}s\|_0)$ pour une constante C_1 . Écrivons ∇ en coordonnées locales $\nabla_i s = \partial_i s + A_i s$. On en déduit une minoration $\|\nabla s\|_0^2 \geq C_2 \|s\|_1^2 - C_3 \|s\|_0 \|s\|_1$ pour des constantes C_2 et C_3 . Il existe C_4 tel que : $C_3 \|s\|_0 \|s\|_1 \leq \frac{1}{2} C_2 \|s\|_1^2 + C_4 \|s\|_0^2$ d'où : $\|\nabla s\|_0^2 \geq \frac{1}{2} C_2 \|s\|_1^2 - C_4 \|s\|_0^2$.

Supposons le résultat vrai pour k . Quitte à utiliser une partition de l'unité, on peut supposer que s est à support dans une carte et donc travailler en coordonnées locales. On a $\|s\|_{k+1} \leq C \sum_i \|\partial_i s\|_k$. On applique l'hypothèse de récurrence et le fait que ∂_i et $[\mathcal{D}, \partial_i]$ sont des opérateurs du premier ordre pour conclure. □

Il est pratique de penser à \mathcal{D} comme à un opérateur non-borné sur l'espace de Hilbert $H = L^2(M, E)$: on appelle opérateur non-borné sur un espace de Hilbert H une application linéaire définie sur un ensemble dense de H , appelé le domaine de l'opérateur, et à valeurs dans H . Soit A un tel opérateur et soit $dom(A)$ son domaine. On appelle graphe de A le sous-espace de $H \oplus H$ défini par : $G_A = \{(x, Ax), x \in dom(A)\}$.

Lemme 5. *La fermeture \overline{G} du graphe G de l'opérateur de Dirac est aussi un graphe et définit donc un opérateur non-borné noté $\overline{\mathcal{D}}$. De plus, le domaine de $\overline{\mathcal{D}}$ est exactement l'espace de Sobolev $W^1(M, E)$.*

Démonstration. Si \overline{G} n'était pas un graphe, il existerait $(0, y) \in \overline{G}$ et une suite (x_j) de fonctions C^∞ avec $y \neq 0$, $x_j \rightarrow 0$, $\mathcal{D}x_j \rightarrow y$ dans $L^2(M, E)$. En particulier, pour tout $s \in L^2(M, E)$, $\langle \mathcal{D}x_j, s \rangle \rightarrow \langle y, s \rangle$ et $\langle x_j, \mathcal{D}s \rangle \rightarrow 0$. Mais \mathcal{D} est autoadjoint d'où $\langle y, s \rangle = 0$ pour tout s d'où $y = 0$, absurde.

Le domaine de $\overline{\mathcal{D}}$ est l'ensemble des $x \in L^2(M, E)$ tel qu'il existe une suite (x_j) dans $C^\infty(M, E)$ telle que $x_j \rightarrow x$ dans $L^2(M, E)$ et $(\mathcal{D}x_j)$ converge dans $L^2(M, E)$. Il résulte des estimées elliptiques qu'il s'agit de $W^1(M, E)$. □

Définition 23. Soient $x, y \in L^2(M, E)$. On dit que l'équation $\mathcal{D}x = y$ est satisfaite en un sens faible si pour tout $s \in C^\infty(M, E)$, $\langle x, \mathcal{D}s \rangle = \langle y, s \rangle$ où \langle, \rangle est le produit scalaire de $L^2(M, E)$.

Définition 24. Un opérateur continu A sur $L^2(M, E)$ est dit régularisant s'il existe un noyau $C^\infty k(m_1, m_2)$ sur $M \times M$, à valeurs dans $\text{Hom}(E_{m_2}, E_{m_1})$, tel que pour tout $s \in L^2(M, E)$: $As(m_1) = \int_M k(m_1, m_2)s(m_2)\text{vol}(m_2)$. En particulier, l'image d'un opérateur régularisant est formée de fonctions C^∞ (d'où le nom de « régularisant »).

Proposition 13. Soient $x, y \in L^2(M, E)$ tels que $\mathcal{D}x = y$ au sens faible, alors $x \in W^1(M, E)$ et $\overline{\mathcal{D}x} = y$.

Démonstration. Par un procédé d'approximation par convolution, on peut construire une famille F_ϵ , $\epsilon \in (0, 1)$, d'opérateurs régularisants autoadjoints tels que (F_ϵ) est une famille bornée d'opérateurs sur $L^2(M, E)$, $([D, F_\epsilon])$ se prolonge en une famille bornée d'opérateurs continus sur $L^2(M, E)$ et pour tout $x, y \in L^2(M, E)$, $\langle F_\epsilon x, y \rangle \rightarrow \langle x, y \rangle$ quand $\epsilon \rightarrow 0$.

Soient $x, y \in L^2(M, E)$ tels que $\mathcal{D}x = y$ au sens faible. Posons $x_\epsilon = F_\epsilon x$, alors x_ϵ est C^∞ et en utilisant le fait que \mathcal{D} et F_ϵ sont autoadjoints, on a : $\langle \mathcal{D}x_\epsilon, s \rangle = \langle y, F_\epsilon s \rangle + \langle x, [F_\epsilon, \mathcal{D}]s \rangle$. On en déduit qu'il existe C tel que pour tout ϵ , $\|\mathcal{D}x_\epsilon\| \leq C$. Par les estimées elliptiques, (x_ϵ) est donc une suite bornée dans W^1 . Par compacité faible de la boule unité d'un espace de Hilbert, on peut supposer quitte à extraire que (x_ϵ) converge faiblement dans W^1 . Mais par le théorème de Rellich, W^1 s'injecte de manière compacte dans $W^0 = L^2$, (x_ϵ) converge donc dans L^2 . Par la dernière propriété des F_ϵ , cette limite est nécessairement x . □

Proposition 14. Soient k entier supérieur ou égal à 1, $x, y \in W^k(M, E)$ et supposons que $\overline{\mathcal{D}x} = y$, alors $x \in W^{k+1}(M, E)$.

Démonstration. On raisonne par récurrence sur k .

Le cas $k = 0$ est la proposition précédente.

Supposons le résultat vrai pour k . Puisque « $\overline{\mathcal{D}x} = y$ » est équivalent à « $\mathcal{D}x = y$ faiblement » qui a un sens local, on peut travailler en coordonnées locales. ∂_i s'étend en un opérateur continu de W^k dans W^{k-1} . Puisque $\mathcal{D}x = y$ faiblement, on a $\mathcal{D}\partial_i x = \partial_i y + [\mathcal{D}, \partial_i]x$ faiblement et le membre de droite appartient à W^{k-1} . Par hypothèse de récurrence, on a $\partial_i x \in W^k$ d'où $x \in W^{k+1}$. □

Théorème 3. $H = L^2(M, E)$ est somme hilbertienne d'un nombre dénombrable d'espaces H_λ , chaque H_λ étant de dimension finie, formé de fonctions C^∞ et espace propre pour \mathcal{D} de valeur propre λ . L'ensemble des valeurs propres λ forme un sous-ensemble discret de \mathbb{R} .

Démonstration. Soit G le graphe de \mathcal{D} . On commence par remarquer que si $J : H \oplus H \rightarrow H \oplus H$ est l'application $(x, y) \rightarrow (y, -x)$, alors on a une décomposition orthogonale : $H \oplus H = \overline{G} \oplus J\overline{G}$. En effet, si (x, y) est dans l'orthogonal de G , alors $x + \mathcal{D}y = 0$ faiblement d'où $y \in W^1$ d'où $(-y, x) \in \overline{G}$, i.e. $(x, y) \in J\overline{G}$.

Soit $x \in H$. $(x, 0)$ se projette sur \overline{G} dans $H \oplus H$, on écrit la projection sous la forme $(Qx, \overline{\mathcal{D}}Qx)$, ce qui définit un opérateur Q sur H . D'après la remarque précédente, $(x, 0) = (Qx, \overline{\mathcal{D}}Qx) + (-\overline{\mathcal{D}}y, y)$ pour un certain $y \in W^1$ d'où : $x = Qx - \overline{\mathcal{D}}y$, $y = -\overline{\mathcal{D}}Qx$ et donc $x = (1 + \overline{\mathcal{D}})^2 Qx$ avec $Qx \in W^2$ d'après la proposition précédente. Il résulte des estimées elliptiques que Q est une application continue de L^2 dans W^2 . D'autre part, Q est clairement autoadjoint et injectif.

Q est un opérateur injectif auto-adjoint de $H = L^2$ qui est compact puisque W^2 s'injecte de manière compacte dans L^2 d'après le théorème de Rellich. D'après le théorème spectral pour

les opérateurs compacts autoadjoints, H se décompose en une somme hilbertienne d'espaces propres de dimension finie pour Q , les valeurs propres formant une suite convergente vers 0. Chaque espace propre est formé de fonctions W^2 puisque Q est à valeurs dans W^2 .

Puisque $(1 + \overline{\mathcal{D}}^2)Q = 1$, un vecteur propre pour Q de valeur propre ρ est aussi vecteur propre pour $\overline{\mathcal{D}}^2$ de valeur propre $(1 - \rho)/\rho$. H est donc somme hilbertienne d'espaces propres de dimension finie pour $\overline{\mathcal{D}}^2$, chacun formé de fonctions W^2 , avec une suite de valeur propre tendant vers $+\infty$. Puisque \mathcal{D}^2 est positif, ses valeurs propres sont positives. Si V est un espace propre pour $\overline{\mathcal{D}}^2$, de valeur propre λ^2 , alors $(\lambda + \overline{\mathcal{D}})V$ et $(\lambda - \overline{\mathcal{D}})V$ sont des espaces propres orthogonaux pour \mathcal{D} de valeurs propres respectives $+\lambda$ et $-\lambda$. Ainsi, H est somme hilbertienne d'espaces propres de dimension finie de \mathcal{D} , chacun étant formé de fonctions W^1 . Si $\mathcal{D}x = \lambda x$, alors par application répétée de la proposition précédente, on voit que $x \in W^k$ pour tout k . Par le théorème de plongement de Sobolev, les espaces propres sont donc formés en fait de fonctions C^∞ . □

Soit $Sp(\mathcal{D})$ le spectre de \mathcal{D} . Si f est une fonction (à valeurs dans \mathbb{C}) bornée sur $Sp(\mathcal{D})$, on peut définir un opérateur continu $f(\mathcal{D})$ sur $L^2(M, E)$: $f(\mathcal{D})$ est l'opérateur agissant par multiplication par $f(\lambda)$ sur l'espace propre de \mathcal{D} associé à la valeur propre λ .

Définition 25. Soit M une variété riemannienne compacte orientée de dimension n et soit E_0, \dots, E_k une suite de fibrés vectoriels sur M , munis de métriques hermitiennes et de connexions compatibles. Pour tout $j = 0, \dots, k-1$ soit $d_j : C^\infty(M, E_j) \rightarrow C^\infty(M, E_{j+1})$ tels que $d_{j+1}d_j = 0$, autrement dit, tels que $(C^\infty(M, E_j), d_j)_j$ soit un complexe. On parle de complexe de Dirac si $E = \oplus_j E_j$ est un fibré de Clifford dont l'opérateur de Dirac associé est $d + d^*$

Exemples

1. Le complexe de de Rham d'après la section sur les fibrés de Clifford.
Pour ceux qui connaissent de la géométrie complexe :
2. Si M est kählérienne, le complexe de Dolbeault.

3.2.3 Développement asymptotique du noyau de la chaleur

Dans toute cette partie, M est une variété riemannienne compacte, E est un fibré de Clifford sur M et \mathcal{D} est l'opérateur de Dirac associé.

Définition 26. L'équation de la chaleur pour \mathcal{D} est l'équation aux dérivées partielles :

$$\frac{\partial s}{\partial t} + \mathcal{D}^2 s = 0$$

L'équation des ondes pour \mathcal{D} est l'équation aux dérivées partielles :

$$\frac{\partial s}{\partial t} - i\mathcal{D}s = 0$$

Remarque : ce qu'on appelle « équation des ondes » est une « racine carrée » de l'équation des ondes traditionnelle. C'est pour faire cette opération qu'on a introduit les opérateurs de Dirac.

Proposition 15. Pour une donnée initiale C^∞ s_0 , les équations de la chaleur et des ondes ont une unique solution C^∞ s_t . La solution existe pour tout $t \geq 0$ dans le cas de l'équation de la chaleur et pour tout $t \in \mathbb{R}$ pour l'équation des ondes. De plus, on a des estimées L^2 de la forme $\|s_t\|_{L^2} \leq \|s_0\|_{L^2}$ pour l'équation de la chaleur et $\|s_t\|_{L^2} = \|s_0\|_{L^2}$ pour l'équation des ondes.

Démonstration. On traite le cas de l'équation de la chaleur, celui de l'équation des ondes est similaire. Si s_t est une solution C^∞ , alors $\partial_t(\|s_t\|^2) = -2\|\mathcal{D}s_t\|^2$ d'où $\|s_t\|^2 \leq \|s_0\|^2$ ce qui prouve l'unicité. Pour l'existence, il suffit de poser $s_t = e^{-t\mathcal{D}^2} s_0$ avec l'opérateur $e^{-t\mathcal{D}^2}$ défini dans la section précédente. \square

Les opérateurs solution des équations de la chaleur et des ondes sont respectivement les opérateurs continus sur $L^2(M, E)$ $e^{-t\mathcal{D}^2}$ et $e^{it\mathcal{D}}$.

On veut montrer que $e^{-t\mathcal{D}^2}$ est un opérateur régularisant.

Lemme 6. *Soit A un opérateur continu autoadjoint sur $L^2(M, E)$, envoyant $L^2(M, E)$ continûment dans l'espace de Banach $C^{r+1}(M, E)$. Alors A^2 est un opérateur à noyau : il existe $k \in C^r$ tel que : $A^2 s(m_1) = \int_M k(m_1, m_2) s(m_2) \text{vol}(m_2)$*

Remarque : k est de classe C^r en tant que section du fibré $(pr_1^* E) \otimes (pr_2^* E^*)$ sur $M \times M$ où pr_1 et pr_2 sont les deux projections de $M \times M$ sur M .

Démonstration. Pour tout $m \in M$, $s \rightarrow As(m)$ est une forme linéaire continue sur $L^2(M, E)$ qui se représente donc par un vecteur $v_m : As(m) = \langle s, v_m \rangle$. Le fait que A envoie $L^2(M, E)$ continûment dans $C^{r+1}(M, E)$ permet de montrer que $m \rightarrow v_m$ est une application de classe C^r de M dans $L^2(M, E)$. On a : $As(m_1) = \int_{m_2} s_{m_2} \bar{v}_{m_1}(m_2) dm_2$. Puisque A est autoadjoint, on a : $v_{m_1}(m_2) = \bar{v}_{m_2}(m_1)$ d'où $As(m_1) = \int_{m_2} s_{m_2} v_{m_2}(m_1) dm_2$ d'où $A^2 s(m_1) = \int_M k(m_1, m_2) s(m_2) dm_2$ avec $k(m_1, m_2) = Av_{m_2}(m_1)$ de classe C^r . \square

Proposition 16. $e^{-t\mathcal{D}^2}$ est un opérateur régularisant.

Démonstration. D'après les estimées elliptiques, $e^{-t\mathcal{D}^2}$ envoie continûment $L^2(M, E)$ dans $W^k(M, E)$ pour tout k . On en déduit par le théorème de plongement de Sobolev que $e^{-t\mathcal{D}^2}$ envoie continûment $L^2(M, E)$ dans $C^r(M, E)$ pour tout r . Puisque $e^{-t\mathcal{D}^2} = (e^{-t\mathcal{D}^2/2})^2$, il suffit d'appliquer le lemme. \square

Le «noyau de la chaleur» k_t est défini par

$$e^{-t\mathcal{D}^2} s(m_1) = \int_M k_t(m_1, m_2) s(m_2) \text{vol}(m_2)$$

$(t, m_1, m_2) \rightarrow k_t(m_1, m_2)$ est une fonction C^∞ de (m_1, m_2) et de $t > 0$.

Définition 27. *Soit $m \in M$ et $\sigma \in E_m$. Une solution fondamentale partant de (m, σ) de l'équation de la chaleur est une section C^∞ w_t de E dépendant de t , qui vérifie l'équation de la chaleur pour $t > 0$ et telle que pour toute section $s \in C^\infty$ de E , on ait : $\langle s, w_t \rangle \rightarrow (s(m), \sigma)$ quand $t \rightarrow 0$ où (\cdot, \cdot) est le produit scalaire sur E_m .*

Le résultat suivant est une simple conséquence des définitions :

Proposition 17. *Soit k_t le noyau de la chaleur, alors $w_t(m') = k_t(m', m)(\sigma)$ est une solution fondamentale de l'équation de la chaleur partant de (m, σ) .*

Définition 28. *Soit f une fonction sur \mathbb{R} . Une série formelle $\sum_{k=0}^{\infty} a_k t^{n_k}$ est un développement asymptotique de f au voisinage de $t = 0$ si $n_k < n_{k+1}$ et si pour tout l , il existe une constante C_l telle que $|f(t) - \sum_{k=0}^l a_k t^{n_k}| \leq C_l t^{l+1}$. Si c'est le cas, on écrira : $f(t) \sim \sum_{k=0}^{\infty} a_k t^{n_k}$.*

Le théorème suivant et son corollaire sont les résultats centraux de cette section :

Théorème 4. *Il existe un développement asymptotique au voisinage de $t = 0$ de la solution fondamentale de l'équation de la chaleur partant de (m, σ) , de la forme :*

$$w_t(m') \sim \frac{1}{(4\pi t)^{n/2}} e^{-\frac{d(m, m')^2}{4t}} \cdot (\eta_0(m') + t\eta_1(m') + t^2\eta_2(m') + \dots)$$

où les η_j sont des sections de E . Les valeurs $\eta_j(m)$ peuvent être calculées à partir d'expressions algébriques faisant intervenir σ , les métriques, les coefficients des connexions et les dérivées de ces données. De plus $\eta_0(m) = \sigma$.

Corollaire 2. *Il existe un développement asymptotique de la restriction à la diagonale de $M \times M$ du noyau de la chaleur k_t de la forme :*

$$k_t(m, m) \sim \frac{1}{(4\pi t)^{n/2}} (\theta_0(m) + t\theta_1(m) + t^2\theta_2(m) + \dots)$$

où les θ_j sont des sections de $\text{End}(E)$. Les valeurs $\theta_j(m)$ peuvent être calculées à partir d'expressions algébriques faisant intervenir les métriques, les coefficients des connexions et les dérivées de ces données. De plus $\theta_0(m) = \text{Id}_{E_m}$.

Pour prouver le théorème, on va construire une «solution asymptotique formelle» par une procédure récursive, entièrement locale et algébrique. Puis on va montrer grâce aux estimées elliptiques que la solution asymptotique formelle est en fait asymptotique à la solution réelle.

Pour démarrer la procédure récursive, il faut une première approximation : c'est ce que donne le lemme suivant :

Lemme 7. *Soit $g^{i,j}(x)$ une métrique riemannienne sur un voisinage de l'origine dans \mathbb{R}^n et soit $f(x, t) = \frac{1}{(4\pi t)^{n/2}} e^{-\text{dist}(0,x)^2/(4t)}$. Alors : $\partial_t f - g^{i,j} \partial_{i,j}^2 f = (\frac{a_1}{t} + a_2)f$ (*) où a_1 et a_2 sont des fonctions C^∞ et de plus $a_1(0) = 0$.*

Démonstration. On se place en coordonnées polaires géodésiques : $r^2 = g_{ij}(x)x^i x^j$, alors : $f(x, t) = \frac{1}{(4\pi t)^{n/2}} e^{-r^2/(4t)}$ d'où : $g^{i,j} \partial_{i,j}^2 f = \frac{1}{(4\pi t)^{n/2}} (\frac{r^2}{4t^2} - \frac{1}{2t} \text{trace}(g^{ij})) e^{-r^2/(4t)}$ et : $\partial_t f = \frac{1}{(4\pi t)^{n/2}} (\frac{r^2}{4t^2} - \frac{n}{2t}) e^{-r^2/(4t)}$. \square

La proposition suivante met en place la procédure récursive. $-\mathcal{D}^2$ est clairement du type de l'opérateur L ci-dessous dans toute carte et pour toute trivialisations de E .

Proposition 18. *Soit $g^{i,j}$ une métrique riemannienne sur un voisinage de l'origine dans \mathbb{R}^n et soient b_i , $i = 1, \dots, n$ et c des fonctions à valeurs matricielles. Soit L l'opérateur différentiel agissant sur les fonctions définies sur \mathbb{R}^n à valeurs vectorielles :*

$$Lu = g^{i,j} \partial_{i,j}^2 u + b^i(x) \partial_i(x) + c(x)u$$

Alors l'équation $\partial_t u = Lu$ a une unique «solution asymptotique» de forme :

$$u(x, t) \sim \frac{1}{(4\pi t)^{n/2}} e^{-\text{dist}(0,x)^2/(4t)} (u_0(x) + tu_1(x) + \dots)$$

pour une valeur initiale $u_0(0)$ donnée.

Ici, «solution asymptotique» signifie que si on tronque la série formelle précédente à l'ordre q pour obtenir une fonction Σ_q alors on a $(\partial_t - L)\Sigma_q = \frac{1}{(4\pi t)^{n/2}} e^{-\text{dist}(0,x)^2/(4t)} t^q r_q(x, t)$ où $r_q(x, t)$ est une fonction C^∞ .

De plus, les valeurs en 0 de u_0 , u_1 ... sont des expressions algébriques finies en la donnée initiale $u_0(0)$ et les valeurs des fonctions g , b_i , c et de leurs dérivées en 0.

Démonstration. On injecte le développement $\frac{1}{(4\pi t)^{n/2}} e^{-\text{dist}(0,x)^2/(4t)}(u_0(x) + tu_1(x) + \dots)$ dans l'équation, on utilise le lemme, on égalise les coefficients des mêmes puissances de t , ce qui donne des équations différentielles vérifiées par les u_i qui se résolvent par récurrence. L'énoncé sur les valeurs en 0 résulte du fait que le calcul des constantes d'intégration à partir des conditions initiales se fait de manière algébrique. \square

Cette série construite sur \mathbb{R}^n peut se construire dans toute carte de M et se recoller en un développement sur tout M en utilisant une partition de l'unité.

