Université de Liège

Faculté des Sciences
Section de Physique

Année académique 1999-2000
2ème Licence en Sciences Physiques

Géométrie et topologie en théorie de jauge

Introduction à la théorie de Chern-Simons

L.Moyaerts

Mémoire présenté
en vue de l'obtention du grade de
licencié en Sciences Physiques

Chapitre 1

Introduction

Le terme de Chern-Simons, que nous allons introduire au cours de ce travail est un exemple très particulier de la relation unissant la mathématique à la physique. Dans ces deux disciplines, il est très courant d'associer à un modèle, à une équation ou encore à une technique de calcul devenue célèbre, le nom du ou des auteurs : équation de Newton, de Schrödinger, de Dirac, inégalités d'Heisenberg, théorie de Yang-Mills, effet Aharonov-Bohm, diagrammes de Feynman, régularisation de t'Hooft et Veltman, critère de Cauchy, espace de Hilbert, tenseur de Riemann-Christoffel, transformation de Fourier, etc. Le cas de Chern-Simons est tout à fait remarquable. Il ne s'agit pas, comme on pourrait le penser *a priori* de deux physiciens théoriciens ayant travaillé sur une nouvelle théorie de jauge, mais bien de deux mathématiciens, spécialisés dans le domaine de la géométrie différentielle moderne. Au-delà de cet aspect purement anecdotique, nous pensons que cela met en évidence un des modes d'interaction qui caractérisent les relations physico-mathématiques. Si nous laissons de côté cette période de l'histoire des sciences, par ailleurs extrêmement intéressante, où mathématique, physique et philosophie sont unies au sein d'une discipline indissociable (pensons simplement aux célèbres *principia* de Newton), nous pouvons trouver, dans un passé moins lointain, essentiellement deux grandes tendances. Tout d'abord, le besoin pour la physique de se donner un matériau et des outils nécessaires à l'expression technique de ses théories amène les mathématiciens à développer de nouveaux concepts, à définir de nouveaux objets et à découvrir au travers de théorèmes les relations qui s'établissent entre eux. On a alors l'image frisant il est vrai parfois un peu le cliché, d'une mathématique rabibochant, consolidant, mettant en forme les tentatives souvent aussi hasardeuses qu'audacieuses assez typiques du travail des physiciens. Cette forme de rapports est très interactive : elle permet à la mathématique d'explorer de nouveaux domaines et en retour, les résultats qui en découlent profitent à la physique sous la forme d'une consolidation de ses bases, d'une clarification ou mieux encore, sous la forme de nouvelles prédictions. Un exemple de clarification nous est donné par la mécanique quantique. Depuis quelques décennies, la forme sous laquelle est présentée dans les livres la mécanique quantique n'a plus rien à voir avec les balbutiements formels qui ont caractérisé son éclosion. Pensons aux deux formes sous lesquelles s'exprimaient cette nouvelle mécanique à ses débuts : d'une part, le formalisme fonctionnel de Schrödinger, basé essentiellement sur l'analyse et d'autre part la "mécanique des matrices" de Bohr et Heisenberg, s'exprimant dans le langage des espaces vectoriels de dimension infinie. L'équivalence des deux représentations a été possible grâce à la théorie des espaces fonctionnels. La seconde image d'épinal est celle du mathématicien déconnecté de la réalité, développant des concepts abstraits et apparemment sans finalité utilitaire. Dans ce cas-là également, il peut y avoir interaction. L'exemple le plus célèbre est peut-être celui de la relativité générale, ou plus exactement le formalisme géométrique sur lequel celle-ci est construite : la géométrie non euclidienne développée par Riemann. C'est un ami et ancien condisciple d'Einstein, alors que celui-ci lui faisait part des problèmes mathématiques rencontrés lors de l'élaboration de sa célèbre théorie, qui a rappelé au grand Albert l'existence du formalisme de Riemann. De cette symbiose entre les idées d'Einstein et la géométrie de Riemann créée près d'un demi-siècle plus tôt est née la théorie peut-être la plus élégante de l'histoire de la physique. L'introduction de la théorie de jauge de Chern-Simons est issue elle aussi des travaux de mathématiciens. Si nous parcourons l'article ([8]) de ces deux auteurs, aucune référence n'est faite au monde de la théorie des champs, qui pourtant va en exploiter les résultats et construire une théorie physique à part entière qui portera leur nom.

Les domaines de la mathématique dans lesquels nous allons évoluer au cours de ce travail sont ceux de la topologie et de la géométrie différentielle. Les rapports entre ces deux disciplines et la physique sont clairs depuis l'invention de la relativité générale. Leur intrusion dans la mécanique quantique et la théorie des champs est par contre beaucoup plus récente. Le cas du monopôle magnétique déjà évoqué par Dirac mis à part, les notions telles que la phase de Berry, l'effet Aharonov-Bohm ou les instantons sont des exemples relativement modernes (nous entendons par là de la seconde moitié de ce siècle) d'incursions géométriques et topologiques dans la physique théorique. Le point de départ de la "géométrisation" des théories de jauge [1] est l'introduction dans la physique des hautes énergies des années cinquante de la théorie de Yang et Mills, qui consiste en une généralisation non abélienne des lois de l'électromagnétisme de Maxwell. A partir de cette idée s'est construit au fil des années le modèle standard dont le succès n'est plus à démontrer. Il est apparu que ces théories peuvent s'exprimer dans un langage emprunté à la géométrie différentielle et à la théorie des groupes, dans le cadre de la théorie des fibrés principaux, dont l'étude topologique a permis de mettre en évidence des propriétés spéciales, souvent très étonnantes, d'origine purement mathématique. C'est d'une étude des classes caractéristiques sur les fibrés principaux menée en 1974 par nos deux mathématiciens qu'est né le terme de Chern-Simons. Celui-ci est apparu originellement dans le cadre de la théorie de la supergravité en 1978 comme un terme "étrange mais nécessaire à la cohérence de la théorie" ([9]). La grande particularité de ce terme est son indépendance vis-à-vis de la métrique, propriété que les physiciens résumeront souvent dans l'appellation "terme topologique". Quelques années plus tard, on mettra en évidence la possibilité de construire à l'aide de celui-ci une théorie de jauge dans un espace de dimension 2+1 ([10]). Celle-ci a permis de développer des aspects nouveaux de la physique dans le plan et s'est imposée dans l'étude de l'effet Hall quantique, en introduisant une description théorique des particules de statistique fractionnaire ([11]).

Aborder les aspects actuels de la théorie de Chern-Simons est une tâche relativement ardue pour un étudiant sortant de la première licence. Ce travail se veut avant tout une introduction à cette théorie ainsi qu'un condensé des concepts mathématiques figurant en toile de fonds des théories de jauge. Nous y aborderons des notions qui ne sont généralement pas évoquées dans le cadre d'un cours de mathématique ou de physique théorique de licence. Outre une introduction aux espaces topologiques et aux notions fondamentales de la géométrie différentielle figurant en appendice, nous nous baserons sur la matière vue aux cours de théorie des champs, d'introduction aux théories de jauge, et de physique des particules. Les sujets qui y sont abordés ne sont pas rappelés. Ce travail s'organise en quatre chapitres. Le premier chapitre prend l'aspect d'un exercice où, afin de nous familiariser avec les objets mathématiques de la géométrie différentielle, nous tenterons de "géométriser" une des premières théories de jauge que la physique a connue, l'électromagnétisme de Maxwell. Sous sa version classique tout d'abord, puis à 2+1 dimensions, ce qui nous permettra d'introduire ensuite le terme de Chern-Simons. Le chapitre suivant constituera la partie mathématique de ce travail. Nous y introduirons quelques notions de topologie algébrique, puis une grande part sera accordée à la notion de fibré principal, évoquée plus haut. Nous appliquerons ensuite, dans le chapitre 4, ce formalisme à la théorie de Yang-Mills, en étudiant tout d'abord les classes caractéristiques introduites originellement par Chern et Simons puis en examinant quelques cas où les aspects topologiques d'une théorie de jauge peuvent être

1. Nous entendrons souvent par théories de jauge les théories rencontrées en physique des hautes énergies, bien que la relativité soit également une théorie de jauge.

mis en évidence. Enfin, dans le dernier chapitre, nous évoquerons le rôle joué par le terme de Chern-Simons dans le cadre de la physique des particules de statistique fractionnaire et terminerons en passant en revue quelques caractéristiques fondamentales de ce terme, dans le cadre d'une théorie des champs classique. Il est à noter que nous n'évoquerons de manière générale ni le problème de la quantification de la théorie de Chern-Simons, ni la traduction dans le langage géométrique de la quantification des théories de jauge, mais que nous en traiterons seulement quelques cas particuliers (angle θ de Yang-Mills, anomalies, opérateur BRS, induction du terme de Chern-Simons par les corrections radiatives, etc.)

Dans le chapitre mathématique, nous avons veillé à rester le plus proche possible du formalisme développé par les mathématiciens, tout en nous accordant quelques libertés, dans le but d'alléger le discours. Peut-être celui-ci paraîtra-t-il encore fort abstrait aux personnes n'ayant pas l'habitude de côtoyer ce genre de formalisme, il est vrai fort touffu, ou bien paraîtra-t-il un peu léger aux personnes pratiquant l'art mathématique désireuses d'en savoir un peu plus sur les applications physiques de la géométrie et de la topologie. Notre intention était de trouver un compromis entre ces deux tendances, et nous espérons que ce but a été atteint.

Bonne lecture.

Chapitre 2

Une théorie de jauge à 2+1 dimensions

2.1 Introduction

Dans ce chapitre, nous nous proposons d'établir un premier contact avec la théorie de Chern-Simons (que nous abrévierons souvent dans la suite par C.S.). Comme le titre l'indique, nous allons montrer qu'il s'agit d'une théorie de jauge à part entière dans un espace-temps à trois dimensions. Afin de rendre le saut dimensionnel moins abrupt, nous repartirons de la théorie classique de Maxwell dans un espace quadri-dimensionnel. Nous en profiterons pour nous familiariser [1] avec les notions de base de la géométrie différentielle, et nous verrons qu'il est possible d'écrire les équations de Maxwell dans un langage aussi élégant que compact. En procédant par analogie, les lois de l'électromagnétisme seront transposées dans un espace tri-dimensionnel, et nous examinerons les nouveautés qui en découlent. Enfin, nous y ajouterons le fameux terme de C.-S. et étudierons avec quelques détails ses propriétés : essentiellement la modification des équations de champ, bien sûr, une analyse des contraintes, et une discussion des jauges fréquemment employées dans cette théorie.

2.2 Electromagnétisme à 3+1 dimensions

Nous nous plaçons dans l'espace-temps de Minkowski, équipé de la métrique habituelle en théorie des champs :

$$\eta_{\mu\nu} = [\mathrm{diag}(1,-1,-1,-1)].$$

Les indices grecs désignent, conformément à l'usage, les composantes spatio-temporelles des grandeurs vectorielles rencontrées ($\mu, \nu, \rho, \ldots = 0, 1, 2, 3$). Les indices latins désigneront toujours, sauf mention explicite du contraire, les composantes spatiales uniquement ($i, j, k, \ldots = 1, 2, 3$). Dans le langage de la géométrie différentielle, nous travaillons dans l'espace $\mathbb{R}^4$ avec les coordonnées cartésiennes (x^0, x^1, x^2, x^3) et la base qui leur est associée :

$$(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}).$$

Les composantes du tenseur métrique sont alors :

$$g_{\mu\nu} = g(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}) = \eta_{\mu\nu}.$$

1. Nous conseillons au lecteur non familier des notations de la géométrie différentielle de consulter l'appendice B, à la fin de ce travail.

Nous adoptons la convention suivante pour le pseudo-tenseur de Levi-Civita :

$$\varepsilon^{0123} = 1.$$

Ceci entraîne directement :

$$\varepsilon^{1230} = -1, \quad \varepsilon_{0123} = -1 \text{ et } \varepsilon_{1230} = 1.$$

A trois (2+1) dimensions, nous posons

$$\varepsilon^{012} = 1$$

et il vient

$$\varepsilon^{120} = \varepsilon_{012} = \varepsilon_{120} = 1.$$

Enfin, nous travaillerons dans le système d'unité utilisé couramment en théorie des champs. Nous posons

$$c = \hbar = 1.$$

2.2.1 Champ électromagnétique et densité de courant

Nous allons dans les lignes qui suivent traduire les lois de l'électromagnétisme dans le langage des formes de la géométrie différentielle. Construisons tout d'abord les objets mathématiques correspondant aux grandeurs physiques que nous connaissons. Sont à transposer :

- les six composantes du champ électromagnétique $(\mathbf{E}, \mathbf{B})$
- les quatre composantes du quadrivecteur densité de courant $(\rho, \mathbf{j})$.

Rappelons que sur un espace vectoriel à n dimensions, le nombre de composantes d'une p-forme est C_n^p. Dans un espace à quatre dimensions, seule la 2-forme possède six composantes. Elle s'écrit, sous sa forme canonique et dans la base duale (dx^0, dx^1, dx^2, dx^3)

$$F = \frac{1}{2} F_{\mu\nu}\, dx^\mu \wedge dx^\nu, \tag{2.1}$$

où la matrice des composantes $F_{\mu\nu}$ est antisymétrique. En développant la somme, il vient

$$F = F_{0i}\, dx^0 \wedge dx^i + \frac{1}{2} F_{ij} dx^i \wedge dx^j.$$

On pose [2]

$$E_i = -F_{i0}, \tag{2.2}$$

$$B_i = -\frac{1}{2}\, \varepsilon_{ijk}\, F_{jk}. \tag{2.3}$$

Inversément, $F_{ij} = -\varepsilon_{ijk} B_k$. On obtient alors

$$F = E_i\, dx^0 \wedge dx^i - \frac{1}{2} \varepsilon_{ijk} B_k dx^i \wedge dx^j.$$

Sous forme matricielle, il vient

$$(F_{\mu\nu}) = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix},$$

2. Remarquons que les champs électriques et magnétiques ne sont pas des morceaux de quadrivecteurs. L'élèvement et l'abaissement des indices n'a donc dans ce cas aucun sens ! Nous placerons systématiquement les indices latins en position inférieure. On a par exemple $= E_i = F^{i0} = -F_{i0}$.

et

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix}.$$

Nous savons qu'à toute forme nous pouvons associer sa forme duale par application de l'étoile de Hodge :

$$^*F = \frac{1}{2}\left[\frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}F^{\alpha\beta}\right] dx^\mu \wedge dx^\nu.$$

Il vient donc successivement

$$\begin{aligned} ^*F &= \frac{1}{2}\varepsilon_{0i\alpha\beta}F^{\alpha\beta}\, dx^0 \wedge dx^i + \frac{1}{4}\varepsilon_{ij\alpha\beta}F^{\alpha\beta}\, dx^i \wedge dx^j \\ &= -\frac{1}{2}\varepsilon_{ijk}F^{jk}\, dx^0 \wedge dx^i + \frac{1}{2}\varepsilon_{ijk}F^{k0}\, dx^i \wedge dx^j \\ &= B_i\, dx^0 \wedge dx^i + \frac{1}{2}\varepsilon_{ijk}E_k\, dx^i \wedge dx^j. \end{aligned}$$

Sous forme matricielle,

$$(^*F_{\mu\nu}) = \begin{pmatrix} 0 & B_1 & B_2 & B_3 \\ -B_1 & 0 & E_3 & -E_2 \\ -B_2 & -E_3 & 0 & E_1 \\ -B_3 & E_2 & -E_1 & 0 \end{pmatrix},$$

et

$$(^*F^{\mu\nu}) = \begin{pmatrix} 0 & -B_1 & -B_2 & -B_3 \\ B_1 & 0 & E_3 & -E_2 \\ B_2 & -E_3 & 0 & E_1 \\ B_3 & E_2 & -E_1 & 0 \end{pmatrix}.$$

Pour le courant, nous avons le choix entre la 1-forme et la 3-forme. Pour une raison qui deviendra évidente dans quelques lignes, nous choisissons de décrire les sources par la 3-forme

$$j = \frac{1}{3!}\, \varepsilon_{\mu\nu\lambda\rho}\, j^\rho\, dx^\mu \wedge dx^\nu \wedge dx^\lambda, \tag{2.4}$$

où nous reconnaissons le quadrivecteur densité de courant j^μ.

2.2.2 Equations de Maxwell

Nous avons à présent tous les outils en main pour écrire dans ce formalisme les équations fondamentales de l'électromagnétisme. Rappelons-les tout d'abord dans leur forme vectorielle habituelle[3]

$$\nabla \times \mathbf{E} + \frac{\partial}{\partial t}\mathbf{B} = 0 \tag{2.5}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.6}$$

$$\nabla \times \mathbf{B} - \frac{\partial}{\partial t}\mathbf{E} = \mathbf{j} \tag{2.7}$$

$$\nabla \cdot \mathbf{E} = \rho \tag{2.8}$$

3. Nous utilisons le système d'unité d'Heaviside-Lorentz, où $\varepsilon_0 = 1$, et les facteurs 4π sont réabsorbés dans la définition des champs. Ce système d'unité permet d'écrire les équations fondamentales de l'électromagnétisme sous leur forme la plus simple.

Les deux premières équations de Maxwell sont obtenues en égalant la dérivée extérieure de F à zéro. En effet,

$$dF = \frac{1}{3!}\left[\partial_\mu F_{\nu\rho} - \partial_\nu F_{\mu\rho} + \partial_\rho F_{\mu\nu}\right] dx^\mu \wedge dx^\nu \wedge dx^\rho, \tag{2.9}$$

et on retrouve la relation

$$\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0, \tag{2.10}$$

bien connue sous le nom d'*identité de Bianchi*.
Il est possible de l'écrire sous une forme plus compacte, faisant intervenir le champ dual. En effet, dF est une 3-forme. Elle s'écrit de manière générique

$$dF = \frac{1}{3!}\varepsilon_{\mu\nu\rho\lambda} G^\lambda\, dx^\mu \wedge dx^\nu \wedge dx^\rho.$$

En identifiant à (2.9), il vient, en passant au champ dual

$$G^\lambda = \partial_\alpha\ {}^*F^{\alpha\lambda},$$

et finalement, (2.10) s'écrit

$$\partial_\mu\ {}^*F^{\mu\nu} = 0. \tag{2.11}$$

Les équations de Maxwell avec source s'écrivent

$$d\ {}^*F = j.$$

Cette relation justifie *a posteriori* notre choix d'une 3-forme pour j. La conservation du courant dérive alors d'une propriété de la dérivée extérieure :

$$0 = d^2 G = dj.$$

En effet,

$$dj = \frac{1}{3!}\,\varepsilon_{\mu\nu\lambda\rho}\,\partial_\varepsilon\, j^\mu\, dx^\varepsilon \wedge dx^\nu \wedge dx^\lambda \wedge dx^\rho,$$

et de là l'équation de continuité

$$\partial_\mu j^\mu = 0. \tag{2.12}$$

En développant $d\ {}^*F$, il vient

$$d\ {}^*F = \frac{1}{3!}[\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu}]\, dx^\mu \wedge dx^\nu \wedge dx^\rho.$$

Identifiant à (2.4), il vient, par un calcul en tout point identique à celui permettant de dériver (2.11),

$$\partial_\mu F^{\mu\nu} = j^\nu. \tag{2.13}$$

En résumé, les deux équations suivantes

$$dF = 0, \tag{2.14}$$

$$d\ {}^*F = j, \tag{2.15}$$

et les propriétés de la dérivée extérieure et de * contiennent les lois fondamentales de l'électromagnétisme. En se rappelant la définition de la codérivée $\delta = {}^*d\ {}^*$, on peut réécrire les deux équations précédentes sous la forme :

$$dF = 0,$$

$$\delta F = {}^*j. \tag{2.16}$$

Mentionnons deux invariants que nous pouvons construire avec F et ${}^*F$:

$$F_{\mu\nu}F^{\mu\nu} = -2(E^2 - B^2), \tag{2.17}$$

$$F_{\mu\nu}\ {}^*F^{\mu\nu} = -2(\mathbf{E}\cdot\mathbf{B}). \tag{2.18}$$

Notons enfin que $\det(F - \lambda I) = \lambda^4 + (B^2 - E^2)\lambda^2 - (\mathbf{E}\cdot\mathbf{B})^2$.

2.2.3 Potentiel et invariance de jauge

Nous savons que les équations de Maxwell homogènes impliquent l'existence d'un potentiel $A^\mu = (\phi, \mathbf{A})$ défini à une transformation de jauge près :

$$A^\mu \longrightarrow A^\mu + \partial^\mu \Lambda. \tag{2.19}$$

L'existence du potentiel est ici assurée par le lemme de Poincaré. En effet, $\mathbb{R}^4$ étant un ouvert étoilé, (2.14) nous permet de construire une 1-forme A telle que $F = dA$. Celle-ci est consistante avec le potentiel habituel A_μ car en tant que 1-forme dans un espace à quatre dimensions, elle possède quatre composantes. L'invariance de jauge prend alors la forme suivante :

$$A \longrightarrow A + d\Lambda, \tag{2.20}$$

pour toute fonction Λ. (2.16) s'écrit maintenant :

$$\delta dA = {}^*j. \tag{2.21}$$

Il reste à faire le lien avec les équations de mouvement habituelles pour le champ A_μ. On a successivement :

$$A = A_\mu dx^\mu,$$

$${}^*A = \frac{1}{3!}\,\varepsilon_{\mu\nu\lambda\rho}\, A^\mu\, dx^\nu \wedge dx^\lambda \wedge dx^\rho,$$

$$d\,{}^*A = \partial_\mu A^\mu dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3$$

et finalement

$${}^*d\,{}^*A = \partial_\mu A^\mu.$$

La condition de Lorentz s'écrit

$$\delta A = 0.$$

Enfin, en se rappelant le lien entre le d'Alembertien et l'opérateur de codérivée, il vient, sous la condition de Lorentz

$$\Box A = (\delta d + d\delta)\, A = \delta dA = {}^*j. \tag{2.22}$$

Les composantes du champ électromagnétique F s'obtiennent en fonction des composantes du potentiel en appliquant la définition de la dérivée extérieure de la 1-forme A :

$$dA = \frac{1}{2}\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) dx^\mu \wedge dx^\nu.$$

D'où

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{2.23}$$

et l'on retrouve les expressions bien connues des champs $\mathbf{B}$ et $\mathbf{E}$ en terme du potentiel :

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad \mathbf{E} = -\nabla A_0 - \frac{\partial}{\partial t}\mathbf{A}.$$

2.2.4 Action et Lagrangien

Nous terminerons ce premier paragraphe en évoquant l'action S, écrite en fonction des formes que nous venons d'introduire. Elle s'écrit

$$S[A] = \int (-\frac{1}{2} F \wedge {}^*F + j \wedge A). \tag{2.24}$$

En développant chacune des grandeurs intervenant dans cette intégrale, on obtient le lagrangien habituel de l'électromagnétisme :

$$S[A] = \int (-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - A^\mu j_\mu) \, dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3,$$

et

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - A^\mu j_\mu. \tag{2.25}$$

Vérifions que S est invariante sous l'action d'une transformation de jauge. On a

$$\begin{aligned} S[A + d\Lambda] - S[A] &= \int j \wedge d\Lambda && (2.26)\\ &= -\int d(j \wedge \Lambda) + \int dj \wedge \Lambda && (2.27)\\ &= 0. && (2.28) \end{aligned}$$

La première intégrale est nulle par le théorème de Stokes et, en supposant les sources à support fini, la seconde par conservation du courant.

2.3 Electromagnétisme à 2+1 dimensions

En suivant un raisonnement très naïf, nous serions amenés à clore ce paragraphe immédiatement, en considérant tous les résultats de la section précédente valables pour des champs électriques et magnétiques à deux composantes. Il n'en n'est rien, car la mathématique tue dans l'oeuf tout espoir d'opérer ce raccourci. En effet, dans le langage des formes sur $\mathbb{R}^3$, il n'existe pas d'objet à quatre composantes, nous empêchant ainsi de construire un champ électromagnétique pour les quatre composantes du champ électrique et magnétique. Cela ne doit pas nous décourager d'opérer par analogie avec les résultats du paragraphe précédent. Nous introduisons donc à nouveau le champ électromagnétique par l'intermédiaire d'une 2-forme F:

$$F = \frac{1}{2} F_{\mu\nu} \, dx^\mu \wedge dx^\nu. \tag{2.29}$$

On aura :

$$\begin{aligned} E_i &= -F_{i0}, && (2.30)\\ B &= -\frac{1}{2} \varepsilon_{ij} F_{ij}, && (2.31) \end{aligned}$$

et, sous forme matricielle,

$$(F_{\mu\nu}) = \begin{pmatrix} 0 & E_1 & E_2 \\ -E_1 & 0 & -B \\ -E_2 & B & 0 \end{pmatrix},$$

et

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E_1 & -E_2 \\ E_1 & 0 & -B \\ E_2 & B & 0 \end{pmatrix}.$$

On remarque donc que le champ magnétique dans le cas de l'électromagnétisme dans le plan est une grandeur à une composante. Il s'agit en l'occurence d'un pseudo-scalaire.

Toujours par analogie, nous écrivons l'identité de Bianchi :

$$dF = 0, \tag{2.32}$$

correspondant à nouveau à l'expression (2.10) dans laquelle, cette fois, les indices symbolisent les valeurs 0,1,2. Ce qui signifie que, des quatre équations correspondant à l'identité de Bianchi à quatre dimensions, nous en avons perdu trois. Il nous reste en effet :

$$\partial_0 F_{12} + \partial_1 F_{20} + \partial_2 F_{01} = 0,$$

s'écrivant en termes des champs

$$-\frac{\partial}{\partial t} B - \varepsilon_{ij}\, \partial_i\, E_j = 0, \tag{2.33}$$

c-à-d

$$\nabla \times \mathbf{E} + \frac{\partial}{\partial t} B = 0. \tag{2.34}$$

En résumé, l'équation de Faraday (3 équations à quatre dimensions) n'est plus qu'une équation scalaire, et la deuxième équation de Maxwell ($\operatorname{div} \mathbf{B} = 0$) a complètement disparu, le champ magnétique ayant perdu son caractère vectoriel.

Nous construisons ensuite l'étoile de Hodge de F. Par définition, on a :

$$^*F = \frac{1}{2}\, \varepsilon_{\mu\nu\rho}\, F^{\mu\nu}\, dx^\rho,$$

et

$$^*F_\mu = \frac{1}{2}\, \varepsilon_{\mu\nu\rho}\, F^{\nu\rho}. \tag{2.35}$$

Attention, l'expression inverse est

$$F_{\mu\nu} = \varepsilon_{\mu\nu\alpha}\ {}^*F^\alpha. \tag{2.36}$$

Le champ dual s'écrit donc

$$\begin{aligned} ^*F &= \frac{1}{2}\varepsilon_{\mu\nu\rho} F^{\nu\rho}\, dx^\mu \\ &= \frac{1}{2}\varepsilon_{0ij} F^{ij}\, dx^0 + \varepsilon_{ij0} F^{j0}\, dx^i \\ &= -B\, dx^0 + \varepsilon_{ij} E_j\, dx^i. \end{aligned}$$

Sous forme vectorielle,

$$^*F_\mu = (-B, E_2, -E_1) \quad \text{et} \quad {}^*F^\mu = \begin{pmatrix} -B \\ -E_2 \\ E_1 \end{pmatrix}.$$

Avec le champ dual, l'identité (2.32) s'écrit cette fois

$$\partial_\mu\ {}^*F^\mu = 0. \tag{2.37}$$

Il reste à construire un objet à trois composantes pour décrire la densité de courant. Nous avons à nouveau le choix entre deux formes : la 1-forme et la 2-forme. Nous adoptons la seconde solution. On pose donc :

$$j = \frac{1}{2}\, \varepsilon_{\lambda\mu\rho}\, j^\rho\, dx^\lambda \wedge dx^\mu. \tag{2.38}$$

En raisonnant une dernière fois par analogie, on est amené à écrire

$$d\,^*F = j, \tag{2.39}$$

qui donne :

$$\partial_\mu F^{\nu\mu} = j^\nu. \tag{2.40}$$

La relation (2.39), qui fait intervenir des 2-formes, cache en fait trois équations que nous pouvons écrire en terme des champs :

$$\nabla \times B - \frac{\partial}{\partial t}\mathbf{E} = -\mathbf{j}, \tag{2.41}$$

$$\nabla \cdot \mathbf{E} = -\rho. \tag{2.42}$$

Remarques :

1. Nous obtenons des équations de Maxwell tout à fait classique, au signe négatif des seconds membres près. Ce changement est tout à fait insignifiant car il est clair le signe du membre dans lequel interviennent les sources dépend du signe donné au terme d'interaction dans le lagrangien : $\pm A^\mu j_\mu$.
2. Par analogie avec les équations de Maxwell habituelles, nous employons à titre figuratif le symbole $\nabla\times$ pour désigner deux opérateurs différents et distincts de l'opérateur rotationnel à trois dimensions :
   - $\varepsilon_{ij}\,\partial_i E_j$ qui rend un (pseudo-)scalaire à partir d'un vecteur,
   - $\varepsilon_{ij}\,\partial_j B$ qui rend un vecteur à partir d'un (pseudo-)scalaire.

Pour terminer ce paragraphe, notons que l'identité de Bianchi peut encore une fois s'écrire de manière très simple en fonction de l'étoile de Hodge de F, cf. (2.11). On a en effet :

$$\begin{aligned} 0 &= dF && (2.43)\\ &= \frac{1}{2}\,\partial_\lambda F_{\mu\nu}\; dx^\lambda \wedge dx^\mu \wedge dx^\nu && (2.44)\\ &= \frac{1}{2}\,\partial^\lambda F^{\mu\nu}\varepsilon_{\lambda\mu\nu}\; dx^0 \wedge dx^1 \wedge dx^2 && (2.45)\\ &= \partial^\lambda\, {}^*F_\lambda\; dx^0 \wedge dx^1 \wedge dx^2. && (2.46) \end{aligned}$$

Nous pouvons de la même manière que dans le premier paragraphe invoquer, par le lemme de Poincaré, l'existence d'un potentiel et effectuer exactement le même développement. Si nous considérons l'action S, nous voyons que

$$S[A] = \int (-\frac{1}{2}F \wedge\ ^*F + A \wedge j).$$

est une intégrale de volume et donc l'intégrand est à nouveau le lagrangien, celui-ci s'écrivant

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + A_\mu j^\mu.$$

2.4 Analyse de contraintes et conditions de jauge

Ce paragraphe s'applique aux cas quadri- et tri-dimensionnels évoqués précédemment. La quantification du champ électromagnétique étant développée en détail dans le cours d'introduction aux théories de jauge, nous ne ferons que rappeler les éléments indispensables à la suite de cet exposé.