Finissons la preuve du théorème. Il faut montrer que la solution asymptotique construite dans la proposition précédente est asymptotique à la vraie solution fondamentale de l'équation de la chaleur partant de (m, σ) avec m le centre du système de coordonnées géodésiques de la proposition précédente et σ la valeur initiale $u_0(0)$. Soit Σ_q la somme des q premiers termes de la solution asymptotique. Pour tout $s \in C^\infty(M, E)$, on a $\langle s, \Sigma_q \rangle \rightarrow \langle s_m, \sigma \rangle$ quand $t \rightarrow 0$. Pour q assez grand, on a pour tout $t > 0$: $(\partial_t + \mathcal{D}^2)\Sigma_q(t) = t^N r(t)$ avec r une section C^∞ de E dépendant continûment de t pour $t \geq 0$. On veut montrer que $|w_t - \Sigma_q(t)| = O(t^{N+1})$. Soit s_t l'unique solution de $(\partial_t + \mathcal{D}^2)s_t = -t^N r(t)$ avec $s_0 = 0$. Alors $\Sigma_q(t) + s_t$ est une solution fondamentale partant de (m, σ) d'où $\Sigma_q(t) + s_t = w_t$. Il faut donc montrer que $|s_t| = O(t^{N+1})$. Par le théorème de plongement de Sobolev, il suffit de montrer qu'il existe C telle que $\|s_t\|_n \leq Ct^{N+1}$. Mais $s_t = \int_0^t e^{-(t-u)\mathcal{D}^2}(-u^N r(u))du$ et $e^{-t\mathcal{D}^2}$ est uniformément borné sur chaque espace de Sobolev grâce aux estimées elliptiques et au fait que $e^{-t\mathcal{D}^2}$ est uniformément borné sur L^2 , d'où le résultat.

En mettant en oeuvre de façon effective la preuve précédente, on peut calculer les premiers coefficients θ_i , par exemple :

Proposition 19. *Pour le laplacien scalaire, le développement asymptotique de noyau de la chaleur commence par les termes : $\theta_0(m) = 1$ et $\theta_1(m) = \frac{1}{6}R(m)$ où $R(m)$ est la courbure scalaire de M au point m .*

La fin de cette section est consacrée au fait de donner un sens à l'expression «trace du noyau de la chaleur». Soient H et H' des espaces de Hilbert de dimension infinie séparables et soient (e_i) et (e_j) des bases hilbertiennes de respectivement H et H' . Soit $A : H \rightarrow H'$ un opérateur continu. A peut être représenté par une «matrice infinie» $c_{i,j}(A) = \langle Ae_i, e_j \rangle$.

Définition 29. *La quantité $\|A\|_{HS}^2 = \sum_{i,j} |c_{i,j}(A)|^2 \in [0, +\infty]$ est indépendante du choix des bases de H et H' .*

Démonstration. Par Parseval, $\|A\|_{HS}^2 = \sum_i \|Ae_i\|^2$, qui est clairement indépendant d'un choix de base de H' . Ce résultat montre d'autre part que $\|A^*\|_{HS}^2$ est indépendant d'un choix de base de H . Mais $\|A\|_{HS}^2 = \|A^*\|_{HS}^2$ car $c_{i,j}(A^*) = \bar{c}_{j,i}(A)$, d'où le résultat. \square

Définition 30. *Un opérateur A tel que $\|A\|_{HS} < +\infty$ est dit de Hilbert-Schmidt et $\|A\|_{HS}$ s'appelle sa norme de Hilbert-Schmidt.*

Les résultats suivants concernant les opérateurs de Hilbert-Schmidt sont relativement faciles et classiques. On ne donnera donc pas les démonstrations (il s'agit essentiellement de manipulations formelles d'analyse hilbertienne).

Proposition 20. *Soit H un espace de Hilbert séparable de dimension infinie. On considère des opérateurs de H dans H .*

1. *La norme de Hilbert-Schmidt provient du produit scalaire $\langle A, B \rangle_{HS} = \sum_{i,j} \bar{c}_{i,j}(A)c_{i,j}(B)$.*
2. *L'ensemble des opérateurs de Hilbert-Schmidt, muni de \langle, \rangle_{HS} , est un espace de Hilbert.*

3. Un opérateur de Hilbert-Schmidt est un opérateur compact.
4. Un produit d'un opérateur de Hilbert-Schmidt par un opérateur continu est un opérateur de Hilbert-Schmidt : l'ensemble des opérateurs de Hilbert-Schmidt est un idéal bilatère de l'anneau des opérateurs continus de H .

Définition 31. Soit H un espace de Hilbert séparable de dimension infinie. Un opérateur continu A de H est dit à trace s'il est produit de deux opérateurs de Hilbert-Schmidt A et B et on définit alors la trace de T comme étant : $Tr(T) = \langle A^*, B \rangle_{HS}$.

Vérifions que $Tr(T)$ ne dépend pas des choix de A et B : soit (e_i) une base hilbertienne de H , on a : $Tr(T) = \sum_{i,j} \bar{c}_{i,j}(A^*)c_{i,j}(B) = \sum_{i,j} c_{j,i}(A)c_{i,j}(B) = \sum_j c_{j,j}(T)$. En particulier :

Proposition 21. Soit T un opérateur autoadjoint à trace, alors $Tr(T)$ est égale à la somme des valeurs propres de T .

Proposition 22. Soient T et B des opérateurs continus de H . Si T est à trace ou si T et B sont Hilbert-Schmidt, alors TB et BT sont à trace et $Tr(TB) = Tr(BT)$.

Récapitulons les différents espaces introduits :

{opérateurs à trace} \subset {opérateurs de Hilbert-Schmidt} \subset {opérateurs compacts} \subset {opérateurs continus}

Il faut penser à cette suite d'espace comme l'analogue non-commutatif de : $l^1 \subset l^2 \subset c_0 \subset l^\infty$. En particulier, les relations de dualité sont identiques.

Remarque(non-nécessaire pour la suite) Mettre en évidence cette analogie est essentielle pour comprendre comment développer un analogue «non-commutatif» de la théorie de la mesure via la théorie des algèbres de von Neumann.

Proposition 23. Soit M une variété riemannienne compacte et A un opérateur à noyau L^2 sur $L^2(M) : Au(m_1) = \int_M k(m_1, m_2)u(m_2)vol(m_2)$, $k \in L^2(M \times M)$. Alors A est Hilbert-Schmidt et $\|A\|_{HS} = \|k\|_{L^2(M \times M)}$

Théorème 5. Si on suppose de plus que $k \in C^\infty(M \times M)$ et que A est un opérateur régularisant, alors A est à trace et $Tr(A) = \int_M k(m, m)vol(m)$

Considérons maintenant des opérateurs agissant sur les sections d'un fibré vectoriel E .

Théorème 6. Soit A un opérateur régularisant sur $L^2(M, E)$ de noyau k . Alors A est à trace et : $Tr(A) = \int_M trk(m, m)vol(m)$ où $tr : S_m \otimes S_m^* \rightarrow \mathbb{C}$ est la trace usuelle sur les endomorphismes de l'espace de dimension finie S_m .

Remarque(non-nécessaire pour la suite) Le développement asymptotique du noyau de la chaleur du laplacien scalaire d'une variété riemannienne est un outil très important en géométrie riemannienne : l'étude des liens existant entre la géométrie de la variété et le spectre du laplacien s'appelle la géométrie spectrale. À titre d'illustration, citons une célèbre formule de Weyl donnant le comportement asymptotique du spectre.

Soit M une variété riemannienne compacte orientée. Le spectre du laplacien scalaire Δ sur M est un ensemble discret de nombres réels positifs $\lambda_0 \leq \lambda_1 \leq \dots$ tendant vers l'infini. L'opérateur $e^{-t\Delta}$ est régularisant, donc à trace et sa trace est donnée par intégration de son noyau sur la diagonale de $M \times M$. Du développement asymptotique du noyau de la chaleur, on déduit donc : $Tr(e^{-t\Delta}) = \sum_i e^{-t\lambda_i} \sim \frac{1}{(4\pi t)^{n/2}}(\Theta_0 + t\Theta_1 + \dots)$ où $\Theta_i = \int_M \theta_i(m)vol(m)$. La connaissance du premier terme $\Theta_0 = Vol(M)$ suffit pour démontrer, à l'aide d'un argument type théorème taubérien :

Théorème 7. Soit $N(\lambda)$ le nombre de valeurs propres inférieures à λ . Alors, quand $\lambda \rightarrow \infty$,

$$N(\lambda) \sim \frac{1}{(4\pi)^{n/2} \Gamma(n/2 + 1)} \text{Vol}(M) \lambda^{n/2}$$

et la j -ième valeur propre vérifie, quand $j \rightarrow \infty$:

$$\lambda_j \sim (4\pi) \frac{\Gamma(n/2 + 1)^{2/n}}{\text{Vol}(M)} j^{2/n}$$

3.2.4 Problème de l'indice

Définition 32. Soit B_1 et B_2 des espaces de Banach. Un opérateur continu $A : B_1 \rightarrow B_2$ est appelé un opérateur de Fredholm si son noyau $\text{Ker}(A)$ est de dimension finie dans B_1 et si son image $\text{Im}(A)$ est fermée et de codimension finie dans B_2 . Si A est un opérateur de Fredholm, l'indice de A est défini par : $\text{ind}(A) = \dim \text{Ker}(A) - \text{codim} \text{Im}(A) = \dim \text{Ker}(A) - \dim \text{coKer}(A)$.

Remarque : On peut montrer que l'hypothèse de finitude de $\dim \text{Ker}(A)$ et $\text{codim} \text{Im}(A)$ implique la fermeture de $\text{Im}(A)$ mais on n'utilisera pas ce résultat par la suite.

Si B_1 et B_2 sont de dimensions finies, alors tout opérateur A de B_1 dans B_2 est Fredholm et d'après le théorème du rang, $\text{ind}(A) = \dim(B_1) - \dim(B_2)$, ce qui ne dépend pas de A ! On s'attend donc à ce que en dimension infinie, l'indice soit un invariant très stable de l'opérateur A . En effet, on peut montrer que :

Proposition 24. La fonction $A \rightarrow \text{ind}(A)$ est constante sur les composantes connexes de l'espace des opérateurs de Fredholm.

Soit M une variété riemannienne compacte orientée de dimension n , soit E un fibré de Clifford sur M et \mathcal{D} l'opérateur de Dirac correspondant. On a vu que \mathcal{D} pouvait être considéré comme un opérateur continu de $W^1(M, E)$ dans $L^2(M, E)$.

Lemme 8. L'opérateur de Dirac \mathcal{D} est un opérateur de Fredholm de $W^1(M, E)$ dans $L^2(M, E)$.

Démonstration. Le théorème spectral pour l'opérateur de Dirac (3.2.2, Théorème 3) montre que $\text{Ker} \mathcal{D}$ est de dimension finie. Pour vérifier le reste de la définition d'un opérateur de Fredholm, il est suffisant de montrer que $\text{Im} \mathcal{D}$ est le supplémentaire orthogonal de $\text{Ker} \mathcal{D}$ dans $L^2(M, E)$. On rappelle que \mathcal{D} est autoadjoint (2.1.5, Proposition 4).

Soit $x \in W^1(M, E)$. Alors il existe une suite (x_j) dans $C^\infty(M, E)$ telle que $x_j \rightarrow x$ dans $W^1(M, E)$ et donc $\mathcal{D}x_j \rightarrow \mathcal{D}x$ dans $L^2(M, E)$. Si $s \in \text{Ker} \mathcal{D}$, on a pour tout j , $\langle \mathcal{D}x_j, s \rangle = \langle x_j, \mathcal{D}s \rangle = 0$ d'où $\langle \mathcal{D}x, s \rangle = 0$. On a donc montré que $\text{Im} \mathcal{D}$ est dans l'orthogonal de $\text{Ker} \mathcal{D}$.

Réciproquement, soit x dans l'orthogonal de $\text{Ker} \mathcal{D}$. Soit f une fonction sur \mathbb{R} vérifiant $f(0) = 0$ et $f(\lambda) = 1/\lambda$ pour tout λ valeur propre non-nulle de \mathcal{D} . Alors $\mathcal{D}f(\mathcal{D})x = x$. Mais $\mathcal{D}f(\mathcal{D})x \in L^2(M, E)$ et le résultat de régularité de 3.3.2, Proposition 13 montre alors que $f(\mathcal{D})x \in W^1(M, E)$ d'où $x \in \text{Im} \mathcal{D}$. □

Il résulte de la preuve que $\text{codim} \text{Im} \mathcal{D} = \dim \text{Ker} \mathcal{D}$: \mathcal{D} est d'indice zéro, ce qui n'est donc pas très intéressant (la réponse ne dépend ni de M , ni de E ...). Pour faire mieux, il faut une donnée géométrique supplémentaire : une graduation.

Définition 33. L'opérateur de Dirac \mathcal{D} est dit gradué s'il provient d'un complexe de Dirac de longueur 2 : plus précisément, on suppose que E s'écrit $E = E^+ \oplus E^-$ et que par rapport à cette décomposition, \mathcal{D} est de la forme : $\begin{pmatrix} 0 & \mathcal{D}_- \\ \mathcal{D}_+ & 0 \end{pmatrix}$. On appelle $\gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ l'opérateur de graduation.

Puisque \mathcal{D} est un opérateur de Fredholm, il en est de même de \mathcal{D}_+ vu comme un opérateur de $W^1(M, E^+)$ dans $L^2(M, E^-)$ et c'est l'indice de cet opérateur qui s'avère être intéressant : on en parlera comme de l'indice de l'opérateur de Dirac gradué \mathcal{D} .

Soit $E = E^+ \oplus E^-$ un fibré de Clifford gradué et soit T un opérateur à trace sur $L^2(M, E)$. La supertrace de T , notée $Tr_S(T)$ est par définition $Tr(\gamma_5 T)$. Si T est un opérateur régularisant de noyau k , on a : $Tr_S(T) = \int tr_S k(m, m) vol(m)$ où la supertrace locale tr_S est par définition $tr(\gamma_5 k)$.

Lemme 9. Soit \mathcal{D} un opérateur de Dirac gradué, alors : $ind(\mathcal{D}) = Tr_S(e^{-t\mathcal{D}^2})$ pour tout $t > 0$.

Démonstration. \mathcal{D} est un opérateur gradué, $E = E^+ \oplus E^-$. Dans une base adaptée à la graduation, il s'écrit :

$$\mathcal{D} = \begin{pmatrix} 0 & \mathcal{D}_- \\ \mathcal{D}_+ & 0 \end{pmatrix}$$

On a donc :

$$\mathcal{D}^2 = \begin{pmatrix} \mathcal{D}_+^* \mathcal{D}_+ & 0 \\ 0 & \mathcal{D}_-^* \mathcal{D}_- \end{pmatrix}$$

Mais \mathcal{D} est autoadjoint (2.1.5, Proposition 4), d'où $\mathcal{D}_- = \mathcal{D}_+^*$. On en déduit que :

$$\mathcal{D}^2 = \begin{pmatrix} \mathcal{D}_+^* \mathcal{D}_+ & 0 \\ 0 & \mathcal{D}_+ \mathcal{D}_+^* \end{pmatrix}$$

On a donc : $Tr_S(e^{-t\mathcal{D}^2}) = Tr(e^{-t\mathcal{D}_+^* \mathcal{D}_+}) - Tr(e^{-t\mathcal{D}_+ \mathcal{D}_+^*})$. Par décomposition spectrale (analogue à 3.2.2, Théorème 3) de $\mathcal{D}_+^* \mathcal{D}_+$ et $\mathcal{D}_+ \mathcal{D}_+^*$ et en utilisant la positivité de ces opérateurs, on a, quand $t \rightarrow +\infty$:

$$Tr(e^{-t\mathcal{D}_+^* \mathcal{D}_+}) = \sum_{\lambda_i \text{ valeur propre de } \mathcal{D}_+^* \mathcal{D}_+} e^{-t\lambda_i} \rightarrow dim Ker(\mathcal{D}_+^* \mathcal{D}_+) = dim Ker \mathcal{D}_+$$

et :

$$Tr(e^{-t\mathcal{D}_+ \mathcal{D}_+^*}) = \sum_{\mu_i \text{ valeur propre de } \mathcal{D}_+ \mathcal{D}_+^*} e^{-t\mu_i} \rightarrow dim Ker(\mathcal{D}_+ \mathcal{D}_+^*) = dim Ker \mathcal{D}_+^*$$

Ainsi, $\lim_{t \rightarrow +\infty} Tr_S(e^{-t\mathcal{D}^2}) = Tr(e^{-t\mathcal{D}_+^* \mathcal{D}_+}) - Tr(e^{-t\mathcal{D}_+ \mathcal{D}_+^*}) = ind \mathcal{D}$.

Pour finir la démonstration, il faut montrer que $Tr_S(e^{-t\mathcal{D}^2}) = Tr(e^{-t\mathcal{D}_+^* \mathcal{D}_+}) - Tr(e^{-t\mathcal{D}_+ \mathcal{D}_+^*})$ est indépendant de t . Or : $\frac{d}{dt} Tr_S(e^{-t\mathcal{D}^2}) = Tr(\mathcal{D}_+ \mathcal{D}_+^* e^{-t\mathcal{D}_+ \mathcal{D}_+^*}) - Tr(\mathcal{D}_+^* \mathcal{D}_+ e^{-t\mathcal{D}_+^* \mathcal{D}_+})$. Il suffit de montrer que cette expression est nulle, ce qui se fait en appliquant deux fois 3.2.3, Proposition 22 :

$$\begin{aligned} Tr(\mathcal{D}_+ \mathcal{D}_+^* e^{-t\mathcal{D}_+ \mathcal{D}_+^*}) &= Tr(\mathcal{D}_+ \mathcal{D}_+^* e^{-\frac{t}{2}\mathcal{D}_+ \mathcal{D}_+^*} e^{-\frac{t}{2}\mathcal{D}_+ \mathcal{D}_+^*}) = Tr(e^{-\frac{t}{2}\mathcal{D}_+ \mathcal{D}_+^*} \mathcal{D}_+ \mathcal{D}_+^* e^{-\frac{t}{2}\mathcal{D}_+ \mathcal{D}_+^*}) \\ &= Tr(\mathcal{D}_+^* e^{-\frac{t}{2}\mathcal{D}_+ \mathcal{D}_+^*} e^{-\frac{t}{2}\mathcal{D}_+ \mathcal{D}_+^*} \mathcal{D}_+) = Tr(\mathcal{D}_+^* e^{-t\mathcal{D}_+ \mathcal{D}_+^*} \mathcal{D}_+) = Tr(\mathcal{D}_+^* \mathcal{D}_+ e^{-t\mathcal{D}_+^* \mathcal{D}_+}) \end{aligned}$$

□

Rappelons le développement asymptotique du noyau k_t de $e^{-t\mathcal{D}^2}$:

$$Tr_S(e^{-t\mathcal{D}^2}) \sim \frac{1}{(4\pi t)^{n/2}} \left(\int tr_S \theta_0 + t \int tr_S \theta_1 + \dots \right)$$

Le lemme précédent dit que $Tr_S(e^{-t\mathcal{D}^2})$ est en fait indépendant de $t > 0$. On a donc :

Proposition 25. *L'indice de l'opérateur de Dirac gradué \mathcal{D} est zéro si $n(= \dim M)$ est impaire et est égal à $\frac{1}{(4\pi)^{n/2}} \int \text{tr}_S \theta_{n/2}$ si n est paire, où $\theta_{n/2}$ est une certaine expression algébrique en les métriques, les coefficients des connexions et les différentes dérivées de ces données.*

Une petite application :

Corollaire 3. *L'indice est multiplicatif vis à vis des revêtements : si \hat{M} est un revêtement à k feuillets de M , et si \hat{E} et $\hat{\mathcal{D}}$ sont les relèvements naturels de E et \mathcal{D} à \hat{M} , alors $\text{ind}(\hat{\mathcal{D}}) = k \cdot \text{ind}(\mathcal{D})$*

Remarque : contrairement à certaines apparences trompeuses, ce résultat n'a rien d'évident si on essaye de partir de la définition de l'indice. Si on admet le théorème de l'indice, il permet d'obtenir la multiplicativité d'invariants topologiques et là encore, on obtient des résultats non-évidents : la multiplicativité de la caractéristique d'Euler-Poincaré est très facile mais démontrer le résultat analogue pour la signature sans utiliser le théorème de l'indice est plus que non trivial (mais possible : c'est un exercice pour le lecteur qui connaîtrait beaucoup de topologie algébrique).

On va finir cette section par deux exemples explicites de calcul d'indice d'opérateur de Dirac. On va ainsi vérifier « empiriquement » un lien entre indice et classes caractéristiques. Le résultat général exprimant cette relation, le théorème de l'indice, sera le sujet de la partie suivante.

Pour ces deux exemples, on adopte la convention « physique » pour la définition des matrices de Dirac et donc de l'opérateur de Dirac : $\{\gamma_i, \gamma_j\} = 2g_{i,j}$ (Voir la remarque qui précède 2.1.5, Définition 10).

Un exemple de calcul direct d'indice d'un opérateur de Dirac

Soit M la sphère S^2 . On note M^+ et M^- les cartes recouvrant M données par la projection stéréographique : M^+ (resp. M^-) est la sphère privée du pôle nord (resp. sud). Pour faire des calculs, on identifiera M^+ à \mathbb{R}^2 et M^- à $\mathbb{R}^2 \cup \{\infty\}$. On note par (x, y) (resp. (r, θ)) les coordonnées cartésiennes (resp. polaires) sur \mathbb{R}^2 . Soit $m \in \mathbb{Z}$, on note E_m le fibré principal sur M de groupe $U(1)$ obtenu en prenant deux fibrés triviaux sur chacun des hémisphères et en les recollant via une application $g : S^1 \rightarrow U(1)$ de degré m au niveau de l'équateur. On définit une connexion A sur E_m en prenant sur M^- la connexion A^- triviale et sur M^+ :

$$A^+ = im \frac{r^2}{r^2 + 1} d\theta = -\frac{im}{x^2 + y^2 + 1} (ydx - xdy)$$

où dans cette formule on identifie $U(1)$ à $\mathbb{R} \cup \{\infty\}$. On vérifie que ces définitions sont compatibles (et donc définissent bien une connexion) en remarquant que à l'infini de \mathbb{R}^2 identifié à M^+ , A^+ se comporte comme $g^{-1}dg$ où $g(\theta) = e^{im\theta}$, i.e. on passe bien de A^+ à A^- par le changement de carte définissant E_m . Le calcul de la courbure F de A s'écrit :

$$F = dA = im \frac{(r^2 + 1)2rdr - r^2 2rdr}{(r^2 + 1)^2} d\theta = im \frac{2rdrd\theta}{(r^2 + 1)^2} = 2im \frac{dxdy}{(x^2 + y^2 + 1)^2}$$

On a donc :

$$\frac{1}{2i\pi} \int_{\mathbb{R}^2} F = m \int_{r=0}^{\infty} \frac{2r}{(r^2 + 1)^2} dr = m$$

En deux dimensions euclidiennes, on prend pour matrices de Dirac :

$$\gamma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \gamma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

L'opérateur de Dirac s'écrit alors :

$$\begin{aligned} \mathcal{D} = \gamma^\mu(\partial_\mu + A_\mu) &= \begin{pmatrix} 0 & \partial_x + i\partial_y + A_x + iA_y \\ \partial_x - i\partial_y + A_x - iA_y & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 2\partial_{\bar{z}} - m\frac{z}{z\bar{z}+1} \\ 2\partial_z + m\frac{\bar{z}}{z\bar{z}+1} & 0 \end{pmatrix} \end{aligned}$$

où $z = x + iy$. Supposons $m \geq 0$. Soit $\begin{pmatrix} \chi \\ 0 \end{pmatrix}$ un spineur de chiralité positive, la condition pour être vecteur propre de valeur propre nulle est donc :

$$2\frac{\partial\chi}{\partial z} + m\frac{\bar{z}}{z\bar{z}+1}\chi = 0$$

ce qui donne :

$$\chi(z, \bar{z}) = \frac{g(\bar{z})}{(1 + z\bar{z})^{m/2}}$$

avec g holomorphe. χ devant être de carré intégrable, g est nécessairement un polynôme de degré inférieur ou égal à $m - 1$ ce qui montre que l'espace des modes nuls de chiralité positive est de dimension m . En écrivant la condition d'être vecteur propre de valeur propre nulle pour un spineur de chiralité négative, on trouve qu'il n'y a pas de mode nul normalisable. Pour $m \geq 0$, on a donc montré que $\text{ind}\mathcal{D} = m$. Si $m < 0$, les rôles joués par la chiralité dans la discussion précédente sont inversés : il y n'y a pas de modes nuls de chiralité positive, $-m$ de chiralité négative, d'où de nouveau $\text{ind}\mathcal{D} = m$. Ainsi, on a vérifié la relation : $\text{ind}\mathcal{D} = m = \frac{1}{2i\pi} \int F = \frac{1}{4i\pi} \int \epsilon^{\mu,\nu} F_{\mu,\nu}$.

Encore un exemple explicite

On prend pour M le tore à deux dimensions paramétré par deux coordonnées x et t , x périodique de période 2π et t périodique de période 1. Soit E un fibré en droites complexes sur M de première classe de Chern égale à un : les sections de E s'identifient aux fonctions $f(x, t)$ 2π -périodiques en x qui vérifient $f(x, t + 1) = e^{-ix}f(x, t)$. L'opérateur de Dirac s'écrit : $\mathcal{D}^+ = \partial_t - i\partial_x + t$, $\mathcal{D}^- = \partial_t + i\partial_x + t$. On cherche à résoudre $\mathcal{D}^+ f = 0$. Pour cela, on développe f en série de Fourier par rapport à x : $f(x, t) = \sum_n a_n(t)e^{inx}$. Les coefficients $a_n(t)$ vérifient chacun une équation différentielle qu'on résout facilement pour obtenir : $a_n(t) = C_n e^{-\frac{(t+n)^2}{2}}$ où les C_n sont des constantes. La condition $f(x, t + 1) = e^{-ix}f(x, t)$ montre qu'en fait C_n ne dépend pas de n . Finalement, on a donc à une constante multiplicative près une unique solution : $f(x, t) = \sum_n e^{-\frac{(n+t)^2}{2} + inx}$. Lorsqu'on cherche à résoudre de même l'équation $\mathcal{D}^- f = 0$, on trouve pour les a_n des exponentielles croissantes et la série définissant formellement f diverge : il n'y a pas de solution non-nulle. Ainsi, on a montré que $\text{ind}(\mathcal{D}) = 1$.

3.3 Théorème de l'indice

On revient à la convention «mathématique» pour la définition des matrices de Dirac et donc de l'opérateur de Dirac : $\{\gamma_i, \gamma_j\} = -2g_{i,j}$ (Voir la remarque qui précède 2.1.5, Définition 10).

3.3.1 Objectif et réduction à un problème local

Dans toute la section, M est une variété spin de dimension paire $n = 2r$. M est alors munie d'un fibré de Clifford canonique S : le fibré spinoriel, et l'opérateur de Dirac \mathcal{D} agit sur les sections de S . \mathcal{D} est gradué par rapport à γ_5 . On rappelle qu'en termes d'une base orthonormée

locale e_1, \dots, e_n , on a $\gamma_5 = i^{n/2}\nu$ où : $\nu = e_1 \dots e_n$. Les opérateurs ν et γ_5 anticommulent avec les e_i et γ_5 a été construit à partir de ν de façon à être une involution : $\gamma_5^2 = 1$.

On veut calculer $ind(\mathcal{D})$ comme une intégrale sur M de classes caractéristiques. D'après 3.2.4, Théorème 3, on a : $ind(\mathcal{D}) = Tr_S(e^{-t\mathcal{D}^2})$ pour tout $t > 0$. D'après 3.2.3, $e^{-t\mathcal{D}^2}$ est un opérateur à noyau k_t et on a un développement asymptotique pour $t \rightarrow 0$:

$$k_t(m, m) \sim \frac{1}{(4\pi t)^{n/2}} (\theta_0(m) + t\theta_1(m) + t^2\theta_2(m) + \dots)$$

Rappelons également que si w_t est une solution fondamentale partant de (m, σ) de $\partial_t s + \mathcal{D}^2 s = 0$, on a pour $t \rightarrow 0$:

$$w_t(m') \sim \frac{1}{(4\pi t)^{n/2}} e^{-\frac{d(m, m')^2}{4t}} (\eta_0(m') + t\eta_1(m') + \dots)$$

On peut voir w_t comme une fonction de σ et pour tout $t > 0$, w_t est alors une section de $S \otimes S_m^*$: dans la suite, on garde ce point de vue. On a alors : $w_t(m) = k_t(m, m)$ et pour tout i , $\theta_i(m) = \eta_i(m)$. Ainsi :

$$ind(\mathcal{D}) = \frac{1}{(4\pi t)^{n/2}} \int tr_S(\theta_{n/2}(m)) vol(m) = \frac{1}{(4\pi t)^{n/2}} \int tr_S(\eta_{n/2}(m)) vol(m)$$

et on sait d'après 3.2.3, Théorème 4 que $\eta_{n/2}(m)$ est une expression algébrique en la courbure $R(m)$ au point m : $\eta_{n/2}(m) = P_{n/2}(R(m))$, où $P_{n/2}$ est une fonction « indépendante » du point $m \in M$. Calculer $P_{n/2}$ est donc un problème purement local.

Soit $m \in M$ et (x_j) un système de coordonnées géodésiques centré sur m . En particulier le point $m \in M$ correspond au point $x_j = 0$ dans \mathbb{R}^n . Soit e_α un repère synchrone local du fibré tangent autour de m qui coïncide en m avec le repère donné par les coordonnées. On a alors au voisinage de m : $g_{j,k}(x) = \delta_{j,k} + O(|x|^2)$ et $\Gamma_{j\alpha}^\beta = -\frac{1}{2}R_{\alpha,j,k}^\beta(0)x^k + O(|x|^2)$. On se rappelle que : $\mathcal{D}^2 = \nabla^* \nabla + \frac{R}{4}$ où R est la courbure scalaire, ce qui s'écrit au voisinage de m : $\mathcal{D}^2 = -g^{jk}(\nabla_j \nabla_k - \Gamma_{j,k}^i \nabla_i) + \frac{R}{4}$. On s'est donc ramené à l'étude de la solution fondamentale w_t partant de 0 de l'équation $\partial_t s + \mathcal{D}^2 s = 0$ définie sur un voisinage U de 0 dans \mathbb{R}^n . On cherche à calculer $\eta_{n/2}(0)$, fonction de la courbure au point m . Quitte à réduire U , on peut trivialisier localement $S \otimes S_m^*$ et donc identifier w_t et $\eta_{n/2}$ à des sections à valeurs dans l'algèbre de Clifford.

Pour résoudre notre problème, il suffirait de savoir calculer les $\eta_i(0) = P_i(R(0))$, $x \in U$, i.e. le développement asymptotique de w_t en 0. La preuve de 3.2.3, Théorème 4 donne une procédure récursive pour calculer les P_i de proche en proche qui est applicable pour les premiers i , mais il est très compliqué d'obtenir une formule générale donnant $P_{n/2}$ pour tout n . Heureusement, on ne cherche que la supertrace de $\eta_{n/2}(0) = P_{n/2}(R(0))$. Puisque $\eta_{n/2}$ est une fonction définie sur U à valeurs dans l'algèbre de Clifford, on pourra utiliser le lemme suivant.

Lemme 10. *Soit e_1, \dots, e_n une base orthonormée de \mathbb{R}^n , si $E \subset \{1, \dots, n\}$, on note $\tilde{E} = \prod_{i \in E} e_i \in Cl(\mathbb{R}^n)$. Soit $c = \sum_E c_E \tilde{E}$ un élément de $Cl(\mathbb{R}^n)$. Si on voit c comme un endomorphisme de la représentation spinorielle, alors sa supertrace est donnée par $(-2i)^{n/2} c_{12\dots n}$.*

Démonstration. Par définition, $tr_S(c) = tr(\gamma_5 c)$. Il faut donc prouver que $tr(c) = 2^{n/2} c_\emptyset$, i.e. que $tr(\tilde{E}) = 2^{n/2}$ si $\tilde{E} = \emptyset$ et 0 sinon. Le premier point est clair car $dim \Delta = 2^{n/2}$. Pour le deuxième, on remarque qu'en tant que représentation, $Cl(\mathbb{R}^n)$ est isomorphe à $\Delta \otimes \Delta^*$. La trace de \tilde{E} sur Δ est donc un multiple de la trace de \tilde{E} agissant par multiplication sur $Cl(\mathbb{R}^n)$ qui est nulle car cette action est une permutation sans point fixe des e_i . \square

Jusqu'à présent, on n'a défini le degré dans l'algèbre de Clifford que modulo 2. La donnée de la base e_α permet de construire comme dans le lemme une base (\tilde{E}) de l'algèbre de Clifford et on pose : $\deg(\tilde{E}) = \text{card}(E)$. Développons $\eta_{n/2}$ sur cette base de l'algèbre de Clifford : pour tout $x \in U$,

$$\eta_{n/2}(x) = \sum_E \eta_{n/2,E}(x) \tilde{E}$$

Alors d'après le lemme :

$$\text{Tr}_S(\eta_{n/2}(x)) = (-2i)^{n/2} \eta_{n/2,\{1,\dots,n\}}(x)$$

On cherche $\eta_{n/2,\{1,\dots,n\}}(0)$ comme fonction de la courbure $R(m)$ en m . L'idée est de modifier \mathcal{D} en un opérateur \mathcal{D}_λ sur U tel que le développement asymptotique de la solution fondamentale de $\partial_t s + \mathcal{D}_\lambda^2 s = 0$ partant de 0 :

$$w_t^\lambda(x) \sim \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}} \sum_{l,E} t^l \eta_{E,l}^\lambda(x) \tilde{E}$$

s'exprime en fonction de celui associé à \mathcal{D} : si on connaît $\eta_{n/2,\{1,\dots,n\}}^\lambda(x)$, on peut en déduire $\eta_{n/2,\{1,\dots,n\}}(x)$. Mais \mathcal{D}_λ sera un opérateur aussi compliqué que \mathcal{D} . Les choses se simplifieront en faisant $\lambda \rightarrow \infty$. On obtiendra un opérateur $L = \mathcal{D}_\infty$, une solution fondamentale :

$$w_t^\infty(x) \sim \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}} \sum_{l,E} t^l \eta_{E,l}^\infty(x) \tilde{E}$$

et on montrera que si on connaît $\eta_{n/2,\{1,\dots,n\}}^\infty(x)$, alors on connaît $\eta_{n/2,\{1,\dots,n\}}(0)$ et que L est un opérateur particulièrement simple : on pourra calculer explicitement sa solution fondamentale partant de 0 et donc calculer $\eta_{n/2,\{1,\dots,n\}}^\infty(x)$.

Dans les deux sections qui suivent, on met en oeuvre le programme qui vient d'être présenté.

3.3.2 Un changement d'échelle bien choisi

Si f est une fonction définie sur un voisinage de $0 \in \mathbb{R}^n$ à valeurs dans l'algèbre de Clifford, homogène de degré k alors on définit l'action de l'opérateur M_λ sur f par : $M_\lambda f(x) = \lambda^{-k} f(\lambda x)$. Si f n'est pas homogène de degré fixé mais seulement une fonction à valeurs dans l'algèbre de Clifford, on définit l'action de M_λ sur f en décomposant f en une somme de fonctions de homogènes $f = \sum_i f_i$, f_i homogène de degré i et on pose $M_\lambda f = \sum_i M_\lambda f_i$.

En tant qu'espace vectoriel, l'algèbre de Clifford est canoniquement isomorphe à l'algèbre extérieure. On peut donc définir la multiplication extérieure sur l'algèbre de Clifford par transport de la multiplication de l'algèbre extérieure. Au niveau des termes de plus haut degré, le produit de Clifford coïncide avec le produit extérieur. On en déduit :

Lemme 11. *Soit c une fonction définie sur un voisinage de $0 \in \mathbb{R}^n$ à valeurs dans l'algèbre de Clifford, homogène de degré k . On considère son action par multiplication à gauche sur l'algèbre de Clifford. Alors l'endomorphisme $\lambda^{-k} M_\lambda^{-1} c M_\lambda$ tend quand $\lambda \rightarrow \infty$ vers l'opération de multiplication extérieure par $c(0)$.*

On définit $\mathcal{D}_\lambda = \frac{1}{\lambda} M_\lambda^{-1} \mathcal{D} M_\lambda$.

Proposition 26. *Quand $\lambda \rightarrow \infty$, \mathcal{D}_λ^2 tend vers l'opérateur L donné par :*

$$L = - \sum_p (\partial_p - \frac{1}{8} \Theta_{p,q} x^k)^2$$

où $\Theta_{p,q}$ est la valeur à l'origine de la 2-forme de courbure de Riemann, pensée comme agissant par multiplication extérieure sur l'algèbre de Clifford, i.e. $\Theta_{p,q} = R_{p,q,\alpha,\beta} e_\alpha e_\beta$.

Démonstration. La connexion spin s'écrit : $\nabla_j = \partial_j + \frac{1}{4}\Gamma_{j,\alpha}^\beta e_\alpha e_\beta = \partial_j - \frac{1}{8}R_{j,k,\alpha,\beta} x^k e_\alpha e_\beta + O(|x|^2 e^2)$ où le symbole e rend compte du degré mis sur l'algèbre de Clifford. Il suffit ensuite de calculer explicitement \mathcal{D}_λ^2 et de prendre la limite $\lambda \rightarrow \infty$. □

Soit w_t^λ la solution fondamentale partant de 0 de l'équation de la chaleur pour l'opérateur \mathcal{D}_λ^2 .

Proposition 27. *Pour tout $t \geq 0$, on a $w_t^\lambda = \lambda^{-n} M_\lambda^{-1} w_{(t/\lambda^2)}$*

Démonstration. On vérifie facilement que $M_\lambda^{-1} w_{t/\lambda^2}$ est solution de l'équation de la chaleur et le facteur λ^{-n} assure que cette solution est bien fondamentale. □

Écrivons le développement asymptotique de w_t sous la forme :

$$w_t \sim \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/(4t)} \sum_{l=0}^{\infty} t^l \eta_l = \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/(4t)} \sum_{l,E} t^l \eta_{l,E} \tilde{E}$$

où on somme sur la base (vectorielle) choisie de l'algèbre de Clifford : $E \subset \{1, \dots, n\}$, $\tilde{E} = \prod_{i \in E} e_i$, $\eta_l = \sum_E \eta_{l,E} \tilde{E}$ et les fonctions $\eta_{l,E}$ sont alors à valeurs scalaires.

D'après la proposition précédente, le développement asymptotique correspondant pour w_t^λ est :

$$w_t^\lambda \sim \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/(4t)} \sum_{l,E} t^l \eta_{l,E}(\lambda^{-1}x) \tilde{E} \lambda^{-2l + \deg \tilde{E}}$$

Or quand $\lambda \rightarrow \infty$, l'opérateur \mathcal{D}_λ^2 tend vers l'opérateur L et donc les coefficients du développement asymptotique du noyau de la chaleur de \mathcal{D}_λ^2 tendent vers les coefficients du développement asymptotique du noyau de la chaleur de L . En faisant $\lambda \rightarrow \infty$ dans la formule précédente, on obtient donc :

Lemme 12. *Soit*

$$w_t^\infty(x) \sim \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/(4t)} \sum_{l=0}^{\infty} t^l \eta_l^\infty(x)$$

le développement asymptotique de la solution fondamentale du noyau de la chaleur pour l'opérateur L . Alors $\eta_{n/2}^\infty(0)$ est un multiple de la forme volume vol , disons égal à $g vol$. De plus g est une fonction de la courbure, $g = g(R)$ et $ind(\mathcal{D}) = (-2i)^{n/2} \int_M g(R)$.

Démonstration. La limite $\lambda \rightarrow \infty$ dans le développement asymptotique de w_t^λ donne comme coefficient de t^0 : $\sum_{E, \deg E = n} \eta_{n/2,E}(0) \tilde{E}$ d'où

$$\eta_{n/2}^\infty(0) = \sum_{E, \deg E = n} \eta_{n/2,E}(0) \tilde{E} = \eta_{n/2, \{1, \dots, n\}}(0) vol = g vol$$

avec $g = \eta_{n/2, \{1, \dots, n\}}$. D'après 3.3.1, on a : $ind(\mathcal{D}) = \int_M Tr_S(\eta_{n/2}(m)) vol(m) = (-2i)^{n/2} \int_M \eta_{n/2, \{1, \dots, n\}} vol(m)$ d'où le résultat grâce à ce qui précède. □

Il ne reste donc plus qu'à calculer la solution fondamentale partant de 0 de l'opérateur L .

3.3.3 Oscillateur harmonique

L'oscillateur harmonique est l'opérateur non-borné : $H = -d^2/dx^2 + a^2x^2, a > 0$ sur $L^2(\mathbb{R})$.

Proposition 28. *L'unique solution fondamentale de $\partial_t u + Hu = 0$ partant de 0 est :*

$$u(x, t) = \sqrt{\frac{a}{2\pi \operatorname{sh}(2at)}} e^{-\frac{ax^2}{2} \operatorname{coth}(2at)}$$

Cette formule s'appelle la formule de Mehler.

Démonstration. On cherche une solution de la forme $u(x, t) = \alpha(t)e^{-\frac{1}{2}\beta(t)x^2}$. On remplace dans l'équation : on obtient des équations différentielles en α et β qu'on sait résoudre. Le choix de normalisation de $u(x, t)$ assure que $u(x, t)$ est bien une solution fondamentale. \square

Remarque(non-nécessaire pour la suite) Le lecteur pourra essayer de retrouver (à une constante multiplicative près) la formule de Mehler par une méthode d'intégration fonctionnelle. (Il s'agit d'évaluer $\langle (x, t) | (0, 0) \rangle = \int D\psi e^{-H\psi}$, ce qui peut se faire par développement autour de la «solution classique» du mouvement vérifiant : $\frac{d^2x}{dt^2} = -2ax$).

Proposition 29. *On rappelle que $L = -\sum_p (\partial_p - \frac{1}{8}\Theta_{p,q}x^k)^2$ où $\Theta_{p,q}$ est la valeur à l'origine de la 2-forme de courbure de Riemann, pensée comme agissant par multiplication extérieure sur l'algèbre de Clifford, i.e. $\Theta_{p,q} = R_{p,q,\alpha,\beta}e_\alpha e_\beta$. Alors la solution fondamentale de l'équation de la chaleur pour L est, en l'origine :*

$$w_t^\infty(0) = \frac{1}{(4\pi t)^{n/2}} \det \left(\frac{t\Theta/4}{\operatorname{sh}(t\Theta/4)} \right)^{1/2} = \frac{1}{(4\pi t)^{n/2}} \hat{A}(-2i\pi t\Theta)$$

On rappelle que la \hat{A} -classe a été définie à la fin de 3.2.1.

Démonstration. On peut voir Θ comme une simple matrice antisymétrique. On peut l'écrire en une base orthonormée comme une matrice diagonale par blocs, chaque bloc étant une matrice 2×2 de la forme : $\begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}$. On peut donc se restreindre au cas bidimensionnel avec Θ de la forme ci-dessus. Il faut montrer que le noyau de la chaleur de L est : $\frac{1}{4\pi t}(it\theta/\operatorname{sh}(it\theta/4))$. On écrit $L = -(\partial_x - (\theta y/8))^2 - (\partial_y - (\theta x/8))^2 = L_0 + L_1$ avec : $L_0 = -(\partial_x^2 + \partial_y^2) - \frac{1}{64}\theta^2(x^2 + y^2)$ et $L_1 = \frac{1}{4}\theta(x\partial_y - y\partial_x)$. L est un opérateur radial, i.e. invariant sous $SO(2)$, il en est donc de même de sa solution fondamentale. D'autre part, L_0 est un opérateur radial et L_1 annule les fonctions radiales : la solution fondamentale de L est donc aussi celle de L_0 . Dans L_0 , on peut séparer les variables et on doit donc étudier $-\partial_x^2 - \frac{1}{64}\theta^2x^2$ qui est un oscillateur harmonique dont la solution fondamentale a été calculée précédemment. \square

On déduit de tout ce qui précède le théorème de l'indice pour l'opérateur de Dirac d'une variété spin, résultat démontré pour la première fois par Atiyah et Singer. On rappelle que la \hat{A} -classe et le caractère de Chern ch ont été définis à la fin de la partie 3.2.1 sur les classes caractéristiques.

Théorème 8. *Soit \mathcal{D} l'opérateur de Dirac sur une variété spin M . Alors :*

$$\operatorname{ind}(\mathcal{D}) = \int_M (\hat{A}(TM))_n$$

où $(\alpha)_n$ signifie qu'on prend la composante de degré n de la forme différentielle α .