Procédons tout d'abord à une analyse de contrainte. Le lagrangien de l'électromagnétisme, que nous désignerons dans la suite par $\mathcal{L}_{\mathrm{M}}$(lagrangien de Maxwell) est

$$\mathcal{L}_{\mathrm{M}} = -\frac{1}{4} F_{\mu\nu}F^{\mu\nu} + A_\mu j^\mu. \tag{2.47}$$

Les impulsions conjuguées se calculent aisément :

$$\Pi^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_\mu)} = F^{\mu 0}.$$

Nous avons donc[4]

$$\Pi^0 = 0, \tag{2.48}$$
$$\Pi^i = F^{i0} = E_i. \tag{2.49}$$

La relation (2.48) nous fournit notre première contrainte.
On obtient le hamiltonien à partir de $\mathcal{H} = \Pi^i \dot{A}_i - \mathcal{L}$:

$$\mathcal{H}_{\mathrm{M,T}} = \frac{1}{2}[\mathrm{E}^2 + \mathrm{B}^2] - \Pi^k \partial^k A_0 - A_\mu j^\mu + \Lambda \Pi^0, \tag{2.50}$$

où nous avons fait apparaître la contrainte primaire et son multiplicateur de Lagrange.

La contrainte secondaire associée à (2.48) est obtenue par la composante 0 des équations de mouvement (2.11):

$$\partial_\mu F^{\mu 0} = \rho, \tag{2.51}$$

qui n'est jamais qu'un camouflage pour la loi de Gauss :

$$\nabla \cdot \mathbf{E} = \rho. \tag{2.52}$$

La condition de stabilité de cette nouvelle contrainte est équivalente à l'identité triviale $0 = 0$.
On a en résumé:

$$\Pi^0 = 0 \quad \dot{\Rightarrow} \quad \partial_k \Pi^k = \rho \quad \dot{\Rightarrow} \quad 0 = 0,$$

et cette théorie possède deux contraintes de première classe, comme il se doit pour une théorie invariante de jauge. Nous avons donc une contrainte sur le degré de liberté (A_0, Π_0) et une combinaison des degrés (A_k, Π_k). Il reste ainsi deux degrés de liberté à quatre dimensions et un seul à trois dimensions.

Il est possible, en ajoutant une tri-(double-)divergence au hamiltonien de faire apparaître la contrainte secondaire :

$$\mathcal{H}_{\mathrm{M,T}} = \frac{1}{2}[\mathrm{E}^2 + \mathrm{B}^2] + A_0[\partial_k \Pi^k - \rho] + \mathbf{A} \cdot \mathbf{j} + \Lambda \Pi^0. \tag{2.53}$$

4. La note du bas de la page 5 est valable également pour les moments Π^μ.

Traitons, pour terminer, deux jauges particulières que nous utiliserons dans la suite. Les développements seront, pour les mêmes raisons, à nouveau très succincts.

1. *Jauge de Coulomb*: La jauge est fixée en imposant $\partial_l A_l = 0$. La contrainte secondaire associée est obtenue à l'aide du hamiltonien total. La stabilité de la condition de jauge entraîne donc :
$$\Delta A_0 - \partial_l \Pi_l = 0, \tag{2.54}$$
qui n'est jamais, en se rappelant la condition de Gauss, que la loi de Poisson de l'électrostatique. Nous avons donc le système de contraintes suivant :
$$\Pi_0 = 0 \;\; \Rightarrow \;\; \partial_l \Pi_l = \rho \tag{2.55}$$
$$\partial_l A_l = 0 \;\; \Rightarrow \;\; \Delta A_0 = \partial_l \Pi_l \tag{2.56}$$
C'est un système de quatre contraintes de seconde classe pour lequel on applique le mécanisme des crochets de Dirac. Tous calculs faits, on obtient
   - les grandeurs de première classe associées :
$$A_k^T = A_k - \partial_k \partial_l \Delta^{-1} A_l, \tag{2.57}$$
$$\Pi_T^k = \Pi^k - \partial^k \partial^l \Delta^{-1} \Pi^l, \tag{2.58}$$
   avec
$$\Delta^{-1} f(x) = -\frac{1}{4\pi} \int d^3y \frac{1}{|\vec{x} - \vec{y}|} f(y);$$
   - les commutateurs à temps égaux
$$[A_k^T(x),\, \Pi_T^l(y)]_{x_0=y_0} = i(\delta_k^l + \partial_k \partial^l \Delta^{-1})\, \delta^{(3)}(\vec{x} - \vec{y}), \tag{2.59}$$
$$[A_k^T(x),\, A_T^l(y)]_{x_0=y_0} = 0, \quad [\Pi_T^k(x),\, \Pi_T^l(y)]_{x_0=y_0} = 0. \tag{2.60}$$

2. *Jauge temporelle*: En imposant $A_0 = 0$, nous obtenons le système de contraintes suivant
$$\Pi_0 = 0 \;\; \Rightarrow \;\; \partial_l \Pi_l = \rho \tag{2.61}$$
$$A_0 = 0 \;\; \Rightarrow \;\; \Lambda = 0 \tag{2.62}$$
On applique le mécanisme des crochets de Dirac sur les contraintes de seconde classe $\Pi_0 = 0$ et $A_0 = 0$, et celles-ci sont réalisées. Il nous reste en main la condition de Gauss, que nous ne pouvons pas réaliser. Le degré de liberté qui lui est associé reste libre mais nous devrons alors imposer cette contrainte sur les états physiques :
$$(\nabla \cdot \mathbf{\Pi} - \rho)|\Psi_{phys} >= 0. \tag{2.63}$$
Pour terminer, réécrivons le hamiltonien dans cette jauge particulière :
$$\mathcal{H}_{\mathrm{M}} = \frac{1}{2}[\mathrm{E}^2 + \mathrm{B}^2] + \mathbf{A} \cdot \mathbf{j}, \tag{2.64}$$
et l'on retrouve la densité d'énergie du champ électromagnétique tout à fait standard. Ayant posé $A_0 = 0$, nous voyons que le crochet qu'il multipliait dans le hamiltonien, c'-à-d la condition de Gauss peut prendre effectivement une valeur quelconque.

Nous avons à présent tout ce qu'il nous faut pour attaquer l'étude du lagrangien de Chern-Simons.

2.5 Introduction du terme de Chern-Simons

Sans nous préoccuper des origines géométriques et topologiques du terme de Chern-Simons, auxquelles nous consacrerons un chapitre entier, nous étudierons ici de manière systématique les conséquences de l'ajout de ce terme dans le lagrangien habituel de l'électromagnétisme. En d'autres termes, nous nous limiterons pour l'instant à la version abélienne (U(1)) de la théorie.

Repartons de la remarque de la section 2.3 relative à l'identité de Bianchi, que nous avions à ce moment réécrite sous la forme :

$$\partial_\mu \ {}^*F^\mu = 0. \tag{2.65}$$

Le 3-vecteur ${}^*F^\mu$ est donc une quantité conservée. Si nous le couplons au champ A_μ, nous savons que l'action modifée reste invariante de jauge (cf section 2.2 : passage de (2.26) à (2.28)). Nous écrivons donc

$$\mathcal{L}_{\mathrm{MCS}} = \mathcal{L}_{\mathrm{M}} + \frac{\kappa}{2} A_\mu \ {}^*F^\mu. \tag{2.66}$$

Au total,

$$\mathcal{L}_{\mathrm{MCS}} = -\frac{1}{4} F_{\mu\nu}F^{\mu\nu} + A^\mu j_\mu + \frac{\kappa}{2} \varepsilon_{\mu\nu\rho} A^\mu \partial^\nu A^\rho \tag{2.67}$$

Ne perdons pas de vue que les composantes de ${}^*F$ ne sont que les composantes des champs électrique et magnétique. On a

$$\begin{aligned} {}^*F^0 &= F_{12} = -B && (2.68)\\ {}^*F^1 &= F_{20} = -E_2 && (2.69)\\ {}^*F^2 &= F_{01} = E_1 && (2.70) \end{aligned}$$

2.5.1 Une théorie "massive" invariante de jauge

Entamons notre étude par une simple analyse dimensionnelle. A trois dimensions[5] , on a

$$[\mathcal{L}] = \mathrm{M}^3, \quad [A_\mu] = \mathrm{M}^{1/2},$$

et on en déduit que la constante de couplage κ a la dimension d'une masse. Le premier effet de l'ajout de notre terme de Chern-Simons est d'introduire une masse dans la théorie de l'électromagnétisme à trois dimensions, indépendamment de tout procédé, comme celui de Higgs, par exemple. Cette propriété et l'origine topologique de ce terme font que la théorie est appelée "topological massive gauge theory" en anglais.

2.5.2 Equations de mouvement

L'expression des équations de champs est, comme on peut le vérifier aisément

$$\partial_\nu F^{\mu\nu} - \frac{\kappa}{2} \varepsilon^{\mu\alpha\beta} F_{\alpha\beta} = j^\mu, \tag{2.71}$$

ou encore

$$\partial_\nu F^{\mu\nu} - \kappa \ {}^*F^\mu = j^\mu. \tag{2.72}$$

Ce résultat peut se retenir facilement en notant que

1. le premier terme correspond aux équations de Maxwell traditionnelles,

5. Dans la littérature, on rencontre également le lagrangien $\mathcal{L} = -\frac{1}{4e^2} F_{\mu\nu}F^{\mu\nu} + \ldots$, où l'on introduit la constante de couplage e^2 de dimension 1 au dénominateur de manière à conserver des champs A^μ de dimension 1. La constante de couplage κ est alors sans dimension. La masse apparaissant dans la théorie est alors κe^2 (voir chapitre 5).

2. le deuxième terme provient de deux contributions; d'une part la contribution du terme de Chern-Simons en tant que couplage d'une source $\frac{\kappa}{2}\ {}^*F_\mu$ avec A^μ(il apparaît ainsi aux côtés de j_μ), d'autre part la contribution provenant des dérivées de A_μ(facteur $\partial_\nu A_\rho$ du terme de C.-S.).

On écrit donc

$$\partial_\nu F^{\mu\nu} - \frac{\kappa}{4}\,\varepsilon^{\mu\alpha\beta}F_{\alpha\beta} = j^\mu + \frac{\kappa}{4}\,\varepsilon^{\mu\alpha\beta}F_{\alpha\beta}.$$

En terme des champs, on a

$$\begin{aligned}\partial_k F^{0k} - \frac{\kappa}{2}B &= \rho, &(2.73)\\ \dot{E}^1 + -\partial_2\mathrm{B} + \kappa E_2 &= j^1, &(2.74)\\ \dot{E}^2 + \partial_1\mathrm{B} - \kappa E_1 &= j^2. &(2.75)\end{aligned}$$

Et finalement,

$$\nabla\cdot\mathbf{E} = -\rho + \kappa\mathrm{B}, \tag{2.76}$$

$$\begin{aligned}\frac{\partial}{\partial t}E^1 - (\nabla\times B)^1 &= (\mathbf{j})^1 - \kappa E_2, &(2.77)\\ \frac{\partial}{\partial t}E^2 - (\nabla\times B)^2 &= (\mathbf{j})^2 + \kappa E_1. &(2.78)\end{aligned}$$

Nous voyons que les champs électrique et magnétique apportent une contribution aux termes de sources.

Il est possible, en manipulant les équations de mouvement, de montrer que la constante de couplage joue effectivement le rôle d'une masse dans la théorie.
Réécrivons (2.71) en fonction de ${}^*F$. Il vient

$$\varepsilon^{\mu\nu\alpha}\partial_\nu\ {}^*F_\alpha + \kappa\ {}^*F^\mu = j^\mu,$$

et après quelques manipulations,

$$[g^{\mu\nu} + \varepsilon^{\mu\nu\beta}\frac{\partial_\beta}{\kappa}]\ {}^*F_\nu = \frac{j^\mu}{\kappa}. \tag{2.79}$$

En appliquant $[g_{\alpha\mu} - \varepsilon_{\alpha\mu\gamma}\frac{\partial^\gamma}{\kappa}]$ aux deux membres de l'équation précédente, il vient (en utilisant, au cours du calcul, l'identité de Bianchi)

$$(\Box + \kappa^2)\ {}^*F_\alpha = \kappa(g_{\mu\alpha} - \varepsilon_{\alpha\mu\gamma}\frac{\partial^\gamma}{\kappa})j^\mu. \tag{2.80}$$

Une autre propriété importante pour la physique du problème s'obtient facilement à partir des équations de champ, plus précisément à partir de la loi de Gauss, que l'on intègre sur les variables d'espace :

$$\int d^2x\ B = \frac{Q}{\kappa} \tag{2.81}$$

Le flux magnétique est ainsi directement relié à la charge électrique et à la constante de couplage.

2.5.3 Analyse de contraintes

Les impulsions conjuguées s'obtiennent très aisément :

$$\Pi^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_\mu)} = F^{\mu 0} + \frac{\kappa}{2}\,\varepsilon^{\rho 0 \mu} A_\rho. \tag{2.82}$$

Explicitement,

$$\Pi^0 = 0 \tag{2.83}$$

$$\Pi^1 = F^{10} + \frac{\kappa}{2} A_2, \quad E^1 = \Pi^1 - \frac{\kappa}{2} A_2 \tag{2.84}$$

$$\Pi^2 = F^{20} - \frac{\kappa}{2} A_1, \quad E^2 = \Pi^2 + \frac{\kappa}{2} A_1 \tag{2.85}$$

La relation (2.83) est une contrainte du système, les deux autres relations n'en sont pas, vu que la dérivée temporelle du champ A y intervient, au travers des termes F^{10} et F^{20}.
La stabilité de la contrainte primaire (2.83) est, comme toujours, donnée par la composante temporelle des équations de mouvement :

$$\partial_\nu F^{0\nu} - \frac{\kappa}{2}\,\varepsilon^{ij} F_{ij} = \rho,$$

c'est à dire

$$\nabla \cdot \mathbf{E} = \rho + \kappa \mathrm{B}. \tag{2.86}$$

En termes des impulsions, on a

$$\nabla \cdot \mathbf{\Pi} - \frac{\kappa}{2} B = \rho. \tag{2.87}$$

2.5.4 Hamiltonien

Le calcul du hamiltonien est assez fastidieux mais néanmoins très instructif. En partant directement de $\mathcal{H} = \Pi^i \dot{A}_i - \mathcal{L}$, on doit aboutir à l'expression suivante :

$$\mathcal{H}_{\mathrm{MCS,T}} = \frac{1}{2}[\Pi^1 - \frac{\kappa}{2} A_2]^2 + \frac{1}{2}[\Pi^2 + \frac{\kappa}{2} A_1]^2 + \Pi_k \partial_k A_0 - A_0 \frac{\kappa}{2} B + \frac{1}{2}\mathrm{B}^2 - A_\mu j^\mu + \Lambda \Pi^0.$$

En utilisant (2.84) et (2.85), il vient

$$\mathcal{H}_{\mathrm{MCS,T}} = \frac{1}{2}[\mathrm{E}^2 + \mathrm{B}^2] + A_0[\partial_k \Pi_k - \frac{\kappa}{2}\mathrm{B} - \rho] + \mathbf{A} \cdot \mathbf{j} + \Lambda \Pi^0 A_0, \tag{2.88}$$

où nous avons, comme lors d'un calcul précédent, fait apparaître la condition de Gauss en ajoutant une double-divergence.

Les conséquences de l'ajout du terme de C.-S. au lagrangien classique sont évidentes :

1. comme nous l'avons déjà signalé , une mise sur le même pied des champs $(\mathbf{E}, \mathrm{B})$ et des sources,
2. une redéfinition du champ électrique au travers des relations (2.84) et (2.85) .

Techniquement parlant, cette théorie de jauge est tout à fait similaire à celle étudiée précédemment.

2.5.5 Etude du terme C.-S. isolé

Examinons à présent ce que devient la théorie lorsqu'on lui enlève le terme cinétique :

$$\mathcal{L}_{\mathrm{CS}} = \frac{\kappa}{2}\,\varepsilon_{\mu\nu\rho}\,A^\mu \partial^\nu A^\rho + A_\mu j^\mu. \tag{2.89}$$

Les équations de mouvement s'écrivent à présent

$$\frac{\kappa}{2}\varepsilon_{\mu\nu\rho}F^{\nu\rho} = -j_\mu, \tag{2.90}$$

ou encore

$$\kappa\ {}^*F^\mu = -j^\mu, \tag{2.91}$$

relation que nous discuterons dans quelques lignes.

La disparition du terme cinétique a des conséquences importantes au niveau des contraintes. En effet, le système ((2.83) - (2.85)) s'écrit à présent

$$\Pi^0 = 0 \quad (K0) \tag{2.92}$$

$$\Pi^1 = \frac{\kappa}{2}A_2 \quad (K1) \tag{2.93}$$

$$\Pi^2 = -\frac{\kappa}{2}A_1 \ (K2) \tag{2.94}$$

La contrainte secondaire de (2.92) est comme toujours donnée par la composante temporelle de (2.90) :

$$\kappa[\partial_1 A_2 - \partial_2 A_1] + \rho = 0 \quad (K3). \tag{2.95}$$

Nous pouvons faire apparaître ces contraintes dans le hamiltonien, qui s'écrit

$$\mathcal{H}_{\mathrm{CS,T}} = \Lambda\Pi^0 + \Lambda_1(\Pi^1 - \frac{\kappa}{2}A_2) + \Lambda_2(\Pi^2 + \frac{\kappa}{2}A_1) - \frac{\kappa}{2}\varepsilon_{\mu i\rho}A^\mu\partial^i A^\rho - A_\mu j^\mu, \tag{2.96}$$

équivalent, à une double-divergence près, à

$$\mathcal{H}_{\mathrm{CS,T}} = \Lambda\Pi^0 + \Lambda_1(\Pi^1 - \frac{\kappa}{2}A_2) + \Lambda_2(\Pi^2 + \frac{\kappa}{2}A_1) - A_0[\kappa[\partial_1 A_2 - \partial_2 A_1] + \rho] + \mathbf{A}\cdot\mathbf{j}. \tag{2.97}$$

On peut à présent calculer les contraintes secondaires de (2.98 - 2.99)

$$\Pi^1 - \frac{\kappa}{2}A_2 = 0 \ \Rightarrow\ \Lambda_2 + \partial^2 A^0 - \frac{j^1}{\kappa} = 0 \tag{2.98}$$

$$\Pi^2 + \frac{\kappa}{2}A_1 = 0 \ \Rightarrow\ \Lambda_1 - \partial^1 A^0 + \frac{j^2}{\kappa} = 0 \tag{2.99}$$

Ces deux relations déterminent les multiplicateurs de Lagrange Λ_1 et Λ_2. Nous avons donc au total un système de quatre contraintes : la contrainte de première classe habituelle (2.92) et un ensemble de trois contraintes de seconde classe (2.93), (2.94) et (2.95) . Nous allons montrer que l'on peut recouvrer une théorie de jauge tout à fait familière. Il suffit pour cela de réaliser les deux contraintes (2.93), (2.94) au moyen des crochets de Dirac. Le calcul est immédiat, car il suffit d'inverser une matrice de 2×2, et l'on obtient

$$K_3^D = K_3 + \partial_2 K_2 + \partial_1 K_1, \tag{2.100}$$

qui donne explicitement

$$K_3^D = \kappa B + \rho - \partial_k \Pi_k - \frac{\kappa}{2}B, \tag{2.101}$$

qui n'est jamais que la condition de Gauss habituelle

$$\nabla \cdot \mathbf{\Pi} - \frac{\kappa}{2} B = \rho. \tag{2.102}$$

En résumé , on réalise les contraintes (2.93) et (2.94) au moyen des crochets de Dirac, et on fixe ensuite la jauge sur les deux contraintes de première classe restantes, de manière tout à fait habituelle : jauge de Coulomb, de Lorentz, ou de Weyl. Nous terminerons cette séquence par le calcul des crochets de Dirac.

Le hamiltonien (2.96) devient

$$\mathcal{H}'_{CS,T} = \Lambda \Pi^0 - A^0[\frac{\kappa}{2} B + \rho] + \Pi^1 \partial_1 A^0 + \Pi^2 \partial_2 A^0 - \frac{2}{\kappa} \Pi^2 j_1 + \frac{2}{\kappa} \Pi^1 j_2, \tag{2.103}$$

et, comme toujours, à une double-divergence près,

$$\mathcal{H}'_{CS,T} = \Lambda \Pi^0 + A^0[\partial_k \Pi_k - \frac{\kappa}{2} B - \rho] - \frac{2}{\kappa}[\Pi^2 j_1 - \Pi^1 j_2]. \tag{2.104}$$

Cette forme a l'avantage de mettre en évidence notre contrainte K_3 modifiée, qui, comme à l'habitude, a A_0 comme multiplicateur de Lagrange. Nous la transformerons une dernière fois, en choisissant d'y faire apparaître les champs (A_0, Π_0) et (A_1, Π_1).

$$\mathcal{H}'_{CS,T} = \Lambda \Pi^0 - A_0[2\partial_1 \Pi^1 + \kappa \partial_2 A^1 + \rho] - A^1 j_1 + \frac{2}{\kappa} \Pi^1 j_2. \tag{2.105}$$

Ecrivons ensuite les crochets de Dirac

$$\begin{aligned}
\{A_1, \Pi^1\}_D &= \frac{1}{2}\, \delta^{(2)}(\vec{x} - \vec{y}) && (2.106)\\
\{A_0, \Pi^0\}_D &= \delta^{(2)}(\vec{x} - \vec{y}) && (2.107)\\
\{A_1, A_2\}_D &= \frac{1}{\kappa}\, \delta^{(2)}(\vec{x} - \vec{y}) && (2.108)\\
\{\Pi^1, \Pi^2\}_D &= \frac{\kappa}{4}\, \delta^{(2)}(\vec{x} - \vec{y}) && (2.109)
\end{aligned}$$

Les autres crochets sont nuls. Remarquons avec (2.107) que ces crochets ne sont pas compatibles avec (2.92). Il faut donc poser une condition de jauge exactement comme nous l'avons fait pour la théorie de Maxwell.

En conclusion, nous sommes parvenus à justifier le titre de ce chapitre : la théorie de Chern-Simons, couplée ou non à la théorie de Maxwell est une théorie de jauge à part entière. Nous reviendrons sur quelques caractéristiques et sur le lien avec la théorie des particules de statistique fractionnaire au cours du chapitre 5.

Chapitre 3

Géométrie, topologie et théories de jauge

3.1 Introduction

Les concepts de base de la topologie et de la géométrie différentielle sont rappelés dans les deux appendices. Nous aborderons au cours de ce chapitre des concepts qui ne sont généralement pas évoqués dans le cadre d'un cours de licence. Quelques notions de topologie algébrique seront introduites : homotopie, groupe fondamental et groupes d'homotopie multi-dimensionnels. Nous envisagerons ensuite brièvement la théorie des groupes et algèbres de Lie sous un éclairage géométrique. Plus précisément, nous envisagerons le groupe de Lie comme une variété et traduirons des concepts familiers (définition de l'algèbre, représentations, ...) dans le langage de la géométrie différentielle. La plus grande partie de ce chapitre sera accordée à une introduction à la théorie des fibrés. Ceux-ci introduisent un formalisme dans lequel peuvent s'exprimer les théories de jauge que l'on rencontre en physique théorique : la théorie de Yang-Mills et la relativité généralisée. Ayant en vue des applications à la théorie des champs, nous nous consacrerons uniquement à l'étude des fibrés principaux, sur lesquels nous définirons, comme on le fait dans le cadre de la géométrie riemannienne, les notions de connexion, courbure et transport parallèle. Enfin, quelques exemples viendront illustrer les concepts théoriques que nous aurons introduits.

3.2 Notions de topologie algébrique - groupes d'homotopie

La topologie algébrique permet d'étudier certaines situations purement topologiques à l'aide de concepts algébriques, essentiellement la *théorie des groupes*. En particulier, le problème de l'homéomorphisme entre deux espaces topologiques, peut être abordé en introduisant des groupes dits d'*homotopie, homologie, ...*

On introduit souvent cette notion par l'examen d'une figure semblable à la suivante :

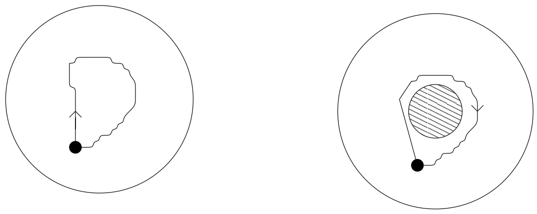


FIG. 3.1 –

Dans le cas de la figure de gauche, on peut intuitivement comprendre qu'il est possible de rétracter le lacet en un point[1]. Il est évident qu'il en va de même pour tout lacet construit à partir de tout point de la figure. Par contre, dans celle de droite, il existe deux types de lacets : le lacet rétractable et celui contenant ce que nous appelerons une *obstruction topologique*. Cet exemple suggère qu'en étudiant les classes d'équivalence de lacets dans une variété, il est possible de déceler des obstructions au sein de cette variété.

3.2.1 Définitions

M désigne un espace topologique dans les paragraphes suivants.

L'étude des chemins, courbes, lacets a été effectuée en détail au cours d'analyse de candidature. Nous introduisons ici des notions nouvelles comme le produit de deux lacets, le lacet inverse, et l'homotopie de deux lacets. Rappelons simplement que la courbe est l'objet physique, que l'on décrit à l'aide d'un paramétrage, le chemin, sous la forme d'applications continues sur un intervalle compact de $\mathbb{R}$. Comme il est possible de paramétrer une courbe de différentes manières, à une courbe peut correspondre plusieurs chemins différents.

Le produit de deux lacets α et β basés en un même point de M est défini par

$$\gamma(t) = \begin{cases} \alpha(2\,t) & \text{si} \quad 0 \leq t \leq \frac{1}{2} \\ \beta(2\,t-1) & \text{si} \quad \frac{1}{2} < t \leq 1 \end{cases}$$

et s'écrit $\gamma = \alpha * \beta$.

Le lacet inverse d'un lacet α est défini par

$$\alpha^{-1}(t) = \alpha(1-t), \quad 0 \leq t \leq 1.$$

Le lacet constant basé en x_0 de M est l'application

$$c(t) = x_0, \quad 0 \leq t \leq 1.$$

Deux lacets α et β sont **homotopes** s'il existe une application continue $H : [0,1] \times [0,1] \to M$ telle que

$$\begin{aligned} H(t,0) &= \alpha(t), \quad 0 \leq t \leq 1, \\ H(t,1) &= \beta(t), \quad 0 \leq t \leq 1, \\ H(0,s) &= H(1,s) = x_0, \quad 0 \leq s \leq 1. \end{aligned}$$

Deux lacets sont donc homotopes s'il est possible de les déformer continuement l'un dans l'autre. On note $\alpha \simeq \beta$. Le mot homotopie est parfois utilisé abusivement pour désigner l'application H.

On montre facilement que l'homotopie est une relation d'équivalence, ce qui nous autorise à travailler avec les classes plutôt qu'avec les éléments eux-mêmes. Nous noterons par $[\alpha]$ la classe d'équivalence de représentant α.

3.2.2 Groupe fondamental

L'ensemble des classes d'homotopie de lacets basés en $x_0 \in M$ se note $\pi_1(M,\, x_0)$. La relation d'homotopie vérifiant la propriété suivante,

$$\alpha_0 \simeq \beta_0,\ \alpha_1 \simeq \beta_1 \Rightarrow \alpha_0 * \alpha_1 \simeq \beta_0 * \beta_1,$$

1. Certaines langues sont plus évocatrices que d'autres pour décrire une même situation. Ainsi en anglais : "a loop which can be *schrunk* to a point"

on peut définir le produit interne pour deux classes $[\alpha]$ et $[\beta]$ de $\pi_1(M, x_0)$ par

$$[\alpha] \circ [\beta] = [\alpha * \beta].$$

$\pi_1(M, x_0)$ est un groupe pour la loi de composition interne que nous venons de définir. On l'appelle **premier groupe d'homotopie** ou **groupe fondamental** de M, basé en x_0.

Le seul problème de cette définition est qu'elle dépend du point x_0 choisi comme base des lacets. Cet handicap peut être éliminé dans le cas d'un espace topologique connexe par arcs. On a le théorème suivant :

Si M est un espace topologique connexe par arcs, alors pour tous x et $y \in M$, les groupes $\pi_1(M, x)$ et $\pi_1(M, y)$ sont isomorphes.

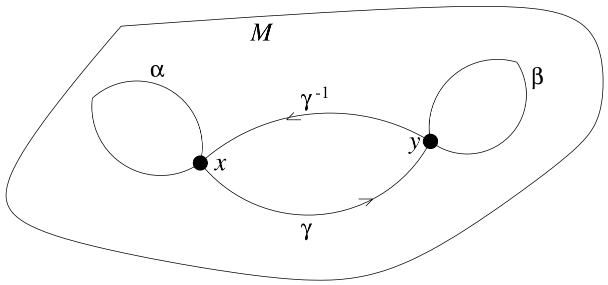


FIG. 3.2 – *La courbe γ reliant les deux points x et y de l'espace connexe par arcs M permet de réaliser une bijection entre les deux groupes fondamentaux de M basés en x et y.*

Comme M est connexe par arcs, on peut utiliser une courbe joignant x à y pour construire une application entre les groupes fondamentaux :

$$\pi_1(M, x) \to \pi_1(M, y), \quad ([\alpha], x) \mapsto ([\gamma^{-1} * \alpha * \gamma], y),$$

et de même dans l'autre sens.

Exemple Sans effectuer aucun calcul, nous pouvons , en partant de la définition du groupe fondamental et par intuition, calculer $\pi_1(S^2)$. Quel que soit le lacet que nous dessinions sur la surface de la sphère, nous pouvons toujours déformer celui-ci vers son point base, en ne rencontrant aucun obstacle. On a donc $\pi_1(S^2) = \{e\}$. De même, on a

$$\pi_1(S^n) = \{e\}, \quad n \geq 2.$$

En quoi le groupe fondamental d'un espace topologique nous permet-il de faire de la topologie? Tout d'abord, nous pouvons *élargir la notion d'homotopie des lacets aux applications d'un espace topologique sur un autre.* Pour deux applications f et g entre deux espaces topologiques, il suffit de trouver une suite continue d'applications f_n intermédiaires qui permettent de passer continument de f à g. On notera $f \simeq g$. Enfin, on peut introduire la notion d'homotopie pour des espaces topologiques : on dit que deux espaces topologiques M et P sont du **même type d'homotopie** s'il existe $f : M \to P$ et $g : P \to M$ telles que $f \circ g \simeq \mathrm{id}|_P$ et $g \circ f \simeq \mathrm{id}|_M$. Cette notion est moins forte que celle d'homéomorphisme : deux espaces homéomorphes sont du même type d'homotopie mais la réciproque n'est pas vraie [2]. On a enfin le théorème suivant :

2. cf. la note du bas de la page 19 et l'exemple suivant

Si M et P sont deux espaces topologiques du même type d'homotopie, alors leurs groupes fondamentaux sont isomorphes. En particulier, les groupes fondamentaux de deux espaces topologiques homéomorphes sont isomorphes.

Il découle de ce théorème que la notion d'homotopie est un invariant topologique.

Enfin, nous introduisons une notion tout à fait fondamentale en topologie, la notion d'espace simplement connexe.

Un espace topologique connexe M est dit **simplement connexe** *si on a $\pi_1(M) = \{e\}$.*

3.2.3 Exemple d'espaces homotopiquement équivalents - Rétraction

Essayons de rendre les notions que nous venons d'introduire un peu plus concrètes. L'exemple peut-être le plus parlant pour illustrer l'homotopie de deux espaces est celui de la rétraction d'un espace sur un autre.

Un sous-espace topologique R connexe par arcs d'un espace topologique connexe par arcs E est un **rétract** *de E, s'il existe une application continue r, appelée* **rétraction** *de E sur R telle $r|_R = \mathrm{id}_R$ ou, de manière équivalente, $r(a) = a$ pour tout $a \in R$ et telle que r soit homotope à id_E.*

L'homotopie d'application entre r et id_E de la définition est réalisée par l'homotopie suivante :

$$\begin{aligned} H(x,0) &= x, \\ H(x,1) &= r(x), \quad \text{pourtout } x \in E \\ H(a,t) &= a, \quad a \in R, \quad t \in [0,1]. \end{aligned}$$

En $t = 0$, H est l'identité sur E et en bout de course H est la rétraction r. Nous pouvons à présent affirmer que E et son rétract R sont du même type d'homotopie, et en déduire immédiatement le résultat suivant.

Si M est un espace topologique connexe par arcs, et R une rétraction de M, alors $\pi_1(M,a) = \pi_1(R,a),\ a \in R$.

Un espace connexe par arcs pour lequel un de ses points est une rétraction, est dit *contractile*. Le groupe fondamental d'un espace contractile est isomorphe au groupe trivial.

Considérons l'espace topologique $\mathbb{R}^n$. L'origine 0 en est un rétract obtenu par la déformation suivante :

$$H(x,t) = (1-t)x, \quad t \in [0,1], \quad x \in \mathbb{R}^n.$$

$\pi_1(\mathbb{R}^n, 0)$ est donc isomorphe au groupe trivial $\{e\}$, c'est un exemple d'espace simplement connexe.

Un deuxième exemple nous est donné par la figure suivante.

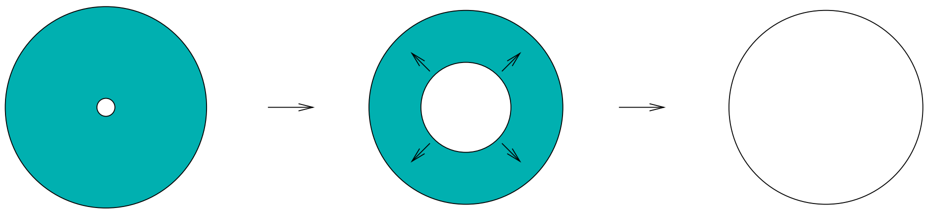


FIG. 3.3 – *De* $\mathbb{R}^2 - \{0\}$ *à* S^1

S^1 est ici une rétraction d'un disque auquel on a retiré l'origine et est obtenu par la déformation :

$$H(x,t) = (1-t)x + t\frac{x}{|x|}.$$

Les deux figures extrêmes possèdent donc le même groupe fondamental, que nous allons à présent apprendre à calculer.

3.2.4 Théorème fondamental

Nous allons montrer qu'il est possible de calculer le groupe fondamental d'un espace topologique en le remplaçant par des figures simples, plus exactement par des polyèdres. Tout d'abord, quelques définitions.

1. *Soient* $x_1, \ldots, x_{m+1}$ *des points distincts dans* $\mathbb{R}^n$*. Ils sont dits* **indépendants** *si les* m *vecteurs*
$$x_2 - x_1,\ \ldots,\ x_{m+1} - x_1$$
sont linéairement indépendants.
2. *Un* **m-simplex** σ^m *est un ensemble de points indépendants de* $\mathbb{R}^n$ *défini par*
$$\sigma^m = \left\{ x = \sum_{i=1}^{m+1} \lambda_i x_i \ \text{ t.q. } \ \lambda_i \geq 0, \sum_{i=1}^{m+1} \lambda_i = 1 \right\}.$$
On écrit souvent $\sigma^m = [x_1, \ldots, x_{m+1}]$*. Les "sommets" des simplex sont appelés des* **vertex** *et les "côtés" les* **faces**.

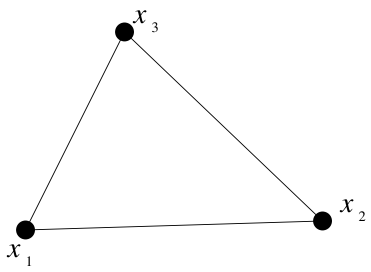


FIG. 3.4 – *Un 3-simplex dans* $\mathbb{R}^2$

3. *Un* **complexe** K *est une collection de simplex* σ^p *satisfaisant les conditions suivantes.*
   - *Si* $\sigma^p \in K$*, alors toutes les faces y appartiennent aussi.*
   - *Si* σ^p *et* $\sigma^q \in K$*, alors soit* $\sigma^p \cap \sigma^q = \emptyset$*, soit ils ont une face en commun.*

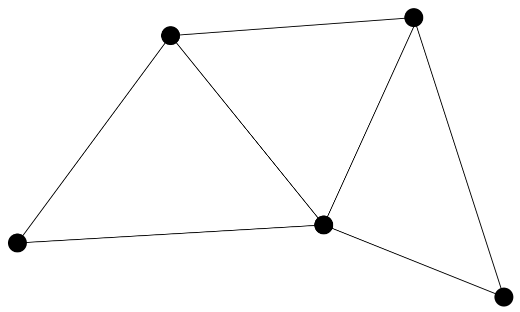


FIG. 3.5 – *Un exemple de complexe*

4. $\dim(K) = \sup_p\{\dim(\sigma^p)\}$.
5. *K est* **connexe par arcs**, *si pour tout couple de vertex u et v de K, il existe une séquence de vertex $v_0, \ldots, v_n$ telle que $v_0 = u$ et $v_n = v$ et que $[v_i, v_{i+1}]$ est un 1-simplex contenu dans K.*
6. *L'espace topologique K muni de la topologie naturelle est le* **polyèdre** *associé à K.*

Nous ne donnons pas l'énoncé mathématique brut du théorème fondamental de calcul mais nous le présentons directement sous la forme d'une méthode.

Le principe de cette méthode est, comme nous l'avons dit, de remplacer l'espace topologique dont on désire calculer le groupe fondamental par un polyèdre qui lui est homéomorphe. Cette opération porte le nom de *triangulation* et si elle possible, l'espace de départ est dit *triangulable*. On considère en fait deux polyèdres : K, obtenu par triangulation et L, un sous-polyèdre unidimensionnel, contractile et contenant tous les vertex de K. A partir de ces deux espaces, on obtient un groupe G dont on obtient les générateurs[3] par les processus suivants :

1. On numérote chaque vertex du polyèdre K et on les désigne par un ensemble de symbôles a_i.
2. On associe un symbole g_{ij} à chaque 1-simplex $[a_i, a_j]$ de K. Si $[a_i, a_j] \in L$, alors on pose $g_{ij} = 1$.
3. Pour chaque 2-simplex $[a_i, a_j, a_k] \in K \backslash L$, on écrit une relation $g_{ij} g_{jk} g_{ik}^{-1} = 1$.

On peut alors montrer que G est isomorphe au groupe fondamental du polyèdre.

Exemple 1 Soit à calculer $\pi_1(S^1)$. On peut se convaincre facilement que cet espace est triangulable par un ensemble de trois 2-simplexes.

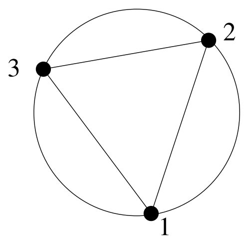


FIG. 3.6 – *Triangulation de S^1*

3. Rappelons que les éléments d'un sous-ensemble $\{x_j\}$ de G sont les générateurs du groupe si tout élément de G peut s'écrire $g = x_1^{i_1} \cdots x_n^{i_n}$. Aux générateurs, on ajoute des contraintes sous forme de *relations* (mot consacré) les unissant. La donnée des générateurs et des relations est la *présentation* de G(ce type de groupe porte le nom de "groupe libre de type fini").

Les deux polyèdres sont donnés par

$$K = [1] \cup [2] \cup [3] \cup [1,2] \cup [2,3] \cup [1,3],$$
$$L = [1,3] \cup [2,3].$$

G contient dans ce cas trois générateurs : g_{12}, g_{13}, g_{23}, mais seul le premier n'est pas inclus dans L. On a donc

$$g_{12} = g,$$
$$g_{13} = g_{23} = 1.$$

Le groupe fondamental du polyèdre et partant, de l'espace S^1 est isomorphe au groupe $\mathbb{Z}$ pour l'addition : $\pi_1(S^1) = \mathbb{Z}$.

Exemple 2 Vérifions que le groupe fondamental d'un disque de rayon unité D^1, espace contractile, est effectivement le groupe trivial. On doit cette fois ajouter au polyèdre précédent le 3-simplexe $[1,2,3]$, ce qui entraîne l'apparition d'une contrainte sur les générateurs

$$g_{12} g_{23} g_{13}^{-1} = 1,$$

qui, combinée aux relations de l'exemple précédent, impose $g_{12} = 1$. D'où la conclusion.

Exemple 3 Soit le disque de rayon unité auquel on a retiré l'origine. Cet espace est par rétraction homotopiquement équivalent à S^1. On en déduit immédiatement que

$$\pi_1(D - \{0\}) = \mathbb{Z}.$$

Il s'agit ici d'un exemple typique où le groupe d'homotopie met en évidence une **obstruction topologique**.

Exemple 4 Traitons pour terminer le ruban de Möbius, que nous retrouverons dans la suite. L'étape la plus délicate, lorsque les espaces deviennent plus complexes, est celle de la triangulation.

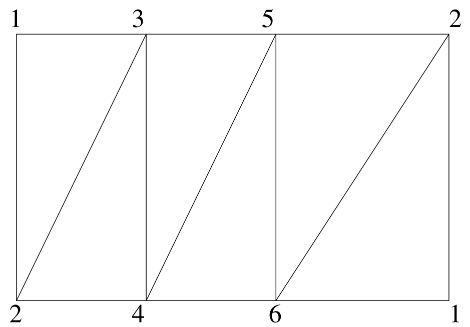


FIG. 3.7 – *Une triangulation du ruban de Möbius*

On a

$$K = [1] \cup [2] \cup [3] \cup [4] \cup [5] \cup [6] \cup [1,2] \cup [1,3] \cup [1,6] \cup [2,3] \cup [2,4] \cup [2,5] \cup [2,6] \cup [3,4] \cup [3,5]$$
$$\cup [4,5] \cup [4,6] \cup [5,6] \cup [1,2,3] \cup [1,2,6] \cup [2,3,4] \cup [2,5,6] \cup [3,4,5] \cup [4,5,6]$$

et

$$L = [1,2] \cup [2,3] \cup [3,4] \cup [4,5] \cup [5,6].$$

Nous laissons au lecteur l'algèbre des générateurs. Il reste en tout et pour tout un seul générateur et on en conclut que $\pi_1(Möbius) = \mathbb{Z}$.

3.2.5 Groupe fondamental d'un produit cartésien

Si nous effectuions le calcul pour le tore, nous trouverions, avec un peu de peine, car la triangulation n'est pas évidente, le résultat suivant

$$\pi_1(T^2) = \mathbb{Z} \oplus \mathbb{Z}.$$

Si par contre nous considérons le tore comme le produit cartésien $S^1 \times S^1$, nous obtenons ce résultat immédiatement, en recourant au théorème suivant

$$\pi_1(A \times B \times \ldots \times Z) = \pi_1(A) \oplus \pi_1(B) \oplus \ldots \oplus \pi_1(Z). \tag{3.1}$$

Le groupe fondamental du cylindre s'obtient de la même manière :

$$\pi_1(S^1 \times [a, b]) = \mathbb{Z} \oplus \{e\} = \mathbb{Z}.$$

3.2.6 Groupes d'homotopie n-dimensionnels

Dans le paragraphe précédent, nous avons classé des espaces topologiques à partir des propriétés d'homotopie de lacets. Le but de cette section est de montrer que l'on peut attribuer, à l'instar de π_1, d'autres groupes à un espace, en étudiant les classes d'homotopie d'autres objets, comme les sphères par exemple.

Soit I l'intervalle $[0,1]$ et E un espace topologique. $I_n = \{x \in \mathbb{R}^n \text{ t.q. } x_i \in I\}$ est par convention le pavé unitaire n-dimensionnel. Une application continue $\alpha : I_n \to E$ est un chemin de dimension n dans E. Si de plus, $\alpha(\partial I_n) = \{x_0\}$, on parle d'un lacet[4] de dimension n de point-base $x_0 \in E$.

A deux dimensions, le lacet bi-dimensionnel peut encore se représenter comme dans la figure suivante,

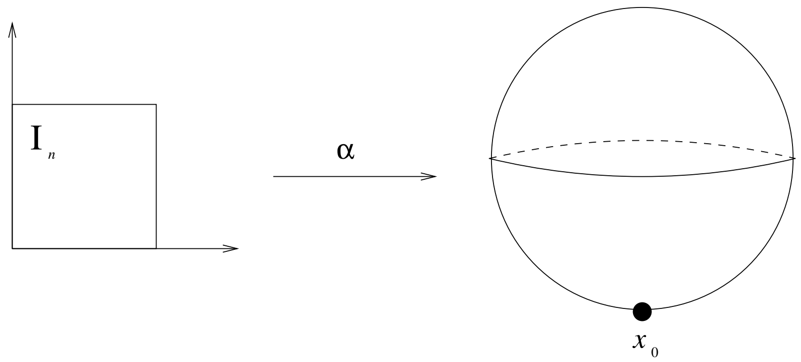


FIG. 3.8 – *Un "lacet" bi-dimensionnel dans* $\mathbb{R}^3$.

On peut définir de la même manière que précédemment le produit de lacets n-dimensionnels de même point-base, le lacet inverse, l'homotopie de deux lacets, montrer qu'il s'agit d'une relation d'équivalence, introduire les classes d'homotopie et finalement aboutir à la définition du $n^{\text{ème}}$ groupe d'homotopie : $\pi_n(E, x_0)$.
Passons en revue les propriétés de ces groupes, en insistant sur les différences par rapport au cas précédent.

– Les groupes $\pi_n(E)$ sont abéliens pour $n > 1$. On a en effet

$$[\alpha] * [\beta] = [\beta] * [\alpha],$$

4. Dans le cas unidimensionnel, on a $\partial I = \{0, 1\}$ et effectivement $\alpha(0) = \alpha(1) = \{x_0\}$.

pour deux lacets n-dimensionnels α et β. Ce résultat se démontre en déformant judicieusement les pavés sous-jacents au premier produit.

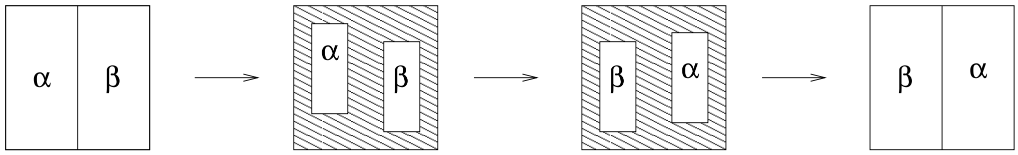


FIG. 3.9 – *Commutativité du produit de deux lacets*

Lors des étapes intermédiaires, les zones hachurées sont envoyées sur le point-base commun aux deux lacets. On peut comprendre que π_1 n'est pas nécessairement abélien car de telles transformations sur la frontière du pavé sont impossibles à une dimension (le pavé unidimensionnel est, rappelons-le, un segment de droite).

- Pour tout couple de points $\{x_0, x_1\}$ d'un espace connexe par arcs E, on a

$$\pi_n(E, x_0) = \pi_n(E, x_1).$$

- Les $n^{\text{ème}}$ groupes d'homotopie de deux espaces du même type d'homotopie (défition analogue à celle du paragraphe précédent) sont isomorphes. En particulier, les $n^{\text{ème}}$ groupes d'homotopie sont des invariants topologiques.
- On conserve la propriété vis-à-vis du produit cartésien :

$$\pi_n(A \times \ldots Z) = \pi_n(A) \oplus \ldots \oplus \pi_n(Z).$$

Exemple Malheureusement, il n'existe pas d'algorithme pour calculer les groupes d'homotopie multidimensionnels. Tout calcul se fait au cas par cas.

Dans le cas des hyper-sphères, notre intuition peut encore nous aider à interpréter le résultat suivant.

$$\pi_n(S^n) = \mathbb{Z}, \quad n \geq 1.$$

Travaillons à deux dimensions. Remarquons tout d'abord que $I^2 \backslash \partial I^2$ est homéomorphe[5] à S^2. Nous pouvons donc le paramétrer à l'aide des coordonnées “sphériques” comme l'indique la figure suivante.

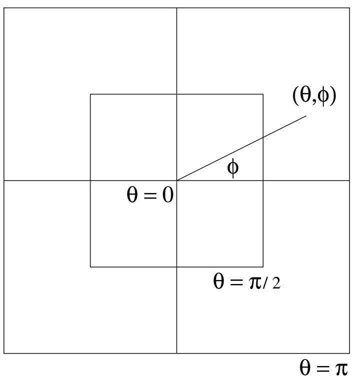


FIG. 3.10 – *Paramétrage “sphérique” du pavé I^2*

5. Rappelons encore une fois que l'homéomorphisme est une bijection, ce qui nous oblige à retirer la frontière du pavé.

Un 2-lacet sera donc donné par une application du type

$$\alpha : S^2 \to S^2, \quad (\theta, \phi) \mapsto (\theta', \phi').$$

Si α envoie S^2 sur un point x_0 de S^2, $[\alpha]$, la classe d'homtopie de α correspond à l'élément neutre du groupe $\pi_2(S^2)$, c-à-d l'élément 0 de $\mathbb{Z}$. Une deuxième possibilité est donnée par $\alpha : \theta' = \theta, \phi' = \phi$. Lorsqu'un petit cancrelat fait le tour de la sphère de départ, le petit cancrelat primé effectue lui aussi un tour sur sa sphère. α appartient à la classe $+1$ de $\pi_2(S^2)$. Si nous considérons ensuite l'application $\alpha : \theta' = \theta$, $\phi' = 2\phi$, le cancrelat primé risque d'attraper le tournis deux fois plus rapidement que son collègue. Nous venons construire la classe $+2$ du groupe d'homotopie. Nous pouvons ainsi mettre en correspondance les classes d'homotopie avec le groupe des entiers en considérant plus généralement l'application $\alpha : \theta' = \theta$, $\phi' = k\phi$.

Nous donnerons d'autres exemples lorsque nous aborderons les propriétés topologiques des groupes de Lie.

3.3 Groupes et algèbres de Lie

Dans l'appendice B, nous introduisons la notion de variété. En dotant une variété d'un produit interne, nous pouvons en faire un groupe, sous les conditions habituelles, et nous obtenons ce que l'on appelle un groupe de Lie.

3.3.1 Groupes de Lie

Définition *Un* **groupe de Lie** G *est une variété différentiable dotée d'une structure de groupe telle que les opérations*

$$\begin{aligned} \cdot \ &: \ G\times G\to G, \quad (g_1,g_2)\mapsto g_1\cdot g_2, \\ ^{-1} \ &: \ G\to G, \quad g\mapsto g^{-1}, \end{aligned}$$

soient différentiables.

Exemples

1. $\mathbb{R}$ est un groupe de Lie pour l'addition.
2. $\mathbb{R}_0^+$ est un groupe de Lie pour la multiplication.
3. S^1 est un groupe de Lie pour les opérations : $\mathrm{e}^{i\theta}\mathrm{e}^{i\varphi}=\mathrm{e}^{i(\theta+\varphi)}$ et $\left(\mathrm{e}^{i\theta}\right)^{-1}=\mathrm{e}^{-i\theta}$. Il s'agit ni plus ni moins du groupe U(1).
4. Le groupe linéaire général des matrices inversibles GL(n,$\mathbb{R}$) (ou GL(n,$\mathbb{C}$)) et ses sous-groupes[6] bien connus : O, SO, U, SU, SL sont des groupes de Lie.
   Etudions plus particulièrement le cas de SU(2). Une matrice 2 sur 2 s'écrit $U=\begin{pmatrix} a & b \\ c & d\end{pmatrix}$, où les éléments sont des nombres complexes. La condition d'unitarité et celle du déterminant imposent les relations
   $$a^*a+c^*c=1, \quad b=-c^*, \quad d=a^*.$$
   D'où une matrice de SU(2) s'écrit de manière tout à fait générale
   $$U=\begin{pmatrix} x_1+ix_2 & -x_3+ix_4 \\ x_3+ix_4 & x_1-ix_2\end{pmatrix},$$
   avec $\sum x_i^2=1$, ce qui nous permet d'identifier SU(2) à S^3. Il s'agit donc d'un groupe connexe et compact.

3.3.2 Représentations d'un groupe de Lie

Définition *Une représentation d'un groupe de Lie est la donnée d'une variété M et d'une application différentiable*

$$G\times M\to M, \quad (g,x)\mapsto \rho_g x,$$

telle que pour tout élément g du groupe G, ρ_g soit un difféomorphisme de M tel que

$$\rho_e=\mathrm{id}_M,$$

$$\rho_g\circ\rho_{g'}=\rho_{gg'}.$$

Une représentation s'appelle aussi *action à gauche*, que l'on abrège par $\rho_g x=gx$. L'*action à droite*[7] se définit de la même manière, sauf que l'on a $\rho_g\circ\rho_{g'}=\rho_{g'g}$ et $\rho_g x=xg$.

6. Un théorème important en théorie des groupes de Lie affirme que tout sous-groupe fermé d'un groupe de Lie est un groupe de Lie
7. Ce n'est pas une représentation

Un peu de vocabulaire relatif aux représentations :
Soit G un groupe de Lie qui agit sur une variété M par $\rho : G \times M \to M$. Cette représentation est dite

- *transitive* si pour tout $p_1,\ p_2 \in M$, il existe $g \in G$ tel que $\rho_g p_1 = p_2$.
- *libre* si, lorsqu'il existe un élément $p \in M$ tel que $\rho_g p = p$ alors g est automatiquement l'élément identité e.
- *effective* si e est le seul élément de G à réaliser l'action triviale sur M.

Comme G est avant tout une variété, on peut envisager les représentations de G sur lui-même :

- *Translation à gauche* : $(g, h) \mapsto L_g h = gh$,
- *Translation à droite* : $(g, h) \mapsto R_g h = hg$,
- *Représentation* adjointe : $(g, h) \mapsto \mathrm{ad}_g h = ghg^{-1}$.

On peut montrer que ces actions sont libres et transitives.

Si la variété M est un espace vectoriel V et l'application ρ est linéaire, on dit que la représentation est *linéaire*. En d'autres termes, il s'agit d'un homéomorphisme de groupes : $G \to \mathrm{GL}(V)$, groupe des transformations linéaires sur l'espace vectoriel V. Une représentation linéaire est dite *irréductible* si V ne contient pas de sous-espace invariant, c-à-d tels que $\rho_g W \subset W$, pour tout $g \in G$.

3.3.3 Algèbre de Lie

Définitions La définition que nous allons donner d'une algèbre de Lie fait appel à la notion de champ vectoriel invariant.

Un champ vectoriel A sur un groupe de Lie G est **invariant (à gauche)** *si, pour tous* $h, g \in G, \quad L_{g*}A|_h = A|_{gh}$.

L'application $_*$ (push-forward) étant linéaire, toute combinaison linéaire de champs vectoriels invariants (c.v.i.)est un champ vectoriel invariant. On montre ensuite, mais nous ne le ferons pas, que le crochet de Lie de deux champs vectoriels invariants est encore un champ vectoriel invariant. Donc l'ensemble des c. v. i. sur G est une sous-algèbre de Lie de l'algèbre des champs vectoriels sur G, $\mathrm{vect}(G)$.

L'ensemble des champs vectoriels invariants à gauche d'un groupe de Lie G, noté $\mathfrak{g}$, muni du crochet de Lie $[\,,\,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ est appelé **algèbre de Lie** *du groupe G.*

Un vecteur $V \in T_eG$ définit un champ vectoriel unique invariant à gauche A_V par

$$A_V|_g = L_{g*}V, \quad g \in G.$$

En effet, $A_V|_{ag} = L_{ag*}V = (L_a L_g)_* V = L_{a*}L_{g*}V = L_{a*}A_V|_g$. Inversément, un champ vectoriel invariant à gauche A définit un vecteur unique $V = A|_e \in T_eG$. On a ainsi un isomorphisme entre les vecteurs tangents à l'identité de G et l'ensemble des champs vectoriels invariants à gauche de G. On retrouve donc la notion plus familière d'algèbre de Lie en tant qu'espace tangent à l'élément identité. On a alors

$$\dim(\mathfrak{g}) = \dim(G).$$

De plus, **tout groupe de Lie G est parallélisable**, car à partir de vecteurs tangents à l'identité de G linéairement indépendants, on peut construire une base pour les vecteurs invariants à gauche en n'importe quel point du groupe.

Les éléments d'une base de $\mathfrak{g}$ sont les *générateurs* de l'algèbre et sont traditionnellement notés $T_a,\ T_b,\ T_c,\ \ldots$. Nous rencontrerons les notations suivantes

$$A = \sum_a A^a T_a, \qquad [T_a, T_b] = f^c_{ab} T_c,$$

pour tout élément A de l'algèbre et où les f^c_{ab} sont les *constantes de structure.*

3.3.4 Représentation d'une algèbre de Lie

Examinons en particulier la représentation adjointe d'une algèbre de Lie.
En notant que $\mathrm{ad}_g e = e$, pour tout $g \in G$, on restreint le push-forward ad_{g*} induit par la représentation adjointe, c-à-d l'application $\mathrm{ad}_{g*} : T_h G \to T_{\mathrm{ad}_g h} G$, à l'élément identité : $h = e$. Explicitement,

$$\mathrm{Ad}_g = \mathrm{ad}_{g*}|_{T_e G} : T_e G \to T_e G.$$

Il s'agit d'une représentation linéaire de G sur $\mathfrak{g}$ vu que $T_e G$, isomorphe à $\mathfrak{g}$, est un espace vectoriel et que l'application est linéaire. On obtient donc la représentation suivante :

$$\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g}.$$

Explicitement, nous avons, pour tout élément V de $\mathfrak{g}$:

$$\mathrm{Ad}_g V = \frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{ad}_g \exp(tV)]\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\exp(tgVg^{-1})\Big|_{t=0} = gV^{-1}g, \quad g \in G.$$

A partir de cette représentation linéaire du groupe, on définit la représentation adjointe de l'algèbre :

$$\mathrm{Ad}_A B = [A, B],$$

obtenue en se limitant à l'espace tangent pour g dans l'expression précédente.

3.3.5 Forme de Maurer-Cartan

Nous avons jusqu'à présent travaillé dans l'espace tangent avec les champs vectoriels. Rien ne nous empêche de considérer des formes et de passer dans la formulation duale.
Etant donné une base $\{V_\mu\}$ de générateurs d'une algèbre de Lie $\mathfrak{g}$ associée à un groupe de Lie G, introduisons une base $\{\theta^\mu\}$ duale de $\{V_\mu\}$. En utilisant la relation (B.3), on obtient l'équation de structure dite de **Maurer-Cartan**

$$\mathrm{d}\theta^\mu = -\frac{1}{2} f^\mu_{\nu\lambda}\, \theta^\nu \wedge \theta^\lambda, \tag{3.2}$$

où $f^\mu_{\nu\lambda}$ sont les constantes de structure. Cela étant, on définit une 1-forme θ, à valeur dans $\mathfrak{g}$:

$$\theta : T_g G \to T_e G, \quad X \mapsto (L_{g^{-1}})_* X = (L_g)_*^{-1} X.$$

Démontrons ensuite les deux relations suivantes

$$\begin{aligned} \theta &= V_\mu \otimes \theta^\mu, && (3.3)\\ \mathrm{d}\theta &+ \frac{1}{2}[\theta, \theta] = 0 && (3.4) \end{aligned}$$

où $\{V_\mu\}$ est une base de $T_e G$ et $\{\theta^\mu\}$ une base de $\Lambda^1(T_g G)$.
Soit $Y = Y^\mu X_\mu \in T_g G$, où $\{X_\mu\}$ est le référentiel associé à $\{V_\mu\}$ par push-forward vers l'élément

g du groupe .
On a

$$\theta(Y) = Y^\mu \theta(X_\mu) = Y^\mu L_{g*}^{-1}(L_{g*}V_\mu) = Y^\mu V_\mu,$$

et bien sûr

$$(V_\mu \otimes \theta^\mu)(Y) = Y^\nu V_\mu \theta^\mu(X_\nu) = Y^\mu V_\mu,$$

et (3.3) est vérifié. La relation (3.4) s'obtient immédiatement si on note que $\mathrm{d}\theta = V_\mu \otimes \mathrm{d}\theta^\mu$, que $[\theta, \theta] = [V_\mu, V_\nu] \otimes \theta^\mu \wedge \theta^\nu$, et en utilisant l'équation de Maurer-Cartan (3.2).
θ porte le nom de **forme de Maurer-Cartan** et réapparaîtra lorsque nous discuterons la courbure sur un fibré, notion que nous introduisons dans la section suivante.

3.4 Fibrés

Nous évoquons maintenant une notion importante de la géométrie différentielle moderne. Nous allons voir qu'elle permet la description d'objets géométriques "palpables", comme le ruban de Möbius, que nous prendrons comme exemple, ou la bouteille de Klein, mais qu'elle laisse la place à un formalisme particulièrement adapté aux théories de jauge de la théorie des champs, au travers d'objets plus abstraits, comme les fibrés principaux. Passons directement à la définition, que nous examinerons en détail à l'aide d'un exemple.

3.4.1 Définition

Un **fibré** est la donnée de

1. Une variété différentiable E, appelée **espace total**.
2. Une variété différentiable M, appelée **espace de base**.
3. Une variété différentiable F, appelée **fibre**.
4. Une surjection $\pi : E \to M$, appelée **projection**. L'image inverse $\pi^{-1}(p) = F_p$ est appelée **fibre au point** p, $p \in M$.
5. Un groupe G, appelé **groupe de structure**, qui est représenté sur F par la gauche.
6. Un recouvrement ouvert $\{U_i\}$ de M et, pour chaque élément U_i du recouvrement un difféomorphisme $\phi_i : U_i \times F \to \pi^{-1}(U_i)$ tel que pour tout $p \in M$ et pour tout $f \in F$, $\pi\phi_i(p, f) = p$. L'application ϕ_i est appelée **factorisation locale**[8], comme ϕ_i^{-1} envoie $\pi_i^{-1}(U_i)$ (partie de l'espace total) sur le produit cartésien $U_i \times F$.
7. Une famille $\{g_{ij}\}$ d'applications telle que si $U_i \cap U_j \neq \emptyset$, $g_{ij} : U_i \cap U_j \to G$, et construite comme suit. Si nous écrivons $\phi_i(p, f) = \phi_{i,p}(f)$ pour $p \in M$ fixé, l'application $\phi_{i,p} : F \to F_p$ est un difféomorphisme. Sur l'intersection non-vide de deux ouverts du recouvrement, $U_i \cap U_j$, on demande que $g_{ij}(p) = \phi_{i,p}^{-1} \circ \phi_{j,p} : F \to F$ soit un élément de G. Donc pour tout $p \in U_i \cap U_j$, $g_{ij}(p)$ est un élément du groupe G agissant sur la fibre F par la gauche. Les fonctions $\{g_{ij}\}$ sont appelées **fonctions de transition**.