Démonstration. D'après la proposition précédente et avec les notations de 3.3.2, on a $\eta_{n/2, \{1, \dots, n\}}^\infty(x) = \frac{1}{(4\pi)^{n/2}} (\hat{A}(-2i\pi\Theta))_n$. De plus, \hat{A} étant défini par une série formelle paire et en utilisant 3.3.2, Lemme 12 :

$$\eta_{n/2, \{1, \dots, n\}}(0) = \frac{1}{(4\pi)^{n/2}} (\hat{A}(2i\pi\Theta))_n(m) = \frac{(2i\pi)^{n/2}}{(4\pi)^{n/2}} (\hat{A}(TM))_n(m) = \frac{(2i)^{n/2}}{4^{n/2}} (\hat{A}(TM))_n(m)$$

d'où :

$$Tr_S(\eta_{n/2}(0)) = (-2i)^{n/2} \eta_{n/2, \{1, \dots, n\}}(0) = \frac{(-2i)^{n/2} (2i)^{n/2}}{4^{n/2}} (\hat{A}(TM))_n(m) = (\hat{A}(TM))_n(m)$$

et :

$$ind(\mathcal{D}) = \int_M Tr_S(\eta_{n/2}(m)) = \int_M (\hat{A}(TM))_n(m)$$

□

On peut généraliser ce résultat au cas de l'opérateur de Dirac d'un fibré spinoriel à coefficients dans un fibré vectoriel hermitien E . E a sa propre connexion et une 2-forme de courbure F . On a alors : $\mathcal{D}^2 = \nabla^* \nabla + \frac{R}{4} + \tilde{F}$ où : $\tilde{F} = \frac{1}{2} F_{\alpha, \beta} e_\alpha e_\beta$. L'opérateur L devient : $L = -\sum_p (\partial_p - \frac{1}{8} \Theta_{p, k})^2 + \tilde{F}(0)$ où $\tilde{F}(0)$ agit par multiplication extérieure. $\tilde{F}(0)$ étant une constante, la solution fondamentale de ce nouveau L est $e^{-t\tilde{F}(0)}$ celle de l'ancien L . On en déduit :

Théorème 9. Soit \mathcal{D}_E l'opérateur de Dirac sur une variété spin M à coefficients dans un fibré vectoriel E . Alors :

$$ind(\mathcal{D}_E) = \int_M (\hat{A}(TM) ch(E))_n.$$

Remarque(non-nécessaire pour la suite) Tout fibré de Dirac sur une variété spin est de la forme $S \otimes E$. Le théorème précédent traite donc le cas de tous les opérateurs de Dirac sur une variété spin. Dans le cas d'une variété non spin, on peut toujours construire localement un fibré spinoriel. La difficulté étant essentiellement locale, on se doute qu'on peut tirer des conséquences du résultat précédent pour des variétés n'admettant pas de structure spin. En fait, Atiyah et Singer ont démontré un théorème de l'indice valable pour tout M et pour tout opérateur elliptique et une façon de l'obtenir est de le déduire du cas des opérateurs de Dirac sur une variété spin.

3.4 Méthode de Fujikawa : application du théorème de l'indice au calcul de l'anomalie chirale

On a donné précédemment un calcul perturbatif de l'anomalie chirale. Fujikawa a compris qu'on pouvait interpréter une anomalie comme une conséquence de la non-invariance de la mesure fermionique dans la formulation par intégration fonctionnelle de la théorie. On se place en signature $(- + \dots +)$ et on adopte la convention «physique» pour la définition des matrices de Dirac et donc de l'opérateur de Dirac : $\{\gamma_i, \gamma_j\} = 2g_{i, j}$ (Voir la remarque qui précède 2.1.5, Définition 10).

On considère l'action effective en présence d'un champ de jauge classique A obtenue après intégration fonctionnelle sur les champs fermioniques :

$$e^{iW[A]} = \int D\psi D\bar{\psi} e^{i \int d^n x \mathcal{L}_m[\psi, \bar{\psi}, D_\mu \psi, D_\mu \bar{\psi}]}$$

où \mathcal{L}_m est la partie de la densité lagrangienne de la théorie décrivant les champs fermioniques. Par exemple, pour un spineur de Dirac sans masse : $\mathcal{L}_m = -\bar{\psi}\not{D}\psi$.

Considérons la transformation :

$$\begin{aligned}\psi(x) &\longrightarrow \psi'(x) = U(x)\psi(x) \\ \bar{\psi}(x) &\longrightarrow \bar{\psi}'(x) = \bar{\psi}(x)\bar{U}(x)\end{aligned}$$

À l'aide de la définition $\bar{\psi} = \psi^\dagger i\gamma_0$, on trouve : $\bar{U}(x) = i\gamma_0 U^\dagger(x) i\gamma_0$. $U(x)$ est une matrice unitaire qui agit sur les indices de la représentation du groupe de jauge et sur les indices de l'algèbre de Clifford. On note \mathcal{U} et $\bar{\mathcal{U}}$ les opérateurs définis par : $\langle x|\mathcal{U}|y\rangle = U(x)\delta(x-y)$ et $\langle x|\bar{\mathcal{U}}|y\rangle = \bar{U}(x)\delta(x-y)$. Formellement, la mesure fermionique se transforme selon :

$$D\psi \longrightarrow D\psi' = \frac{1}{\text{Det}(\mathcal{U})} D\psi$$

et

$$D\bar{\psi} \longrightarrow D\bar{\psi}' = \frac{1}{\text{Det}(\bar{\mathcal{U}})} D\bar{\psi}$$

La situation est radicalement différente suivant que U est une transformation chirale ou non.

1. U est une transformation unitaire non chirale, i.e. ne faisant pas intervenir γ_5 : $U(x) = e^{i\epsilon^\alpha(x)t_\alpha}$ avec $t_\alpha^\dagger = t_\alpha$ et $[\gamma^\mu, t_\alpha] = 0$. γ_0 commutant avec $U(x)$, on obtient $\bar{U}(x) = U^{-1}(x)$: la mesure fermionique est invariante.
2. U est une transformation unitaire chirale : $U(x) = e^{i\epsilon^\alpha(x)t_\alpha\gamma_5}$ avec $t_\alpha^\dagger = t_\alpha$ et $[\gamma^\mu, t_\alpha] = 0$. Cette fois-ci, γ_0 anticommute avec $U(x)$ d'où $\bar{U}(x) = U(x)$. La mesure fermionique varie donc d'un facteur $(\text{Det}(\mathcal{U}))^{-2}$ qui en général n'a aucune raison de valoir 1. On définit l'anomalie $\mathcal{A}_\alpha(x)$ par : $(\text{Det}\mathcal{U})^{-2} = e^{i\int dx\epsilon^\alpha(x)\mathcal{A}_\alpha(x)}$. Autrement dit, \mathcal{A} mesure la variation locale infinitésimale de la mesure fermionique sous une transformation chirale infinitésimale.

Formellement,

$$(\text{Det}\mathcal{U})^{-2} = e^{-2\text{Tr}(\log(\mathcal{U}))}$$

et

$$\text{Tr}(\log(\mathcal{U})) = \int d^n x \langle x|\text{tr}\log\mathcal{U}|x\rangle = \int d^n x \delta(0)\text{tr}(t_\alpha\gamma_5)$$

d'où $\mathcal{A}_\alpha(x) = -2\delta(0)\text{tr}(t_\alpha\gamma_5)$ ce qui bien sûr n'est pas bien défini : $\text{tr}(t_\alpha\gamma_5) = 0$ et moralement $\delta(0) = \infty$, on a donc affaire à une forme indéterminée. Pour lever l'indétermination, il faut régulariser les choses, ce qui peut se faire en posant par exemple :

$$\text{Tr}(\log(\mathcal{U})) = \lim_{t \rightarrow 0} \int d^n x \text{tr}(\langle x|\epsilon^\alpha(x)\gamma_5 t_\alpha e^{-t(i\not{D})^2}|x\rangle)$$

Le choix du régulateur n'est pas sans importance : pour que le résultat final ait un sens physique, il faut que l'expression ci-dessus soit invariante de jauge, ce qui est bien le cas. Or, on a vu que :

$$\lim_{t \rightarrow 0} \int d^n x \text{tr}(\langle x|\gamma_5 e^{-t(i\not{D})^2}|x\rangle) = \text{ind}(\not{D})$$

et d'après le théorème de l'indice, on a : $\text{ind}(\not{D}) = \int_M \hat{A}(M) \text{ch}(E)$. Ainsi :

$$\mathcal{A}_\alpha(x) = -2\text{tr}(t_\alpha(\hat{A}(M)\text{ch}(E))_n)$$

où $(\alpha)_n$ signifie qu'on prend la composante de degré n de la forme différentielle α .

Considérons maintenant le cas où l'action classique est invariante sous une transformation axiale globale de groupe $G : U(x) = e^{i\epsilon^\alpha t_\alpha \gamma_5}$. Rappelons que le cas de la symétrie chirale pour un spineur de Dirac sans masse correspond à $G = U(1)$. Il résulte du fait que l'action classique $S = \int d^n x \mathcal{L}_m$ est invariante sous cette transformation que si on considère la transformation locale $U(x) = e^{i\epsilon^\alpha(x) t_\alpha \gamma_5}$, alors $\delta_{\epsilon^\alpha(x)} S$ est de la forme

$$\delta_{\epsilon^\alpha(x)} S = - \int d^n x J_{5,\alpha}^\mu(x) \partial_\mu \epsilon^\alpha(x) = \int d^n x \partial_\mu J_{5,\alpha}^\mu(x) \epsilon^\alpha(x)$$

$J_{5,\alpha}^\mu(x)$ est le courant associé à la symétrie de l'action et vérifie dans la théorie classique $\partial_\mu J_{5,\alpha}^\mu(x) = 0$ le long des solutions des équations du mouvement. On a déjà rencontré ce courant plus haut et on avait défini la présence d'une anomalie par la non-annulation de $\partial_\mu \langle J_{5,\alpha}^\mu(x) \rangle_A$ dans la théorie quantique. Il ne reste plus qu'à établir le lien avec l'anomalie $\mathcal{A}_\alpha(x)$ pour avoir la formule générale donnant $\partial_\mu \langle J_{5,\alpha}^\mu(x) \rangle_A$.

Pour cela, on rappelle que :

$$e^{iW[A]} = \int D\psi D\bar{\psi} e^{i \int d^n x \mathcal{L}_m[\psi, \bar{\psi}, D_\mu \psi, D_\mu \bar{\psi}]}$$

et on fait le changement de variable

$$\psi(x) \longrightarrow \psi'(x) = U(x)\psi(x)$$

$$\bar{\psi}(x) \longrightarrow \bar{\psi}'(x) = \bar{\psi}(x)\bar{U}(x)$$

dans l'intégrale, on en déduit :

$$\begin{aligned} e^{iW[A]} &= \int D\psi D\bar{\psi} e^{i \int d^n x \mathcal{L}_m[\psi, \bar{\psi}, D_\mu \psi, D_\mu \bar{\psi}]} \\ &= \int D\psi D\bar{\psi} e^{i \int d^n x \epsilon^\alpha(x) \mathcal{A}_\alpha(x)} e^{i \int d^n x \mathcal{L}_m[\psi, \bar{\psi}, D_\mu \psi, D_\mu \bar{\psi}] + i \int d^n x \partial_\mu J_{5,\alpha}^\mu(x) \epsilon^\alpha(x)} \\ &= \int D\psi D\bar{\psi} e^{i \int d^n x \mathcal{L}_m[\psi, \bar{\psi}, D_\mu \psi, D_\mu \bar{\psi}]} [1 + i \int d^n x \epsilon^\alpha(x) (\mathcal{A}_\alpha(x) + \partial_\mu J_{5,\alpha}^\mu(x)) + O(\epsilon^2)] \end{aligned}$$

d'où : $\partial_\mu \langle J_{5,\alpha}^\mu(x) \rangle_A = -\mathcal{A}_\alpha(x)$. On en déduit donc :

$$\partial_\mu \langle J_{5,\alpha}^\mu(x) \rangle_A = -\mathcal{A}_\alpha(x) = 2tr(t_\alpha(\hat{A}(TM)ch(E))_n)$$

Dans le cas où $\hat{A}(TM)$ est trivial, i.e. dans le cas où il est possible de négliger les effets de la géométrie intrinsèque de M , on obtient :

$$\partial_\mu \langle J_{5,\alpha}^\mu(x) \rangle_A = -\mathcal{A}_\alpha(x) = 2tr(t_\alpha(ch(E))_n) = 2 \frac{1}{r!(4i\pi)^r} \epsilon^{\mu_1 \dots \mu_{2r}} tr(t_\alpha F_{\mu_1, \mu_2} \dots F_{\mu_{2r-1}, \mu_{2r}})$$

Ce qui permet de retrouver les résultats obtenus de manière perturbative, d'établir leur validité au niveau non-perturbatif et de les généraliser en dimension paire $n = 2r$ quelconque.

Pour résumer, l'anomalie chirale apparaît comme une non-invariance de la mesure fermionique intervenant dans la formulation par intégration fonctionnelle de la théorie quantique des champs. Pour calculer cette anomalie, il faut adopter une procédure de régularisation qui conduit à voir l'anomalie comme la quantité locale apparaissant dans le théorème de l'indice.

4 Anomalie de jauge

4.1 Quantification des théories de jauge et anomalies

4.1.1 Quantification : de la symétrie de jauge à la symétrie BRST

Dans cette section 4.1, ce qu'on note A et F sont en fait A^{ph} et F^{ph} .

On rappelle que la donnée géométrique d'une théorie de jauge (non-abélienne) comprend un groupe de Lie G compact, une variété riemannienne M , un fibré principal P de groupe G sur M , une représentation linéaire \mathcal{R} de G sur un espace vectoriel de dimension finie, le fibré vectoriel E sur M associé à P et à \mathbb{R} . Une connexion A sur P est alors une 1-forme sur P à valeurs dans $Lie(G)$ et sa courbure F est une 2-forme sur P à valeurs dans $Lie(G)$. Une telle connexion induit une connexion sur E , notée $A^{\mathbb{R}}$, de courbure $F^{\mathbb{R}}$. La détermination dynamique de la théorie repose sur la donnée d'une densité lagrangienne $\mathcal{L}(A, \psi)$. Les variables dynamiques sont : le champ de jauge A , qui est une connexion sur P et le champ de matière ψ qui est une section de E (en pratique M est munie d'une structure spin et E est un produit tensoriel du fibré spinoriel de M par un autre fibré vectoriel : ψ est donc un champ fermionique). Le lagrangien de la théorie classique est une fonction de ces variables dynamiques. Ce qui précède permet de bien définir une théorie de jauge (non-abélienne) classique : les équations d'Euler-Lagrange déduites du lagrangien sont des équations aux dérivées partielles non-linéaires de type Yang-Mills et étudier cette théorie, c'est étudier les solutions de ces équations (ce qui est un problème mathématique en général plus que non-trivial même dans le cas le plus simple où il n'y a pas de champ de matière).

Il s'agit maintenant de quantifier la théorie classique. Pour simplifier, on prend $M = \mathbb{R}^4$ pour le reste de cette section. Formellement, la procédure de quantification par intégration fonctionnelle consiste, pour calculer une fonction de corrélation, à remplacer la trajectoire classique donnée par les équations d'Euler-Lagrange de la théorie classique par une somme pondérée sur toute les trajectoires possibles, le poids associé à une trajectoire étant proportionnel à $e^{i \int \mathcal{L} d^4x}$. Dans le cas qui nous intéresse, celui des théories de jauge, on devrait donc intégrer sur tous les champs de jauge et de matière possibles. Mais il faut faire attention au fait que deux configurations de champs de jauge qui diffèrent d'une transformation de jauge sont physiquement identiques : l'intégration portant sur les champs de jauge ne doit pas se faire sur l'espace \mathcal{U} de toutes les connexions mais sur le quotient de \mathcal{U} par le groupe \mathcal{G} des transformations de jauge. Explicitement, il faut donc mettre dans l'intégrale fonctionnelle un terme fixant la jauge. On se place dans la jauge dite axiale : $A_3^\alpha = 0$. Les variables canoniques sont alors A_1^α et A_2^α (A_0^α est une fonction de ces deux champs via les équations du mouvement). Si O_1, \dots, O_N sont des observables de la théorie, on écrit donc :

$$\langle 0|TO_1\dots O_N|0 \rangle = C \int DA_\mu^\alpha D\psi_l \left(\prod_{x,\alpha} \delta(A_3^\alpha) \right) O_1 \dots O_N e^{i \int \mathcal{L}[A_\mu, \psi] d^4x}$$

où C est une constante de normalisation. Cette expression est-elle invariante de jauge ? C'est le cas du facteur comprenant le lagrangien par hypothèse : la théorie classique est invariante de jauge. Il n'y a pas non plus de problème pour la « mesure » DA_μ^α (si on fait une transformation de jauge, le jacobien associé est égal à 1 : c'est une conséquence de l'antisymétrie des constantes de structure). En revanche, il faut faire l'hypothèse qu'il en est de même pour la mesure fermionique $D\psi_l$. Si ce n'est pas le cas, l'intégrale fonctionnelle n'est pas invariante de jauge et il faut donc s'arrêter : les conséquences qu'on pourrait en dériver n'ont pas de signification physique (si on persiste à poursuivre le développement de la théorie, on trouve que les modes longitudinaux du champ de jauge ne se découplent pas des modes transversaux : on perd l'unitarité des matrices de transition, des vecteurs de l'espace des états de la théorie de norme négative apparaissent...)

Pour développer une théorie cohérente, on suppose donc que l'intégrale fonctionnelle ci-dessus est invariante de jauge. Sous la forme précédente, le rôle privilégié que fait jouer le facteur fixant la jauge à certaines coordonnées d'espace-temps rend non explicite l'invariance sous le groupe des transformations de Lorentz. Heureusement, il se trouve que la forme précise du facteur fixant la jauge n'a pas grande importance. Plus précisément, la formule précédente peut s'écrire :

$$\langle 0|TO_1\dots O_N|0 \rangle = C \int DA_\mu^\alpha \left(\prod_{x,\alpha} \delta(A_3^\alpha) \right) G[A_\mu^\alpha]$$

où : $G[A_\mu^\alpha] = \int D\psi_l O_1 \dots O_N e^{i \int \mathcal{L}[A_\mu, \psi] d^4x}$ est invariant de jauge. On a donc affaire à une intégrale de la forme

$$I_G = C \int DA_\mu^\alpha B[f(A_\mu)] Det \mathcal{F}[A_\mu^\alpha] G(A_\mu)$$

où $B[f(A_\mu)]$ est une fonction fixant la jauge et $\mathcal{F}_{\alpha,\beta}(x,y) = \frac{\delta f_\alpha(A^\epsilon(x))}{\delta \epsilon^\beta(y)}|_{\epsilon=0}$ où $A^\epsilon(x)$ est la transformée de jauge de paramètre $\epsilon(x)$ de $A(x)$. Soit \mathcal{U} l'espace des connexions : \mathcal{U} est partitionné en les classes de connexions équivalentes de jauge. L'équation $f(A_\mu) = 0$ définit une hypersurface H dans \mathcal{U} intersectant une unique fois chacune de ces classes. Le facteur $Det \mathcal{F}$ permet de définir la mesure sur H induite par la mesure DA_μ sur \mathcal{U} . Dans la formule précédente, on a $B[f] = (\prod_{x,\alpha} \delta(f^\alpha(x)))$, $f^\alpha(A) = A_3^\alpha$, $f_\alpha(A^\epsilon(x)) = \partial_3 \epsilon$ et $\mathcal{F}_{\alpha,\beta}(x,y) = \delta_{\alpha,\beta} \frac{\partial}{\partial x^3} \delta^4(x-y)$. Le résultat suivant, dû à Faddeev et Popov, montre que ces choix de B et de f ne sont pas importants :

Proposition 30. *Si $G[A_\mu]$ est une fonctionnelle invariante de jauge, alors I_G ne dépend pas de f et ne dépend de B que par un facteur global.*

Démonstration. On note A_μ^g la transformée de jauge de A_μ sous un g : $x \rightarrow g(x) \in G$. I_G , DA_μ et $G[A_\mu]$ étant invariants de jauge, on a :

$$I_G = C \int DA_\mu B[f(A_\mu^g)] Det(\mathcal{F}(A_\mu^g)) G[A_\mu]$$

pour toute transformation de jauge g . On va maintenant moyenner sur les g possibles : soit ρ une fonction telle que $\int \rho(g) Dg = C_0 < +\infty$. Alors :

$$C_0 I_G = \int DA_\mu G[A_\mu] H[A_\mu]$$

où :

$$H[A_\mu^g] = \int Dg \rho(g) B[f(A_\mu^g)] Det(\mathcal{F}[A_\mu^g]),$$

$$\mathcal{F}[A_\mu^g](x,y) = \frac{\delta f[A^{\tilde{g}(g,\epsilon)}]}{\delta \epsilon(y)}|_{\epsilon=0} = \int d^4z \frac{\delta f[A^{\tilde{g}(x)}]}{\delta \tilde{g}(z)}|_{\tilde{g}=g} \frac{\delta \tilde{g}(g,\epsilon;z)}{\delta \epsilon(y)}|_{\epsilon=0}$$

Soit $\mathcal{G}[g](z,x) = \frac{\delta \tilde{g}(g,\epsilon;z)}{\delta \epsilon(y)}|_{\epsilon=0}$ et choisissons $\rho[g] = \frac{1}{Det \mathcal{G}}$. Alors :

$$H[A_\mu] = \int Dg B[f(A_\mu^g)] Det(\delta f(A_\mu^g)/\delta g) = \int Df B[f] = C_B$$

qui est indépendant de f □

Ainsi, pour toute fonction f fixant la jauge et pour toute fonction B , on a :

$$\langle 0|TO_1\dots O_N|0 \rangle = \frac{C}{C_B} \int DA_\mu^\alpha B[f(A_\mu)] Det(\mathcal{F}[A_\mu]) \int D\psi_l O_1 \dots O_N e^{i \int \mathcal{L}[A_\mu, \psi] d^4x}$$

Il est donc possible de faire un choix de terme fixant la jauge manifestement invariant de Lorentz : $f_\alpha = \partial_\mu A_\alpha^\mu$, $B[f] = e^{-\frac{i}{2\xi} \int d^4x f_\alpha(x) f_\alpha(x)}$ où $\xi \in \mathbb{R}$. Introduire ce facteur $B[f]$ dans l'intégrale fonctionnelle a le même effet que d'ajouter un terme supplémentaire à la densité lagrangienne : $\mathcal{L}_{g.f} = -\frac{1}{2\xi} f_\alpha f_\alpha$. Calculons $Det\mathcal{F}$ associé à ce choix. On a :

$$f_\alpha(A_\mu^\epsilon) = \partial_\mu(\partial^\mu \epsilon_\alpha + C_{\beta,\gamma}^\alpha A_\mu^\gamma \epsilon^\beta)$$

d'où :

$$\mathcal{F}_{\alpha,\beta}(x,y) = \frac{\delta f_\alpha((A_\mu^\epsilon)(x))}{\delta \epsilon^\beta(y)} = \frac{\partial}{\partial x^\mu} \left(\frac{\partial}{\partial x_\mu} \delta_{\alpha,\beta} + C_{\beta,\gamma}^\alpha A_\mu^\gamma(x) \right) \delta^4(x-y)$$

De façon générale, on peut écrire :

$$Det\mathcal{F} = \int D\omega D\omega^* e^{i \int d^4x d^4y \omega_\alpha^* \mathcal{F}_{\alpha,\beta}(x,y) \omega_\beta(y)}$$

où ω_α , ω_α^* sont des champs scalaires réels indépendants qui anticommulent, en particulier, ils violent la relation spin-statistique. Le champ ω_α s'appelle un champ fantôme et le champ ω_α^* un champ anti-fantôme. Ici, on a donc :

$$Det\mathcal{F} = \int D\omega D\omega^* e^{i \int \mathcal{L}_g d^4x}$$

avec :

$$\mathcal{L}_g = -\partial_\mu \omega_\alpha^* \partial^\mu \omega_\alpha + C_{\beta,\gamma}^\alpha \partial_\mu \omega_\alpha^* A_\mu^\gamma \omega_\beta.$$

On peut maintenant considérer la théorie de densité lagrangienne

$$\mathcal{L}_{mod} = \mathcal{L} + \mathcal{L}_{g.f} + \mathcal{L}_g$$

et la quantifier perturbativement : bien sûr, il y a maintenant des champs fantômes dans les diagrammes de Feynman. Dans le cas d'une théorie abélienne comme l'électrodynamique, on a $C_{\beta,\gamma}^\alpha = 0$ et les champs fantômes ne sont couplés à aucun autre champ, ils ne donnent donc qu'une constante globale dans l'intégrale fonctionnelle. Mais en général, les champs fantômes sont couplés au champ de jauge via le terme : $C_{\beta,\gamma}^\alpha \partial_\mu \omega_\alpha^* A_\mu^\gamma \omega_\beta$. L'espace de Hilbert obtenu (ou plutôt qu'on obtiendrait si on savait faire une quantification non-perturbative correcte) possède des états contenant des champs fantômes. Mais ces états n'ont rien de physique, les champs fantômes ne s'observent pas expérimentalement, ce ne sont que des artefacts permettant de faire des calculs dans une jauge fixée. Cela signifie que l'espace de Hilbert obtenu est trop gros : il contient des états non-physiques. Il s'agit alors de trouver un moyen de retrouver l'espace de Hilbert des états physiques à partir de cet espace trop gros. On appelle nombre fantôme d'un état le nombre algébrique de particules fantômes, i.e. on compte +1 pour une particule fantôme et -1 pour une particule anti-fantôme. Il résulte de la structure du terme de couplage que le nombre fantôme d'un état est conservé par la dynamique.