8. En anglais, **locale trivialisation**

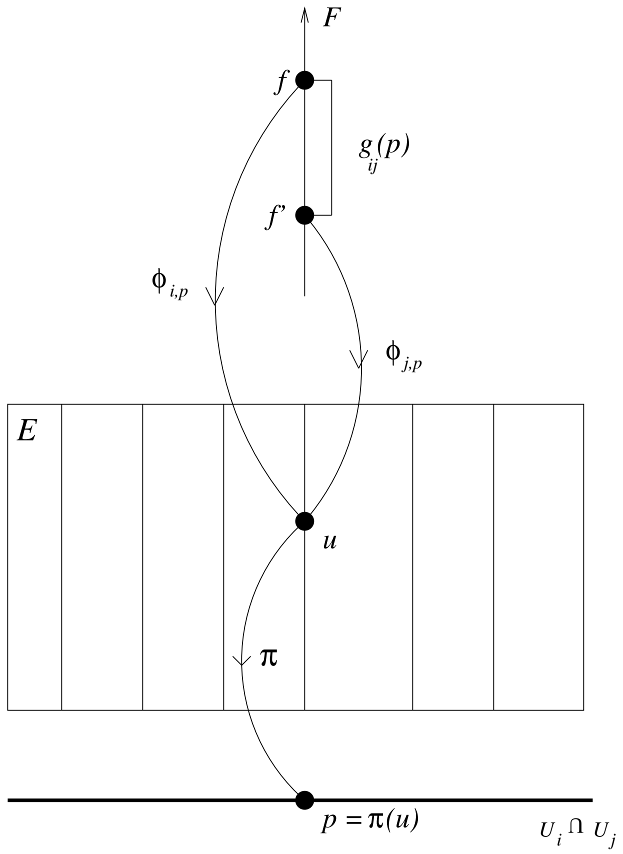


FIG. 3.11 – *Illustration graphique de la définition - En chaque point de l'espace de base M (pour lequel on se donne un recouvrement), est attaché une copie de la variété F, la fibre. Les applications ϕ_i^{-1}, attachées chacune à un ouvert du recouvrement, "factorisent" localement l'espace total E en $U_i \times F$. Sur l'intersection de deux ouverts, représentée ici, deux factorisations sont possibles. La manière dont elles sont reliées sont matérialisées par les fonctions de transition g_{ij} appartenant au groupe de structure G. En d'autres termes, la fibre F sert d'espace de représentation pour G.*

Avant de passer aux exemples, quelques remarques sur les fonctions de transition du point 7.

- Soit un point u de l'espace total E, tel que $\pi(u) = p$ dans M. Si p appartient à l'intersection $U_i \cap U_j$ de deux ouverts du recouvrement de l'espace de base M, on a deux factorisations possibles de E "au-dessus" du point p : $u = \phi_j(p, f)$ ou $u = \phi_i(p, f')$. La relation entre les éléments de la fibre au-dessus de p est assurée par la fonction de transition $g_{ij}(p)$, telle que $u = \phi_j(p, f) = \phi_i(p, g_{ij}(p)f)$. Le rôle des fonctions de transition est d'indiquer comment les factorisations locales $U_i \times F$ et $U_j \times F$ doivent être "collées" [9] l'une à l'autre.
- Les fonctions de transition vérifient les relations suivantes

$$\begin{aligned} g_{ij}(p) &= g_{ji}(p)^{-1}, \quad p \in U_i \cap U_j, \\ g_{ij}(p)g_{jk}(p) &= g_{ik}(p), \quad p \in U_i \cap U_j \cap U_k. \end{aligned}$$

9. A nouveau, la langue anglaise est plus suggestive pour décrire la situation : "The functions of transition indicate how the local trivialisations are to be *glued together*."

Exemple : le cylindre et le ruban de Möbius

Ces deux exemples de fibrés ont l'avantage de pouvoir être visualisés et même confectionnés à moindre frais, à l'aide d'un simple ruban de papier. Ils nous permettront donc concrétiser chacun des éléments de la définition précédente, et d'insister sur le rôle prépondérant des fonctions de transition dans la construction de l'espace total.

L'espace de base est, dans les deux cas, la variété S^1. On écrira

$$E \xrightarrow{\pi} S^1$$

La fibre F est le segment [-1,1]. Un recouvrement ouvert possible de S^1 est donné par $U_1 =]0, \frac{3\pi}{2}[$ et $U_2 =]-\pi, \frac{\pi}{2}[$. Ces deux ouverts se chevauchent dans les régions indiquées sur la figure suivante.

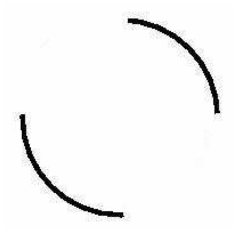


FIG. 3.12 – *Zones de chevauchement*

A désignera la zone $]\pi, \frac{3\pi}{2}[$ et B la zone $]0, \frac{\pi}{2}[$. A ce stade, si nous passons à la partie bricolage, nous avons en main deux rubans de papier de même longeur, qui sont nos ouverts U_1 et U_2. Ces deux rubans correspondent au produit cartésien d'un "trois-quart" de cercle et d'un segment, autrement dit un bout de cylindre. Il nous reste à présent à les coller pour confectionner nos objets. Les zones de collage sont évidemment les zones où nos ouverts se chevauchent. Ce n'est pas la première zone de collage (soit A) qui va déterminer si nous fabriquons un cylindre ou un ruban de Möbius. Mathématiquement, cela revient à écrire

$$g_{12}(\theta) : t \mapsto t, \quad \theta \in A, \; t \in F,$$

$$\phi_1^{-1}(u) = (\theta, t), \quad \phi_2^{-1}(u) = (\theta, t).$$

Par contre, pour la zone B, nous avons deux possibilités :

1. *Cylindre* : Les bouts sont collés comme en A,

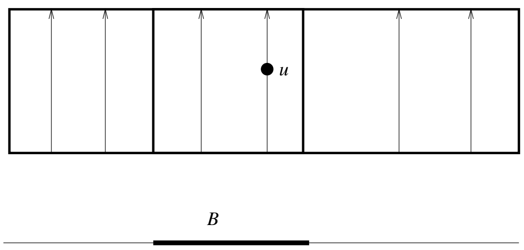


FIG. 3.13 – *Collage de la zone B, option cylindre*

et à nouveau

$$g_{12}(\theta) : t \mapsto t, \quad \theta \in B, \; t \in F,$$

$$\phi_1^{-1}(u) = (\theta, t), \quad \phi_2^{-1}(u) = (\theta, t).$$

2. *Ruban de Möbius*: Les bouts sont collés de la manière suivante,

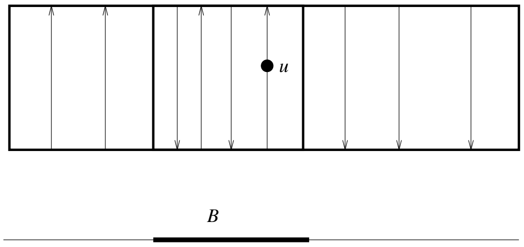


FIG. 3.14 – *Collage de la zone B option ruban de Möbius*

ce qui correspond à prendre

$$g_{12}(\theta) : t \mapsto -t, \quad \theta \in B,\ t \in F,$$

$$\phi_1^{-1}(u) = (\theta, t), \quad \phi_2^{-1}(u) = (\theta, -t).$$

et donne la figure bien connue

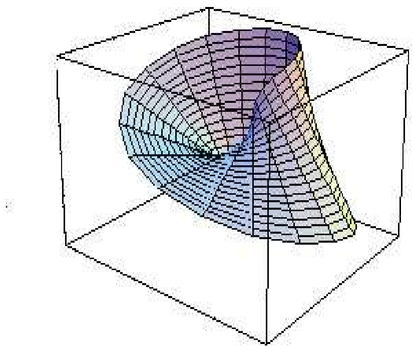


FIG. 3.15 – *Le ruban reconstitué*

Nous voyons que dans le cas du cylindre les fonctions de transitions valent l'identité partout, ce qui correspond au groupe de structure trivial: $G = \{e\}$. Le cylindre est donc **globalement un produit cartésien** de deux variétés: $S^1 \times [-1, 1]$. On dit que l'on a affaire à un **fibré trivial**. Dans le cas du ruban, nous avons un groupe à deux éléments: e pour la zone A, et $g : t \mapsto -t$ pour la zone B. Donc $G = \{e, g\}$ avec $g^2 = e$: il s'agit d'un groupe cyclique à deux élements, le groupe d'inversion. Le ruban n'est que *localement* un produit cartésien.

Pour résumer grossièrement, envisager une variété en tant que fibré revient à déterminer dans quelle mesure celle-ci diffère d'un produit cartésien.

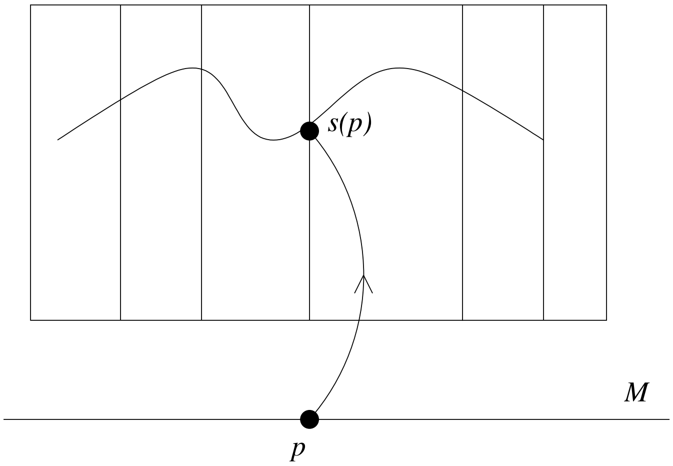
3.4.2 Section

Soit $E \xrightarrow{\pi} M$ *un fibré. Une* **section** *est une application*

$$s : M \to E$$

différentiable telle que $\pi \circ s = \mathrm{id}|_M$.

Elle envoie donc un élément de l'espace de base p sur un élément de la fibre au-dessus de p: $F_p = \pi^{-1}(p)$. Si elle est définie seulement sur un sous-espace $U \subset M$, on parle de *section locale*.



3.4.3 Classification des fibrés

Fibrés vectoriels

Un **fibré vectoriel** $E \xrightarrow{\pi} M$ de dimension n est un fibré dont la fibre est un espace vectoriel de dimension n. Il s'ensuit que le groupe de structure, représenté sur F, espace vectoriel est GL(n,$\mathbb{R}$)(ou GL(n,$\mathbb{C}$)). En effet, les fonctions de transition, agissant sur F, transforment isomorphiquement (ce sont des difféomorphismes) un espace vectoriel en un autre et tombent de ce fait dans le domaine des transformations linéaires.

1. L'exemple le plus important de fibré vectoriel est celui de **fibré tangent**. Si M est l'espace de base, le fibré tangent associé à M est défini par
$$TM = \bigcup_{p\in M} T_pM.$$
La fibre au-dessus d'un point p de M est l'espace tangent T_pM. Déterminons le groupe de structure. Soit u un point de TM tel que $\pi(u) = p \in U_i \cap U_j$ (intersection de deux ouverts du recouvrement de M). Si $\{x^\mu\}$ et $\{y^\mu\}$ sont les coordonnées du point p dans les systèmes attachés à U_i et U_j respectivement, le vecteur V correspondant à u dans T_pM s'exprime $V = V^\mu \partial/\partial x^\mu|_p = \tilde{V}^\mu \partial/\partial y^\mu|_p$, de telle sorte que les factorisations locales sont
$$\phi_i^{-1}(u) = (p, \{V^\mu\}), \quad \phi_j^{-1}(u) = (p, \{\tilde{V}^\mu\}).$$
Les fonctions de transition relient les coordonnées de la fibre par
$$V^\mu = G^\mu_\nu(p)\tilde{V}^\nu,$$
où $\{G^\mu_\nu(p)\} = \{(\partial x^\mu/\partial y^\nu)_p\} \in$ GL(dim(M),$\mathbb{R}$).
2. L'équivalent "dual" de TM est T^*M, défini par
$$T^*M = \bigcup_{p\in M} T_p^*M,$$
et porte le nom de **fibré cotangent**. Le groupe de structure est à nouveau GL(dim(M),$\mathbb{R}$).

Le formalisme des fibrés tangents et cotangents est utilisé en mécanique classique (formulation lagrangienne et hamiltonienne - description de l'espace des phases) ainsi qu'en relativité généralisée.

Fibrés principaux

Un fibré est dit **principal** lorsque la fibre est identique au groupe de structure. On le note souvent $P \xrightarrow{\pi} M$. Le groupe de structure est donc représenté sur lui-même, toujours par la gauche. Nous savons que, dans ce cas, la translation à droite est également une représentation. Etant donné un recouvrement $\{U_i\}$ de M, et une factorisation ϕ_i, l'**action à droite** de G sur $\pi^{-1}(U_i)$ est définie par $\phi_i^{-1}(ua) = (p, g_i a)$, que l'on note très souvent

$$ua = \phi_i(p, g_i a),$$

pour $a \in G$ et $u \in \pi^{-1}(p)$. Etudions cette action particulière.
Tout d'abord, elle ne dépend pas de la factorisation. En effet, si $p \in U_i \cap U_j$,

$$ua = \phi_j(p, g_j a) = \phi_j(p, g_{ji}(p) g_i a) = \phi_i(p, g_i a).$$

Notons ensuite que $\pi(ua) = \pi(u)$. Agir sur un fibré par la droite en un point de l'espace total revient à parcourir la fibre à laquelle ce point appartient. De plus, comme l'action à droite d'un groupe sur lui-même est transitive, on peut écrire, si $\pi(u) = p$

$$\pi^{-1}(p) = \{ua | a \in G\}.$$

Autrement dit, on peut construire toute la fibre à partir d'un de ses points à l'aide de cette opération.
Enfin, l'action à droite nous permet de relier deux sections locales sur un recouvrement. Définissons tout d'abord une section particulière : celle-ci envoie un point de l'espace de base vers un point de la fibre associée de coordonnée e (élément identité du groupe). Explicitement :

$$s_i(p) = \phi_i(p, e).$$

On peut ainsi désigner tout point u de F_p, par $\phi_i(p, g) = \phi_i(p, e)g = s_i(p)g$. Cette factorisation spéciale, reliée à la section locale est la **factorisation canonique**. Cela étant, sur l'intersection de deux ouverts, nous avons la situation suivante :

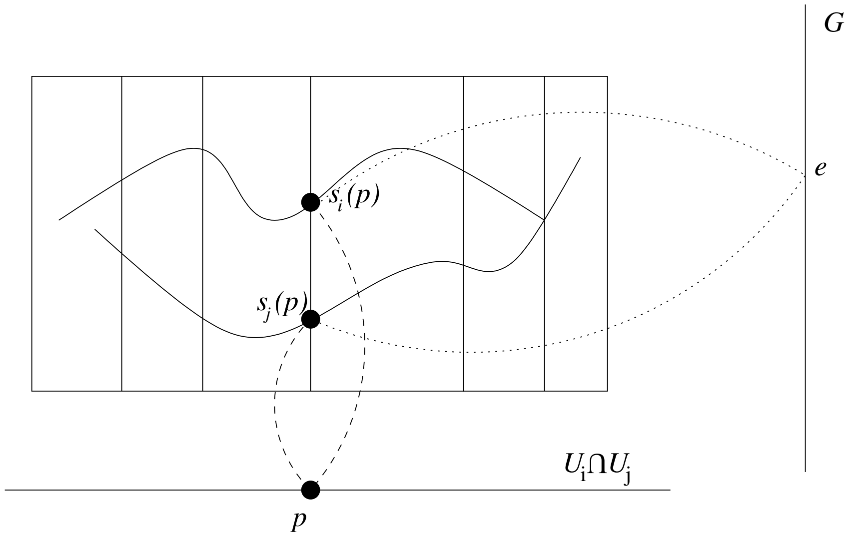


FIG. 3.16 – *Représentation de deux factorisations canoniques locale sur un recouvrement*

Les deux factorisations ϕ_i et ϕ_j envoient l'élément identité sur deux points distincts du fibré. Nous pouvons les relier par une action à droite du type

$$s_j(p) = s_i(p)\gamma(p).$$

Montrons que $\gamma(p)$ n'est rien d'autre que la fonction de structure $g_{ij}(p)$. On a, pour tout $g \in G$,

$$g_{ij}(p)g = \phi_{i,p}^{-1} \circ \phi_{j,p}(g) = \phi_{i,p}^{-1}(s_j(p)g) = \phi_{i,p}^{-1}(s_i(p)\gamma(p)g) = \gamma(p)g,$$

d'où on tire $g_{ij}(p) = \gamma(p)$. Il vient donc finalement

$$s_j(p) = s_i(p)g_{ij}(p), \tag{3.5}$$

et l'on voit que les fonctions de structure, qui agissent naturellement par la gauche, relient cette fois deux sections locales par action à droite.

Cas où un fibré est trivial ...

Terminons par deux résultats importants, permettant d'affirmer la trivialité d'un fibré.

Un fibré $E \xrightarrow{\pi} M$ est trivial si M est simplement connexe.

Par exemple, tous les fibrés construits sur $\mathbb{R}^n$ sont triviaux. Ce résultat relie ainsi la trivialité éventuelle d'un fibré à la *topologie de son espace de base.*

Un fibré principal est trivial si et seulement s'il admet une section globale.

La démonstration de ce résultat est très simple. Si P est trivial, tout élément $u \in P$ peut s'écrire (p, g), $p \in M$, $g \in G$. Dès lors, $s(p) = (p, e)$ est par exemple une section globale. Inversément, supposons que s soit une section globale sur P. Comme l'action à droite est une représentation transitive et libre de G, tout élément $u \in P$ peut s'écrire $u = s(p)g$, avec $p \in M$ et $g \in G$. Donc tout élément du fibré est la donnée d'un élément de l'espace de base et d'un élément de la fibre, d'où la conclusion.

3.5 Connexion sur les fibrés

La notion de connexion est généralement évoquée au cours de relativité générale, où l'on aborde la notion de connexion métrique. Les théories de jauge de la théorie des champs s'inscrivent dans le langage des fibrés principaux sur lesquels nous allons définir des notions similaires de connexion, de courbure, de dérivée covariante. Il existe deux manières de définir la connexion sur un fibré principal, les deux sont bien entendu équivalentes.

Soit u un élément d'un fibré principal $P \xrightarrow{\pi} M$ sur G et soit G_p la fibre au-dessus du point $p = \pi(u)$. Le **sous-espace vertical** *V_uP est un sous-espace de T_uP, tangent à G_p en u.*

Grâce à l'action à droite, nous pouvons nous donner un moyen de construction de V_uP.
Soit A un élément de $\mathfrak{g}$, algèbre de Lie associée au groupe G. L'action à droite

$$u \exp(tA)$$

sur u définit une courbe de paramètre t dans la fibre G_p. Il suffit alors de prendre le vecteur tangent à cette courbe au point u pour avoir un élément de T_uP tangent à la fibre G_p en u, donc un élément de V_uP. Si cette construction est effectuée en chaque point de P, on obtient un champ vectoriel, noté $A^\sharp$, appelé *champ vectoriel fondamental* généré par l'élément A de l'algèbre de Lie. En tout point u de P, on obtient un isomorphisme :

$$\sharp : \mathfrak{g} \to V_uP, \quad A \mapsto A^\sharp$$

entre l'algèbre de Lie et le sous-espace vertical. Un champ vectoriel X de l'espace vertical V_uP peut donc s'écrire $X = A^\sharp$. Nous dirons dans la suite que l'élément de l'algèbre de Lie A *génère* le champ vectoriel X.

L'action de connecter les différentes fibres entre elles revient, soit à construire l'espace manquant, l'espace **horizontal**, soit à contruire explicitement un objet mathématique appelé **connexion**. Les deux procédés sont bien entendu équivalents, comme nous allons le montrer.

Définition

Une **connexion** *sur un fibré principal P est une séparation de T_uP en les sous-espaces verticaux V_uP et H_uP telle que*

1. $T_uP = H_uP \oplus V_uP$,
2. *Un champ vectoriel différentiable X sur P est séparé en champs vectoriels différentiables X^H et X^V,*
3. $H_{ug} = R_{g*}H_uP$, $u \in P$, $g \in G$.

Le premier point de cette définition est clair, le deuxième est quelque peu technique, mais sans importance vraiment géométrique, la troisième stipule que tous les sous-espaces horizontaux de la fibre à laquelle appartient le point u peuvent être générés à partir de H_uP par l'application R_{g*} induite par l'action à droite.

La séparation est réalisée en matérialisant la connexion dans un nouvel objet qui porte le même nom : la 1-forme connexion.

Définition

Une **connexion** *sur P est une 1-forme $\omega \in \Lambda^1(P, \mathfrak{g})$ telle que*

1. $\omega(A^\sharp) = A$ $(A \in \mathfrak{g})$,
2. $R_g^*\omega = \mathrm{Ad}_{g^{-1}}\omega$.

La première condition fait de la connexion ω une projection de T_uP sur l'espace vertical V_uP isomorphe à $\mathfrak{g}$. Grâce à ω, on peut définir l'espace horizontal

$$H_uP = \{X \in T_uP | \omega(X) = 0\}.$$

Il reste à montrer que cette définition est cohérente avec la précédente. Il est clair que ω réalise une séparation de T_uP, seule la troisième assertion de la première définition nécessite une vérification. Nous devons montrer que $R_{g*}H_uP = H_{ug}P$. Soit $X \in H_uP$. On a

$$\omega(R_{g*}X) = R_g^*\omega(X) = g^{-1}\omega(X)g = 0,$$

par hypothèse. Donc $R_{g*}X \in H_{ug}P$.

La connexion ω est un objet défini globalement sur P. Cherchons à présent une réalisation de celui-ci sur l'espace M.

Soit $\{U_i\}$ un recouvrement de M et soit σ_i une section locale définie sur chaque U_i. Une **connexion locale** *associée à une connexion ω est définie par*

$$\mathcal{A}_i = \sigma_i^*\omega, \tag{3.6}$$

et appartient à $\Lambda^1(U_i, \mathfrak{g})$.

Le théorème suivant permet d'effectuer le trajet inverse. Etant donné une collection $\mathcal{A}_i$ de connexions locales, est-il possible de contruire une connexion globale?

Etant donné $\mathcal{A}_i \in \Lambda^1(U_i, \mathfrak{g})$, *et une section locale* σ_i, *il existe une 1-forme connexion* ω *telle que* $\mathcal{A}_i = \sigma_i^* \omega$. *On a, par exemple,*

$$\omega_i = g_i^{-1} \pi^* \mathcal{A}_i g_i + g_i^{-1} \mathrm{d}_P g_i, \tag{3.7}$$

où d_P *est la dérivée extérieure sur* P *et* g_i *provient de la factorisation canonique :* $\phi^{-1}(u) = (p, g_i)$, *pour* $u = \sigma_i(p) g_i$.

Il reste cependant à s'assurer que la connexion ainsi obtenue satisfasse $\omega_i = \omega_j$ sur $U_i \cap U_j$, pour que la séparation en espace horizontal et vertical soit unique sur P et ne dépende pas du recouvrement de M. Pour réaliser cette condition, les connexions locales doivent satisfaire à une transformation particulière.

Pour dériver la loi de transformation des connexions locales, nous devons recourir au lemme suivant.

Soit P *un fibré principal construit sur* M *et* G. $\sigma_i(\sigma_j)$ *est une section locale sur* U_i (U_j). *Pour* $p \in U_i \cap U_j$, *et* $X \in T_p M$, *on a*

$$\sigma_{j*} X = R_{g_{ij}*}(\sigma_{i*} X) + (g_{ij}^{-1} \mathrm{d} g_{ij}(X))^\sharp.$$

Quelques mots d'explication s'imposent. Par l'intermédiaire des sections locales σ_{ij}, nous associons deux points du fibré, plus précisément deux points de la fibre au-dessus d'un point $p \in U_i \cap U_j$. Nous savons que ces deux points ne coïncident pas nécessairement (cf. fig (3.16)). A chacune de ces applications, on peut comme toujours associer une autre application σ_*, agissant sur les espaces tangents. Le vecteur X, tangent en p à l'espace de base, est envoyé vers le fibré, mais vers deux endroits différents, relativement aux deux sections. Le lemme précédent nous apprend comment relier l'action de la section locale σ_j sur X à celle de la section σ_i. Le premier terme est relativement évident, car nous savons depuis la relation (3.5) que les deux points vers lesquels pointent les deux sections locales sont reliées par la fonction de structure $g_{ij}(p)$. Il suffit alors d'appliquer la version de l'action à droite dans l'espace tangent, via le push-forward $_*$: $R_{g_{ij}*}$. La correction donnée par le deuxième terme, grandeur fondamentalement verticale, provient du fait que l'"écart" séparant les deux sections locales (voir fig. 3.16), en des termes plus précis la fonction de structure g_{ij}, varie d'un point à un autre de l'espace de base. En fait, la relation obtenue n'est jamais qu'une application de la formule de Leibnitz. En effet, si nous considérons la courbe $\gamma(t)$ dont X est le vecteur tangent en p, point où nous choisissons $t = 0$, nous avons par définition

$$\begin{aligned} \sigma_{j*} X &= \left. \frac{d}{dt} \left(\sigma_i(\gamma(t)) g_{ij}(\gamma(t)) \right) \right|_{t=0} \\ &= \left\{ g_{ij}(\gamma(t)) \frac{d}{dt} \sigma_i(\gamma(t)) \right\}_{t=0} + \left\{ \sigma_i(\gamma(t)) \frac{d}{dt} g_{ij}(\gamma(t)) \right\}_{t=0}. \end{aligned}$$

Le deuxième terme subit alors plusieurs transformations, qui ne sont pas fondamentalement intéressantes pour l'exposé, l'important étant de comprendre son origine. Le résultat (3.7) s'obtient en appliquant la 1-forme ω sur cette relation. On a successivement

$$\begin{aligned} \sigma_j^* \omega(X) &= R_{g_{ij}}^* \omega(\sigma_{i*} X) + g_{ij}^{-1} \mathrm{d} g_{ij}(X) \\ &= g_{ij}^{-1} \omega(\sigma_{i*} X) g_{ij} + g_{ij}^{-1} \mathrm{d} g_{ij}(X) \end{aligned}$$

où l'on a utilisé les deux conditions de la seconde définition de la connexion. Cette équation se réduit enfin à

$$\mathcal{A}_j = g_{ij}^{-1}\mathcal{A}_i g_{ij} + g_{ij}^{-1}\mathrm{d}g_{ij}, \tag{3.8}$$

que nous pouvons dès à présent associer dans notre esprit aux transformations de jauge de la théorie des champs.

3.6 Transport parallèle

Comme dans le cas des variétés riemanniennes de la relativité générale, le transport parallèle sera défini le long d'une courbe dans le fibré principal P. Celle-ci est obtenue à partir d'une courbe de M par ce que l'on appelle un lift horizontal.

Soit P un fibré principal construit sur l'espace M et le groupe de Lie G. Soit $\gamma : [0,1] \to M$ une courbe dans M. Une courbe $\tilde{\gamma} : [0,1] \to P$ est dit **lift horizontal** *de la courbe γ si $\pi\tilde{\gamma} = \gamma$ et si le vecteur tangent à $\tilde{\gamma}(t)$ appartient à $H_{\tilde{\gamma}(t)}P$ pour toute valeur de t.*

On a ensuite le résultat suivant, que nous allons démontrer, avant d'aborder la notion de transport parallèle.

Soit $\gamma : [0,1] \to M$ une courbe dans M et soit $u_0 \in \pi^{-1}(\gamma(0))$. Il existe un et un seul lift horizontal $\tilde{\gamma}(t)$ dans P tel que $\tilde{\gamma}(0) = u_0$.

Passons donc à la construction du lift d'une courbe. Nous travaillerons dans un ouvert U_i contenant γ, sur lequel nous avons une section σ_i. Notre lift $\tilde{\gamma}$ peut s'exprimer par

$$\tilde{\gamma}(t) = \sigma_i(\gamma(t))g_i(\gamma(t)),$$

grâce à la factorisation canonique. Tout se ramène à la détermination de $g_i(t)$.
Désignons ensuite par $X(t)$ un vecteur tangent à la courbe au point $\gamma(t)$. Son correspondant $\tilde{X}(t)$ est tangent à $\tilde{\gamma}$ en $\tilde{\gamma}(t)$. D'après la définition du lift, $\tilde{X}(t)$ est horizontal pour tout t, donc $\omega(\tilde{X}(t)) = 0$. La relation entre $\tilde{X}$ et $\sigma_{i*}X$ est semblable à celle du lemme du paragraphe précédent :

$$\tilde{X} = g_i(t)^{-1}\sigma_{i*}Xg_i(t) + [g_i(t)^{-1}\mathrm{d}g_i(X)]^\sharp.$$

Appliquant comme la dernière fois la 1-forme connexion ω aux deux membres de cette équation, on obtient

$$0 = g_i(t)^{-1}\omega(\sigma_{i*}X(t))g_i(t) + g_i(t)^{-1}\mathrm{d}g_i(X(t)),$$

pour toute valeur du paramètre t. On transforme cette relation en l'expression suivante

$$\frac{\mathrm{d}g_i(t)}{\mathrm{d}t} = -\omega(\sigma_{i*}X(t))g_i(t),$$

si l'on se rappelle que $\mathrm{d}g_i(X) = \frac{\mathrm{d}g_i(\gamma(t))}{\mathrm{d}t}$. En faisant intervenir la connexion locale, elle s'écrit finalement

$$\frac{\mathrm{d}g_i(t)}{\mathrm{d}t} = -\mathcal{A}_i(X(t))g_i(t) \tag{3.9}$$

La construction du lift se ramène donc à la résolution d'une équation différentielle qui, si on lui ajoute la condition initiale $g_i(0) = e$, possède une solution unique.
Sous forme infinitésimale, cette équation s'écrit

$$g_i(t+\varepsilon) = [1 - \varepsilon\,\mathcal{A}_i(X(t))]g_i(t).$$

Par itérations successives, on obtient

$$g_i(1) = \lim_{n\to\infty}[1 - \frac{1}{n}\mathcal{A}_i(X(\frac{n-1}{n}))]\dots[1 - \frac{1}{n}\mathcal{A}_i(X(\frac{1}{n}))]. \tag{3.10}$$

Rappelons ici que $\mathcal{A}_i$ est une 1-forme à valeurs dans $\mathfrak{g}$. Si $\mathfrak{g}$ n'est pas abélien, rien n'indique que $\mathcal{A}_i(X(t+\varepsilon))$ commute avec $\mathcal{A}_i(X(t))$, et précaution doit être prise dans une expression du type (3.10). On écrira dès lors celle-ci sous la forme

$$g_i(1) = \mathrm{Pexp}\left(-\int_\gamma \mathcal{A}_i\right), \tag{3.11}$$

avec (P pour path-ordered product)

$$\mathrm{Pexp}\left(-\int_\gamma \mathcal{A}_i\right) = \sum_{j=0}^{\infty}(-1)^j\int_0^1 dt_j\int_0^{t_j} dt_{j-1}\dots\int_0^{t_2} dt_1\mathcal{A}_i(X(t_j))\dots\mathcal{A}_i(X(t_2))\mathcal{A}_i(X(t_1)) \tag{3.12}$$

Nous avons à présent tout en main pour aborder la notion de transport parallèle sur un fibré principal.

Soient $\gamma : [0,1] \to M$ une courbe dans M et $u_0 \in \pi^{-1}(\gamma(0))$ un point de la fibre au-dessus de l'origine de la courbe. Comme il existe un et un seul lift horizontal $\tilde{\gamma}$ de γ passant par u_0, il existe un et un seul point $u_1 = \tilde{\gamma}(1) \in \pi^{-1}(\gamma(1))$. Ce point u_1 est appelé **transport parallèle** de u_0 le long de $\tilde{\gamma}$. Le transport parallèle désigne également l'action d'amener u_0 sur u_1 et on dit que u_0 est transporté parallèlement sur u_1. Le transport parallèle est alors dans cette acception l'application

$$\Gamma(\tilde{\gamma}) : \pi^{-1}(\gamma(0)) \to \pi^{-1}(\gamma(1)), \quad u_0 \mapsto u_1 = \sigma_i(\gamma(1))\mathrm{Pexp}\left(-\int_\gamma \mathcal{A}_i\right). \tag{3.13}$$

3.7 Courbure

Pour introduire la notion de courbure, nous devons d'abord construire l'équivalent de la dérivée covariante sur un fibré principal. Comme toute dérivée extérieure, elle doit agir sur des formes définies sur P : $\varphi \in \Lambda^r(P)$. Ces formes agissent elles-mêmes sur des vecteurs tangents au fibré :

$$\varphi : \overbrace{TP\otimes\dots\otimes TP}^{r} \to V.$$

Cependant, nous savons que T_uP en un point $u \in P$ se décompose en H_uP et V_uP si l'on se donne une connexion. Un vecteur $X \in T_uP$ s'écrit $X = X^H + X^V$.

Soient $\varphi \in \Lambda^r(P)$ et $X_1, \dots, X_{r+1} \in T_uP$. La **dérivée covariante** *de φ est définie par*

$$\mathrm{D}\varphi(X_1,\dots,X_{r+1}) = \mathrm{d}_P\varphi(X_1^H,\dots,X_{r+1}^H). \tag{3.14}$$

La **courbure** *Ω est simplement la dérivée covariante de la 1-forme connexion ω*

$$\Omega = \mathrm{D}\omega. \tag{3.15}$$

Il s'agit d'une 2-forme de $\Lambda^2(P,\mathfrak{g})$.
Nous donnons sans démonstration deux propriétés de la 2-forme courbure. La première est l'analogue de la deuxième condition de la définition de la 1-forme connexion, la deuxième est l'importante *équation de structure de Cartan.*

1. On a $R_a^* = a^{-1}\Omega a, \quad a \in G.$

2. Si $X, Y \in T_u P$, on a
$$\Omega(X,Y) = \mathrm{d}_P\omega(X,Y) + [\omega(X), \omega(Y)]$$
qui s'écrit également
$$\Omega = \mathrm{d}_P\omega + \frac{1}{2}[\omega,\omega] \tag{3.16}$$
ou encore
$$\Omega = \mathrm{d}_P\omega + \omega \wedge \omega. \tag{3.17}$$

Interprétation géométrique

Montrons tout d'abord que l'expression de la courbure est particulièrement simple lorsque celle-ci est évaluée sur des vecteurs horizontaux. On a, en effet, par définition ou en utilisant l'équation de Cartan
$$\Omega(X,Y) = \mathrm{d}_P\omega(X,Y)$$
qui s'écrit, par (B.3),
$$\Omega(X,Y) = -\omega([X,Y]), \tag{3.18}$$
pour $X, Y \in H_u P$. Considérons ensuite un parallélogramme $OPQR$ dans M. Les coordonnées des sommets s'écrivent dans un système de coordonnées attaché à une carte de M :

$$\begin{array}{rcl} O & = & (0,0,\ldots,0) \\ P & = & (\varepsilon,0,\ldots,0) \text{ ou } P = \varepsilon V \quad (V = \dfrac{\partial}{\partial x^1}) \\ Q & = & (\varepsilon,\delta,\ldots,0) \text{ ou } Q = \varepsilon V + \delta W \quad (W = V = \dfrac{\partial}{\partial x^2}) \\ R & = & (0,\delta,\ldots,0) \text{ ou } R = \delta W \end{array}$$

$OPQR$ est une courbe fermée γ dans M. Examinons le lift horizontal $\tilde{\gamma}$ dans P. Soient $X, Y \in H_u P$ des vecteurs tangents au lift tels que $\pi_* X = \varepsilon V$ et $\pi_* Y = \delta W$. On a alors
$$\pi_*([X,Y]) = \varepsilon\delta[V,W] = \varepsilon\delta[\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}] = 0,$$
ce qui montre que $[X,Y]$ est une grandeur verticale. Nous avons donc une courbe fermée γ dans M, pour laquelle nous avons un “résidu vertical” exprimé par ce commutateur. La courbe $\tilde{\gamma}$ n'est donc pas fermée dans P : le point initial et le point final appartiennent à la même fibre, celle située au-dessus de O, mais ne coïncident pas. Le commutateur de deux champs vectoriels étant un champ vectoriel, il existe $A \in \mathfrak{g}$ tel que
$$[X,Y] = A^\sharp,$$
avec la notation introduite précédemment pour les champs vectoriels sur le fibré P. La courbure Ω, par son action sur les deux vecteurs X et Y,
$$\Omega(X,Y) = -\omega([X,Y]) = A,$$
nous donne l'élément de l'algèbre de Lie A générant le champ vectoriel $[X,Y]$.

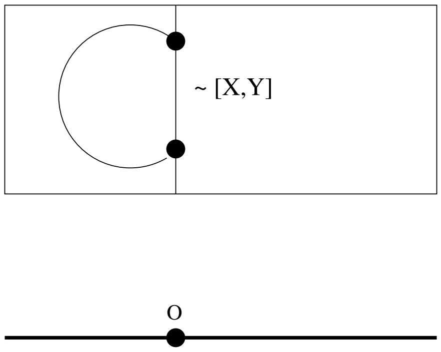


FIG. 3.17 – *La courbure sur un fibré principal, exprime comme dans le cas des variétés riemanniennes le défaut de parallélisme d'un vecteur transporté parallèlement le long d'une courbe fermée.*

De manière plus générale, nous pouvons envisager l'action de la 2-forme courbure Ω. En tant que 2-forme, Ω agit sur deux champs vectoriels de l'espace TP tangent au fibré. En un point u de P, soient X et Y deux vecteurs de T_uP. Ces deux vecteurs peuvent se décomposer en deux parties. L'une, tangente à la fibre à laquelle appartient u, est un élément de V_uP, l'autre appartient à l'espace horizontal H_uP. La courbure, en tant que dérivée covariante, agit uniquement sur la partie horizontale des vecteurs X et Y :

$$\Omega(X, Y) = D\omega(X, Y) = d_P(X^H, Y^H).$$

Les deux vecteurs X^H et Y^H, en tant que vecteurs horizontaux, peuvent être associés à deux vecteurs V et W de T_pM ($\pi(u) = p$, $\pi_*(X) = V$, $\pi_*(Y) = W$), sur lesquels on peut construire, comme précédemment, un parallélogramme. La courbure $\Omega(X, Y)$ rend alors l'élément de l'algèbre de Lie générant le commutateur $[X^H, Y^H] \in V_uP$.

Courbure locale

Nous pouvons, comme pour la connexion, définir une courbure locale au moyen d'une section :

$$\mathcal{F} = \sigma^*\Omega.$$

L'équation de Cartan (3.17) s'écrit, en fonction de la connexion et de la courbure locales

$$\mathcal{F} = \mathrm{d}\mathcal{A} + \mathcal{A} \wedge \mathcal{A}, \tag{3.19}$$

ou bien

$$\mathcal{F} = \mathrm{d}\mathcal{A} + \frac{1}{2}[\mathcal{A}, \mathcal{A}]. \tag{3.20}$$

Attention cependant, lorsqu'on évalue cette forme sur des vecteurs de TM, il vient

$$\mathcal{F}(X, Y) = \mathrm{d}\mathcal{A}(X, Y) + [\mathcal{A}(X), \mathcal{A}(Y)]. \tag{3.21}$$

Si nous passons à présent aux composantes des formes $\mathcal{A}$ et $\mathcal{F}$, on obtient

$$\mathcal{F}_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu + [\mathcal{A}_\mu, \mathcal{A}_\nu]. \tag{3.22}$$

Si l'on développe ces grandeurs sur une base de $\mathfrak{g}$,

$$\mathcal{A}_\mu = A_\mu{}^\alpha T_\alpha, \quad \mathcal{F}_{\mu\nu} = F_{\mu\nu}{}^\alpha T_\alpha \tag{3.23}$$

en utilisant $[T_\alpha, T_\beta] = f_{\alpha\beta}{}^\gamma T_\gamma$, il vient finalement

$$F_{\mu\nu}{}^\alpha = \partial_\mu A_\nu{}^\alpha - \partial_\nu A_\mu{}^\alpha + f_{\beta\gamma}{}^\alpha A_\mu{}^\beta A_\nu{}^\gamma. \tag{3.24}$$

Le théorème suivant nous apprend comment les courbures locales se transforment lorsque l'on passe d'un ouvert du recouvrement de M à l'autre.

Soient U_i et U_j deux ouverts de M tels que $U_i \cap U_j \neq \emptyset$, $\mathcal{F}_i$ et $\mathcal{F}_j$ les courbures locales correspondantes. Sur $U_i \cap U_j$, on a

$$\mathcal{F}_j = g_{ij}^{-1} \mathcal{F}_i g_{ij}, \tag{3.25}$$

où g_{ij} sont les fonctions de transition sur $U_i \cap U_j$.

La preuve consiste simplement à introduire la transformation (3.8) dans l'expression des courbures locales.

Remarque Signalons que $\mathcal{F} = 0$ sur un ouvert U ne signifie pas automatiquement que $\mathcal{A}$, la connexion dont il dérive soit identiquement nulle sur U. En effet, $\mathcal{F}$ est également nul lorsque la connexion est dite *purement de jauge*, c-à-d lorsque

$$\mathcal{A} = g^{-1} \mathrm{d}g. \tag{3.26}$$

Le moment est venu de rappeler que nous avions introduit une 1-forme particulière sur un groupe de Lie, la forme de Maurer-Cartan. Celle-ci nous permet d'envisager une deuxième interprétation de la notion de courbure sur un fibré principal. Il suffit pour cela de se rappeler la relation (3.4), que l'on peut à présent interpréter, en comparant à (3.17) comme étant la courbure associée à la 1-forme de Maurer-Cartan. Celle-ci étant nulle en l'occurence, la courbure d'une connexion ω exprime comment celle-ci diffère d'une forme de Maurer-Cartan([5], p.454).

Terminons par une relation que nous avons déjà rencontrée dans le chapitre précédent, plus exactement dans la section consacrée aux équations de Maxwell. Evaluons l'expression $\mathrm{D}\Omega$. Par définition de la dérivée covariante, on a

$$\mathrm{D}\Omega(X, Y, Z) = \mathrm{d}_P\Omega(X^H, Y^H, Z^H).$$

Il suffit de calculer $\mathrm{d}_P\Omega$ en utilisant la formule de Leibniz

$$\mathrm{d}_P\Omega = \mathrm{d}_P\omega \wedge \omega - \omega \wedge \mathrm{d}_P\omega, \tag{3.27}$$

pour se rendre compte que l'on doit inévitablement évaluer ω sur des vecteurs horizontaux. Il vient donc

$$\mathrm{D}\Omega = 0, \tag{3.28}$$

qui est la version globale de l'**identité de Bianchi**.
Localement, on a

$$\mathrm{d}\mathcal{F} = \mathrm{d}\mathcal{A} \wedge \mathcal{A} - \mathcal{A} \wedge \mathrm{d}\mathcal{A} = \mathcal{F} \wedge \mathcal{A} - \mathcal{A} \wedge \mathcal{F},$$

que l'on réécrit

$$\mathrm{d}\mathcal{F} + \mathcal{A} \wedge \mathcal{F} - \mathcal{F} \wedge \mathcal{A} = \mathrm{d}\mathcal{F} + [\mathcal{A}, \mathcal{F}] = 0. \tag{3.29}$$

En posant

$$\mathcal{D}\eta = \mathrm{d}\eta + [\mathcal{A}, \eta], \tag{3.30}$$

pour tout $\eta \in \Lambda^p(M, \mathfrak{g})$, il vient

$$\mathcal{D}\mathcal{F} = 0. \tag{3.31}$$

La connexion, contenue implicitement dans la définition de la dérivée covariante globale, au travers de X^H, apparaît explicitement dans l'expression locale (3.30).

Pour terminer ce paragraphe, remarquons qu'il est également possible de faire apparaître la courbure au niveau local de deux manières

1. En calculant le transport parallèle (3.13) le long d'une courbe fermée. Voir [6], p.484 pour la démonstration.
2. En déterminant le commutateur de deux dérivées covariantes

$$[\mathcal{D}_\mu, \mathcal{D}_\nu]^\alpha = \mathcal{F}^\alpha_{\mu\nu}.$$

3.8 Théories de jauge - Attribution des rôles

Le lecteur familier de la théorie de jauge de Yang et Mills aura bien entendu reconnu plusieurs grandeurs et lois de transformation connues. Nous allons dans cette section résumer rapidement, car nous laisserons une grande place aux exemples, les liens entre les objets mathématiques présentés jusqu'ici et les grandeurs physiques familières.

Espace physique et groupe de jauge Nous l'avons déjà évoqué à quelques reprises, l'espace physique, en d'autres termes l'espace-temps, que nous doterons d'une topologie triviale ou plus évoluée (voir les exemples ci-dessous) est identifié à l'espace de base M d'un fibré principal P. Le groupe de jauge, le groupe des transformations laissant une théorie invariante est le groupe de structure G du fibré, et également sa fibre, puisque P est principal. En chaque point de l'espace physique, on attache une copie du groupe de jauge et, suivant la topologie de cet espace physique et la manière dont les ouverts du recouvrement se rattachent les uns aux autres, le fibré ainsi obtenu est plus ou moins éloigné du produit cartésien $M \times G$. Tout ceci se précisera dans les exemples de la section suivante.

Potentiel et champ de Yang-Mills Le potentiel est identifié aux formes locales $\mathcal{A}_i = \sigma_i \omega$. Le champ de Yang-Mills correspond aux $\mathcal{F}_i = \sigma_i \Omega$. Enfin, la dérivée covariante que l'on emploie en théorie des champs pour introduire l'interaction portée par le champ de jauge correspond, on peut s'en douter, à l'opération (3.30), version locale de la dérivée covariante (3.15).

Transformations de jauge Les conditions de raccordement (3.8) et (3.25) des potentiels sur l'intersection de deux ouverts correspondants sont appelées transformations de jauge, mais pas dans le sens qu'on leur connaît habituellement. Nous sommes familiers des théories de jauge sur des espaces de base triviaux, le plus souvent l'espace-temps de Minkowski. Dans ce cas, il n'y a pas de fonctions de transition. Les transformations de jauge habituelles proviennent du fait qu'une section définie sur un ouvert U_i n'est pas unique : par action à droite sur la première, on peut toujours en définir une deuxième

$$\sigma_2(p) = \sigma_1(p) g(p),$$

où p est un point de M et g une fontion sur M à valeur dans G.

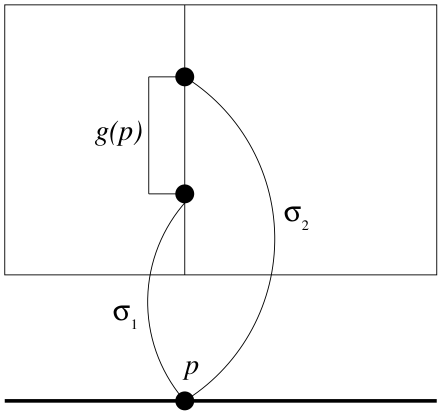


FIG. 3.18 – *La section σ, permettant d'associer des objets locaux aux objets globaux n'est pas unique*

Comme nous avons deux sections sur un même ouvert, le potentiel n'est pas défini univoquement à partir de la 1-forme connexion. On a $\mathcal{A}_1 = \sigma_1^*\omega$ et $\mathcal{A}_2 = \sigma_2^*\omega$. La relation entre les deux est donnée par

$$\mathcal{A}_2 = g^{-1}\mathcal{A}_1 g + g^{-1}\mathrm{d}g, \tag{3.32}$$

que l'on dérive de la même manière que (3.8). La non-univocité de la section correspond aux degrés de liberté de jauge et cette transformation est la transformation de jauge habituelle. Les relations (3.8) et (3.25) interviendront dans des espaces où il n'est pas possible de définir une section globale. De même, la transformation de la courbure locale s'écrit

$$\mathcal{F}_2 = g^{-1}\mathcal{F}g. \tag{3.33}$$

Champs Il nous reste un objet à introduire sans lequel nous serions bien en peine de faire de la théorie des champs. Nous allons voir sous quelle forme on introduit le champ de matière[10] dans la théorie des fibrés.

Un **champ** *sur un fibré principal P, d'espace de base M et de groupe de structure G est une application $\varphi : P \to V$, où V est un espace vectoriel sur lequel G est représenté par ρ, telle que*

$$\varphi(pg) = \rho(g^{-1})\varphi(p), \quad p \in P,\ g \in G,$$

Nous avons déjà rencontré cette propriété, appelée *équivariance*. Tout d'abord, nous l'avons imposée dans la définition de la 1-forme connexion (deuxième point), sous une forme particulière, la représentation adjointe n'étant qu'une représentation linéaire spéciale ($V = \mathfrak{g}$, $\rho = \mathrm{Ad}$), ensuite constatée dans le cas de la courbure, toujours sous cette forme de la représentation adjointe. Nous l'imposons à nouveau ici. Elle garantit qu'en connaissant la valeur dans V du champ en un point du fibré, on peut trouver sa valeur en un autre point appartenant à la même fibre, simplement par action à droite et sa représentation sur V. En d'autres termes, la valeur de champ dépend essentiellement de la projection $\pi^{-1}(p)$, donc du point de l'espace physique, garantissant qu'il s'agit bien d'un champ au sens habituel.

Les objets utilisés en théorie des champs ne sont pas ces champs globaux et, comme pour les connexions et courbures locales, on définit

$$\phi_i = \sigma_i^*\varphi = \varphi \circ \sigma_i : U_i \to V. \tag{3.34}$$

10. Nous effectuons cette précision car l'expression de champ est déjà utilisé pour caractériser la courbure locale : le champ de jauge

Sous l'action des deux types de transformation de jauge, on a

$$\phi_j(p) = \rho(g_{ij}(p)^{-1})\phi_i(p), \quad \text{pour } p \in U_i \cap U_j, \tag{3.35}$$

et

$$\phi^{'}(p) = \rho(g(p)^{-1})\phi(p), \tag{3.36}$$

où $g(p)$ est l'action à droite réalisant le changement de section.

Le tableau ci-dessous reprend sous forme de résumé les différents objets que nous avons traités jusqu'ici.

Nom	Formes sur P	Espace	Formes locales ($U_i \subset M$)	Espace	Relations
Champ	$\varphi : P \to V$	$\Lambda^0(P, V)$	$\phi_i : U_i \to V$	$\Lambda^0(U_i, V)$	$\mathcal{D}\phi_i = \mathrm{d}\phi_i + [\mathcal{A}_i, \phi_i]$
Connexion	$\omega : P \to \mathfrak{g}$	$\Lambda^1(P, \mathfrak{g})$	$\mathcal{A}_i : U_i \to \mathfrak{g}$	$\Lambda^1(U_i, \mathfrak{g})$	$\mathcal{F}_i = \mathrm{d}\mathcal{A}_i + \frac{1}{2}[\mathcal{A}_i, \mathcal{A}_i]$
Courbure	$\Omega : P \to \mathfrak{g}$	$\Lambda^2(P, \mathfrak{g})$	$\mathcal{F}_i : U_i \to \mathfrak{g}$	$\Lambda^2(U_i, \mathfrak{g})$	$\mathcal{D}\mathcal{F}_i = 0$

Constante de couplage Nous avons redérivé mathématiquement les objets principaux de la théorie des champs. Les constantes de couplage, de nature purement physique, n'apparaissent donc pas. Les connexions et courbures physiques s'obtiennent en effectuant les substitutions

$$\mathcal{A}_\mu = ig\mathcal{A}_\mu^{\text{phys}},$$

$$\mathcal{F}_{\mu\nu} = ig\mathcal{F}_{\mu\nu}^{\text{phys}}.$$

La dérivée covariante devient par exemple

$$\mathcal{D}_\mu\phi = \partial_\mu\phi + ig[A_\mu^{\text{phys}}, \phi],$$

et une transformation de jauge

$$\mathcal{A}_\mu^{\text{phys}} \to h^{-1}\mathcal{A}_\mu^{\text{phys}}h - \frac{i}{g}\,h^{-1}\partial_\mu h.$$

3.9 Exemples

Nous allons à présent animer tous les objets mathématiques que nous avons introduits, en examinant quelques exemples de la physique dans lesquels ils interviennent. Nous étudierons les deux "cas d'école" que sont le monopôle de Dirac et l'effet Arahonov-Bohm. En guise d'avant-propos du chapitre précédent, nous introduirons brièvement la théorie de Yang-Mills.

3.9.1 Monopôle magnétique - I

Nous étudierons ce premier exemple assez en détail, d'une part pour son importance historique, d'autre part parce qu'il nous permet de mettre en pratique, presque sous la forme d'un exercice, les notions théoriques abordées jusqu'ici.
Commençons tout d'abord par examiner un fibré principal de base S^2 et de fibre $U(1)$.
Considérons la sphère S^3, que nous pouvons envisager comme un sous-ensemble de $\mathbb{C}^2$:

$$S^3 = \{(z_1, z_2) \in \mathbb{C}^2 \text{ t.q. } |z_1|^2 + |z_2|^2 = 1\},$$

et l'application $\pi : S^3 \to S^2$ définie par

$$\pi(z_1, z_2) = (\overline{z_1}z_2 + \overline{z_2}z_1, -i(\overline{z_1}z_2 - \overline{z_2}z_1), |z_1|^2 - |z_2|^2),$$

appelée *application de Hopf.*
Il n'est pas évident, au vu de cette définition, que cette application agit vers la sphère S^2. Cela le devient si on paramétrise S^3 par

$$z_1 = \cos(\theta/2)e^{i\psi_1},\ z_2 = \sin(\theta/2)e^{i\psi_2},$$

avec $0 \leq \theta \leq \pi$ et $\psi_1, \psi_2 \in \mathrm{I\!R}$. Notre application devient alors

$$\pi(\cos(\theta/2)e^{i\psi_1},\ \sin(\theta/2)e^{i\psi_2}) = (\sin(\theta)\cos(\psi_2 - \psi_1), \sin(\theta)\sin(\psi_2 - \psi_1), \cos(\theta)).$$

Nous voyons que $\pi(z_1, z_2) = \pi(z_1', z_2')$ ssi $z_j' = e^{i\alpha}z_j$ $(j = 1, 2)$, avec $\alpha \in \mathrm{I\!R}$.
Nous pouvons dès lors réinterpréter S^3 comme un fibré principal construit sur l'espace de base S^2 et de fibre standard $U(1)$. Choisissons de recouvrir S^2 de la manière suivante :

$$U_1 = S^2 - \{(0, 0, -1)\},\ U_2 = S^2 - \{(0, 0, 1)\}. \tag{3.37}$$

Ensuite, établissons la fonction de transition sur l'intersection $U_1 \cap U_2$ des ouverts du recouvrement. Rappelons que celle-ci est construite sur les factorisations locales ϕ_i (cf. point 7 de la définition des fibrés). Dans le cas qui nous occupe, choisissons notre application ϕ_i en écrivant :

$$\phi_i^{-1} : \pi^{-1}(U_i) = \{(z_1,\, z_2) \in S^3 | z_i \neq 0\} \to S^2 \times U(1),\ (z_1, z_2) \mapsto (p, \frac{z_i}{|z_i|}),\ i = 1, 2.$$

Soient ensuite le point $p = (\sin(\theta)\cos(\varphi),\, \sin(\theta)\sin(\varphi),\, \cos(\theta)) \in U_1 \cap U_2$ et $\alpha \in \mathrm{I\!R}$. On a

$$\phi_{1,p}(e^{i\alpha}) = (\cos(\theta/2)e^{i\alpha},\, \sin(\theta/2)e^{i(\varphi+\alpha)}),$$

et donc

$$(\phi_{2,p}^{-1} \circ \phi_{1,p})(e^{i\alpha}) = e^{i\varphi}e^{i\alpha}. \tag{3.38}$$

Nous obtenons ainsi la fonction de transition recherchée

$$g_{12} : U_1 \cap U_2 \to U(1), \quad (\sin(\theta)\cos(\varphi),\, \sin(\theta)\sin(\varphi),\, \cos(\theta)) \mapsto e^{i\varphi}.$$

La relation (3.38) nous montre que le groupe de structure est effectivement représenté sur lui-même, comme il se doit pour un fibré principal. Il s'agit d'une représentation par la gauche (cf. définition du fibré principal). Pour l'action à droite, nous avons

$$R_{e^{i\alpha}}(z_1,\, z_2) = (z_1 e^{i\alpha},\, z_2 e^{i\alpha}),\ (z_1, z_2) \in S^3,\ \alpha \in \mathrm{I\!R}.$$

Appliquons à présent tous ces préliminaires au cas du monopôle de Dirac. On se place dans l'espace euclidien de dimension 3, et on suppose l'existence d'une charge magnétique placée à l'origine. Pour décrire cette situation, Dirac utilise le potentiel

$$A = i\frac{m}{2}\frac{1}{r(z - r)}(x\,dy - y\,dx), \tag{3.39}$$

en utilisant les coordonnées cartésiennes (x, y, z) et $r = \sqrt{x^2 + y^2 + z^2}$. La constante m est réelle et arbitraire. En calculant la dérivée extérieure de la forme A, nous obtenons le champ électromagnétique

$$F = i\frac{m}{2}\frac{1}{r^3}(x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy),$$

et donc[11]

$$\mathbf{B} = -\frac{m}{2e}\frac{\mathbf{r}}{r^3}, \tag{3.40}$$

ce qui vérifie bien sûr l'équation

$$\nabla \cdot \mathbf{B} = 2\pi \frac{m}{e}\delta(\mathbf{r}). \tag{3.41}$$

Les équations de Maxwell sont vérifiées pour $r \neq 0$.
Remarquons que (3.39) n'est pas défini sur la "demi-droite" $z = r$, que l'on appelle "corde de Dirac". On peut se demander pourquoi les singularités se situent seulement sur la partie positive de l'axe des z. En fait, le potentiel

$$A = i\frac{m}{2}\frac{1}{r(z+r)}(x\,dy - y\,dx), \tag{3.42}$$

convient également pour décrire la même situation, et dans ce cas, les singularités se situent sur la partie négative. Ces deux potentiels, décrivant le même problème physique, sont reliés par une transformation de jauge. Démontrons-le, en passant tout d'abord aux coordonnées sphériques. Le potentiel (3.39) s'écrit

$$A = i\frac{m}{2}(-1 - \cos\theta)d\varphi. \tag{3.43}$$

Considérons ensuite la transformation de jauge

$$A' = A + \gamma^{-1}d\gamma, \tag{3.44}$$

où γ est une fonction définie sur l'espace physique, à valeur dans $U(1)$. Si nous prenons

$$\gamma(x) = e^{im\varphi}, \tag{3.45}$$

nous obtenons

$$A' = A = i\frac{m}{2}(1 - \cos\theta)d\varphi, \tag{3.46}$$

qui est l'équivalent de la forme (3.42) du potentiel. Nous obtenons ainsi la célèbre condition de quantification de Dirac, car pour que (3.45) soit bien définie, nous devons imposer $m \in \mathbb{Z}$.
Il nous reste à montrer que la théorie du monopôle de Dirac est une théorie sur un fibré $(S^2, U(1))$. Le fait de prendre $U(1)$ comme groupe de jauge et, partant, comme fibre ne pose pas de problème, puisque nous parlons d'électromagnétisme. Il peut cependant paraître plus curieux de travailler sur la sphère S^2, alors que nous avons précisé au départ que nous nous placions dans l'espace euclidien habituel. Rappelons-nous que nous avons placé un monopôle en l'origine et de ce fait rendu les équations de Maxwell valides seulement sur le complémentaire de l'origine, variété homotopiquement équivalente à la sphère S^2. Celle-ci nécessite un recouvrement de deux ouverts, d'où la présence de deux potentiels dans la théorie. Le potentiel (3.39) n'est pas défini au pôle nord, nous le notons A_2 par référence à l'ouvert U_2 défini plus haut. Le potentiel (3.42) est baptisé A_1, et défini sur l'ouvert U_1 ne contenant pas le pôle sud. Il existe une connexion globale $\mathcal{A}$ définie sur l'espace total S^3, telle que

$$\begin{aligned}
\sigma_1^*\mathcal{A} &= A_1 = i\frac{m}{2}(1 - \cos(\theta))d\varphi \\
\sigma_2^*\mathcal{A} &= A_2 = i\frac{m}{2}(-1 - \cos(\theta))d\varphi
\end{aligned}$$

11. cf. remarque du paragraphe précédent, relative à la constante de couplage

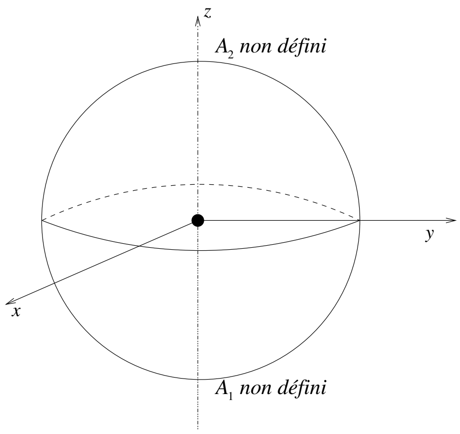


FIG. 3.19 –

Ces deux connexions locales sont reliées par la transformation de jauge

$$g_{12}(x) = e^{im\varphi}, \quad x \in U_1 \cap U_2, \quad m \in \mathbb{Z}.$$

Nous voyons ici qu'il existe plusieurs transformations de jauge possibles, plus exactement une infinité dénombrable correspondant aux valeurs du paramètre entier m. Comme γ envoie un élément de $U_1 \cap U_2 \simeq S^1$ vers le groupe $U(1) \simeq S^1$, cet entier s'interprète comme un élément du groupe fondamental de S^1

$$\pi_1(S^1) = \mathbb{Z}.$$

La quantification de la charge magnétique est donc d'origine purement topologique.