On introduit un champ auxiliaire h_α permettant de remplacer le terme $-\frac{1}{2\xi} f_\alpha f_\alpha$ par $\frac{\xi}{2} h_\alpha h_\alpha + h_\alpha f_\alpha$ et par une intégration fonctionnelle sur Dh^α . On s'est donc ramené à une théorie contenant les champs A_μ^α , ψ , ω_α , ω_α^* et h_α de densité lagrangienne :

$$\mathcal{L}_n = \mathcal{L} + \frac{\xi}{2} h_\alpha h_\alpha + h_\alpha f_\alpha + \omega_\alpha^* \rho_\alpha$$

où :

$$\rho_\alpha(x) = \partial_\mu \partial^\mu \omega_\alpha(x) + C_{\beta,\gamma}^\alpha \partial_\mu (A_\beta^\mu(x) \omega_\gamma(x)) = \int d^4y \mathcal{F}_{\alpha,\beta}(x,y) \omega_\beta(y)$$

Cette densité lagrangienne n'est pas invariant de jauge puisque par construction on a fixé la jauge. Néanmoins, il reste une trace de la symétrie de jauge du lagrangien initial, c'est la symétrie BRST. On note s l'opérateur associé. L'opérateur s agit sur les champs de matières et les champs de jauge comme une transformation de jauge infinitésimale mais avec le paramètre ϵ_α remplacé par ω_α . Autrement dit : $sA_\mu = \partial_\mu\omega - i[A_\mu, \omega]$, $s\psi = i\omega\psi$, $s\psi^\dagger = i\psi^\dagger\omega$. On impose à s de vérifier : $s(FG) = (sF)G \pm F(sG)$ avec $+$ si F est bosonique et $-$ si F est fermionique. L'action de s sur les champs fantômes est définie par : $s\omega = i\omega\omega$, $s\omega^* = -h$, $sh = 0$ de telle façon que $s^2A_\mu = 0$, $s^2\psi = 0$, $s^2\omega = 0$ et $s^2\omega^* = 0$, i.e. $s^2 = 0$. Puisque \mathcal{L} est invariant de jauge, \mathcal{L} est invariant BRST, i.e. $s\mathcal{L} = 0$. \mathcal{F} étant défini comme une variation de jauge de f , on a $\rho_\alpha = sf_\alpha$ d'où $s\rho = 0$. Ainsi :

$$s\mathcal{L}_n = s\left(\frac{\xi}{2}h_\alpha h_\alpha + h_\alpha f_\alpha + \omega_\alpha^* \rho_\alpha\right) = h_\alpha(sf_\alpha) + (s\omega_\alpha^*)\rho_\alpha = h_\alpha\rho_\alpha - h_\alpha\rho_\alpha = 0.$$

En fait, on a : $\mathcal{L}_n = \mathcal{L} + s\Psi$ où $\Psi = -\omega_\alpha^* f_\alpha - \frac{\xi}{2}\omega_\alpha^* h_\alpha$ ce qui rend manifeste l'invariance de \mathcal{L}_n sous la symétrie BRST.

L'opérateur s fait augmenter le nombre fantôme de $+1$ et vérifie $s^2 = 0$. Si F est une fonctionnelle en les champs $A_\mu, \psi, \omega, \omega^*, h$ (F est une expression polynomiale en les champs et leurs dérivées ou au pire une série formelle), on dit que F est BRST-fermée si $sF = 0$, autrement dit si BRST-invariante, et BRST-exacte s'il existe G tel que $F = dG$. On appelle cohomologie BRST de degré k l'espace des fonctionnelles de la forme précédente de nombre fantôme k fermées, modulo l'espace des fonctionnelles de ce type exactes. \mathcal{L} étant un invariant de jauge, $s\mathcal{L} = 0$, i.e. \mathcal{L} est BRST-fermé. De plus, \mathcal{L} est de nombre fantôme 0. Le terme fixant la jauge et le terme contenant les champs fantômes donnent une contribution $s\Psi$ avec Ψ de nombre fantôme -1 . Ainsi, \mathcal{L}_n est dans la même classe de cohomologie BRST que \mathcal{L} . Soit Q l'opérateur de charge BRST fermionique, défini par $s\Phi(x) = i[Q, \Phi(x)]_\mp$ où $[\cdot, \cdot]_-$ est le commutateur si Φ est un champ bosonique et $[\cdot, \cdot]_+$ est l'anticommutateur si Φ est un champ fermionique. Q est un opérateur qui fait augmenter de $+1$ le nombre fantôme et qui vérifie $Q^2 = 0$. Q agit sur l'espace de Hilbert issu de la quantification, celui-ci contenant des états non-physiques.

On peut alors montrer qu'on retrouve l'espace des états physiques en prenant l'espace des classes de cohomologie de l'opérateur Q de nombre fantôme 0, i.e. l'espace des états annulés par Q modulo les états dans l'image de Q . Pour plus de détails, voir un cours de théorie quantique des champs.

Le résultat suivant sera utilisé pour la détermination de l'anomalie de jauge :

Théorème 10. *Une classe de cohomologie BRST de nombre fantôme 0 possède un représentant qui est une fonctionnelle invariante de jauge en les champs A_μ et ψ .*

Démonstration. Soit I un représentant d'une classe de cohomologie BRST de nombre fantôme 0, en particulier $sI = 0$. A priori, I est une fonctionnelle en les champs $A_\mu, \psi, \omega, \omega^*$ et h . On peut écrire : $I = \sum_{N=0}^{+\infty} I_N$ où I_N contient tous les termes avec un nombre de champs ω^* ou h exactement égal à N . Puisque $s\omega^* = -h$ et $sh = 0$, s préserve chacun des I_N d'où $sI_N = 0$ pour tout N . Soit $t = \omega_\alpha^* \delta / \delta h_\alpha$. Puisque $s + h_\beta \delta / \delta \omega_\beta^*$ ne contient pas de ω^* ou de h , on a : $st + ts = -(\omega_\alpha^* \delta / \delta \omega_\alpha^* + h_\alpha \delta / \delta h_\alpha)$. Appelons \tilde{N} l'opérateur $-(st + ts)$, alors : $\tilde{N}I_N = NI_N$ d'où, si $N \neq 0$, $I_N = s(-\frac{1}{N}tI_N)$ ce qui montre que I_0 est dans la même classe de cohomologie BRST que I . Or I_0 ne contient pas de ω^* , ni de h et puisqu'il est de nombre fantôme 0, il ne contient pas non plus de ω . □

4.1.2 Anomalie de jauge

On considère une théorie de jauge et on va étudier la possibilité d'une anomalie associée à la symétrie de jauge. On appelle A le champ de jauge, ψ le champ fermionique, $S[A, \psi, \bar{\psi}]$ l'action classique et $\Gamma[A, \psi, \bar{\psi}]$ l'action quantique effective. On définit l'anomalie de jauge $\mathcal{A}(x)$ à partir de la variation de la mesure fermionique sous une transformation de jauge infinitésimale de paramètre $\epsilon(x)$:

$$D\psi D\bar{\psi} \longrightarrow D\psi D\bar{\psi} e^{i \int d^n x \epsilon(x) \mathcal{A}(x)}$$

On s'attend à interpréter l'anomalie comme une non-invariance de l'action quantique effective.

En fait, il faut faire attention : par définition d'une théorie de jauge, S est invariant sous une transformation de jauge mais ce n'est pas le cas de Γ même en absence d'anomalie. En effet, comme expliqué dans la section précédente, il ne faut pas prendre l'action classique S pour définir l'intégrale fonctionnelle sur $DA D\psi D\bar{\psi}$ mais l'action classique plus un terme fixant la jauge qui brise l'invariance de jauge. Pour retrouver l'interprétation de l'anomalie comme la non-invariance d'une action, il faut considérer une action intermédiaire entre S et Γ : on définit $\tilde{\Gamma}[A]$ comme étant l'action effective obtenue après intégration sur les champs fermioniques :

$$e^{i\tilde{\Gamma}[A]} = \int D\psi D\bar{\psi} e^{iS_{mat}[A, \psi, \bar{\psi}]}$$

où S_{mat} est la partie de l'action classique décrivant les champs de matières. Pour poursuivre la quantification et intégrer sur DA , il faut considérer l'action $S_{jauge}[A] + \tilde{\Gamma}[A]$ mais pour développer la théorie de la section précédente, il est essentielle que $\tilde{\Gamma}$ soit une fonction de A invariante de jauge. La variation de $\tilde{\Gamma}[A]$ sous une transformation infinitésimale s'écrit :

$$\delta_\epsilon \tilde{\Gamma}[A] = \int d^n x \epsilon(x) \mathcal{A}(x)$$

ce qui montre bien que l'invariance de jauge de $\tilde{\Gamma}[A]$ est équivalent à l'annulation de l'anomalie $\mathcal{A}(x)$.

L'existence d'une anomalie de jauge empêche de mener à bien de façon cohérente le programme de quantification décrit dans la section précédente, c'est donc un critère de cohérence interne qui doit être vérifié pour qu'une théorie soit physiquement acceptable.

Comme l'anomalie chirale, l'anomalie de jauge est associée à la non-conservation d'un courant. On définit un courant en posant $J_\alpha^\mu = \delta S_{mat} / \delta A_\mu^\alpha$ et on vérifie facilement que

$$\frac{\delta \tilde{\Gamma}}{\delta A_\mu^\alpha(x)} = \langle J_\alpha^\mu(x) \rangle_A$$

où $\langle . \rangle_A$ est l'amplitude vide-vide calculée en présence du champ classique A et que

$$D_\mu \frac{\delta \tilde{\Gamma}}{\delta A_\mu^\alpha(x)} = -\mathcal{A}_\alpha(x)$$

d'où :

$$(D_\mu \langle J_\alpha^\mu(x) \rangle_A)_\alpha = -\mathcal{A}_\alpha(x)$$

4.1.3 Conditions de Wess-Zumino

On a vu que :

$$\delta_\epsilon \tilde{\Gamma}[A] = \int d^n x \epsilon^\alpha(x) \mathcal{A}_\alpha(x)$$

ce qu'on peut réécrire formellement :

$$\mathcal{A}_\alpha(x) = -(D_\mu \frac{\delta}{\delta A_\mu(x)})_\alpha \tilde{\Gamma}[A]$$

L'anomalie apparaissant comme la dérivée d'une fonctionnelle, on se doute que sa forme n'est pas quelconque. Soit

$$\mathcal{G}_\alpha(x) = -(D_\mu \frac{\delta}{\delta A_\mu(x)})_\alpha = -\frac{\partial}{\partial x_\mu} \frac{\delta}{\delta A_\mu^\alpha(x)} - C_{\alpha\delta\gamma} A_\mu^\delta(x) \frac{\delta}{\delta A_\mu^\gamma(x)}$$

Autrement dit, $\mathcal{G}_\alpha(x)$ est un opérateur et : $\mathcal{A}_\alpha(x) = \mathcal{G}_\alpha(x) \tilde{\Gamma}[A]$. Un simple calcul utilisant le fait que les constantes de structure $C_{\alpha\beta\gamma}$ vérifient l'identité de Jacobi montre que :

$$[\mathcal{G}_\alpha(x), \mathcal{G}_\beta(y)] = \delta(x-y) C_{\alpha\beta\gamma} \mathcal{G}_\gamma(x)$$

ce qui appliqué à $\tilde{\Gamma}[A]$ donne :

$$\mathcal{G}_\alpha(x) \mathcal{A}_\beta(y) - \mathcal{G}_\beta(y) \mathcal{A}_\alpha(x) = \delta(x-y) C_{\alpha\beta\gamma} \mathcal{A}_\gamma(x)$$

Cette formule s'appelle la contrainte de Wess-Zumino.

4.2 Calcul de l'anomalie de jauge

4.2.1 Théorie des formes de Chern-Simons

Pour commencer, quelques calculs explicites :

Rappel : $A = A_\mu dx^\mu$, $F = dA + A^2$.

Lemme 13.

$$tr(F^2) = d(tr(AdA + \frac{2}{3}A^3))$$

Démonstration. Pour une fois, on va faire le calcul de façon très détaillée :

$$(*) = d(tr(AdA + \frac{2}{3}A^3)) = tr(dAdA + \frac{2}{3}dAA^2 - \frac{2}{3}AdAA + \frac{2}{3}A^2dA)$$

or $dA = F - A^2$ d'où :

$$(*) = tr(F^2 - A^2F - FA^2 + A^4 + \frac{2}{3}[FA^2 - A^4 - AFA + A^4 + A^2F - A^4])$$

Or :

$$\begin{aligned} tr(A^4) &= (tr(A_\kappa A_\lambda A_\mu A_\nu)) dx^\kappa \wedge dx^\lambda \wedge dx^\mu \wedge dx^\nu = -(tr(A_\kappa A_\lambda A_\mu A_\nu)) dx^\lambda \wedge dx^\mu \wedge dx^\nu \wedge dx^\kappa \\ &= -(tr(A_\lambda A_\mu A_\nu A_\kappa)) dx^\lambda \wedge dx^\mu \wedge dx^\nu \wedge dx^\kappa = -tr(A^4) \end{aligned}$$

d'où $tr(A^4) = 0$. De même, on montre que $tr(FA^2) = -tr(AFA) = tr(A^2F)$. Ainsi :

$$(*) = tr(F^2 - A^2F - FA^2 + \frac{2}{3}(FA^2 - AFA + A^2F)) = tr(F^2) - \frac{1}{3}tr(FA^2 + 2AFA + A^2F) = tr(F^2)$$

□

De même, on montre :

Lemme 14.

$$\text{tr}(F^3) = d(\text{tr}(A(dA)^2 + \frac{3}{2}A^3dA + \frac{3}{5}A^5))$$

On va voir que ces formules ne sont pas le fait du hasard. Soit $P_j(F) = \text{tr}(F^j)$ le polynôme invariant intervenant dans la définition du caractère de Chern. On a vu que $P_j(F)$ est une $2j$ -forme fermée donc localement exacte. Soit U_i un ouvert de carte contractile et A_i, F_i les restrictions de la connexion et de la courbure à U_i . Sur chacun des U_i , on peut donc écrire : $P_j(F) = dQ_{2j-1}(A_i, F_i)$. Q_{j-1} , qui n'est définie que localement, s'appelle une forme de Chern-Simons : c'est une $(2j-1)$ -forme. Soit \tilde{P}_j la polarisation de P_j (on rappelle que $\tilde{P}_j(X_1, \dots, X_j)$ est $1/j!$ fois le coefficient de $t_1 \dots t_j$ dans $P_j(t_1 X_1 + \dots + t_j X_j)$). On a vu dans la partie sur les classes caractéristiques que : $Q_{2j-1}(A, F) = j \int_0^1 \tilde{P}_j(A, F_t, \dots, F_t)$ où $F_t = tdA + t^2 A^2$. Ici, $\tilde{P}(X_1, \dots, X_j) = \text{tr}(X_1 \dots X_j)$, d'où :

$$\begin{aligned} Q_{2j-1}(A, F) &= j \int_0^1 \text{tr}(AF_t^{j-1})dt = j \int_0^1 \text{tr}(A(tdA + t^2 A^2)^{j-1})dt \\ &= j \int_0^1 t^{j-1} \text{tr}(A(F + (t-1)A^2)^{j-1})dt \end{aligned}$$

Remarquons que Q_{2j-1} n'est défini qu'à l'addition d'une forme exacte près. Dans la suite, ce qu'on appelle Q_{2j-1} est l'expression donnée ci-dessus. Au moyen de cette formule, on retrouve facilement les résultats des deux lemmes précédents :

$$Q_3 = 2 \int_0^1 \text{tr}(A(tdA + t^2 A^2))dt = \text{tr}(AdA + \frac{2}{3}A^3) = \text{tr}(AF - \frac{1}{3}A^3)$$

$$Q_5 = 3 \int_0^1 \text{tr}(A(tdA + t^2 A^2)^2)dt = \text{tr}(AdAdA + \frac{3}{2}A^3dA + \frac{3}{5}A^5) = \text{tr}(AF^2 - \frac{1}{2}A^3F + \frac{1}{10}A^5)$$

Plus généralement, on va écrire une formule reliant $S(A_1, F_1)$ et $S(A_0, F_0)$ où $S(A, F)$ est un polynôme en A et en F , $A_t = A_0 + t(A_1 - A_0)$, $F_t = dA_t + A_t^2$. Soit l_t l'opérateur agissant sur les polynômes en A_t et en F_t défini formellement par $l_t A_t = 0$, $l_t F_t = dt(A_1 - A_0)$ et par la condition que l_t soit une antidérivation vis à vis du produit extérieur des formes différentielles.

Proposition 31. *Soit $S(A, F)$ un polynôme en A et en F , alors :*

$$(dl_t + l_t d)S(A_t, F_t) = dt \frac{\partial}{\partial t} S(A_t, F_t)$$

d'où la formule d'homotopie :

$$S(A_1, F_1) - S(A_0, F_0) = (dk_{01} + k_{01}d)S(A_t, F_t)$$

où l'opérateur d'homotopie k_{01} est défini par : $k_{01}S(A_t, F_t) = \int_0^1 l_t S(A_t, F_t)dt$.

Démonstration. On a :

$$(dl_t + l_t d)A_t = l_t dA_t = l_t(F_t - A_t^2) = l_t F_t = dt \frac{\partial A_t}{\partial t}$$

de même, et grâce à l'identité de Bianchi $dF_t + [A_t, F_t] = 0$:

$$(dl_t + l_t d)F_t = d(dt(A_1 - A_0)) + l_t(-A_t F_t + F_t A_t)$$

$$= dt[d(A_1 - A_0) + A_t(A_1 - A_0) + (A_1 - A_0)A_t]$$

or

$$\begin{aligned} \frac{\partial F_t}{\partial t} &= d(A_1 - A_0) + (A_1 - A_0)A_0 + A_0(A_1 - A_0) + 2t(A_1 - A_0)^2 \\ &= d(A_1 - A_0) + (A_1 - A_0)A_t + A_t(A_1 - A_0) \end{aligned}$$

On a donc obtenu le résultat désiré dans les cas $S(A, F) = A$ et $S(A, F) = F$. On en déduit le résultat général en remarquant que les opérateurs de chacun des deux membres sont des dérivations. \square

Remarque On peut retrouver les formules donnant les classes de Chern-Simons à partir de ce résultat. Il suffit en effet de prendre $S(A, F) = \text{tr}(F^j)$, $A_1 = A$, $A_0 = 0$ et d'utiliser le fait que $\text{tr}(F^j)$ est fermée pour avoir : $\text{tr}(F^j) = d(k_{01}\text{tr}(F^j))$. Ainsi, $Q_{2j-1}(A, F) = k_{01}\text{tr}(F^j) = \int_0^1 l_t \text{tr}(F_t^j) dt = j \int_0^1 \text{tr}(AF_t^{j-1}) dt$.