3.9.2 Effet Aharonov-Bohm

Nous ne pouvons clôre ce chapitre sans évoquer cette expérience célèbre dans laquelle la topologie d'un espace physique se manifeste d'une manière étonnante. Notre but ici n'est pas d'étudier l'effet Arahonov-Bohm en détail. Notre présentation est volontairement simplifiée. De nombreux développements plus détaillés se trouvent aisément dans la littérature.
Considérons la situation suivante :

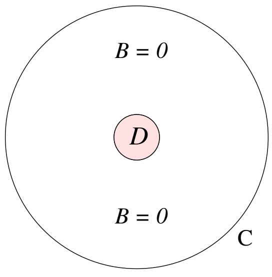


FIG. 3.20 –

Un solénoïde, supposé de longueur infinie, est placé le long de l'axe z. La figure ci-dessus est une coupe dans un plan x, y. Le champ magnétique est nul à l'extérieur. Soit enfin une courbe fermée C autour du solénoïde et S une surface encerclée par C. Le flux magnétique au travers de S est

$$\Phi = \int_S \mathbf{B} \cdot d\mathbf{S}.$$

En utilisant la formule de Stokes, nous serions tenté d'écrire

$$\Phi = \oint_C \mathbf{A}_C \cdot d\mathbf{l},$$

où A_C est tel que $\mathbf{B} = \nabla \times \mathbf{A}_C$. Ceci n'est évidemment pas possible. En effet, en tout point de la courbe C, le champ magnétique est rigoureusement nul, ce qui entraîne que $\mathbf{A}_C$ est purement de jauge. En d'autres termes, ce potentiel est obtenu par une transformation de jauge du potentiel nul et donc $\mathbf{A}_C = \nabla\lambda$, ce qui est impossible car il viendrait

$$\Phi = \oint_C \mathbf{A}_C \cdot d\mathbf{l} = \oint_C \nabla\lambda \cdot d\mathbf{l} = 0.$$

L'erreur que nous avons commise est d'avoir voulu définir le potentiel A_C de manière globale. Nous travaillons en effet sur la variété S^1. Le recouvrement de celle-ci requérant deux ouverts, définir un potentiel global n'est donc pas possible. Considérons deux potentiels locaux $\mathbf{A}_C^1$ et $\mathbf{A}_C^2$ correspondant aux deux ouverts du recouvrement. Ceux-ci s'écrivent

$$\mathbf{A}_C^1 = \nabla\lambda_1,$$

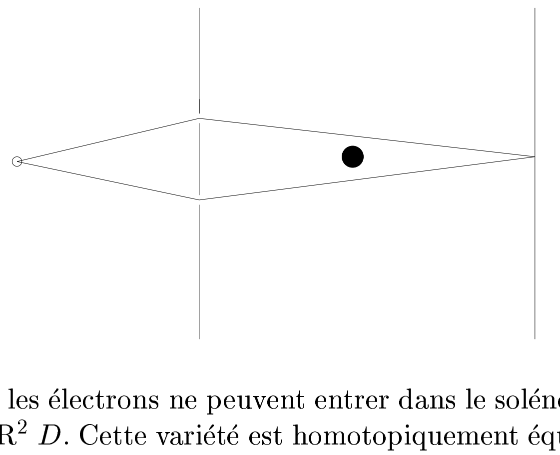
et

$$\mathbf{A}_C^2 = \nabla\lambda_2.$$

La forme des deux potentiels garantit toujours que $\mathbf{B} = 0$ en tout point de la courbe C, comme il se doit, mais cette fois

$$\Phi = \int_{U_1} \mathbf{A}_C^1 \cdot d\mathbf{l} + \int_{U_2} \mathbf{A}_C^2 \cdot d\mathbf{l} \neq 0. \tag{3.47}$$

Le plus étonnant est que cette non-trivialité du potentiel, que nous avons déduite théoriquement à l'aide de notions géométriques, peut-être mise en évidence expérimentalement. Considérons une expérience classique de diffraction d'électrons, que nous schématisons par une simple interférence de Young.



Si nous admettons que les électrons ne peuvent entrer dans le solénoïde, la région de l'espace qui leur est accessible est $\mathbb{R}^2$ D. Cette variété est homotopiquement équivalente à S^1. Il est donc impossible de définir univoquement un potentiel sur tout l'espace. Les deux parties de la fonction d'onde associée à un électron traversent donc deux régions où les potentiels sont différents. Lors de la superposition, il vient

$$\psi = e^{i\lambda_1}\psi_1 + e^{i\lambda_2}\psi_2 = e^{i\lambda_1}(\psi_1 + e^{i(\lambda_2-\lambda_1)}\psi_2).$$

Il est montré expérimentalement que la figure d'interférences obtenue peut être modifiée en faisant varier l'intensité du champ magnétique dans le solénoïde, ce qui est tout à fait remarquable puisque les électrons traversent une région où ce champ magnétique est nul ! Dans la littérature, beaucoup d'auteurs résument la situation en affirmant que dans le cas d'un espace possédant une topologie non triviale, le potentiel prend une importance tout aussi grande que le champ lui-même. C'est peut-être une conclusion rapide, car la grandeur pertinente est ici donnée par le flux (3.47) et non par les deux potentiels locaux.

3.9.3 Théorie de Yang-Mills

Avec ce dernier exemple, nous anticipons sur le prochain chapitre. Rappelons-nous, tout d'abord, le chapitre précédent où nous avions reformulé l'électromagnétisme de Maxwell dans le formalisme de la géométrie différentielle. Dans le paragraphe 2.2.4, nous avions évoqué l'action $S[A]$ qui s'écrit dans ce langage

$$S[A] = \int (-\frac{1}{2} F \wedge {}^*F), \tag{3.48}$$

où $F = dA$, et dans laquelle on retrouve le lagrangien bien connu

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}.$$

La théorie de Maxwell est une théorie de jauge sur l'espace de base $\mathbb{R}^4$ et de groupe de jauge $U(1)$.
L'idée de la théorie de Yang-Mills est de construire une théorie de jauge sur une variété M et un groupe de Lie quelconque par analogie avec le cas de l'électromagnétisme. Dans ce cas, l'action prend une forme tout à fait semblable à (3.48) :

$$S[\mathcal{A}] = \int_M \mathrm{tr}(\mathcal{F} \wedge {}^*\mathcal{F}), \tag{3.49}$$

où $\mathcal{F} = d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}$. Le lagrangien est alors

$$\mathcal{L} = -\frac{1}{4} F^a_{\mu\nu} F^{\mu\nu}_a, \tag{3.50}$$

où l'on a développé les composantes de la courbure $\mathcal{F}$ sur une base des générateurs de l'algèbre associée au groupe de jauge :

$$\mathcal{F}_{\mu\nu} = F^a_{\mu\nu} T_a,$$

et utilisé la convention $\mathrm{tr}(T_a T_b) = -\frac{1}{2}\delta_{ab}$.
Rappelons que pour la courbure $\mathcal{F}$, nous possédons une relation d'origine purement géométrique

$$\mathcal{D}\mathcal{F} = d\mathcal{A} + [\mathcal{A}, \mathcal{F}] = 0, \tag{3.51}$$

qui s'écrit en termes des composantes, dans l'espace de Minkowski (cf. 2.22),

$$(\mathcal{D}_\mu {}^*F^{\mu\nu})_a = \partial_\mu {}^*F^{\mu\nu}_a + f_{abc} A^b_\mu {}^*F^{\mu\nu}_c = 0. \tag{3.52}$$

D'autre part, les équations de champs s'écrivent maintenant

$$\mathcal{D} {}^*\mathcal{F} = 0, \tag{3.53}$$

ce qui donne

$$\mathcal{D}_\mu \mathcal{F}^{\mu\nu} = 0, \tag{3.54}$$

que l'on peut encore décomposer sur l'algèbre du groupe

$$\mathcal{D}_\mu F^{\mu\nu}_a = 0. \tag{3.55}$$

Ces quelques points techniques vont être exploités au cours du chapitre suivant, dans lequel nous allons montrer que cette théorie de Y.M. possède des propriétés topologiques non triviales.

Chapitre 4

Aspects topologiques de la théorie de Yang-Mills

4.1 Introduction

Armés des outils mathématiques introduits au cours du chapitre précédent, nous allons examiner succinctement quelques cas importants où des propriétés topologiques et géométriques apparaissent en théorie des champs. Il nous faudra dans un premier temps introduire la notion de classes caractéristiques, ou *classes de Chern*, avec lesquelles nous réexaminerons le cas du monopôle magnétique. Nous tenterons ensuite de faire apparaître la structure géométrique sous-jacente à une théorie de Y.M. $SU(2)$. Dans le cas d'une théorie euclidienne (utilisée dans les approches semi-classiques de la théorie des champs), nous verrons que la non-trivialité de la structure en fibré peut être établie en toute rigueur, sous forme de résultats mathématiques. Une des manifestations les plus connues de cette non-trivialité, le cas des instantons de Y.M. sera évoquée. Dans la suite et la fin de ce chapitre, nous reviendrons en espace de Minkowski et, de manière moins générale que pour le cas euclidien, nous examinerons au cas par cas plusieurs problèmes posés lors de la quantification des théories de jauge non abéliennes. Dans le formalisme canonique, tout d'abord, où nous évoquerons l'angle θ de Y.M.. Ensuite, nous envisagerons quelques problèmes posés dans le langage fonctionnel, avec une étude des anomalies d'Adler-Bell-Jackiw qui nous permettra d'introduire la notion d'index d'un opérateur. Enfin, nous terminerons par une introduction à l'opérateur BRS et les ghosts de Faddeev-Popov dans le langage de la géométrie différentielle.

4.2 Classes caractéristiques des fibrés

4.2.1 Polynômes invariants

Définition *Soient G un groupe de Lie, et $\mathfrak{g}$, son algèbre de Lie. Une application*

$$I : \overbrace{\mathfrak{g} \times \mathfrak{g} \times \cdots \times \mathfrak{g}}^{j} \to \mathbb{R} \ (ou\ \mathbb{C}),$$

linéaire et symétrique en ses j arguments est **invariante** *si pour tout $\mathfrak{g} \in G$ et tous $A_1, A_2, \ldots, A_j \in \mathfrak{g}$, on a*

$$I(\mathrm{Ad}_g A_1, \mathrm{Ad}_g A_2, \ldots, \mathrm{Ad}_g A_j) = I(A_1, A_2, \ldots, A_j). \tag{4.1}$$

Si nous introduisons la notation suivante :

$$P(A) = I(\overbrace{A, A, \ldots, A}^{j}),$$

l'objet ainsi défini est un **polynôme invariant sur** G. L'ensemble des polynômes invariants sur G de degré j est noté $I^j(G)$.

Exemple Pour tout élément F d'une algèbre de Lie $\mathfrak{g}$ (plus exactement une forme à valeurs dans $\mathfrak{g}$), l'expression

$$P(F) = \mathrm{tr}(F \wedge F)$$

est un polynôme invariant sur $\mathfrak{g}$, comme on peut le vérifier aisément.

Particularisons les notations que nous venons d'introduire au cas des polynômes invariants construits sur des formes. Soit $A \in \Lambda^r(M, \mathfrak{g})$. Si nous l'écrivons $A = \eta T$, où $\eta \in \Lambda^r(M)$ et $T \in \mathfrak{g}$, nous aurons

$$I(A_1, \ldots, A_j) = \eta_1 \wedge \cdots \wedge \eta_j \, I(T_1, \ldots, T_j).$$

4.2.2 Théorème fondamental

Passons aux propriétés des polynômes invariants.

Théorème *Soit P un fibré principal construit sur la variété M et le groupe de Lie G. Ce fibré est doté d'une connexion ω et partant d'une courbure Ω dont les expressions locales sont $\mathcal{A}$ et $\mathcal{F}$. Si P est un polynôme invariant, on a les deux propriétés suivantes :*

1. *$dP(\mathcal{F}) = 0$,*
2. *si ω' est une seconde connexion sur P, de courbure associée Ω', dont les formes locales sont $\mathcal{A}'$ et $\mathcal{F}'$, alors $P(\mathcal{F}') - P(\mathcal{F})$ est une forme exacte.*

Avant d'entamer la démonstration, remarquons que nous n'avons pas précisé la section locale grâce à laquelle nous obtenons $\mathcal{A}$ et $\mathcal{F}$, ni sur quel ouvert de recouvrement de M celles-ci sont définies. En effet, vu (3.25) (resp. (3.33)), et le fait que le polynôme P est invariant, celui-ci est indépendant du recouvrement (resp. de la définition de la section).
Montrons tout d'abord que $P(\mathcal{F})$ est une forme fermée. Nous utilisons pour cela les deux propriétés suivantes.

1. Si $T_1, \cdots, T_j \in \mathfrak{g}$, alors on a pour tout $X \in \mathfrak{g}$

$$\sum_{i=1}^{j} I(T_1, \ldots, [X, T_i], \ldots, T_j) = 0. \tag{4.2}$$

   Cette relation s'obtient [1] en dérivant les deux membres de l'identité suivante par rapport au paramètre t

$$I(\exp(-tX)\, T_1 \exp(\mathrm{tX}), \ldots, \exp(-tX)\, T_j \exp(tX)) = I(T_1, \ldots, T_j),$$

   et en l'évaluant en $t = 0$.

2. Pour $A_i = \eta_i T_i$, où η_i sont des 2-formes et $T_i \in \mathfrak{g}$, nous avons

$$I(A_1, \ldots, A_j) = \eta_1 \wedge \cdots \wedge \eta_j I(T_1, \ldots, T_j),$$

1. Le lecteur non convaincu peut dériver très simplement cette expression sur l'exemple suivant : $I(A_1, A_2) = \frac{1}{2}\mathrm{tr}(A_1 A_2 + A_2 A_1)$.

et

$$\begin{aligned} dI(A_1, \dots, A_j) &= d(\eta_1 \wedge \cdots \wedge \eta_j) I(T_1, \dots, T_j) \\ &= \sum_{i=1}^{j} \eta_1 \wedge \cdots \wedge d\eta_i \wedge \cdots \wedge \eta_j I(T_1, \dots, T_j) \\ &= \sum_{i=1}^{j} I(A_1, \dots, dA_i, \dots, A_j). \end{aligned}$$

Au total, il vient

$$\begin{aligned} dP(\mathcal{F}) &= dI(\mathcal{F}, \dots, \mathcal{F}) \\ &= \sum_{i=1}^{j} I(\mathcal{F}, \dots, d\mathcal{F}, \dots, \mathcal{F}) \\ &= \sum_{i=1}^{j} (I(\mathcal{F}, \dots, d\mathcal{F}, \dots, \mathcal{F}) + I(\mathcal{F}, \dots, [\mathcal{A}, \mathcal{F}], \dots, \mathcal{F})) \\ &= \sum_{i=1}^{j} I(\mathcal{F}, \dots, \mathcal{D}\mathcal{F}, \dots, \mathcal{F}) \\ &= 0 \end{aligned}$$

en recourant aux deux propriétés précédentes et à l'identité de Bianchi.

Passons à présent à la seconde partie du théorème. Définissons tout d'abord un potentiel $\mathcal{A}_t$, qui nous servira à interpoler les deux connexions $\mathcal{A}$ et $\mathcal{A}'$ par

$$\mathcal{A}' = \mathcal{A} + t\theta, \quad t \in [0, 1], \quad \theta = \mathcal{A}' - \mathcal{A}.$$

La courbure associée est

$$\mathcal{F}_t = \mathcal{F} + t\mathcal{D}\theta + t^2 \theta \wedge \theta.$$

Ensuite, nous pouvons écrire

$$\begin{aligned} P(\mathcal{F}') - P(\mathcal{F}) &= \int_0^1 \frac{d}{dt} P(\mathcal{F}_t)\, dt \\ &= j \int_0^1 I\left(\frac{d}{dt}\mathcal{F}_t, \mathcal{F}_t, \dots, \mathcal{F}_t\right) dt \end{aligned}$$

et tout revient à montrer que l'intégrand est une différentielle exacte.
D'une part, nous avons

$$I\left(\frac{d}{dt}\mathcal{F}_t, \mathcal{F}_t, \dots, \mathcal{F}_t\right) = I(\mathcal{D}\theta, \mathcal{F}_t, \dots, \mathcal{F}_t) + 2tI(\theta \wedge \theta, \mathcal{F}_t, \dots, \mathcal{F}_t).$$

D'autre part, il vient

$$\begin{aligned} dI(\theta, \mathcal{F}_t, \dots, \mathcal{F}_t) &= I(d\theta, \mathcal{F}_t, \dots, \mathcal{F}_t) - (j-1) I(\theta, d\mathcal{F}_t, \dots, \mathcal{F}_t) \\ &= I(\mathcal{D}\theta, \mathcal{F}_t, \dots, \mathcal{F}_t) - (j-1) I(\theta, \mathcal{D}\mathcal{F}_t, \dots, \mathcal{F}_t) \\ &= I(\mathcal{D}\theta, \mathcal{F}_t, \dots, \mathcal{F}_t) - (j-1) t I(\theta, [\mathcal{F}_t, \theta], \dots, \mathcal{F}_t) \end{aligned}$$

où l'on utilisé successivement la propriété 1 du point précédent, pour remplacer les dérivées extérieures par des dérivées covariantes, et l'identité suivante ($\mathcal{D}_t\mathcal{F}_t = 0$)

$$\begin{aligned} 0 &= d\mathcal{F}_t + (\mathcal{A} + t\theta) \wedge \mathcal{F}_t - \mathcal{F}_t \wedge (A + t\theta) \\ &= \mathcal{D}\mathcal{F}_t - t[\mathcal{F}_t, \theta]. \end{aligned}$$

Nous utilisons à nouveau la propriété 1 (dans laquelle $T_1 = X = \theta$) pour écrire

$$I([\theta, \theta], \mathcal{F}_t, \ldots, \mathcal{F}_t) - (j-1)I(\theta, [\mathcal{F}_t, \theta], \mathcal{F}_t, \ldots, \mathcal{F}_t) = 0$$

et en déduire

$$I\left(\frac{d}{dt}\mathcal{F}_t, \mathcal{F}_t, \ldots, \mathcal{F}_t\right) = dI(\theta, \mathcal{F}_t, \ldots, \mathcal{F}_t).$$

Nous obtenons alors le résultat annoncé

$$P(\mathcal{F}') - P(\mathcal{F}) = d\left(j\int_0^1 I(\mathcal{A}' - \mathcal{A}, \mathcal{F}_t, \ldots, \mathcal{F}_t)\, dt\right), \tag{4.3}$$

connu sous le nom de **formule de Chern-Simons**.

Exemple Appliquons cette formule sur l'exemple du paragraphe précédent. On a

$$\begin{aligned} \mathrm{tr}(\mathcal{F}' \wedge \mathcal{F}') - \mathrm{tr}(\mathcal{F} \wedge \mathcal{F}) &= 2\, d\int_0^1 \{\mathrm{tr}((\mathcal{A}' - \mathcal{A}) \wedge \mathcal{F}) + t\mathrm{tr}((\mathcal{A}' - \mathcal{A}) \wedge \mathcal{D}(\mathcal{A}' - \mathcal{A})) \\ &+ t^2\mathrm{tr}((\mathcal{A}' - \mathcal{A}) \wedge (\mathcal{A}' - \mathcal{A}) \wedge (\mathcal{A}' - \mathcal{A}))\}\, dt. \end{aligned}$$

En posant $\mathcal{A} = 0$, et en rebaptisant $\mathcal{A}'$ par $\mathcal{A}$, il vient, en réalisant l'intégration,

$$\mathrm{tr}(\mathcal{F} \wedge \mathcal{F}) = d[\mathrm{tr}(\mathcal{A} \wedge d\mathcal{A} + \frac{2}{3}A \wedge A \wedge A)]. \tag{4.4}$$

Ecrivons enfin cette expression en termes des composantes des formes qui y interviennent. L'argument de la dérivée extérieure est une 3-forme, qui s'écrit de manière générique

$$K = \frac{1}{3!}\varepsilon_{\lambda\mu\rho\sigma}K^\lambda dx^\mu \wedge dx^\rho \wedge dx^\sigma.$$

De l'autre côté, nous avons

$$\mathcal{A} \wedge d\mathcal{A} + \frac{2}{3}\mathcal{A} \wedge \mathcal{A} \wedge \mathcal{A} = \underbrace{(\frac{1}{2}\mathcal{A}_\mu(\partial_\rho A_\sigma - \partial_\sigma A_\rho) + \frac{2}{3}\mathcal{A}_\mu\mathcal{A}_\rho\mathcal{A}_\sigma)}\ dx^\mu \wedge dx^\rho \wedge dx^\sigma,$$

et nous devons antisymétriser l'accollade, de manière à l'identifier avec l'expression précédente. Après quelques lignes d'algèbre, on obtient à quatre dimensions les quatre composantes de K :

$$K^\nu = -\varepsilon^{\nu\mu\rho\sigma}(\mathcal{A}_\mu\partial_\rho A_\sigma + \frac{2}{3}\mathcal{A}_\nu\mathcal{A}_\rho\mathcal{A}_\sigma). \tag{4.5}$$

L'identité (4.4) s'écrit finalement

$$\frac{1}{2}\mathrm{tr}(\mathcal{F}_{\mu\nu}\ {}^*\mathcal{F}^{\mu\nu}) = \partial_\mu[\varepsilon^{\nu\mu\rho\sigma}\mathrm{tr}(\mathcal{A}_\mu\ \partial_\rho\mathcal{A}_\sigma + \frac{2}{3}\mathcal{A}_\mu\mathcal{A}_\rho\mathcal{A}_\sigma)]. \tag{4.6}$$

4.2.3 Classes caractéristiques - classes de Chern

Les objets que nous venons d'étudier permettent une caractérisation des fibrés principaux, comme nous pourrons le constater dans les exemples qui vont suivre. Le paragraphe précédent nous a appris que ces objets sont des formes fermées. Il n'est dès lors pas totalement absurde de les envisager en tant que membres d'un groupe de cohomologie sur le fibré (cf. annexe B) ou un des espaces qui le constituent. Les classes de cohomologie sont appelées des **classes caractéristiques**. Plus précisément, un théorème établit que pour un fibré construit sur une variété M et de groupe de structure G, le lien entre l'ensemble des polynômes invariants de degré j sur G et le groupe de cohomologie $H^{2j}(M)$ est un homomorphisme, appelé **homomorphisme de Weyl**. Insistons sur le fait que cet homomorphisme a lieu vers le groupe des classes de cohomologie de la variété M, c-à-d de l'espace de base, et non pas vers celui du fibré lui-même. Nous allons à présent aborder quelques aspects plus concrets de la théorie des classes caractéristiques et voir ce que ces objets peuvent nous apporter dans le cadre d'une théorie de jauge.

Définition *Soit E un fibré construit sur l'espace de base M, de dimension m et de groupe de structure* $\mathrm{GL}(m,\mathbb{C})$ *(ou un de ses sous-groupes). Etant donné une connexion ω et sa courbure associée Ω, de formes locales $\mathcal{A}$ et $\mathcal{F}$ respectivement, la relation*

$$\det(t\mathbf{1} + \frac{i\mathcal{F}}{2\pi}) = \sum_{j=0}^{m} c_j(\mathcal{F}) t^{m-j},$$

définit la $j^{\text{è}}$ classe de Chern c_j sous la forme du polynôme invariant $c_j(\mathcal{F}) \in I^j(\mathrm{GL}(m,\mathbb{C}))$.

Cette définition peut paraître fort peu naturelle. En fait, elle possède une forme connue pour le lecteur familier de la théorie des algèbres de Lie. Elle permet de construire d'autres grandeurs caractéristiques à partir des générateurs, comme l'opérateur de Casimir d'une algèbre semi-simple, qui s'obtient en développant une expression analogue :

$$\det(t\mathbf{1} + a_i T_i) = \sum_{j=0}^{m} P_{m-j}(a_i) t^j,$$

dans laquelle les générateurs T_i ont pris la place de la courbure $\mathcal{F}$.

Il reste à présent à déterminer explicitement la forme de ces polynômes. Supposons que la matrice $\mathcal{F}$ soit diagonale : $\mathcal{F} = \mathrm{diag}(\lambda_1, \ldots , \lambda_m)$. Le déterminant se développe alors selon

$$\begin{aligned}
\det(t\mathbf{1} + \frac{i\mathcal{F}}{2\pi}) &= \left(t + \frac{i\lambda_1}{2\pi}\right) \cdots \left(t + \frac{i\lambda_m}{2\pi}\right) \\
&= t^m + t^{m-1}\frac{i}{2\pi}\left(\sum_i \lambda_i\right) + t^{m-2}\left(\frac{i}{2\pi}\right)^2 \left\{\frac{(\sum_i \lambda_i)^2 - \sum_i \lambda_i^2}{2}\right\} \\
&+ \cdots + \left(\frac{i}{2\pi}\right)^m \lambda_1\lambda_2\cdots\lambda_m.
\end{aligned}$$

On obtient finalement les polynômes suivants

$$c_0(\mathcal{F}) = 1, \tag{4.7}$$

$$c_1(\mathcal{F}) = \frac{i}{2\pi}\mathrm{tr}(\mathcal{F}), \tag{4.8}$$

$$c_2(\mathcal{F}) = \left(\frac{i}{2\pi}\right)^2 \left\{\mathrm{tr}(\mathcal{F})\cdot\mathrm{tr}(\mathcal{F}) - \mathrm{tr}(\frac{\mathcal{F}\wedge\mathcal{F})}{2}\right\}, \tag{4.9}$$

$$\vdots$$

$$c_m(\mathcal{F}) \quad = \quad \left(\frac{i}{2\pi}\right)^m \det(\mathcal{F}). \tag{4.10}$$

Examinons directement sur un exemple les informations que ces nouveaux objets sont susceptibles de nous fournir. Nous reprenons le cas du monopôle magnétique évoqué dans le chapitre précédent.

4.2.4 Monopôle magnétique - II

Rappelons qu'il s'agit d'un fibré construit sur la variété S^2 et le groupe $U(1)$. Il est dans ce cas particulier très simple de contruire les classes caractéristiques. Seules la classe triviale c_0 et la classe c_1 subsistent, les autres étant identiquement nulles. Vu le caractère abélien du groupe de jauge, les traces disparaissent, et il nous reste

$$\begin{aligned} c_0(\mathcal{F}) &= 1, \\ c_1(\mathcal{F}) &= \frac{i}{2\pi}\mathcal{F}. \end{aligned}$$

Nous pouvons constater que la classe non triviale c_1 est, aux coefficients multiplicatifs près, fondamentalement le champ électromagnétique. Si nous intégrons $c_1(\mathcal{F})$ sur notre espace de base S^2, nous allons obtenir par le théorème de Gauss la charge magnétique, que nous avions appelée m (cf. chapitre précédent) :

$$\int_{S^2} c_1(\mathcal{F}) = \frac{i}{2\pi}\int_{S^2} \mathcal{F} = -m. \tag{4.11}$$

En tant qu'intégrale d'une classe caractéristique sur la variété de base, m est appelé **nombre de Chern**. Nous constatons que ce nombre m apparaît dans plusieurs situations apparamment différentes. Il est, à la fois,

- la charge magnétique introduite dans le potentiel de Dirac (3.39),
- le nombre entier caractérisant la fonction de transition (3.45) et partant la transformation de jauge (3.44),
- le résultat de l'intégration de la classe caractéristique c_1 du fibré.

Ce petit état des lieux nous montre de manière concrète la façon dont une classe de Chern peut caractériser un fibré. Rappelons encore une fois que la manière avec laquelle les éléments d'un fibré sont assemblés est matérialisée par les fonctions de transition (se rappeler l'exemple du ruban de Möbius). Le nombre m n'est rien d'autre qu'une sorte de signature de la construction.

Il est temps de résumer et de rassembler les éléments disparates que nous avons récoltés. L'étude du monopôle magnétique se développe, pour des raisons déjà exposées, dans le formalisme d'un fibré principal $(S^2, U(1))$. La variété de base est telle qu'il est impossible de définir le potentiel électromagnétique de manière univoque. Les deux potentiels locaux, définis sur les deux cartes nécessaires pour décrire la variété S^2, sont reliés par une transformation de jauge, correspondant à la fonction de transition (3.45), dans laquelle m est le paramètre (réel) de la transformation. Cette dernière, ciment de la construction du fibré principal comme nous venons de le rappeler, tombe dans le domaine des classes d'homotopie $\pi_1(S^1) \simeq \mathbb{Z}$, en tant qu'application de S^1 (intersection des deux cartes) vers $U(1) \simeq S^1$ (groupe de jauge). Le paramètre de transformation ne prend que des valeurs entières, et la charge magnétique est quantifiée. Il existe ainsi une infinité dénombrable de manières de réaliser un fibré $(S^2, U(1))$:

1. Un simple produit cartésien $S^2 \times U(1)$, correspondant à la théorie de Maxwell sans monopôle : la charge m est nulle, l'espace de base est l'espace de Minkowski sans obstruction topologique, la fonction de transition (3.45) est la transformation identité et il est possible de définir un potentiel sur la totalité de l'espace de base [2].
2. Une infinité de fibrés non triviaux construits sur un espace de base avec obstruction topologique, le monopôle , dont la force m varie d'un fibré à l'autre par valeurs entières.

Nous avons beaucoup insisté sur cet exemple car, comme nous allons le voir dans la suite, nous appliquerons la même démarche pour une théorie de jauge non abélienne et constaterons que les idées fondamentales sont identiques.

4.3 Aspects topologiques d'une théorie $SU(2)$

Précisons tout d'abord quelques points de notations et de conventions. Dans le cas où aucune confusion n'est possible, nous employerons les notations introduites dans la section (3.9.3): les indices grecs désignent les indices d'espace-temps, ceux-ci variant entre 0 et 4, des indices latins (a, b, c, ...) l'algèbre du groupe de jauge (indices de couleurs). Les constantes f^c_{ab} désignent les constantes de structure de l'algèbre. Ses éléments seront représentés par des lettres rondes et leurs composantes dans la base des générateurs par des lettres droites.

4.3.1 Cadre géométrique

Espace de base

L'espace de travail pour une théorie de Y.M. est l'espace-temps à quatre dimensions de Minkowski. Afin de faire apparaître les aspects topologiques d'une théorie de jauge construite sur ce dernier, nous allons le considérer dans un cadre particulier et, il faut le dire, très peu familier. Pour cela, partons de considérations très simples. Nous savons que la droite des nombres réels $\mathbb{R}$ peut être mis en bijection avec la variété S^1 de la manière suivante :

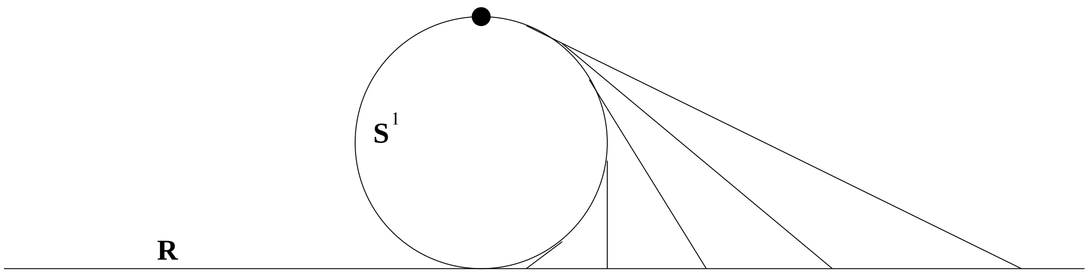


FIG. 4.1 – *Compactification de l'ensemble* $\mathbb{R}$

La droite $\mathbb{R}$ est immergée dans le plan, et “repliée” en un cercle, comme indiqué sur la figure. A chaque point du cercle, on associe la droite tangente, et l'intersection de celle-ci avec la droite originale. Ce petit procédé nous donne la bijection annoncée. Les deux infinis sont représentés par un point, le pôle nord. On écrit

$$\mathbb{R} \cup \{\infty\} = S^1.$$

2. Ne pas confondre le caractère univoque et le caractère global d'une connexion. Un potentiel défini sur la totalité de l'espace de base l'est à une transformation de jauge près (cf. liberté du choix de la section dans la théorie du fibré principal).

Cette opération porte le nom de **compactification**, que l'on peut réaliser de manière générale à n dimensions. Dans la suite, nous considérerons l'espace de Minkowski compactifié, c-à-d la variété $S^4 = \mathbb{R}^4 \cup \{\infty\}$, que nous allons prendre comme variété de base d'un fibré principal. Cette variété S^4 nécessite comme toujours dans le cas d'une sphère un atlas constitué par deux cartes, fondamentalement les deux hémisphères dont l'intersection ("équateur") est la sphère S^3.
Cette compactification de l'espace nécessite cependant un petit commentaire. En effet, nous savons qu'une théorie de jauge sur l'espace de Minkowski, libre de toute obstruction topologique, est une théorie sur un fibré trivial car l'espace de base est simplement connexe. Il est donc possible d'y définir globalement une connexion. Après compactification, l'espace de base n'est non seulement plus topologiquement trivial, mais nous devons de plus envisager deux potentiels locaux correspondant aux deux cartes. Ce point est un peu délicat, et nous allons tenter de le clarifier quelque peu dans la suite.

Fonctions de transition

Comme nous venons de le signaler, nous devons envisager deux connexions locales. Celles-ci sont reliées par une transformation de jauge par l'intermédiaire, comme toujours, des fonctions de transition. On a, avec des notations évidentes,

$$\mathcal{A}_N = g^{-1}(x)\mathcal{A}_S g(x) + g^{-1}(x)dg(x) \tag{4.12}$$

où la fonction de transition $g(x)$ est définie sur l'équateur S^3 et à valeurs dans le groupe de jauge G. La fonction $g(x)$ tombe dès lors dans les classes du troisième groupe d'homotopie de G. Or $G = SU(2)$ est aussi la variété S^3. On a donc

$$\pi_3(SU(2)) = \mathbb{Z}.$$

Toutes les propriétés que nous allons mettre en évidence restent valable pour un produit direct (cf. (3.1)), par exemple pour $SU(2) \times U(1)$) et pour tout groupe contenant $SU(2)$ comme sous-groupe (comme $SU(3)$).

Classes de Chern

En se rappelant que les éléments de l'algèbre de Lie $\mathfrak{su}(2)$ sont de trace nulle, nous obtenons facilement le seul polynôme invariant caractéristique :

$$c_2(\mathcal{F}) = \frac{1}{8\pi^2}\mathrm{tr}(\mathcal{F} \wedge \mathcal{F}), \tag{4.13}$$

objet que nous connaissons bien car nous l'avons étudié en détail, à dessein bien entendu, dans la section 4.2.2 .

4.3.2 Solution self-duale

Si nous juxtaposons les relations (3.51) et (3.53), nous pouvons nous rendre compte qu'il existe une classe de champs $\mathcal{A}$ satisfaisant automatiquement les équations de champs (3.53), celle pour laquelle on a

$$\mathcal{F} = \lambda\ {}^*\mathcal{F}, \tag{4.14}$$

où λ est un nombre réel ou complexe. Rappelons que l'étoile de Hodge fait intervenir le tenseur métrique. La valeur de λ est donc déterminée par la métrique de l'espace sur lequel on travaille. Ainsi, comme $\mathcal{F} = \lambda^2\ {}^{**}\mathcal{F}$, on a

$$\mathcal{F} = \begin{cases} \lambda^2\mathcal{F} & \text{pour la métrique euclidienne} \\ -\lambda^2\mathcal{F} & \text{pour la métrique de Minkowski} \end{cases} \tag{4.15}$$

Il vient donc, dans le cas euclidien,

$$\mathcal{F} = \pm\ {}^*\mathcal{F} \tag{4.16}$$

et, dans le cas de l'espace de Minkowski,

$$\mathcal{F} = \pm i\ {}^*\mathcal{F}.$$

Cette dernière relation requiert donc une algèbre $\mathfrak{g}$ telle que $i\mathfrak{g} = \mathfrak{g}$, relation non satisfaite par les algèbres associées aux groupes compacts que l'on rencontre en physique. Nous travaillerons donc en métrique euclidienne avec les champs satisfaisant la relation (4.16). Ces champs particuliers sont dits champs **self-duaux**. Cette condition de self-dualité autorise une résolution exacte du problème de Y.M., ce qui nous permettra de montrer que la structure géométrique sous-jacente est bien celle du fibré principal décrit dans les lignes précédentes.
Dans le cas de champs self-duaux, l'action (euclidienne) s'écrit (cf. (3.49))

$$S[\mathcal{A}] = \int -\mathrm{tr}(\mathcal{F} \wedge \mathcal{F}). \tag{4.17}$$

Nous voyons apparaître le nombre de Chern associé à (4.13), mais cette particularité ne nous intéresse pas dans l'immédiat. Nous la considérons pour le moment comme un accident.
Pour que l'action $S[\mathcal{A}]$ soit finie, nous devons nous assurer que l'intégrale a un sens. Un premier ansatz que nous pouvons poser est d'exiger que le potentiel $\mathcal{A}$ se comporte de manière suivante à l'infini :

$$\mathcal{A}_\mu(x) = g^{-1}(x)\,\partial_\mu g(x), \tag{4.18}$$

où $g(x)$ est une fonction définie sur la sphère S^3, à valeur dans le groupe de jauge. Ceci garantit l'annulation de $\mathcal{F}$ à l'infini, mais nullement la convergence de l'intégrale. Nous devrons vérifier qu'il en sera bien ainsi lorsque nous aurons la solution.
Nous cherchons une solution du type

$$\mathcal{A}_\mu(x) = f(r^2) g^{-1}(x)\,\partial_\mu g(x), \quad r^2 = \sum_i x_i^2, \tag{4.19}$$

avec

$$\lim_{r\to\infty} f(r^2) = 1.$$

Cherchons tout d'abord une expression pour la fonction $g(x)$. Cette fonction associe un point de la sphère S^3 d'équation $x^2 + y^2 + z^2 + t^2 = 1$ à un élément de $SU(2)$ dont on se rappelera l'expression p. 29. L'application la plus simple que nous pouvons imaginer est une simple bijection entre les deux variétés :

$$\begin{aligned} g : U_N \cap U_S \to SU(2), \quad (x,y,z,t) \quad &\mapsto \quad \begin{pmatrix} t - iz & y - ix \\ -y - ix & t + iz \end{pmatrix} \\ &= \quad \frac{1}{r}(t\mathbf{1} - i(x\sigma_1 + y\sigma_2 + z\sigma_3)), \end{aligned} \tag{4.20}$$

où $r = \sqrt{x^2 + y^2 + z^2 + t^2}$, et σ_i sont les matrices de Pauli. Si nous posons $\tau^0 = \mathbf{1}$, $\tau^i = i\sigma^i$, l'expression précédente s'écrit

$$g(x) = \frac{x_\alpha \tau^\alpha}{r}. \tag{4.21}$$

Cette fonction de transition est une simple identité de S^3 sur elle-même. Il est relativement évident que cette application appartient à la première classe d'homotopie de $\pi_3(S^3)$.
Réinjectant cette solution dans (4.19), calculant $\mathcal{F}_{\mu\nu}$, et exprimant la condition de self-dualité, on obtient (λ est une constante)

$$\mathcal{A}_\mu(x) = \frac{1}{r^2 + \lambda^2}(x_\alpha \tau_\alpha)^\dagger \tau_\mu, \tag{4.22}$$

ou, pour retrouver l'ansatz initial,

$$\mathcal{A}_\mu(x) = \frac{r^2}{r^2+\lambda^2}\, g^{-1}(x)\, \partial_\mu g(x). \tag{4.23}$$

Analysons directement cette solution particulière des équations de champ. Elle est régulière à l'origine, et converge vers zéro à l'infini comme r^{-1}. Le champ $\mathcal{F}$ convergera fondamentalement en r^{-2} , l'intégrand dans (4.17) en r^{-4} et l'intégrale n'est pas finie. Comme nous allons le constater, ceci n'est qu'un demi échec. Si nous effectuons une transformation de jauge sur le potentiel obtenu, que nous noterons $\mathcal{A}_1$,

$$\mathcal{A}_{2,\mu} = g^{-1}\mathcal{A}_{1,\mu} g + g^{-1}\partial g,$$

nous obtenons

$$\mathcal{A}_{2,\mu} = \frac{\lambda^2}{\lambda^2+r^2} g^{-1}\partial_\mu g.$$

On vérifie que ce potentiel décroît en r^{-3}, ce qui garantit la convergence de l'intégrale, mais il est singulier à l'origine.

Nous nous retrouvons donc dans une situation analogue à celle rencontrée dans l'étude du monopôle magnétique, où nous avions deux potentiels dont l'un était singulier à l'un des deux pôles de S^2 et le second à l'autre pôle. Il est donc possible de décrire cette situation dans la géométrie du paragraphe précédent. L'espace euclidien $\mathbb{R}^4$ est compactifié et nous choisissons de placer l'infini au pôle sud. $\mathcal{A}_1$ (rebaptisé $\mathcal{A}^N$) est le potentiel local défini sur l'hémisphère nord, ce qui est licite puisqu'il est régulier à l'origine, et $\mathcal{A}_2$ ($\mathcal{A}^S$) est défini sur l'hémisphère sud, assurant la convergence de (4.17), que l'on réécrit

$$S[A] = -\int_{U_N} \mathrm{tr}(\mathcal{F}^N \wedge \mathcal{F}^N) - \int_{U_S} \mathrm{tr}(\mathcal{F}^S \wedge \mathcal{F}^S). \tag{4.24}$$

Grâce à la solution analytique des équations de Y.M. dans le cas des champs self-duaux, nous avons pu faire apparaître la structure géométrique sous-jacente que nous avons évoquée précédemment.

On peut alors se demander s'il est possible de le montrer dans le cas d'un champ de Yang-Mills général. La réponse est affirmative mais uniquement dans le cas de l'action euclidienne. Un résultat mathématique ([7]) montre effectivement qu'un champ de Yang-Mills sur une variété quadri-dimensionnelle munie d'une métrique euclidienne, donnant lieu à une énergie finie, ne peut posséder de singularité. S'il en possède, celles-ci sont apparentes et peuvent être levées par une transformation de jauge. *Un champ de Yang-Mills, d'énergie finie, pour un fibré construit sur* $\mathbb{R}^4$ *peut être étendu à un champ régulier pour un fibré construit sur la compactification* S^4 *de* $\mathbb{R}^4$. En des termes plus techniques, le théorème donne explicitement le comportement du champ à l'infini :

Si $\mathcal{F}$ *est un champ de Yang-Mills sur le complémentaire d'une boule dans* $\mathbb{R}^4$*, alors il existe un nombre réel* C *tel que*

$$|F(x)| \le |x|^{-4} C \quad (|F|^2 = |F^a_{\mu\nu} F_a^{\mu\nu}|),$$

pour tout point $x \in \mathbb{R}^4$ *tel que* $|x| \ge 2$.

Avec une convergence en r^{-4}, l'action est finie. La variété sur laquelle ce champ est défini, le complémentaire d'une boule dans $\mathbb{R}^4$, est identifiée à la carte de la sphère S^4 contenant le pôle sud, que nous avons associé à l'infini. Le théorème montre enfin qu'il est possible de définir un potentiel sur la boule de rayon unité, que nous identifions à la carte contenant le pôle nord, soit l'origine, dont la singularité en ce point peut s'éliminer par une transformation de jauge.

4.3.3 Nombre de Chern d'une transformation de jauge

Nous allons dans ce paragraphe calculer le nombre de Chern associé à la classe caractéristique (4.13) :

$$k = \frac{1}{8\pi^2}\int_{S^4} \mathrm{tr}(\mathcal{F}\wedge\mathcal{F}). \tag{4.25}$$

Dans le cas du monopôle, il nous avait rendu le nombre entier caractérisant l'homotopie de la fonction de transition (3.45). C'est un résultat que nous allons à nouveau vérifier ici.
Opérons tout d'abord une transformation sur la relation (4.25). Rappelons que nous travaillons avec les deux connexions locales $\mathcal{A}^N$ et $\mathcal{A}^S$, leurs courbures associées $\mathcal{F}^N$ et $\mathcal{F}^S$, et les transformations de jauge suivantes

$$\mathcal{A}^S = g^{-1}\mathcal{A}^N g + g^{-1}dg, \tag{4.26}$$
$$\mathcal{F}^S = g^{-1}\mathcal{F}^N g, \tag{4.27}$$

avec des notations claires. On a successivement, en utilisant le résultat (4.4) et le théorème de Stokes

$$\begin{aligned}\int_{S^4}\mathrm{tr}(\mathcal{F}\wedge\mathcal{F}) &= \int_{U^N}\mathrm{tr}(\mathcal{F}^N\wedge\mathcal{F}^N)+\int_{U^S}\mathrm{tr}(\mathcal{F}^S\wedge\mathcal{F}^S)\\ &= \int_{S^3}\mathrm{tr}(\mathcal{A}^N\wedge d\mathcal{F}^N-\frac{1}{3}A^N\wedge A^N\wedge A^N)\\ &- \int_{S^3}\mathrm{tr}(\mathcal{A}^S\wedge d\mathcal{F}^S-\frac{1}{3}A^S\wedge A^S\wedge A^S).\end{aligned}$$

Le signe moins dans la dernière expression provient des orientations différentes des deux hémisphères. Il faut ensuite remplacer la connexion et la courbure $\mathcal{A}^S$ et $\mathcal{F}^S$ en injectant dans l'expression précédente les transformations de jauge (4.26) et (4.27). Après une page d'algèbre, dans laquelle on se rappelera les propriétés de commutation des formes, ainsi que la formule de Leibnitz (voir appendice), il vient

$$\int_{S^4}\mathrm{tr}(\mathcal{F}\wedge\mathcal{F}) = \frac{1}{3}\int_{S^3}\mathrm{tr}(g^{-1}dg\wedge g^{-1}dg\wedge g^{-1}dg) - \int_{S^3} d(\mathrm{tr}(g^{-1}\mathcal{A}^N\wedge dg)).$$

La seconde intégrale est nulle et il reste au total

$$k = \frac{1}{24\pi^2}\int_{S^3}\mathrm{tr}(g^{-1}dg\wedge g^{-1}dg\wedge g^{-1}dg). \tag{4.28}$$

Dans la littérature, nous rencontrons ce nombre sous l'appellation "winding number of the gauge transformation". Celui-ci, comme nous allons le voir compte le nombre de fois que la variété S^3 recouvre (to wind signifie enrouler) le groupe $SU(2)$, exactement comme l'entier m du monopôle désignait le nombre de fois que le cercle S^1 recouvrait le groupe $U(1)$. Il s'agit bien sûr du nombre entier caractérisant l'homotopie de la fonction de transition $g(x)$ et partant les transformations de jauge (4.26) et (4.27). Il nous reste à calculer k pour la transformation de jauge (4.21). Le résultat suivant

$$g^{-1}dg = \underbrace{\frac{1}{r^2}\left[-x_0\,dx^i + x_i\,dx^0 + \varepsilon_{lki}\,x_l\,dx^k\right]}\,\tau^i$$

s'obtient en notant que

$$g^{-1} = \frac{x_0\tau^0 + x_i\tau^i}{r},$$

et

$$\tau^i\tau^j = \frac{1}{2}\left[\{\tau^i,\tau^j\}+[\tau^i,\tau^j]\right] = -2\varepsilon_{ijk}\tau^k.$$

En désignant l'accolade par ξ_i, l'expression précédente s'écrit $g^{-1}dg = \xi_i \tau^i$. En injectant dans l'expression du nombre k (4.28), on obtient

$$k = \frac{1}{24\pi^2} \int_{S^3} (\xi_i \wedge \xi_j \wedge \xi_k)\, \mathrm{tr}[\tau^i\, \tau^j\, \tau^k].$$

Comme $(\xi_i \wedge \xi_j \wedge \xi_k) = \varepsilon_{ijk}(\xi_1 \wedge \xi_2 \wedge \xi_3)$, $\mathrm{tr}[\tau^i\, \tau^j\, \tau^k] = 2\,\varepsilon^{ijk}$, et $\varepsilon_{ijk}\varepsilon^{ijk} = 6$, il vient

$$k = \frac{1}{2\pi^2} \int_{S^3} (\xi_1 \wedge \xi_2 \wedge \xi_3).$$

Il reste alors à effectuer la contraction des formes en les explicitant et, après un travail relativement fastidieux, nous obtenons

$$k = \frac{1}{2\pi^2}\, \frac{1}{r^4} \int_{S^3} \frac{1}{6} \varepsilon^{\alpha\beta\gamma\delta} x_\alpha dx_\beta \wedge dx_\gamma \wedge dx_\delta$$

(avec la convention[3] $\varepsilon^{0123} = -1$). Les efforts déployés lors du calcul vont être récompensés. En effet, l'intégrale obtenue est extrêmement sympathique. La variété S^3 est la frontière d'une boule dans un espace de dimension 4. Nous pouvons appliquer la formule de Stokes, que nous pouvons trouver dans l'appendice, mais que nous rappelons ici pour des raisons purement esthétiques :

$$\int_K d\varphi = \int_{\partial K} \varphi,$$

où K est une variété de bord ∂K. Dans notre cas, la forme φ à intégrer est

$$\frac{1}{r^4} \frac{1}{3!} \varepsilon^{\alpha\beta\gamma\delta} x_\alpha dx_\beta \wedge dx_\gamma \wedge dx_\delta.$$

Il vient successivement

$$\begin{aligned} \frac{1}{6r^4} \int_{S^3} \varepsilon^{\alpha\beta\gamma\delta}\, x_\alpha\, dx_\beta \wedge dx_\gamma \wedge dx_\delta &= \frac{1}{6r^4} \int_{V^4} \varepsilon^{\alpha\beta\gamma\delta}\, \partial^\mu x_\alpha\, dx_\mu \wedge dx_\beta \wedge dx_\gamma \wedge dx_\delta \\ &= \frac{4}{r^4} \int_{V^4} dx_0\, dx_1\, dx_2\, dx_3, \end{aligned}$$

en notant que $dx_\alpha \wedge dx_\beta \wedge dx_\gamma \wedge dx_\delta = \varepsilon_{\alpha\beta\gamma\delta}\, dx_0 \wedge dx_1 \wedge dx_2 \wedge dx_3$, $\varepsilon_{\alpha\beta\gamma\delta}\varepsilon^{\alpha\beta\gamma\delta} = 24$. Le volume d'une sphère de rayon r à 4 dimensions est $\frac{\pi^2 r^4}{2}$. Au total,

$$k = \frac{1}{2\pi^2} \frac{4}{r^4} \frac{\pi^2 r^4}{2} = 1$$

La seconde classe de Chern intégrée sur la variété S^4 rend finalement l'information suivante : la fonction de transition (4.21) opère un recouvrement simple du groupe $SU(2)$, ce qui est bien le cas, puisque nous avons construit la fonction (4.21) en ce sens. En considérant une fonction de transition

$$g_k(x) = g(x)^k,$$

avec k entier, on peut montrer que $\int_{S^4} c_2(\mathcal{F}) = k$. En d'autres termes, la fonction g_k opère k fois le recouvrement du groupe. Enfin, il est trivial que $g_0 = \mathbf{1}$ possède un winding number identiquement nul. Comme dans le cas du monopôle, ceci signifie simplement que le fibré est trivial.

3. Notons que, comme nous travaillons en espace euclidien, le déterminant du tenseur métrique est positif et $\varepsilon_{0123} = \varepsilon^{0123} = -1$.

4.3.4 Instantons de Yang-Mills

Dans le formalisme des intégrales de chemin de la théorie des champs, la valeur moyenne du vide est donnée par

$$< 0|0 >= \int [D\phi] e^{-S_E(\phi, \partial_\mu \phi)},$$

où S_E est l'action euclidienne. Les valeurs du champ qui contribuent le plus à cette valeur sont celles qui réalisent des minima de l'action euclidienne. Dans le cas de la théorie $SU(2)$, nous allons montrer qu'il existe précisément des minima de l'action (3.49), d'origine purement topologique. En posant

$$\mathcal{F}^+ = \frac{1}{2}(\mathcal{F} + {}^*\mathcal{F}), \quad \mathcal{F}^- = \frac{1}{2}(\mathcal{F} - {}^*\mathcal{F}),$$

nous pouvons décomposer le champ $\mathcal{F}$ en une partie self-duale et une partie anti self-duale : $\mathcal{F} = \mathcal{F}^+ + \mathcal{F}^-$. L'action (3.49) s'écrit alors

$$\begin{aligned} S &= -\int \mathrm{tr}\left[(\mathcal{F}^+ + \mathcal{F}^-) \wedge {}^*(\mathcal{F}^+ + \mathcal{F}^-)\right] \\ &= -\int \mathrm{tr}\left[(\mathcal{F}^+ \wedge {}^*\mathcal{F}^+) - (\mathcal{F}^- \wedge {}^*\mathcal{F}^-)\right] \end{aligned}$$

Rappelons que dans l'ensemble des p-formes $\Lambda^p(M)$ sur une variété M, il est possible, grâce à l'étoile de Hodge, de définir un produit scalaire : $< \varphi, \psi >= \int_M \varphi \wedge {}^*\psi$. Dans le cas présent, les formes appartiennent à l'ensemble $\Lambda^2(M, \mathfrak{su}(2))$. Il est facile de montrer que

$$< \varphi, \psi >= \int_M \mathrm{tr}(\varphi \wedge {}^*\psi)$$

est un produit scalaire pour $\Lambda^2(M, \mathfrak{s}u(2))$. Il est possible d'introduire une norme définie positive (rappelons que nous travaillons en espace euclidien) [4] :

$$\|\varphi\|^2 = - < \varphi, \varphi > .$$

L'action S s'écrit finalement

$$S = \|\mathcal{F}^+\|^2 + \|\mathcal{F}^-\|^2.$$

Si nous effectuons la même opération pour $c_2(\mathcal{F})$, nous obtenons

$$8\pi^2 k = \|\mathcal{F}^+\|^+ - \|\mathcal{F}^-\|^2.$$

Pour une paire de nombres réels a et b, on a l'inégalité $|a^2 - b^2| \leq a^2 + b^2$. Il vient donc

$$S \geq 8\pi^2 |k|,$$

Pour une théorie de jauge $SU(2)$ non triviale ($k \neq 0$), il existe des minima locaux de l'action euclidienne, que l'on appelle des **instantons** ([4]). Ces minima sont réalisés lorsque S vaut précisément $8\pi^2|k|$, c-à-d lorsque ${}^*\mathcal{F} = \frac{k}{|k|}\mathcal{F}$, donc dans le cas des solutions (anti) self-duales, que l'on appelle aussi solutions en instantons.

4. Rappelons que nous prenons pour $\mathfrak{s}u(2)$ les générateurs $T_i = \frac{i}{2}\sigma_i$, avec σ_i les matrices de Pauli. On a alors $\mathrm{tr}(T_i T_j) = -\frac{1}{2}\delta_{ij}$.

4.3.5 Formalisme canonique

Nous revenons en espace de Minkowski dans la suite et la fin de ce chapitre. Nous ne tenterons pas, comme nous l'avons fait dans le cas euclidien, de mettre en évidence de manière générale la structure géométrique de la théorie. Nous allons plutôt examiner quelques cas particuliers où les conséquences topologiques de cette structure se manifestent.

Précisons tout d'abord un point de notation. Comme

$$[\mathcal{D}_\mu \phi]^a = \underbrace{(\partial_\mu \delta^a_b + g\, f^a_{cb}\, A^c_\mu)}\, \phi^b,$$

nous poserons

$$[\mathcal{D}_\mu \phi]^a = D^a_{\mu b} \phi^b, \tag{4.29}$$

afin d'alléger les notations. Rappelons l'expression du lagrangien de Yang-Mills

$$\mathcal{L}_{YM} = -\frac{1}{4} F^a_{\mu\nu} F^{\mu\nu}_a. \tag{4.30}$$

Les équations de mouvement s'écrivent

$$D^\mu_{ab} F^b_{\mu\nu} = 0. \tag{4.31}$$

Les moments canoniquement conjugués aux champs A^μ_a sont simplement $\Pi^\mu_a = F^{\mu 0}_a$. Nous voyons apparaître une contrainte primaire

$$\Pi^0_a = 0, \tag{4.32}$$

analogue à (2.48). La contrainte secondaire associée à (4.32) est donnée par la composante 0 des équations de mouvement. Il vient donc la chaîne de contraintes suivante

$$\Pi^0_a = 0 \quad \Rightarrow \quad D^a_{kb} \Pi^b_k = 0, \tag{4.33}$$

deux contraintes de première classe, signature de l'invariance de jauge. La contrainte secondaire est l'analogue non abélien de la loi de Gauss de l'électromagnétisme. La densité hamiltonienne est donnée par

$$\mathcal{H}_{YM,T} = \frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) - A^a_0\, (D^a_{kb} \Pi^b_k) + \Lambda^a \Pi^0_a,$$

où les champs $\mathbf{E}$ et $\mathbf{B}$ sont définis comme en électromagnétisme (cf. (2.2) et (2.3)) :

$$E^a_i = F^{i0}_a, \qquad B^a_i = -\frac{1}{2} \varepsilon^{ijk} F^{jk}_a.$$

Nous choisissons de travailler dans la jauge temporelle en posant

$$A^0_a = 0,$$

pour toute valeur de l'indice a. La stabilité de cette condition de jauge entraîne l'annulation des multiplicateurs de Lagrange Λ_a. (cf. (2.62)). Le hamiltonien prend alors la forme classique

$$H = \int \frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2)\, d^3 r.$$

Il s'agit, comme dans le cas abélien d'une jauge de classe II. Seule la contrainte sur les degrés de liberté non-physiques (A^0_a, Π^0_a) est réalisée, la condition de Gauss n'est pas réalisée. Celle-ci est alors réintégrée en imposant que seuls les états annihilés par la condition de Gauss relèvent de la théorie de Yang-Mills :

$$D^a_{kb} E^b_k\, |\psi_{\text{phys}}> = 0. \tag{4.34}$$

Examinons d'un peu plus près cette condition. Remarquons tout d'abord qu'il subsiste, malgré une condition de jauge, une symétrie dans la théorie. En effet, ayant posé $A_a^0 = 0$, il subsiste une invariance de jauge sur les composantes spatiales du potentiel vecteur. En d'autres termes, la théorie est invariante sous la transformation

$$A_i \to g^{-1} A_i g + g^{-1} \nabla_i g, \tag{4.35}$$

et sous sa version infinitésimale :

$$\delta A_a^k = -D_{ab}^k \Lambda_b.$$

Par le théorème de Noether, on obtient les générateurs de la transformation. Il s'agit précisément de $G_a = \int d^3 r D_{kb}^a E_k^b$. La condition de Gauss est donc le générateur de la symétrie subsistante dans la théorie. On peut également vérifier qu'il commute avec le hamiltonien :

$$[H, G_a] = 0,$$

pour toute valeur de a. Nous pouvons représenter, du moins formellement, la transformation de jauge (4.35) par $\mathfrak{G} = e^{iG_a}$. La condition (4.34) peut alors se réécrire :

$$\mathfrak{G} |\psi >_{\text{phys}} = |\psi >_{\text{phys}} . \tag{4.36}$$

Remarquons que tout ce qui vient d'être exposé peut s'appliquer au cas de la théorie abélienne évoquée au chapitre 2. Il s'agit d'une caractéristique de la jauge temporelle. Nous n'avons par contre pas encore tenu compte du caractère particulier d'une transformation de jauge du type (4.35), plus exactement de l'homotopie éventuellement non triviale de la fonction g. Dans quelle mesure la relation (4.36) est-elle valable pour une transformation quelconque ? Si nous la réexprimons par

$$\mathfrak{G}_0 |\psi >_{\text{phys}} = |\psi >_{\text{phys}}, \tag{4.37}$$

est-il possible de la généraliser à une transformation finie $\mathfrak{G}_n$ construite sur une fonction g quelconque de la classe d'homotopie n ? Comme il s'agit d'une transformation de symétrie, nous pouvons raisonnablement nous attendre à l'apparition d'un facteur de phase dépendant précisément de la classe d'homotopie de g :

$$\mathfrak{G}_n |\psi >_{\text{phys}} = e^{in\theta} |\psi >_{\text{phys}} . \tag{4.38}$$

Nous allons voir dans la suite comment il est possible d'étudier plus précisément cette ambiguité de la théorie, dont l'origine est manifestement topologique.