Par construction, les classes caractéristiques sont invariantes de jauge : ce n'est pas le cas des formes de Chern-Simons. Soit g une transformation de jauge, alors : $A \rightarrow A^g = g^{-1}(A + d)g$, $F \rightarrow F^g = g^{-1}Fg$. On pose $A_t^g = t(g^{-1}Ag + g^{-1}dg)$, $F_t = t dA + t^2 A^2$ et $F_t^g = dA_t^g + (A_t^g)^2 = g^{-1}F_t g$. En particulier, on a : $A_0^g = g^{-1}dg$, $A_1^g = A^g$, $F_0^g = 0$ et $F_1^g = F^g$. D'après la formule d'homotopie et la remarque précédente :

$$Q_{2j-1}(A^g, F^g) - Q_{2j-1}(g^{-1}dg, 0) = d(k_{01}Q_{2j-1}(A_t^g, F_t^g)) + k_{01}dQ_{2j-1}(A_t^g, F_t^g) = d\alpha_{2j-2} + Q_{2j-1}(A, F)$$

où : $\alpha_{2j-2} = k_{01}Q_{2j-1}(A_t + dg.g^{-1}, F_t)$ est une $(2j - 2)$ -forme. On en déduit finalement le comportement de Q_{2j-1} sous une transformation de jauge :

$$Q_{2j-1}(A^g, F^g) - Q_{2j-1}(A, F) = Q_{2j-1}(g^{-1}dg) + d\alpha_{2j-2}$$

$Q_{2j-1}(g^{-1}dg)$ peut se calculer :

$$\begin{aligned} Q_{2j-1}(g^{-1}dg) &= \int_0^1 \text{tr}(g^{-1}dg[(t^2 - t)(g^{-1}dg)^2]^{j-1}) dt = \text{tr}((g^{-1}dg)^{2j-1}) \int_0^1 (t^2 - t)^{j-1} dt \\ &= \text{tr}((g^{-1}dg)^{2j-1}) \frac{(-1)^{j-1} ((j-1)!)^2}{(2j-1)!} \end{aligned}$$

Grâce à la formule exprimant k_{01} , on peut calculer les α_{2j-2} . Par exemple :

$$\alpha_2 = \int_0^1 l_t \text{tr}[(A_t + dg.g^{-1})F_t - \frac{1}{3}(A_t + dg.g^{-1})^3] = -\text{tr}(dg.g^{-1}A) dt$$

Considérons le cas particulier d'une transformation de jauge infinitésimale engendrée par v , i.e. $g(x) = \exp(v(x))$ et on fait les calculs au premier ordre en v . On note par un δ la variation infinitésimale sous cette transformation. On a : $\delta A = dv + [A, v]$, $\delta F = [F, v]$. On a $d\delta Q_{2j-1} = \delta dQ_{2j-1} = 0$ car $dQ_{2j-1} = \text{tr}(F^j)$ qui est invariant de jauge. δQ_{2j-1} est donc localement exacte, ce qui permet de définir $Q_{2j-2}^1(v, A, F)$ localement par la condition : $\delta Q_{2j-1} = dQ_{2j-2}^1$. Si on spécialise les formules établies précédemment pour une transformation de jauge générale au cas d'une transformation infinitésimale, on voit qu'on peut choisir :

$$Q_{2j-2}^1 = j(j-1) \int_0^1 (1-t) \text{tr}(vd(AF_t^{j-2})) dt$$

Q_{2j-2}^1 n'est défini qu'à l'addition d'une forme exacte et d'une forme invariante de jauge près. Dans la suite, ce qu'on appelle Q_{2j-2}^1 est l'expression ci-dessus. Par exemple :

$$Q_2^1 = \text{tr}(vdA)$$

et

$$Q_4^1 = \text{tr}(vd(AdA + \frac{1}{2}A^3))$$

Résumons ce qui précède : la théorie des classes caractéristiques revient à introduire des polynômes en la courbure P_j , on obtient des $2j$ -formes définies globalement sur M , fermées : $dP_j = 0$, et invariantes de jauge : $\delta P_j = 0$. La condition de fermeture permet de définir localement une $(2j-1)$ -forme Q_{2j-1} par $P_j = dQ_{2j-1}$. Les Q_{2j-1} s'appellent les formes de Chern-Simons (on parle parfois de classes caractéristiques secondaires). L'invariance de jauge des P_j se traduit par l'existence locale d'une $(2j-2)$ -forme Q_{2j-2}^1 vérifiant $\delta Q_{2j-1} = dQ_{2j-2}^1$. On est donc passé d'une $2j$ -forme P_j à une $(2j-1)$ -forme Q_{2j-1} puis à une $(2j-2)$ -forme Q_{2j-2}^1 : les relations existantes entre ces formes différentielles s'appellent les équations de descente.

Remarque(non-nécessaire pour la suite). Les formes de Chern-Simons apparaissent souvent dans les situations à trois dimensions :

1. En mathématique : soit (M, g) une variété riemannienne compacte orientée de dimension trois. On s'intéresse au fibré P des repères orthonormés orientés sur M et à la connexion A sur P déduite de la connexion de Levi-Civita sur TM . La quantité

$$Q_3 = \text{tr}(AdA + \frac{2}{3}A^3)$$

n'est définie a priori que localement du fait de la non-invariance de jauge. Mais on peut montrer que M est parallélisable, i.e. TM et donc P sont trivialisables. Un choix d'une base s de sections permet alors de donner un sens à

$$\Phi(s) = \frac{1}{16\pi^2} \int_M Q_3.$$

On peut montrer qu'un autre choix de trivialisations t donne un $\Phi(t)$ qui ne diffère de $\Phi(s)$ que par un entier. Ceci permet de définir un invariant $J(M, g) \in \mathbb{R}/\mathbb{Z}$ de (M, g) . On peut alors démontrer que $J(M, g)$ est en fait un invariant de la classe conforme de g et que g est un point critique de la fonctionnelle $g \rightarrow J(M, g)$ si et seulement si g est localement conformément plate. Dans ce cadre, la forme de Chern-Simons apparaît comme un outil de compréhension de la géométrie riemannienne de dimension trois. Pour plus de détails, voir un appendice du livre de Chern [7].

2. En physique : on considère une théorie de jauge de groupe $U(1)$ sur une variété de base de dimension trois de densité lagrangienne : $\mathcal{L} = -\frac{1}{4}F_{\mu,\nu}F^{\mu,\nu} + \frac{1}{4}mA \wedge F$. Autrement dit, on a ajouté un multiple de Q_3 à la densité lagrangienne standard d'une théorie de jauge $U(1)$. Il résulte des propriétés de Q_3 que l'action est bien invariante sous une transformation de jauge. Les équations du mouvement (classique) conduisent à : $(\partial^\mu \partial_\mu + m^2)(*F) = 0$, i.e. $*F$ est un champ vectoriel massif de masse m . En trois dimensions, on peut donc avoir un terme de masse invariant de jauge donné par la 3-forme de Chern-Simons. C'est ce qui se produit de manière effective en électrodynamique quantique à trois dimensions : un calcul à une boucle fait apparaître le terme de Chern-Simons dans l'action effective décrivant la propagation du photon.

3. À l'interface mathématiques/physique : soit M une variété orientée de dimension trois et G un groupe de Lie compact simple. Witten a eu l'idée de ne pas considérer la 3-forme de Chern-Simons comme un terme supplémentaire à l'action de Yang-Mills mais comme une action en tant que tel. Plus précisément, il considère l'action : $S = \frac{k}{4\pi} \int_M Q_3$. Un argument similaire à celui de 1) montre qu'on peut donner un sens à cette expression modulo $2\pi k$. Pour construire une théorie quantique des champs, il suffit que e^{iS} soit bien défini, ce qui est possible si k est entier. Ce qui est remarquable est le fait que l'action de cette théorie ne dépend d'aucun choix de métrique sur M : on parle de théorie topologique des champs. Witten a montré qu'on pouvait donner un sens à la théorie quantique des champs associée à S et que les fonctions de corrélation de cette théorie sont liées à des invariants en théorie des noeuds. Le lecteur intéressé pourra lire le livre d'Atiyah [4] ou l'article de Witten [13].

4.2.2 Résolution des contraintes de Wess-Zumino

Soit $\tilde{\Gamma}[A]$ l'action effective obtenue après intégration sur les champs fermioniques. L'anomalie de jauge $\mathcal{A}_\alpha(x, A)$ exprime la variation de $\tilde{\Gamma}[A]$ sous une transformation de jauge de paramètre infinitésimale $\epsilon^\alpha(x)$: $\delta_\epsilon \tilde{\Gamma}[A] = \int d^n x \epsilon^\alpha(x) \mathcal{A}_\alpha(x, A)$. Soit s l'opérateur de la symétrie BRST : $s\mathcal{A}_\mu = D_\mu \omega$ (s agit comme une transformation de jauge dans laquelle le champ fantôme $\omega^\alpha(x)$ remplace $\epsilon^\alpha(x)$). On a alors :

$$s\tilde{\Gamma}[A] = \int d^n x (D_\mu \omega(x))^\alpha \frac{\delta \tilde{\Gamma}}{\delta A_\mu^\alpha(x)} = \int d^n x \omega(x)^\alpha \mathcal{G}_\alpha(x) \tilde{\Gamma}$$

après intégration par parties. Ainsi $s\tilde{\Gamma}[A] = \int d^n x \omega^\alpha(x) \mathcal{A}_\alpha(x, A)$, quantité qu'on va appeler $\mathcal{A}[\omega, A]$. $s^2 = 0$ d'où $s\mathcal{A}[\omega, A] = s^2 \tilde{\Gamma}[A] = 0$. Rappelons que $s\omega = i\omega\omega$, i.e. $s\omega^\alpha = -\frac{1}{2} C_{\alpha,\beta,\gamma} \omega^\beta \omega^\gamma$ et que s est une antidérivation : $s(\omega^\alpha \mathcal{A}_\alpha) = (s\omega^\alpha) \mathcal{A}_\alpha - \omega^\alpha (s\mathcal{A}_\alpha)$ d'où :

$$\begin{aligned} s\mathcal{A}[\omega, A] &= \int d^n x \left\{ -\frac{1}{2} C_{\alpha,\beta,\gamma} \omega^\beta(x) \omega^\gamma(x) \mathcal{A}_\alpha(x, A) - \omega^\alpha(x) \int d^n y \omega^\beta(y) \mathcal{G}_\beta(y) \mathcal{A}_\alpha(x, A) \right\} \\ &= \int \int d^n x d^n y \left(-\frac{1}{2} \omega^\alpha(x) \omega^\beta(x) \right) \left\{ C_{\alpha,\beta,\gamma} \mathcal{A}_\gamma(x, A) \delta^n(x-y) + \mathcal{G}_\beta(y) \mathcal{A}_\alpha(x, A) - \mathcal{G}_\alpha(x) \mathcal{A}_\beta(y, A) \right\} \end{aligned}$$

ainsi $s\mathcal{A}[\omega] = 0$ si et seulement si la condition de Wess-Zumino est vérifiée. On peut donc interpréter la condition de Wess-Zumino comme une conséquence de la symétrie BRST de la théorie : $\mathcal{A}[\omega, A]$ est une fonctionnelle BRST-fermée de nombre fantôme un. Supposons qu'il existe $F[A]$ une fonctionnelle locale (i.e. intégrale d'un polynôme en $A(x)$ et ses dérivées au même point x) de nombre fantôme zéro telle que $\mathcal{A}[\omega, A] = sF[A]$, on pourrait ajouter un contre-terme $-F$ à l'action classique $\Delta S = -F$ et alors on aurait : $\Delta \tilde{\Gamma} = -F$ d'où $s(\tilde{\Gamma} + \Delta \tilde{\Gamma}) = s(\tilde{\Gamma} - F) = \mathcal{A} - sF = 0$ d'où une annulation de l'anomalie. On dira qu'une anomalie est « vraie » si elle ne peut pas être éliminée par addition d'un contre-terme local. Autrement dit, une anomalie « vraie » correspond à une classe non-triviale de cohomologie BRST de nombre fantôme un sur l'espace des fonctionnelles locales.

On va maintenant calculer l'anomalie de jauge. On rappelle qu'on se place sur (M, g) variété lorentzienne de dimension paire $n = 2r$. Soit \mathcal{A} la fonctionnelle donnant l'anomalie. On a défini précédemment $\mathcal{A}[\omega, A]$ et on a vu que la condition de Wess-Zumino s'écrivait $s\mathcal{A}[\omega, A] = 0$ où s est la symétrie BRST avec la condition pour $\mathcal{A}[\omega, A]$ de ne pas être BRST-exacte si on veut une « vraie » anomalie. On redéfinit les champs fantômes en posant $w = -i\omega$.

Rappelons les équations de descente établies dans la partie sur les formes de Chern-Simons, reformulées dans le langage de la cohomologie BRST : $P_{r+1} = tr(F^{r+1})$, $dP_{r+1} = sP_{r+1} = 0$,

$P_{r+1} = dQ_{2r+1}$, $sQ_{2r+1} = dQ_{2r}^1$. Bien entendu, pour que ces équations aient un sens, il faut que M soit plongé dans une variété \tilde{M} de dimension $2r + 2$.

Supposons que Q_{2r}^1 et sQ_{2r}^1 sont globalement définis sur M . On a alors :

Théorème 11. $\int_M Q_{2r}^1(\omega, A)$ est un représentant de la cohomologie $BRST$ en nombre fantôme un et vérifie donc la condition de Wess-Zumino. De plus, l'anomalie est de la forme :

$$\mathcal{A}[\omega, A] = C \int_M Q_{2r}^1(\omega, A) + s(\dots)$$

où C est une constante. Si $C \neq 0$, alors l'anomalie est «vraie».

Démonstration. Rappelons que la cohomologie $BRST$ est définie à partir d'espaces de fonctionnelles locales en $A, \psi, \omega, \omega^*$ et h .

Montrons que $s \int Q_{2r}^1 = 0$. On a $sQ_{2r+1} = dQ_{2r}^1$ d'où $0 = s(sQ_{2r+1}) = s(dQ_{2r}^1) = -d(sQ_{2r}^1)$. On choisit l'espace \tilde{M} de dimension $2r + 2$ de telle façon que toute $2r$ -forme fermée soit exacte. Alors on peut écrire : $sQ_{2r}^1 = d\alpha_{2r-1}^1$ d'où $s \int_M Q_{2r}^1 = \int_M d\alpha_{2r-1}^1 = 0$. Ainsi, $\int_M Q_{2r}^1$ est $BRST$ -fermé.

Montrons que Q_{2r}^1 n'est pas $BRST$ -exact. Supposons que $Q_{2r}^1 = s\alpha_{2r}$ avec α_{2r} une fonctionnelle en les champs de jauge uniquement. Alors $sQ_{2r+1} = dQ_{2r}^1 = ds\alpha_{2r} = -sd\alpha_{2r}$ d'où $s(Q_{2r+1} + d\alpha_{2r}) = 0$. Or on a vu que la cohomologie $BRST$ de degré 0 est constituée de fonctionnelles invariantes de jauge et d'autre part il n'existe pas de forme de degré impair invariante de jauge. Ainsi, $Q_{2r+1} + d\alpha_{2r} = 0$ d'où $dQ_{2r+1} = 0$. Mais $dQ_{2r+1} = P_{r+1}$ qui est non-trivial en cohomologie $BRST$: sinon, on aurait $P_{r+1} = s\alpha_{2r+1}$, α_{2r+1} contiendrait au moins un champ anti-fantôme ω^* et puisque $s\omega^*$ contient h , P_{r+1} devrait contenir h , ce qui n'est pas.

Montrons que $\int_M Q_{2r}^1$ n'est pas $BRST$ exact. Supposons que $\int_M Q_{2r}^1 = s \int_M \beta_{2r} = \int_M sQ_{2r}^1$. Alors $Q_{2r}^1 + d\gamma_{2r-1} = s\beta_{2r}$. Mais Q_{2r}^1 n'est défini qu'à l'addition d'une forme exacte près. Si on pose $\tilde{Q}_{2r}^1 = Q_{2r}^1 + d\gamma_{2r-1}$, alors \tilde{Q}_{2r}^1 serait $BRST$ -exact ce qui est impossible d'après le paragraphe précédent. \square

La constante C dépend du champ chiral considéré. Pour un fermion de Weyl de chiralité positive, en dimension $n = 4$ et en signature lorentzienne, on verra que $C = \frac{i}{24\pi^2}$.

4.2.3 Cas du Modèle Standard

On a vu que l'existence d'une anomalie pour une symétrie de jauge dans une théorie se traduisait par une action effective non-invariante de jauge d'où une perte d'unitarité, l'existence d'états de norme négative ... Pour avoir une théorie cohérente, les contributions à l'anomalie des différents fermions chiraux présents doivent donc s'annuler.

Si R est une représentation du groupe de jauge G et si les t_α sont les représentants associés des générateurs de l'algèbre de Lie de G , on pose :

$$D_{\alpha,\beta,\gamma}^R = tr_R(t_\alpha \{t_\beta, t_\gamma\}) = str(t_\alpha t_\beta t_\gamma),$$

où str signifie trace symétrisée.

On considère une théorie en dimension $n = 4$ d'espace-temps, de groupe de jauge G , contenant des fermions gauches (i.e. de chiralité positive) $\psi_i(L)$ se transformant suivant des représentations R_i^L de G et des fermions droits (i.e. de chiralité négative) ψ_i^R se transformant suivant des représentations R_i^R de G . L'anomalie de jauge totale s'écrit alors :

$$\mathcal{A}_\alpha = -\frac{1}{24\pi^2} \partial_\mu (A_\nu^{ph,\beta} \partial_\rho A_\sigma^{ph,\gamma} - \frac{i}{4} A_\nu^{ph,\beta} [A_\rho^{ph}, A_\sigma^{ph}] \gamma) \times \left(\sum_i D_{\alpha,\beta,\gamma}^{R_i^L} - \sum_j D_{\alpha,\beta,\gamma}^{R_j^R} \right)$$

Remarquons qu'on peut inclure formellement les fermions non-chiraux dans cette expression car un fermion non-chiral peut s'écrire comme la somme d'un fermion droite et d'un fermion gauche, les deux se transformant sous la même représentation de G , d'où (comme prévu) une contribution nulle à l'anomalie.

Soit ψ un champ de particules chiral, se transformant suivant une représentation R de G , de générateurs t_α . Le champ d'antiparticules associé s'écrit : $\psi^c = i\gamma^0 C\psi^*$ où C est la matrice de conjugaison de charge. On a $C(\gamma^\mu)^T = -\gamma^\mu C$ d'où $C\gamma_5^T = \gamma_5 C$ (attention, ceci n'est valable qu'en dimension divisible par 4 : un signe supplémentaire apparaîtrait ici en dimension congrue à 2 modulo 4). D'où : $C\gamma_5^* = \gamma_5 C$. On en déduit que $\gamma_5\psi^c = \mp\psi^c$ si $\gamma_5\psi = \pm\psi$. Autrement dit, le passage au champ d'antiparticules associé renverse la chiralité : si ψ est gauche (resp. droit) alors ψ^c est droit (resp. gauche). D'autre part, ψ^c se transforme suivant la représentation de G de générateurs infinitésimaux $-t_\alpha$. En particulier, dans le calcul de l'anomalie, on peut remplacer les champs de particules droites par des champs d'antiparticules gauches et donc ne considérer que des fermions gauches.

Attention : Quand, dans l'expression de l'anomalie, on somme sur les différents types de fermions, il ne faut pas inclure particules et antiparticules séparément : les deux sont décrits par le même champ fermionique ψ ou de façon équivalente par ψ^c . Les calculs précédents montrent que la contribution à l'anomalie ne dépend du point de vue adopté.

Pour connaître les théories présentant une anomalie de jauge, il faut étudier les groupes de jauge G et les représentations R de G telles que $D_{\alpha,\beta,\gamma}^R \neq 0$. Remarquons que puisque l'anomalie ne fait intervenir que les t_α , il faut étudier en fait les algèbres de Lie et les représentations d'algèbre de Lie. Soit R une telle représentation, donnée par les t_α^R , la représentation conjuguée \bar{R} est définie par $t_\alpha^{\bar{R}} = -(t_\alpha^R)^* = -(t_\alpha^R)^T$ (de telle façon que $e^{i\epsilon^\alpha t_\alpha^{\bar{R}}} = (e^{i\epsilon^\alpha t_\alpha^R})^*$). Si R est isomorphe à \bar{R} , on parle de représentation réelle si on peut choisir les t_α^R imaginaires pures et antisymétriques et de représentation pseudo-réelle sinon. Dans ce cas,

$$D_{\alpha,\beta,\gamma}^R = \text{str}(t_\alpha^R t_\beta^R t_\gamma^R) = \text{str}((t_\alpha^R)^T (t_\beta^R)^T (t_\gamma^R)^T) = -\text{str}(t_\alpha^R t_\beta^R t_\gamma^R) = -D_{\alpha,\beta,\gamma}^R$$

d'où $D_{\alpha,\beta,\gamma}^R = 0$.

Remarque : En dimension $n = 2r$, il faut étudier $D_{\alpha_1, \dots, \alpha_r}^R$ et si n est divisible par 4, l'argument précédent reste valable mais pas si n est congru à 2 modulo 4.

Si on a un groupe G dont toutes les représentations de $\text{Lie}(G)$ vérifient $R \sim \bar{R}$, on n'a donc pas d'anomalie pour une théorie de jauge de groupe G en dimension $n = 4$.

Exemples

1. $SO(2n+1)$, $n \geq 1$ (en particulier $SU(2)$), $SO(4n)$, $n \geq 2$, $Sp(2n)$, $n \geq 3$, G_2 , F_4 , E_7 , E_8 et les produits directs de ses groupes : toutes les représentations sont réelles ou pseudo-réelles et donc vérifient $D_{\alpha,\beta,\gamma} = 0$.
2. $SO(4n+2)$, $n \geq 1$, E_6 : $D_{\alpha,\beta,\gamma} = 0$ pour toute représentation bien qu'ils admettent des représentations ni réelles ni pseudo-réelles.
3. $SU(n)$, $n \geq 3$ et $U(1)$ ont des représentations avec $D_{\alpha,\beta,\gamma} \neq 0$.

D'où :

Théorème 12. *Si le groupe de jauge est $G = G_1 \times \dots \times G_k$ avec chaque G_i simple ou $U(1)$, alors pour avoir des représentations avec $D_{\alpha,\beta,\gamma} \neq 0$, il faut qu'au moins un des G_i soit un $SU(n)$ pour un $n \geq 3$ ou un $U(1)$.*

Pour obtenir un résultat plus précis, il faut examiner les différents triplets de bosons de jauge pouvant conduire à $D_{\alpha,\beta,\gamma} \neq 0$. G_s désigne un groupe de Lie simple (les t_α correspondants vérifient $\text{tr}(t_\alpha) = 0$).

- $U(1)_A - U(1)_B - U(1)_C$. Contribution $tr(t_A t_B t_C) = \sum_i q_{i,A} q_{i,B} q_{i,C}$ où les $q_{i,X}$ sont les charges $U(1)$ des différents champs fermioniques gauches et t_X est le générateur associé à $U(1)_X$ (les indices A, B et C distinguent les différents facteurs $U(1)$ dans le groupe de jauge global).
- $U(1)_A - U(1)_B - G_s$. Contribution proportionnelle à $tr(t_\alpha)$ donc nulle où t_α est le générateur associé à G_s .
- $U(1) - G_s - G_{s'}$ ou $G_s - G_s - G_{s'}$ ou $G_s - G_{s'} - G_{s''}$ avec $G_s \neq G_{s'}$ et $G_{s''} \neq G_{s'}$. Contribution proportionnelle à $tr(t_\alpha)$ donc nulle où t_α est le générateur associé à $G_{s'}$.
- $G_s - G_s - G_s$. Contribution non-nulle si et seulement si $G_s = SU(n)$, $n \geq 3$.

On va appliquer ces résultats au cas du Modèle Standard. On appelle ainsi la théorie de jauge décrivant les interactions électromagnétique, faible et forte construite à partir des données expérimentales. Son groupe de jauge est $SU(3) \times SU(2) \times U(1)$. Il résulte de ce qui précède qu'il est nécessaire d'étudier la présence d'anomalies pour les triplets de jauge suivants :

$$SU(3) - SU(3) - SU(3),$$

$$SU(3) - SU(3) - U(1),$$

$$SU(2) - SU(2) - U(1),$$

$$U(1) - U(1) - U(1).$$

Pour ce faire, il faut connaître le contenu fermionique du Modèle Standard. Les fermions se répartissent en trois familles identiques du point de vue des couplages aux champs de jauge, on dit qu'on a trois générations. Il suffit donc de vérifier la condition d'absence d'anomalie pour une de ces générations. La première (dans l'ordre croissant des masses) comprend, en termes de champs chiraux gauches :

- le neutrino électronique gauche $\nu_{e,L}$ et l'électron gauche e_L qui forment à eux deux un vecteur qui se transforme sous la représentation standard de $SU(2)$. Chacun est un singlet (i.e. se transforme trivialement) sous $SU(3)$ et possède une hypercharge faible de $1/2$ où l'hypercharge faible désigne la charge $U(1)$.
- l'antiélectron gauche $(e_R)^c$ qui est un singlet sous $SU(2)$, $SU(3)$ et a une hypercharge faible de -1 .
- le quark up gauche u_L et le quark down d_L qui forment à eux deux un vecteur qui se transforme sous la représentation standard de $SU(2)$. Chacun se transforme sous la représentation standard de $SU(3)$ et possède une hypercharge faible de $-1/6$.
- l'antiquark up gauche $(u_R)^c$ qui est un singlet sous $SU(2)$ se transforme sous la représentation conjuguée de la représentation standard de $SU(3)$ et a une hypercharge faible de $2/3$.
- l'antiquark down gauche $(d_R)^c$ qui est un singlet sous $SU(2)$ se transforme sous la représentation conjuguée de la représentation standard de $SU(3)$ et a une hypercharge faible de $-1/3$.

Le Modèle Standard prévoit de plus l'existence d'un champ scalaire appelé champ de Higgs responsable d'une brisure spontanée de la symétrie $SU(2) \times U(1)$ et dont le couplage aux champs fermioniques est à l'origine de la masse des fermions. Il y a donc deux cas à distinguer :

1. Le cas de haute énergie : la symétrie $SU(2) \times U(1)$ n'est pas spontanément brisée, les fermions ne sont pas couplés au champ de Higgs et sont donc sans masse. Étudions les différentes anomalies possibles :
 - $SU(3) - SU(3) - SU(3)$: la représentation globale de $SU(3)$ est réelle d'où pas d'anomalie.
 - $SU(3) - SU(3) - U(1)$: seuls les quarks et les antiquarks contribuent, les contributions des représentations standard et conjuguée de la standard de $SU(3)$ sont identiques, la

contribution des hypercharges faibles se met donc en facteur et : $2 \times (-1/6) + 2/3 + (-1/3) = 0$ d'où pas d'anomalie.

$SU(2) - SU(2) - U(1)$: seuls les deux doublets sous $SU(2)$ contribuent, l'hypercharge faible se met en facteur et : $1/2 + 3 \times (-1/6) = 0$ d'où pas d'anomalie.

$U(1) - U(1) - U(1)$: $2 \times (1/2) + (-1) + 2 \times 3 \times (-1/6) + 3 \times (2/3) + 3 \times (-1/3) = 0$ d'où pas d'anomalie.

Conclusion : pas d'anomalie de jauge.

2. Le cas de basse énergie : le champ de Higgs, un doublet $SU(2)$ de champs scalaires complexes $\phi = (\phi^+, \phi^0)$ vérifie $\langle \phi^+ \rangle = 0$ et $\langle \phi^0 \rangle = v \neq 0$, ce qui réduit la symétrie $SU(2) \times U(1)$ à un sous-groupe $U(1)_{em}$ diagonalement plongé dans $SU(2) \times U(1)$. La symétrie de jauge restante $U(1)_{em}$ est celle de l'électromagnétisme et la charge associée est la charge électrique qui est déterminée par le couplage initial à $SU(2) \times U(1)$. Le groupe de jauge restant est $SU(3) \times U(1)_{em}$. Les quarks formant une représentation globale réelle de $SU(3) \times U(1)_{em}$, il n'y a pas d'anomalie $SU(3) - SU(3) - SU(3)$ ni $SU(3) - SU(3) - U(1)$ et les quarks ne contribuent pas à l'anomalie $U(1) - U(1) - U(1)$. Le couplage au champ de Higgs rend l'électron massif : il ne contribue donc à aucune anomalie. Le neutrino étant un singlet sous $SU(3) \times U(1)_{em}$, lui aussi ne contribue à aucune anomalie.

Conclusion : pas d'anomalie de jauge.

À toute échelle d'énergie, le Modèle Standard est donc exempt d'anomalie de jauge, ce qui est heureux pour la cohérence de la théorie. Les conditions d'absence d'anomalies de jauge sont les seules relations non-triviales reliant les nombres quantiques (les différentes charges) des quarks et des leptons.

4.3 La construction d'Alvarez-Gaumé et Ginsparg

4.3.1 But de la construction

On a vu que : $\mathcal{A}[\omega, A] = C \int d^n x Q_{2r}^1(\omega, A) + s(\dots)$.

On va donner une construction géométrique permettant de retrouver ce résultat, de déterminer la constante C et d'interpréter l'anomalie comme une conséquence de la non trivialité topologique de l'espace des orbites de jauge de la théorie. Ce qui suit est tiré de l'article d'Alvarez-Gaumé et Ginsparg [1].

Commençons par rappeler le cadre. On raisonne en signature euclidienne $(++\dots+)$ et on considère donc une variété M riemannienne compacte spin de dimension $n = 2r$ et on adopte la convention «physique» pour la définition des matrices de Dirac et donc de l'opérateur de Dirac : $\{\gamma_i, \gamma_j\} = 2g_{i,j}$ (Voir la remarque qui précède 2.1.5, Définition 10). Soit $S = S^+ \oplus S^-$ le fibré des spineurs sur M , G un groupe de Lie compact semi-simple simplement connexe, P un fibré principal de groupe de G , ρ une représentation linéaire de dimension finie de G et soit $\rho(P)$ le fibré vectoriel associé à P et à ρ . On notera par A une connexion sur P et par ψ une section de $S^+ \otimes \rho(P)$.

Autrement dit, physiquement, on considère une théorie contenant un fermion spineur de Weyl et un boson de jauge. L'opérateur de Dirac est donné par : $\mathcal{D} = \gamma^\mu (D_\mu + A_\mu)$ où les γ^μ sont les matrices de Dirac et on rappelle qu'on pose $\gamma_5 = i^{2r} \gamma^1 \dots \gamma^{2r}$. On n'a pas mis la contribution de la connexion spinorielle car dans la suite on ne va pas s'intéresser aux effets de la géométrie intrinsèque (i.e. de la métrique) de M , ce qui est légitime pour une variété de \hat{A} -classe triviale. On prendra $M = S^{2r}$ qui vérifie bien $\hat{A}(TS^{2r}) = 1$.

La dynamique de la théorie physique est alors déterminée par la donnée de l'action $\tilde{\Gamma}[A]$ qui s'écrit formellement :

$$e^{-\tilde{\Gamma}[A]} = \int D\psi D\bar{\psi} e^{-\int d^{2r}x \bar{\psi} i\mathcal{D}_+\psi}$$

où $\mathcal{D}_+ = \mathcal{D}P_+$ avec $P_{\pm} = \frac{1 \pm \gamma_5}{2}$ les opérateurs de chiralité.

Problème : comment donner un sens à cette intégrale fonctionnelle ?

On a envie de poser : $e^{-\tilde{\Gamma}[A]} = \det(i\mathcal{D}_+)$. La première objection qui vient à l'esprit est : $i\mathcal{D}_+$ est un opérateur qui agit sur un espace de dimension infinie, le produit de ses valeurs propres ne va pas converger... Comme on l'a vu précédemment, ceci n'est en général pas gênant à condition de régulariser proprement les choses. Mais ici, la situation est beaucoup plus grave et le fait d'être en dimension infinie ne joue pas de rôle : $i\mathcal{D}_+$ n'est pas un endomorphisme ! En effet $i\mathcal{D}_+ : C^\infty(M, S^+ \otimes \rho(P)) \rightarrow C^\infty(M, S^- \otimes \rho(P))$, i.e. envoie spineurs de chiralité positive sur spineurs de chiralité négative. Parler de valeur propre ou de déterminant pour cet opérateur n'a pas de sens. On va donc changer notre définition et poser, toujours formellement :

$$e^{-\tilde{\Gamma}[A]} = \int D\psi D\bar{\psi} e^{-\int d^{2r}x \bar{\psi} i\hat{D}\psi}$$

où $\hat{D} = \mathcal{D}_+ + \mathcal{D}_-$ est cette fois bien un endomorphisme et on pose : $e^{-\tilde{\Gamma}[A]} = \det(i\hat{D})$ (correctement régularisé). Il est important de remarquer que \hat{D} décrit essentiellement la même physique que \mathcal{D}_+ , on a juste ajouté un spineur de chiralité négative non couplé au champ de jauge. $i\hat{D}$ n'est pas hermitien mais il est elliptique et on peut considérer son spectre $(\lambda_n)_n$. Malheureusement, contrairement à \mathcal{D} , \hat{D} n'est pas invariant de jauge : c'est donc également le cas de son spectre, de son déterminant et donc de $\tilde{\Gamma}[A]$, d'où la présence d'une anomalie. L'existence de l'anomalie ne signifie pas qu'on s'y est mal pris pour définir la théorie physique, c'est une conséquence inéluctable de la situation physique étudiée : un spineur couplé différemment au champ de jauge suivant sa chiralité. La dissymétrie présente initialement a pour conséquence une action quantique effective non invariante de jauge. Il ne reste plus qu'à quantifier cette non invariance, i.e. à calculer cette anomalie.

On remarque tout d'abord que $|\det(i\hat{D})|$ est invariant de jauge, en effet :

$$\det(i\hat{D})\det((i\hat{D})^\dagger) = \det(i\mathcal{D}_+ i\mathcal{D}_-)\det(i\partial_- i\partial_+)$$

À une constante près, on a donc : $|\det(i\hat{D})| = \sqrt{\det(i\mathcal{D})}$. La définition du module de $\det(i\hat{D})$ ne pose donc pas de problème. En revanche, on va voir que l'argument varie au cours d'une transformation de jauge et est donc à l'origine d'une anomalie.

Soit \mathcal{U} l'espace des connexions sur P . L'espace \mathcal{U} est topologiquement trivial : c'est un espace vectoriel donc un espace contractile. Soit \mathcal{G} le groupe des transformations de jauge de P fixant un point donné (on parle de groupe des transformations de jauge pointé). Alors \mathcal{G} agit librement sur \mathcal{U} et le quotient \mathcal{U}/\mathcal{G} représente l'espace des configurations physiquement distinctes du champ de jauge (en fait, modulo l'action de G qui fait changer le point fixé). Contrairement à ce qui se passe pour \mathcal{U} , la topologie de \mathcal{U}/\mathcal{G} est en générale non triviale. Le déterminant $\det(i\hat{D})$ (convenablement régularisé) est une fonction bien définie sur \mathcal{U} , i.e. une section du fibré trivial en droites complexes sur \mathcal{U} . Mais n'étant pas invariant sous \mathcal{G} , il ne peut pas définir une fonction sur \mathcal{U}/\mathcal{G} mais seulement une section d'un fibré en droites complexes généralement non trivial sur \mathcal{U}/\mathcal{G} . On reviendra plus loin sur ce point.

4.3.2 Développement adiabatique

On rappelle qu'on suppose $M = S^{2r}$ et de plus on suppose P trivial. Soit $A \in \mathcal{U}$ tel que $i\mathcal{D}(A)$ soit sans mode nul. On considère dans \mathcal{U} une boucle passant par A , autrement dit une

famille à un paramètre de transformations de jauge : $A^\theta = g^{-1}(\theta)Ag(\theta) + g^{-1}(\theta)d_x g\theta$ où $g : S^1 \times M \rightarrow G$ vérifie $g(0, x) = g(2\pi, x) = 1$. Ces conditions aux limites font qu'on peut voir g comme une application de S^{2r+1} dans G ou encore comme une boucle $\theta \rightarrow g(\theta, \cdot)$ dans \mathcal{G} . On va essayer de calculer la variation de $\det(i\hat{D})$ lorsqu'on parcourt la boucle $\theta \rightarrow A^\theta$. On définit $w(\cdot, A) : S^1 \rightarrow S^1$ par $\det(i\hat{D}(A^\theta)) = \sqrt{\det(i\mathcal{D}(A))}e^{iw(\theta, A)}$ et on veut calculer le degré de w :

$$d = \frac{1}{2\pi} \int_0^{2\pi} \frac{\partial w}{\partial \theta}(\theta, A) d\theta$$

On commence par définir $A^{t, \theta} = tA^\theta$ famille à deux paramètres de connexions qui décrit dans \mathcal{U} un disque \mathcal{D} dont le bord est la boucle $\theta \rightarrow A^\theta$. $(t, \theta) \rightarrow \det(i\hat{D}^{t, \theta})$ est une fonction sur \mathcal{D} et on suppose que ses zéros dans \mathcal{D} sont isolés et simples. La restriction de l'argument de $\det(i\hat{D}^{t, \theta})$ à un petit cercle autour d'un des zéros de $\det(i\hat{D}^{t, \theta})$ définit une application de ce cercle à valeurs dans S^1 . On appellera indice du zéro le degré de cette application. Il est clair que le degré de w est égal à la somme sur les zéros de $\det(i\hat{D}^{t, \theta})$ contenus dans \mathcal{D} des indices. En fait, on va voir que les indices des zéros sont reliés aux modes nuls d'un opérateur de Dirac en dimension $2r + 2$.

L'idée est relativement naturelle : on va considérer le produit $\mathcal{D} \times M$ de façon à avoir une copie de M et de P muni de la connexion $A^{t, \theta}$ pour chaque couple (t, θ) . Plus précisément, on considère $S^2 \times M$ en prenant pour carte \mathcal{D} paramétré par (t, θ) sur S^2 privée d'une petite calotte contenant le pôle sud. On désigne par (s, θ) un système de coordonnées au voisinage du pôle sud. On définit un champ de jauge sur $S^2 \times M$ par :

$$\mathcal{A}(x, t, \theta) = A^{t, \theta} = tg^{-1}(A + d_x g)\mathcal{A}(x, s, \theta) = A - sd_\theta g g^{-1}$$

\mathcal{A} est une connexion sur un fibré principal de groupe G sur $S^2 \times M$ défini par la fonction de recollement $g(\theta, x)$. $S^2 \times M$ est une variété spin qui possède donc un opérateur de Dirac :

$$i\mathcal{D}_{2r+2} = \sum_{a=1}^{2r+2} i(\partial_a + \mathcal{A}_a)\Gamma^a.$$

Une représentation explicite des matrices de Dirac Γ^a est donnée par : $\Gamma^\mu = \sigma_1 \otimes \gamma^\mu$, $\Gamma^{2r+1} = \sigma \otimes 1$, $\Gamma^{2r+2} = \sigma_1 \otimes \gamma^5$, $\Gamma_5 = i^{r+1} \prod_{a=1}^{2r+2} \Gamma^a = \sigma_3 \otimes 1$ où les σ_i sont les matrices de Pauli :

$$\begin{aligned} \sigma_1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ \sigma_3 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

Le premier objectif est d'étudier les modes nuls de $i\mathcal{D}_{2r+2}$. L'idée est d'introduire un petit paramètre $\epsilon > 0$ permettant un traitement perturbatif. On pose :

$$i\mathcal{D}_{2r+2}^\epsilon = \frac{1}{\epsilon} \sum_{\mu=1}^{2r} D_\mu \Gamma^\mu + i\mathcal{D}_2$$

où $i\mathcal{D}_2 = \sum_i D_i \Gamma^i$ et où i décrit les deux coordonnées supplémentaires. Étant donné que $i\mathcal{D}_{2r+2}^\epsilon \Gamma_5 = -\Gamma_5 i\mathcal{D}_{2r+2}^\epsilon$, on a $\text{ind}(i\mathcal{D}_{2r+2}^\epsilon) = \text{ind}(i\mathcal{D}_{2r+2})$. On va s'intéresser à la limite $\epsilon \rightarrow 0$ qui correspond à prendre les dimensions de M très petites devant celles des deux dimensions

supplémentaires. Les modes nuls de $i\mathcal{D}_{2r+2}^\epsilon$ sont aussi ceux de « l'hamiltonien » (ce nom a des raisons évidentes) :

$$H_\epsilon = (i\mathcal{D}_{2r+2}^\epsilon)^2 = \frac{1}{\epsilon^2}(iD_\mu\Gamma^\mu)^2 + (iD_i\Gamma^i)^2 + \frac{i^2}{\epsilon}\Gamma^i\Gamma^\mu(D_iD_\mu - D_\mu D_i)$$

ψ est une section d'un fibré non trivial sur $S^2 \times M$: $\psi(x, t, \theta)$ et $\tilde{\psi}(x, s, \theta)$, définies sur chacune des deux cartes, sont reliées par : $\psi(x, 1, \theta) = g(\theta, x)\tilde{\psi}(x, 1, \theta)$. On commence par l'étude de $\psi(x, t, \theta)$. On a alors :

$$H_\epsilon = \frac{1}{\epsilon^2}1 \otimes (i\mathcal{D}_{2r}^{t,\theta})^2 - (\partial_i)^2 + \frac{1}{\epsilon}\Gamma^i\Gamma^\mu i\partial_i(iA_\mu^{t,\theta})$$

où : $i\mathcal{D}_{2r}^{t,\theta} = iD_\mu(A^{t,\theta})\gamma^\mu$.

Si toutes les valeurs propres de $(i\mathcal{D}_{2r}^{t,\theta})^2$ étaient strictement positives, alors les valeurs propres de H_ϵ seraient en $1/\epsilon^2$ et H_ϵ n'aurait pas de mode nul pour ϵ petit. Si on pense à H_ϵ comme à un hamiltonien, à un potentiel, on voit bien que ses fonctions propres seront concentrées près des points où $\det((i\mathcal{D}_{2r}^{t,\theta})^2) = 0$. L'introduction du paramètre ϵ a justement pour but d'exacerber cet effet : quand ϵ devient petit, le potentiel devient partout très élevé excepté aux voisinages immédiats des zéros de $(i\mathcal{D}_{2r}^{t,\theta})^2$. Les «barrières» séparant les différents «puits» sont alors si élevées qu'on peut considérer les situations en les différents zéros comme des problèmes indépendants et traiter le problème au voisinage d'un zéro de manière perturbative. Soit $\lambda(t, \theta)$ une valeur propre de $i\mathcal{D}_{2r}^{t,\theta}$ variant continûment avec (t, θ) (c'est le cas génériquement sur des domaines assez petits) : $i\mathcal{D}_{2r}^{t,\theta} = \lambda(t, \theta)\psi_\lambda^{t,\theta}(x)$. Soit (t_0, θ_0) tel que $\lambda(t_0, \theta_0) = 0$. Soit ϕ_i un système de coordonnées tel que le zéro ait pour coordonnées $\phi_i = 0$. $\gamma_5\psi^{\phi_1, \phi_2}$ est vecteur propre pour la valeur propre $-\lambda(\phi_1, \phi_2)$: on a donc en fait deux modes qui passent par zéro en $\phi_i = 0$. En $\phi_i = 0$, l'espace des modes nuls est engendré par les deux états $\psi_\pm = \frac{1 \pm \gamma_5}{2}\psi^{\phi_i=0}(x)$.

On va appliquer la théorie des perturbations non dégénérées, i.e. on recherche les solutions perturbées dans cet espace de dimension deux : $\psi(x, t, \theta) = f_+(t, \theta)\psi_+^{t,\theta} + f_-(t, \theta)\psi_-^{t,\theta}$. Près de $\phi_i = 0$, la perturbation de $i\mathcal{D}_{2r}^{\phi_1, \phi_2}$ est (on approche les opérateurs par leur valeur en $\phi_i = 0$) :

$$\delta(i\mathcal{D}_{2r}^{\phi_1, \phi_2}) = \sum_j (i\partial_j \mathcal{A})\phi_j$$

où $\partial_j \mathcal{A} = \frac{\mathcal{A}^{\phi_1, \phi_2}}{\partial \phi_j}|_{\phi_i=0}$ qui s'écrit, dans la base (ψ_+, ψ_-) :

$$\begin{pmatrix} 0 & \sum_i z_i^* \phi_i \\ \sum_i z_i \phi_i & 0 \end{pmatrix}$$

où $z = \langle \psi_-, i\partial_i \mathcal{A} \psi_+ \rangle$. Une valeur approchée de $\lambda(\phi_i)$ est donc donnée par : $\lambda(\phi_i)^2 = (\sum_i z_i \phi_i)(\sum_i z_i^* \phi_i) = |\sum_i z_i \phi_i|^2$, $\lambda(\phi_i) = \pm |z_1 \phi_1 + z_2 \phi_2|$. Quitte à faire une rotation des coordonnées ϕ_1 et ϕ_2 , on peut supposer $Re(z_1^* z_2) = 0$ et alors $\lambda(\phi_i)^2 = |z_1|^2 \phi_1^2 + |z_2|^2 \phi_2^2$.

D'autre part, on a : $\Gamma^{2r+1} = \sigma_2 \otimes 1$, $\Gamma^{2r+2} = \sigma_1 \otimes \gamma_5$, $\Gamma^\mu = \sigma_1 \otimes \gamma^\mu$ d'où :

$$\Gamma^{2r+1}\Gamma^\mu = (\sigma_2 \otimes 1)(\sigma_1 \otimes \gamma^\mu) = (\sigma_2\sigma_1 \otimes \gamma^\mu) = -i\sigma_3 \otimes \gamma^\mu$$

$$\Gamma^{2r+2}\Gamma^\mu = (\sigma_1 \otimes \gamma_5)(\sigma_1 \otimes \gamma^\mu) = 1 \otimes \gamma_5\gamma^\mu$$

Le dernier terme de H_ϵ en $(\phi_i = 0)_i$ se réécrit alors :

$$\frac{1}{\epsilon}\Gamma^i\Gamma^\mu i\partial_i(iA_\mu) = \frac{1}{\epsilon}(\sigma_3 \otimes \frac{\partial(i\mathcal{A})}{\partial \phi_1} + 1 \otimes i\gamma_5 \frac{\partial(i\mathcal{A})}{\partial \phi_2})$$

En ne conservant que les termes en $1/\epsilon$, on obtient finalement l'expression approchée de H_ϵ au voisinage de $(\phi_i = 0)_i$ dans la base (ψ_+, ψ_-) :

$$\left(-\frac{\partial^2}{\partial\phi_1^2} - \frac{\partial^2}{\partial\phi_2^2} + \frac{1}{\epsilon^2}(|z_1|^2\phi_1^2 + |z_2|^2\phi_2^2)\right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{\epsilon} \begin{pmatrix} 0 & \chi z_1^* + iz_2^* \\ \chi z_1 - iz_2 & 0 \end{pmatrix}$$

où $\chi = \pm 1$ dépend du fait de savoir si on suit un mode de chiralité positive ou négative. Les valeurs propres de H_ϵ sont donc :

$$\frac{1}{\epsilon}|z_1| + \frac{1}{\epsilon}|z_2| \pm \sqrt{|z_1|^2 + |z_2|^2 + 2\chi Im(z_1^* z_2)}$$

Dans cette approximation, H_ϵ ne peut avoir un mode nul que si $|z_1||z_2| = \chi Im(z_1^* z_2)$. Autrement dit, l'existence d'un mode nul de H_ϵ de chiralité donnée concentré près d'un zéro de $det(i\mathcal{D}_{2r}^{t,\theta})$ fixe le signe de $Im(z_1^* z_2)$ ce qui donne une information sur le comportement de $i\mathcal{D}_{2r}^{t,\theta}$ au voisinage de ce point. Inversement, on voit qu'au voisinage d'un zéro de $det(i\mathcal{D}_{2r}^{t,\theta})$, on peut construire un mode de valeur propre zéro (en tout cas à l'ordre un en ϵ près) de chiralité positive ou négative selon le signe de $Im(z_1^* z_2)$. Il reste à relier ces résultats à $det(i\hat{D})$. On rappelle que par définition $i\hat{D} = i\mathcal{D}_+ + i\partial_-$ et que, puisque $|det(i\hat{D})| = \sqrt{det(i\mathcal{D})}$, $det(i\hat{D})$ s'annule exactement aux mêmes points que $det(i\mathcal{D})$. D'autre part, on remarque que, au signe près, $det(i\hat{D}) = det(i\partial_- i\mathcal{D}_+)$. En effet, décomposons les vecteurs propres de $i\hat{D}$ en leurs parties chirales : $i\hat{D}\phi = \lambda\phi$, $\phi = \phi_+ + \phi_-$, $i\mathcal{D}_+\phi_+ = \lambda\phi_+$, $i\partial_-\phi_- = \lambda\phi_-$ d'où $i\partial_-(i\mathcal{D}_+\phi_+) = \lambda^2\phi_+$ si $\lambda \neq 0$. Puisque $\{\gamma_5, i\hat{D}\} = 0$, les valeurs propres non-nulles de $i\hat{D}$ viennent par paires $\pm\lambda$ d'où le résultat. Soit donc $N = i\partial_- i\mathcal{D}_+$. Les zéros de $det(i\hat{D})$ sont les mêmes que ceux de $det(N)$. On se place donc au voisinage d'un de ces zéros et on va faire un calcul perturbatif. Il convient de faire attention : N n'est pas hermitien et il faut donc adapter un petit peu la méthode des perturbations traditionnelle. On écrit $N = N_0 + \delta N$ avec N_0 l'opérateur non-perturbé. Soit e_i les vecteurs propres de N_0 et f_i ceux de N_0^\dagger , soit λ_i et $\bar{\lambda}_i$ les valeurs propres associées. Alors la variation au premier ordre des valeurs propres est donnée par : $\delta\lambda_i = \langle f_i, \delta N e_i \rangle / \langle f_i, e_i \rangle$.

Soit z la valeur propre de $i\partial_- i\mathcal{D}_+^{\phi_1, \phi_2}$ qu'on suit au voisinage de $\phi_i = 0$, on a à une constante près qui n'a pas d'importance pour l'indice du zéro :

$$z = \langle (i\partial_-)^{-1} \psi_- | \delta(i\partial_- i\mathcal{D}_+) \psi_+ \rangle = \langle \psi_- | \delta(i\mathcal{A}) \psi_+ \rangle$$

On écrit $z = x + iy$, $z_i = x_i + iy_i$ et $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$. Localement, on a donc affaire à une similitude : l'indice du zéro étudié est donc $+1$ ou -1 selon qu'on ait affaire à une similitude directe ou indirecte. Pour trancher entre les deux possibilités, il suffit de regarder le déterminant :

$$det \begin{pmatrix} 0 & \chi z_1^* + iz_2^* \\ \chi z_1 - iz_2 & 0 \end{pmatrix} = x_1 y_2 - x_2 y_1 = Im(z_1^* z_2)$$

En remettant tout ensemble, on voit qu'on a obtenu une correspondance univoque entre les modes nuls de chiralité positive et négative de $i\mathcal{D}_{2n+2}^\xi$ et les indices respectivement $+1$ ou -1 des zéros de $det(i\hat{D})$. Remarquons que ce qui précède démontre que l'indice d'un zéro de $det(i\hat{D})$ est nécessairement $+1$ ou -1 ce qui n'était nullement évident a priori. Le résultat obtenu étant la conclusion d'un raisonnement perturbatif à l'ordre dominant en $1/\epsilon$, on peut légitimement se demander s'il est valable en toute généralité. C'est le cas : en effet, au voisinage d'un zéro de $det(i\hat{D})$, on a vu qu'on pouvait construire un mode de H_ϵ de chiralité prescrite par le signe de $\frac{Im(z_1^* z_2)}{|z_1||z_2|}$ et de valeur propre nulle à l'ordre un en ϵ près. Il est possible que du fait de corrections d'ordre plus élevé ou d'effets non perturbatifs, cette valeur propre soit en réalité non nulle mais

alors le mode construit est associé à un autre de mode de valeur propre opposée. Par hypothèse, la valeur propre ne diverge pas quand ϵ tend vers zéro, le deuxième mode est donc concentré au voisinage d'un zéro de $\det(i\hat{D})$ d'indice opposé à celui du zéro initialement considéré. Ainsi, on peut rassembler par paires les modes qui ne sont qu'approximativement nuls : les deux zéros associés aux membres d'une paire étant d'indices opposés l'égalité entre le nombre total de modes nuls comptés avec chiralité et le nombre de zéros comptés avec indice est vraie de manière exacte, indépendante de toute approximation.

Sur la carte décrivant le voisinage du pôle sud de S^2 , le terme dominant de H_ϵ quand ϵ tend vers zéro est :

$$\frac{1}{\epsilon^2} \mathbf{1} \otimes (i\mathcal{D}_{2r}(A))^2 = \frac{1}{\epsilon^2} \mathbf{1} \otimes [i(\partial_\mu + A_\mu)\gamma^\mu]^2$$

Dans la limite $\epsilon \rightarrow 0$, le seul mode nul possible est $\tilde{\psi} = 0$. Cette solution se recolle bien avec les fonctions $\psi(x, t, \theta)$ car ces dernières sont, comme on l'a vu, concentrées autour des zéros de $\det(i\hat{D})$. Finalement, on a obtenu :

$$d = \frac{1}{2\pi} \int_0^{2\pi} \frac{\partial w}{\partial \theta}(\theta, A) d\theta = \text{ind}(i\mathcal{D}_{2r+2}^\epsilon) = \text{ind}(i\mathcal{D}_{2r+2})$$

D'après le théorème de l'indice :

$$\text{ind}(i\mathcal{D}_{2r+2}) = \frac{1}{(r+1)!(2i\pi)^{r+1}} \int_{S^2 \times S^{2r}} \text{tr}(F^{r+1})$$

Rappelons les équations de descente : $\text{tr}(F^{r+1}) = dQ_{2r+1}$, $\delta Q_{2r+1} = dQ_{2r}^1$. On en déduit :

$$\text{ind}(i\mathcal{D}_{2r+2}) = \frac{1}{(r+1)!(2i\pi)^{r+1}} \left[\int_{S^2_+ \times S^{2r}} dQ_{2r+1}(A_+, F_+) + \int_{S^2_- \times S^{2r}} dQ_{2r+1}(A_-, F_-) \right]$$

où l'indice + (resp. -) réfère à la carte supérieure (resp. inférieure). La partie S^2 de la connexion étant triviale sur la carte inférieure, la deuxième intégrale est nulle. Ainsi :

$$\text{ind}(i\mathcal{D}_{2r+2}) = \frac{1}{(r+1)!(2i\pi)^{r+1}} \int_{S^1 \times S^{2r}} Q_{2r+1}(A_+, F_+)$$

Or : $A_+(x, 1, \theta) = A^\theta(x) + d\theta g^{-1}(x, \theta) \partial_\theta g(x, \theta)$ et $F_+(x, 1, \theta) = g^{-1}(x, \theta)(dA + A^2)g(x, \theta) = dA^\theta + A^\theta A^\theta = F^\theta$. Soit $v^\theta(x) = g^{-1}(x, \theta) \partial_\theta g(x, \theta)$, alors :

$$\int_{S^1 \times S^{2r}} Q_{2r+1}(A_+, F_+) = \int_{S^1 \times S^{2r}} [Q_{2r+1}(A^\theta + d\theta v^\theta, F^\theta) - Q_{2r+1}(A^\theta, F^\theta)]$$

La deuxième forme à intégrer ne contient pas de $d\theta$ et donne donc une contribution nulle à l'intégrale. En utilisant la deuxième équation de descente, on a alors :

$$\text{ind}(i\mathcal{D}_{2r+2}) = \frac{1}{(r+1)!(2i\pi)^{r+1}} \int_{S^1} d\theta \int_{S^{2r}} Q_{2r}^1(v^\theta, A^\theta, F^\theta)$$

On en déduit :

$$\delta_v \tilde{\Gamma}[A] = \frac{-1}{(r+1)!(2i\pi)^r} \int_{S^{2r}} Q_{2r}^1(v, A, F)$$

Ce résultat est valable en signature euclidienne. Si on passe en signature lorentzienne, on a :

$$\delta_v \tilde{\Gamma}[A] = \frac{(-1)^{r+1}}{(r+1)!(2\pi)^r} \int_{S^{2r}} Q_{2r}^1(v, A, F)$$

Les deux formules précédentes résolvent complètement le problème de la détermination de l'anomalie de jauge dans le cas étudié. L'argument qu'on a présenté pour établir le lien entre l'anomalie de jauge en dimension n et l'indice d'un opérateur de Dirac en dimension $n + 2$ n'est valable que pour la sphère S^{2r} mais on peut montrer grâce à une généralisation du théorème de l'indice à des familles d'opérateurs que c'est en fait valable pour tout M . En revanche, la formule précédente utilisée pour $ind(i\mathcal{D}_{2r+2})$ n'est valable que lorsqu'on a négligé les effets de la géométrie intrinsèque de M (plus précisément, cette formule n'est valable que si la \hat{A} -classe de $M \times S^2$ est triviale). Le cas général s'interprète en termes physiques comme une anomalie qui est à la fois de jauge et gravitationnelle : de même que l'anomalie de jauge est reliée à une non-invariance de jauge de $\hat{\Gamma}[A]$ sous une transformation de jauge infinitésimale, l'anomalie gravitationnelle est reliée à une non-invariance de $\hat{\Gamma}[A]$ sous un changement local de coordonnées. Pour plus de détails sur les anomalies gravitationnelles, voir l'article de Witten et Alvarez-Gaumé [3]. Le lecteur pourra en particulier comparer le calcul perturbatif de l'anomalie gravitationnelle qui y est fait avec la démonstration du théorème de l'indice présentée dans la partie 3).

4.3.3 Anomalie : un problème de géométrie classique ou quantique ? Un saut dans l'infini

Dans ce qui précède, on a obtenu des obstructions topologiques conduisant à la présence d'une anomalie. Mais une anomalie est un phénomène quantique alors que les obstructions en termes d'indice d'opérateurs agissant sur les sections de fibrés semblent relever de la géométrie classique. En fait, ce qu'on devrait appeler «géométrie quantique» est la géométrie de l'espace de dimension infinie des orbites de jauge \mathcal{U}/\mathcal{G} puisque c'est sur cet espace qu'on intègre pour définir la théorie quantique. On va donc commencer par établir quelques propriétés de cet espace dans le cas où $M = S^n$.

On a déjà dit que \mathcal{U} était topologiquement trivial car contractile mais que ce n'était pas le cas du quotient \mathcal{U}/\mathcal{G} . Vérifions le au niveau de l'homotopie. On peut voir \mathcal{U} comme un fibré de base \mathcal{U}/\mathcal{G} de fibre \mathcal{G} . On a donc une suite exacte d'homotopie associée :

$$\dots \longrightarrow \pi_{i+1}(\mathcal{U}/\mathcal{G}) \longrightarrow \pi_i(\mathcal{G}) \longrightarrow \pi_i(\mathcal{U}) \longrightarrow \pi_i(\mathcal{U}/\mathcal{G}) \longrightarrow \pi_{i-1}(\mathcal{G}) \longrightarrow \dots$$

Or, pour tout $i > 0$, $\pi_i(\mathcal{U}) = 0$ d'où $\pi_i(\mathcal{U}/\mathcal{G}) = \pi_{i-1}(\mathcal{G})$. De plus $\pi_i(\mathcal{G}) = \pi_{i+n}(G)$ d'où $\pi_i(\mathcal{U}/\mathcal{G}) = \pi_{i+n-1}(G)$. La non-trivialité de l'espace des orbites provient donc de la non-trivialité des groupes d'homotopie de G .

Revenons à l'anomalie de jauge. On a vu qu'elle pouvait s'interpréter comme un problème de définition d'un déterminant. $det(i\hat{D})$ est une fonction définie sur \mathcal{U} à valeur dans \mathbb{C} , i.e. une section d'un fibré trivial sur \mathcal{U} . La non-invariance de jauge de $det(i\hat{D})$ empêche de le définir comme une fonction sur \mathcal{U}/\mathcal{G} : on n'obtient qu'une section d'un fibré en droites complexes sur \mathcal{U}/\mathcal{G} . Classifier les fibrés en droites complexes est équivalent à classifier les fibrés principaux de groupe $U(1)$ associés. Les fibrés de groupe $U(1)$ sur S^m sont classifiés (topologiquement) par $\pi_{m-1}(U(1))$, qui n'est non-nul que pour $m = 2$, $\pi_1(U(1)) = \mathbb{Z}$. La non-trivialité des fibrés en droites complexes sur \mathcal{U}/\mathcal{G} peut donc s'étudier en regardant les restrictions aux sphères de dimension deux non-contractiles plongées dans \mathcal{U}/\mathcal{G} . Or on a vu au paragraphe précédent que $\pi_2(\mathcal{U}/\mathcal{G}) = \pi_1(\mathcal{G}) = \pi_{n+1}(G)$. \mathcal{U}/\mathcal{G} possède donc des sphères de dimension deux non-contractiles si et seulement si $\pi_{n+1}(G)$ est non-trivial. La condition $\pi_{n+1}(G) \neq 0$ est donc suffisante pour qu'il y ait une anomalie. Mais elle n'est pas nécessaire (pour s'en convaincre, prendre $G = U(1)$) : on peut avoir une variation locale de $det(i\hat{D})$ sous des transformations de jauge même si l'intégrale de sa variation le long de toute boucle de \mathcal{G} est nulle. Cette remarque nous donne l'occasion de préciser la relation existante entre les anomalies et la topologie : aussi bien l'anomalie chirale que l'anomalie de jauge sont des fonctions sur M , ce sont des grandeurs locales. Elles contiennent

donc plus d'information que leur intégrale sur M qui elle a une interprétation topologique via le théorème de l'indice. Ceci explique que l'approche naturelle du théorème de l'indice en lien avec les anomalies est une approche locale : c'est le cas de la démonstration présentée dans la section 3. D'un point de vue purement mathématique, la version locale du théorème de l'indice n'apporte pas grand chose par rapport aux approches topologiques (basées sur la théorie du cobordisme ou sur la K-théorie) dans le cas où on travaille sur une variété M compacte mais elle est beaucoup plus souple et est essentielle pour des généralisations du théorème de l'indice au cas de variétés de base non-compactes et/ou avec bord.

Pour résumé, les anomalies ne sont pas des grandeurs topologiques : elles contiennent une information algébrique non-triviale, à savoir la forme précise de l'anomalie. Mais via leur intégrale elles sont des mesures de non-trivialité topologique :

1. Anomalie chirale : mesure de la non-trivialité de la topologie d'un G -fibré principal sur M
2. Anomalie de jauge : mesure de la non-trivialité d'un fibré déterminant sur le produit de M par une variété de dimension deux, ou, ce qui est équivalent, sur un sous-espace à deux paramètres de l'espace des orbites de jauge.

On va finir en citant sans démonstration quelques résultats concernant la topologie d'objets de dimension infinie qui mettent en perspective la théorie de l'indice. Dans ce mémoire, on s'est principalement intéressé à l'indice d'opérateurs différentiels agissant sur un espace de Hilbert \mathcal{H} , en général non-bornés. Quitte à remplacer \mathcal{D} par $\mathcal{D}(1 + \mathcal{D}^* \mathcal{D})^{-1/2}$, on peut se ramener à l'étude d'opérateurs continus sur \mathcal{H} et on obtient ainsi des opérateurs de Fredholm. On appelle \mathcal{F} l'espace des opérateurs de Fredholm agissant sur \mathcal{H} muni de la topologie induite par celle de l'espace des opérateurs continus $\mathcal{L}(\mathcal{H})$. L'indice est une fonction définie sur \mathcal{F} à valeurs dans \mathbb{Z} : $ind: A \rightarrow dimKerA - dimKerA^*$. On a déjà dit que l'indice était une quantité stable par déformation continue : ind est donc constante sur les composantes connexes de \mathcal{F} . En fait, on a mieux : on peut montrer (théorème d'Atiyah-Jänich) que les composantes connexes de \mathcal{F} sont exactement les ensembles d'opérateurs de Fredholm d'indice fixé, autrement dit, on a un isomorphisme $\pi_0(\mathcal{F}) = \mathbb{Z}$ induit par ind . L'indice est une quantité associée à un opérateur, i.e. à une famille d'opérateurs à zéro paramètres.

Pour considérer des familles d'opérateurs à un paramètre, on introduit l'espace \mathcal{F}^1 des opérateurs de Fredholm auto-adjoints. Par construction, $ind = 0$ sur \mathcal{F}^1 . On peut montrer que $\mathcal{F}^1 = \mathcal{F}_+^1 \cup \mathcal{F}_-^1 \cup \mathcal{F}_*^1$ avec \mathcal{F}_+^1 l'espace des opérateurs de \mathcal{F} de spectre essentiellement positif, i.e. avec un nombre fini de valeurs propres négatives et \mathcal{F}_-^1 l'espace des opérateurs de \mathbb{F} de spectre essentiellement négatif, i.e. avec un nombre fini de valeurs propres positives. \mathcal{F}_+^1 et \mathcal{F}_-^1 sont contractiles, ce qui n'est pas le cas de \mathcal{F}_*^1 qui contient le contenu topologique intéressant. Les éléments de \mathcal{F}_*^1 ont un spectre contenu dans aucune des demi-droites positive ou négative, ce qui est par exemple le cas d'un opérateur de Dirac mais pas du laplacien. Ceci a pour conséquence un effet topologique non-trivial. On considère une boucle $t \rightarrow \mathcal{D}_t$ dans \mathcal{F}_*^1 paramétrée par $[0, 1[$ et on fixe une valeur propre de \mathcal{D}_0 qu'on suit par continuité le long de la boucle. Lorsque $t \rightarrow 1$, on tend vers une valeur propre de \mathcal{D}_0 qui n'est pas nécessairement celle dont on est parti, il peut se produire un décalage et on appelle flot spectral de la famille (\mathcal{D}_t) le nombre de valeurs propres dont on s'est décalé. C'est une quantité bien définie au sens où elle ne dépend pas du choix de la valeur propre initiale : c'est tout le spectre qui se translate lorsqu'on fait le tour de la boucle. On peut montrer que le flot spectral est invariant par déformation continue d'une boucle dans \mathcal{F}_*^1 d'où une application $\pi_1(\mathcal{F}_*^1) \rightarrow \mathbb{Z}$ qui est en fait un isomorphisme.

On commence donc à voir apparaître une «dualité» entre \mathcal{F} et l'indice d'une part et \mathcal{F}_*^1 et le flot spectral d'autre part. Pour l'expliciter, on introduit U le groupe unitaire infini défini

comme la limite inductive des groupes unitaires $U(N)$ pour les inclusions naturelles de $U(N)$ dans $U(N + 1)$. Il est assez facile de voir que à p fixé, les groupes d'homotopie $\pi_p(U(N))$ se stabilisent pour N assez grand à un groupe qui est donc aussi $\pi_p(U)$. En revanche, le résultat suivant est beaucoup plus profond :

Théorème 13. *Périodicité de Bott*

1. $\pi_p(U) = \mathbb{Z}$ si p est impair
2. $\pi_p(U) = 0$ si p est pair.

En fait, on a l'équivalence d'homotopie $U \sim \Omega^2(U)$ où, si X est un espace topologique, $\Omega(X)$ est l'espace des courbes fermées sur X avec un point base fixé et $\Omega^2(X) = \Omega(\Omega(X))$.

On pourrait se demander : quel rapport avec les opérateurs de Fredholm ? On peut montrer les équivalences d'homotopie : $\mathcal{F}_*^1 \sim U$ et $\mathcal{F} \sim \Omega(U)$. En particulier, $\mathcal{F}_*^1 \sim \Omega^2(\mathcal{F}_*^1)$ et $\mathcal{F} \sim \Omega^2(\mathcal{F})$, ce qu'on peut interpréter comme la raison conceptuelle qui se cache derrière la répétition dans le comportement des opérateurs de Dirac lorsqu'on passe de la dimension n à la dimension $n + 2$.

Références

- [1] Alvarez-Gaumé and Ginsparg. The topological meaning of non-abelian anomalies. *Nucl. Phys. B*, 3 :449–474, 1984.
- [2] Alvarez-Gaumé and Vazquez-Mozo. Introductory lectures on quantum field theory. 2010. arXiv :hep-th/0510040.
- [3] Alvarez-Gaumé and Witten. Gravitational anomalies. *Nucl. Phys. B*, B234(269), 1984.
- [4] Atiyah. *The geometry and physics of knots*. Cambridge University Press, 1990.
- [5] Berline, Getzler, and Vergne. *Heat kernels and Dirac operators*. Grundlehren der mathematischen Wissenschaften, 298. Berlin, New-York, Paris : Springer, 1992.
- [6] Bilal. Lecture on anomalies. 2008. arXiv :hep-th/0802.0634.
- [7] Chern. *Complex manifold without potential theory*. Universitext.
- [8] Connes and Marcolli. *Noncommutative geometry, quantum fields and motives*. Colloquium Publications, American Mathematical Society. Providence, R.I American Mathematical Society, 2008.
- [9] Coquereaux. *Espaces fibrés et connexions*. Centre de Physique Théorique, Luminy-Marseille, 2002.
- [10] Nakahara. *Geometry, topology and physics*. Graduate student series in physics. Bristol, Philadelphia : A.Hilger, 1990.
- [11] Nash. *Differential topology and quantum field theory*. London : Academic Press, 1991.
- [12] Roe. *Elliptic operators, topology and asymptotic methods*. Pitman research notes in mathematics series. Harlow : Longman Scientific and Technical, 1988.
- [13] Witten. Quantum field theory and the Jones polynomial. *Comm. Math. Phys*, 121(3) :351–399, 1989.