4.3.6 Angle θ de la théorie de Y.M.

Il existe un autre moyen de voir qu'il existe une ambiguité. Supposons tout d'abord que les états relevant de la théorie sont invariants sous toutes les transformations, quelle que soit l'homotopie de g. Nous devons pour cela modifier la théorie. Nous la construisons sur le lagrangien suivant :

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu} + \theta \, \frac{1}{32\pi^2} \; {}^*F_{\mu\nu}^a F_a^{\mu\nu} \tag{4.39}$$

Le terme ajouté au lagrangien n'est rien d'autre que la seconde classe de Chern de $\mathfrak{su}(2)$, écrite en termes des composantes (cf. (4.5)). Nous savons que ce terme supplémentaire est une dérivée totale. La seule différence que cet ajout induit est une redéfinition des moments canoniquement conjugués :

$$\Pi_a^\mu = F_a^{\mu 0} - \frac{\theta}{16\pi^2} \varepsilon^{0\mu\rho\sigma} F_{\rho\sigma}^a.$$

On peut vérifier immédiatement que la contrainte primaire $\Pi_a^0 = 0$ subsiste, et en explicitant les équations de mouvement, que la chaîne complète de contraintes (4.33) est également conservée. Cette théorie est donc une théorie de jauge comparable à la précédente. Ceci est heureux puisque nous savons que les classes de Chern sont constituées de polynômes invariants; le terme ajouté est, bien sûr, invariant de jauge. Nous travaillons à nouveau dans la jauge temporelle et fixons $A_0^a = 0$, pout tout a. Le hamiltonien est identique à celui du paragraphe précédent, à la redéfinition près des moments conjugués :

$$\Pi_a^i = E_i^a + \frac{\theta}{8\pi^2} B_i^a,$$

qui donne

$$\mathcal{H} = \frac{1}{2}[(\Pi_a^i - \frac{\theta}{8\pi^2} B_i^a)^2 + (B_i^a)^2].$$

Passons à présent dans le formalisme fonctionnel, où les vecteurs d'état s'écrivent sous la forme de fonctionnelles d'ondes : $\Psi[A]$, et les moments conjugués sous la forme de dérivées fonctionnelles. Dans ce formalisme, l'équation de Schrödinger s'écrit

$$\int d^3r \frac{1}{2}\left[(-i\frac{\delta}{\delta A_a^i} - \frac{\theta}{8\pi^2} B_i^a)^2 + (B_i^a)^2\right]\Psi[A] = E\Psi[A].$$

Rappelons que nous travaillons sous l'hypothèse d'états invariants sous toute transformation de jauge :

$$\mathfrak{G}_n \Psi[A] = \Psi[A].$$

Nous allons montrer qu'il est possible de recouvrer un hamiltonien sous sa forme habituelle, en éliminant le paramètre θ qui s'y trouve, par une redéfinition des états physiques. Il est facile de voir, après un calcul il est vrai un peu fastidieux, que les états

$$\Psi'[A] = e^{i\theta W[A]}\Psi[A], \tag{4.40}$$

satisfont une équation de Schrödinger sans paramètre, sous la condition

$$\frac{\delta W[A]}{\delta A_a^i(\mathbf{r})} = \frac{1}{8\pi^2} B_i^a(\mathbf{r}).$$

L'intégration de cette équation fonctionnelle donne précisément

$$W[A] = \frac{1}{8\pi^2}\int d^3r\, \varepsilon^{ijk}\,\mathrm{tr}(\mathcal{A}_i\partial_j\mathcal{A}_k + \frac{2}{3}\mathcal{A}_i\mathcal{A}_j\mathcal{A}_k).$$

Il reste à examiner le comportement des états $\Psi'[A]$ sous l'action d'une transformation de jauge (4.35).

$$W[A] \to W[A] + n,$$

où n est le nombre de Chern (ou d'enroulement (winding number)) de la transformation :

$$n = \int d^3r\, \varepsilon^{ijk}\,\mathrm{tr}(g^{-1}\partial_i g\, g^{-1}\partial_k g\, g^{-1}\partial_k g),$$

que nous avons appris à calculer dans un paragraphe précédent. Finalement, nous réobtenons explicitement la relation (4.38) :

$$\mathfrak{G}_n \Psi'[A] = e^{in\theta}\Psi'[A],$$

que nous avions supposée précédemment.

Il est possible d'obtenir expérimentalement des contraintes sur ce paramètre θ. Celui-ci doit être extrêmement petit, de l'ordre de 10^{-9}.

4.4 Anomalies

Rappelons tout d'abord brièvement ce que ce l'on entend par le concept d'anomalie. Une théorie possède une anomalie lorsqu'une symétrie présente au niveau classique est violée lors de la quantification. Une symétrie classique est typiquement une symétrie du lagrangien donnant lieu par le théorème de Noether à la conservation d'un courant.
La présence d'anomalies a des conséquences très importantes pour une théorie : elles rendent la théorie non renormalisable. Leur étude est donc d'un intérêt capital. Adler, Bardeen et Zumino ont mis en évidence, sous forme de deux théorèmes, des propriétés remarquables relatives aux anomalies :

1. L'absence d'anomalies à l'ordre le plus bas de la théorie des perturbations assure l'absence d'anomalies à tout ordre.
2. Dans une même théorie, toutes les anomalies sont reliées entre elles.

Nous allons montrer que la présence d'anomalies est effectivement totalement indépendante du traitement perturbatif de la théorie, ainsi que l'origine topologique de la violation de la symétrie classique.

Le cas le plus connu est celui de l'anomalie d'Adler-Jackiw, apparaissant dans la théorie de l'électrodynamique de fermions non massiques. Le lagrangien s'écrit

$$\mathcal{L} = \bar{\psi}(i\gamma \cdot D)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{4.41}$$

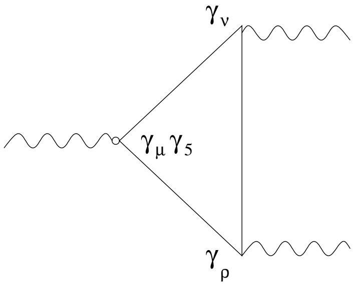
Celui-ci est invariant sous la transformation chirale

$$\psi \to e^{i\,\alpha(x)\gamma^5}\psi,$$

donnant lieu, par le théorème de Noether, à la conservation du courant

$$j^5_\mu = \bar{\psi}\gamma_\mu\gamma^5\psi. \tag{4.42}$$

L'anomalie apparaît lors du calcul des corrections à une boucle, plus précisément dans le calcul de l'amplitude suivante .



L'amplitude calculée par les règles de Feynman habituelles fait intervenir une intégrale divergente, dont la régularisation, quelle qu'elle soit, entraîne la violation des identités de Ward correspondant à (4.42). Le calcul suivant, dû à Fujikawa ([19]) dérive l'anomalie dans le formalisme fonctionnel, montrant ainsi que la violation se situe effectivement au niveau de la quantification elle-même, et non au niveau du calcul perturbatif. Il faut donc étudier l'effet d'une transformation chirale sur l'intégrale fonctionnelle

$$Z = \int [D\psi][D\bar{\psi}]e^{i\int d^4x\bar{\psi}(i\gamma\cdot D)\psi}. \tag{4.43}$$

Sous les transformations

$$\psi(x) \to (1 + i\alpha(x)\gamma^5)\psi(x),$$
$$\bar{\psi}(x) \to \bar{\psi}(x)(1 + i\alpha(x)\gamma^5),$$

on obtient

$$\int d^4x\bar{\psi}(i\gamma\cdot D)\psi \to \int d^4x[\bar{\psi}(i\gamma\cdot D)\psi + \alpha(x)\partial_\mu(\bar{\psi}\gamma_\mu\gamma^5\psi)].$$

Au niveau classique, le calcul s'arrête ici et on conclut à la conservation du courant (4.42). Nous devons, au niveau quantique, examiner l'effet de la transformation sur la mesure. Si nous développons les champs fermioniques sur la base des états propres ϕ_i de l'opérateur $i\gamma\cdot D$,

$$\psi = \sum_i a_i\phi_i, \quad \bar{\psi} = \sum_i \hat{a}_i\hat{\phi}_i,$$

avec $(i\gamma\cdot D)\phi_i = \lambda_i\phi_i$, et $a_i, \hat{a}_i$ des nombres de Grassman, nous pouvons écrire la mesure comme

$$[D\psi][D\bar{\psi}] = \Pi_m da_m d\hat{a}_m.$$

Sous la transformation chirale, les coefficients du développement deviennent

$$a'_m = \sum_n(\delta_{mn} + i\int d^4x\, \alpha(x)\phi^\dagger_m(x)\gamma^5\phi_n(x))a_n = \sum_n(\delta_{mn} + C_{mn})a_n.$$

La mesure se transforme alors comme

$$[D\psi][D\bar{\psi}] \to (\det(1+C))^{-2}[D\psi][D\bar{\psi}],$$

où l'*inverse* du déterminant apparaît, vu que l'on intègre sur des variables de Grassman. On a d'autre part

$$\det(1+C) = e^{\mathrm{tr}(\ln(1+C))} \simeq e^{\mathrm{tr}(C)},$$

vu que les éléments de C sont infinitésimaux. Il vient finalement

$$Z \to \int[D\psi][D\bar{\psi}]\exp\left[i\int d^4x\left\{\bar{\psi}(i\gamma\cdot D)\psi + \alpha(x)[\partial^\mu j^5_\mu - 2\sum_n \phi^\dagger_n(x)\gamma^5\phi_n(x)]\right\}\right]. \tag{4.44}$$

On a donc, au niveau quantique,

$$\partial^\mu j^5_\mu = 2\sum_n \phi^\dagger_n(x)\gamma^5\phi_n(x). \tag{4.45}$$

Le second membre de cette relation peut encore s'écrire

$$\sum_n \phi^\dagger_n(x)\gamma^5\phi_n(x) =< x|\mathrm{tr}(\gamma^5)|x >= 0,$$

et la conservation de la symétrie serait assurée au niveau quantique. Cependant, rappelons que nous avons extrait (4.45) comme intégrand dans l'expression (4.44). Or il apparaît que l'intégrale $\int d^4x\alpha(x)\sum_n \phi^\dagger_n(x)\gamma^5\phi_n(x)$ est divergente. Il faut donc la régulariser :

$$\begin{aligned}\sum_n \phi^\dagger_n(x)\gamma^5\phi_n(x) &\to \lim_{M\to\infty}\sum_n \phi^\dagger_n(x)\gamma^5\phi_n(x)e^{\left(\frac{\lambda_i}{M}\right)^2}, \qquad (4.46)\\ &= \lim_{M\to\infty} < x|\mathrm{tr}\left(\gamma^5 e^{\frac{(i\gamma\cdot D)^2}{M^2}}\right)|x >\end{aligned}$$

Après une page d'algèbre des matrices de Dirac [5], on obtient

$$
\begin{aligned}
\lim_{M\to\infty}\sum_n \phi_n^\dagger(x)\gamma^5\phi_n(x)e^{\left(\frac{\lambda_i}{M}\right)^2} &= -\frac{e^2}{32\pi^2}\varepsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu} \\
&= -\frac{e^2}{16\pi^2}F_{\mu\nu}\,{}^*F^{\mu\nu},
\end{aligned}
\tag{4.47}
$$

où l'on reconnaît une expression se rapprochant de la seconde classe de Chern $c_2(F)$. D'autre part, nous allons donner une seconde expression pour cette intégrale, ce qui nous permettra de lui donner une interprétation physique. Remarquons tout d'abord que si $|\phi_i>$ est un état propre de $i\gamma\cdot D$ pour la valeur propre λ_i, $\gamma^5|\phi_i>$ est un état propre de valeur propre $-\lambda_i$. Cela se montre aisément en utilisant la propriété d'anticommutation de γ^5 avec toutes les matrices de Dirac. Il vient donc

$$
\begin{aligned}
\int d^4x \lim_{M\to\infty}\sum_n \phi_n^\dagger(x)\gamma^5\phi_n(x)e^{\left(\frac{\lambda_i}{M}\right)^2} &= \lim_{M\to\infty}\sum_n <\phi_n|\gamma^5 e^{(\frac{i\gamma\cdot D}{M})^2}|\phi_n> \\
&= \lim_{M\to\infty}\sum_n <\phi_n|\gamma^5|\phi_n> e^{(\frac{\lambda_n}{M})^2} \\
&= 0,
\end{aligned}
$$

pour toutes les valeurs non nulles de l'indice n. La contribution vient exclusivement des zéro-modes de l'opérateur $i\gamma\cdot D$ que l'on classe suivant leur valeur propre relativement à l'opérateur γ^5 :

$$\gamma^5|0,i>_\pm = \pm|0,i>_\pm .$$

Il vient donc

$$
\begin{aligned}
\lim_{M\to\infty}\sum_n <\phi_n|\gamma^5|\phi_n> e^{(\frac{\lambda_n}{M})^2} &= \sum_i {}_+<0,i|0,i>_+ - \sum_i {}_-<0,i|0,i>_- \\
&= N_R - N_L,
\end{aligned}
$$

et finalement

$$\int d^4x\partial^\mu j_\mu^5 = N_R - N_L. \tag{4.48}$$

Dans la théorie classique, cette relation implique que la différence entre le nombre fermions polarisés à gauche et à droite est conservé dans n'importe quel processus. Nous pouvons d'ailleurs la dériver à l'aide d'un raisonnement plus simple, en calculant la variation de la charge q_5 associée au courant j_μ^5 lors d'un processus :

$$\Delta q_5 = \int_{-\infty}^{\infty} dt\,\dot{q_5} = \int d^4x\,\partial^\mu j_\mu^5.$$

Ceci a posé pendant un certain temps un problème en physique des particules. Il s'agit du problème dit "$U(1)$". Le lagrangien de la chromodynamique quantique, construit sur des champs de Yang-Mills et des fermions non massiques, possède un grand degré de symétrie, dont la symétrie chirale. La brisure spontanée de symétrie, donnant une masse aux particules, brise également la symétrie chirale mais pas totalement, car il subsiste une invariance par rotation chirale sur tous les champs fermioniques non massifs de la théorie. Celle-ci donne lieu à la relation (4.42), avec la conséquence (4.48), non observée dans la nature. De plus, par le théorème de Goldstone, il doit exister un boson non massif, également non observé. La découverte de l'anomalie d'Adler a donné un certain espoir de résolution du problème, tué dans l'oeuf lorsqu'on

5. Le calcul est effectué en détail dans ([6]), p.664 et sq. ainsi que dans ([3]), p.436 et sq.

a réalisé que le membre de droite de (4.47) était une dérivée totale, pour laquelle on négligeait le terme de surface. Il a fallu attendre Belavin et Polyakov, avec la découverte des instantons, pour réaliser que $\int d^4x\partial^\mu j^5_\mu$ n'est pas automatiquement nulle. Dans le cas d'une théorie de jauge non-abélienne, nous savons en effet que, à des coefficients multiplicatifs près,

$$\int d^4x\,\mathrm{tr}(\,{}^*\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu}) = n,$$

avec n entier. D'autre part, si nous étudions les anomalies chirales d'une théorie de jauge non abélienne (voir [3] pour les détails et (4.49) ci-dessous)), nous allons aboutir précisément à

$$\int d^4x\partial^\mu j^5_\mu \propto \int d^4x\,\mathrm{tr}(\,{}^*\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu}) = n,$$

où le caractère entier de n est heureux puisque cette expression, rappelons-le, désigne une différence de nombre de particules. Il est enfin possible de montrer que dans la théorie de jauge du paragraphe précédent, pour laquelle le lagrangien fait intervenir l'angle θ, la symétrie chirale est automatiquement brisée. En effet, rappelons encore une fois que la seconde classe de Chern est une dérivée totale. On peut écrire, avec les notations de la relation (4.5), un courant conservé

$$J^5_\mu = j^5_\mu - \frac{1}{2\pi^2}K_\mu,$$

où l'on a tenu compte de l'anomalie. La charge Q_5 s'écrit alors

$$Q_5 = q_5 + 2W(\mathbf{A}),$$

et n'est pas invariante de jauge, car sous une transformation de jauge, elle change de deux fois le nombre d'enroulement n. Nous avons trois opérateurs à considérer : le hamiltonien H, $\mathfrak{G}_n$, générateur des transformations de jauge, et Q_5. H commute avec les deux autres, mais ces derniers ne commutent pas entre eux

$$[\mathfrak{G}_n, Q_5] = 2n\mathfrak{G}_n.$$

Les trois opérateurs ne peuvent pas être diagonalisés simultanément. Avec H, on diagonalise $\mathfrak{G}_n$, comme dans la relation (4.38), ce qui entraîne que les états propres ne sont pas invariants sous la transformation de chiralité générée par Q_5.
Signalons enfin que la relation

$$N_R - N_L = -\frac{e^2}{4\pi^2}\int d^4x\ \ {}^*F^{\mu\nu}F_{\mu\nu},$$

est une version particulière d'un théorème mathématique, le théorème d'Atiyah-Singer, relatif à la notion d'index. L'*index* d'un opérateur D est défini par

$$\mathrm{ind}D = \dim\,\ker D - \dim\,\ker D^\dagger,$$

où $D^\dagger$ est l'opérateur adjoint de D. Dans le cas présent, le rôle de l'opérateur D est joué par $i\gamma\cdot D_R$, défini en décomposant les spineurs en composantes R et L :

$$\psi = \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix}, \quad i\gamma\cdot D = \begin{pmatrix} 0 & i\gamma\cdot D_L \\ i\gamma\cdot D_R & 0 \end{pmatrix}.$$

D'autre part, on introduit une nouvelle espèce de polynôme : les caractères de Chern, défini par

$$\mathrm{ch}_j(\mathcal{F}) = \frac{1}{j!}\mathrm{tr}\left(\frac{i\mathcal{F}}{2\pi}\right)^j,$$

ce qui donne, en termes des classes de Chern qui nous sont familières

$$\begin{aligned} \mathrm{ch}_1(\mathcal{F}) &= c_1(\mathcal{F}), \\ \mathrm{ch}_2(\mathcal{F}) &= \frac{1}{2}[c_1(\mathcal{F})^2 - 2c_2(\mathcal{F})]. \end{aligned}$$

Le théorème s'écrit alors dans le cas présent

$$\mathrm{ind}(i\gamma \cdot D_R) = \int \mathrm{ch}_2(F).$$

Dans le cas d'une théorie $SU(2)$,

$$\mathrm{ch}_2(\mathcal{F}) = -c_2(\mathcal{F}). \tag{4.49}$$

Retenons de tout ceci l'origine topologique des anomalies chirales, celle-ci allant dans le sens des deux théorèmes rappelés au début de ce paragraphe, ainsi que le rôle joué par la régularisation. Dans le calcul perturbatif, non rappelé ici, la contribution à la violation des identités de Ward proviennent de la régularisation des intégrales divergentes, de même que la contribution (4.47) à la non conservation du courant chiral provient de la régularisation (4.46).

4.5 Opérateur BRS

La quantification d'une théorie nécessite la fixation d'une jauge. C'est particulièrement clair lorsqu'on aborde la quantification dans le formalisme fonctionnel. En effet, l'expression

$$Z = \int [DA] e^{-i\int d^4x \frac{1}{4} F_{\mu\nu}F^{\mu\nu}},$$

dans le cas de la quantification du champ électromagnétique, n'est pas bien définie. L'intégration sur toutes les configurations $A_\mu = \frac{1}{e}\partial_\mu \alpha(x)$ est divergente puisque l'intégrand vaut 1. La cause en est l'invariance sous une transformation de jauge, qui entraîne une intégration redondante sur une infinité continue de configurations de champ équivalentes.
Dans le cadre d'une théorie de Yang-Mills générale, nous pouvons poser le problème de la manière suivante. Si nous désignons par $\mathfrak{U}$ l'espace de toutes les configurations des potentiels de jauge sur S^4, et par $\mathfrak{G}$ l'ensemble de toutes les transformations de jauge (plus précisément l'ensemble des applications $g : S^4 \to G$, où G est le groupe de jauge) sur S^4, l'espace des configurations physiques doit être identifié à l'espace quotient $\mathfrak{U}/\mathfrak{G}$, qui nous permet de répartir en classes les configurations équivalentes par une transformation de jauge de G.
Soit $a \in \mathfrak{U}$ un représentant de la classe $[a] \in \mathfrak{U}/\mathfrak{G}$. Soit également

$$\mathcal{A} = g^{-1}ag + g^{-1}dg,$$

un élément de $[a]$. On peut montrer, mais cela semble relativement naturel, que $\mathfrak{U}$ est en fait un fibré sur $\mathfrak{U}/\mathfrak{G}$ de fibre $\mathfrak{G}$. Si nous désignons par δ la dérivée extérieure sur ce fibré, examinons son action sur le potentiel ci-dessus :

$$\begin{aligned} \delta\mathcal{A} &= -g^{-2}\delta g a g + g^{-1}\delta a g + g^{-1} a \delta g - g^{-2}\delta g d g + g^{-1} d(\delta g) \\ &= -g^{-1}\delta g \mathcal{A} + g^{-1}\delta a g + \mathcal{A} g^{-1}\delta g + d[g^{-1}\delta g], \end{aligned}$$

où l'on a supposé la commutation des opérateurs de dérivation extérieure, et utilisé $d(\delta g) = dgg^{-1}\delta g + gd(g^{-1}\delta g)$. Il vient finalement

$$\delta\mathcal{A} = g^{-1}\delta a g + \mathcal{D}(g^{-1}\delta g).$$

Le premier terme représente la partie "horizontale" de $\delta\mathcal{A}$, tandis que le second, qui ne fait pas intervenir la classe $[a]$ d'où nous sommes partis, est la dérivée "verticale", le long de la fibre $\mathfrak{U}$.

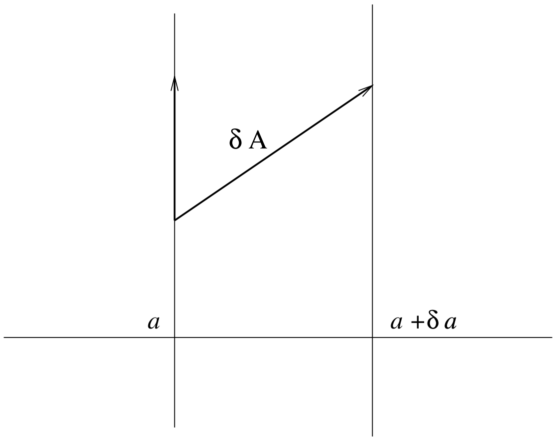


FIG. 4.2 – *L'opération BRS sur le potentiel A est la partie verticale, le long de la fibre* $\mathfrak{U}$ *de la transformation* δ.

On définit l'opérateur BRS par

$$\mathcal{S}\mathcal{A} = \mathcal{D}(\omega),$$

où $\omega = g^{-1}\delta g|_{fibre}$, est le ghost de Faddeev-Popov. On a enfin

$$\mathcal{S}\omega = -\omega \wedge \omega.$$

Ces transformations sont les transformations BRST bien connues

$$\begin{aligned} \delta A^a_\mu &= D^{ac}_\mu c^c \delta\lambda, \\ \delta c^a &= -\frac{1}{2} f^{abc} c^b c^c \delta\lambda. \end{aligned}$$

Il est toutefois à noter que le caractère grassmannien du ghost ω n'apparaît pas directement.

Chapitre 5

Chern-Simons et la physique dans le plan

5.1 Introduction

Dans ce dernier chapitre, nous revenons à l'étude que nous avions entamée au cours du chapitre deux et consacrée à la théorie de jauge de Chern-Simons. Nous allons dans les quelques pages qui suivent évoquer quelques caractéristiques de cette théorie. Tout d'abord, nous nous intéresserons aux propriétés de symétrie des systèmes de particules identiques dans un espace à 2+1 dimensions. Nous allons voir que la statistique, comme on l'appelle en mécanique quantique, de tels systèmes est radicalement différente du cas quadrimensionnel : il est mathématiquement possible d'envisager l'existence de particules ne répondant pas au clivage habituel fermions-bosons. Autrement dit, il s'agit de particules possédant une statistique dite "fractionnaire", que l'on a appelées *anyons* (pour any statistic). Nous montrerons en nous limitant au cas non relativiste qu'une mécanique quantique de particules de statistique fractionnaire se formule de manière naturelle à l'aide de la théorie de jauge de C.S.. Nous étudierons ensuite brièvement quelques particularités d'une théorie des champs à 2+1 dimensions : algèbres de Poincaré et de Clifford, propagateur et induction du terme de C.S. lors de la renormalisation de l'électrodynamique. Pour cloturer ce chapitre, nous envisagerons succintement le cas de la théorie non-abélienne.

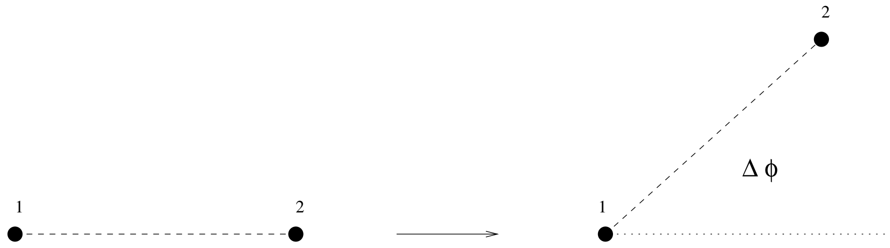
5.2 Introduction à l'étude des particules de statistique fractionnaire

5.2.1 Statistique fractionnaire à deux dimensions

A l'aide d'un raisonnement très simple, puis par des considérations de topologie algébrique faisant appel aux notions vues dans le chapitre 3, nous allons constater que le spin d'une particule amenée à évoluer dans un plan n'est pas quantifié en des valeurs entières ou demi-entières, comme dans le cas tri-dimensionnel.
Tout d'abord, le fait qu'à deux dimensions, le spin peut être arbitraire n'est pas trop surprenant. En effet, le groupe des rotations $SO(n)$ est abélien pour $n = 2$ et non abélien pour $n > 2$. Il n'y a dès lors aucune relation de commutation et aucune restriction sur les valeurs propres possibles du spin n'est imposée par l'algèbre.
Considérons ensuite une situation très simple : un système constitué par deux particules ponctuelles identiques dans un espace de dimension quelconque. La notion de statistique est usuellement introduite en relation avec le signe acquis par la fonction d'onde totale lors de l'échange des deux particules. Si celle-ci change de signe lors de l'échange , elle décrit un système de fermions,

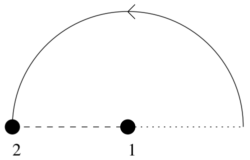
sinon elle se rapporte à un système de bosons. Donnons une définition plus générale de la statistique d'un système de deux particules identiques. Soit $\psi(1,2)$ la fonction d'onde du système. Pour un système où le moment angulaire est défini, supposons que la particule 2 effectue une rotation autour de 1 d'un angle azimutal $\Delta\varphi$; la fonction d'onde devient

$$\psi(1,2) \to e^{i\nu\Delta\varphi}\psi(1,2).$$

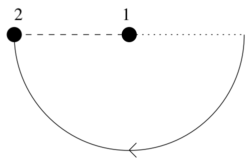


La phase acquise lors de la transformation dépend du paramètre ν appelé généralement *statistique* du système. Reprenons le cas de l'échange des deux particules. Il existe deux possibilités d'effectuer cette transformation :

1. La particule 2 tourne d'un angle $\Delta\varphi = \pi$ autour de 1, puis une translation rigide ramène le centre de masse à sa position originale. La fonction d'onde acquiert une phase $\exp{(i\nu\pi)}$.



2. La particule 2 tourne d'un angle $\Delta\varphi = -\pi$ autour de 1, puis de nouveau une translation rigide. La fonction d'onde acquiert une phase $\exp{(-i\nu\pi)}$.



Si la dimension d de l'espace est telle que $d \geq 3$, les deux transformations sont équivalentes, car il suffit d'utiliser une troisième dimension pour sortir du plan et amener la première figure sur la seconde. La phase acquise par la fonction d'onde est donc identique dans les deux cas, ce qui entraîne $\nu = 0$ ou 1. A trois dimensions et plus, la statistique ne peut être arbitraire, ou, autrement dit, il ne peut y avoir que des fermions ou des bosons.
Si par contre $d = 2$, les deux situations ne sont pas équivalentes car il n'est pas possible de passer d'une transformation à une autre de manière continue. Il n'existe alors aucune contrainte sur le paramètre ν.
D'une manière plus formelle, mais néanmoins élégante, nous allons montrer que ceci peut être déduit de considérations simples de mécanique quantique et de topologie. Nous désignons par M_N^d la variété des configurations de N particules identiques dans un espace de dimension d. Dans le formalisme des intégrales de chemin de la mécanique quantique (formalisme de Feynman), l'amplitude de transition du système d'une configuration q au temps t à une configuration q' au temps t' est donnée par

$$< q', t'|q, t >= \int [Dq] e^{i\int_t^{t'} d\tau L[q(\tau),\dot{q}(\tau)]} = K(q', t'; q, t).$$

La fonction d'onde de la configuration finale s'écrit alors

$$\psi(q',t') = \int_{M_N^d} dq\, K(q',t';q,t)\psi(q,t).$$

Nous nous intéresserons aux boucles dans M_N^d et nous posons donc $q' = q$. Nous savons que, pour une variété donnée, nous pouvons répertorier les différentes boucles de cette variété dans les différentes classes du groupe d'homotopie π_1. L'amplitude de transition K peut donc être décomposée en une somme sur les classes du groupe d'homotopie $\pi_1(M_N^d)$:

$$K(q,t';q,t) = \sum_{\alpha\in\pi_1(M_N^d)} \chi(\alpha)K_\alpha(q,t';q,t), \tag{5.1}$$

avec

$$K_\alpha(q,t';q,t) = \int [Dq_\alpha] e^{i\int_t^{t'} d\tau\, L(q_\alpha(\tau),\dot{q}_\alpha(\tau))},$$

$q_\alpha(\tau)$ une boucle de la classe d'homotopie $[\alpha]$ et $\chi(\alpha)$ un poids attaché à cette classe. La relation

$$K(q'',t'';q,t) = \int_{M_N^d} dq'\, K(q'',t'';q',t')K(q',t';q,t)$$

impose la condition

$$\chi(\alpha_1)\chi(\alpha_2) = \chi(\alpha_1 \cdot \alpha_2),$$

pour toutes classes de courbes α_1 et α_2. Cette relation nous montre dans quelle mesure mécanique quantique et topologie se tiennent la main. Le facteur d'amplitude χ_α doit appartenir à une représentation de dimension 1 du groupe fondamental des configurations du système.

Pour terminer, il nous reste à déterminer le groupe d'homotopie de la variété M_N^d. Celle-ci est l'espace de configuration d'un système de N particules dans un espace euclidien à d dimensions, $\mathbb{R}^d$. Une configuration du système est donnée par un élément de $\mathbb{R}^{N\,d}$. Si nous supposons que les particules ne peuvent s'interpénétrer, nous devons retirer l'ensemble

$$\Delta = \left\{ (\mathbf{r}_1, \cdots, \mathbf{r}_N) \in \mathbb{R}^{N\,d} : \mathbf{r}_i = \mathbf{r}_j \text{ pour } i \neq j \right\}.$$

De plus, si nous admettons que les particules identiques sont indiscernables, nous devons identifier deux configurations différant par l'ordre des particules. Il faut donc effectuer le quotient par le groupe de permutation de N éléments S_N. Au total, on a

$$M_N^d = \frac{\mathbb{R}^{N\,d} - \Delta}{S_N}.$$

Il reste alors à consulter un traité de topologie algébrique[1] pour obtenir

$$\pi_1(M_N^d) = \left\{ \begin{array}{ll} S_N & \text{si } d \geq 3 \\ B_N & \text{si } d = 2 \end{array} \right. .$$

Le groupe B_N (B pour *braid group*) est une généralisation du groupe discret de permutation S_N, qui comprend S_N comme sous-groupe. Nous voyons donc à nouveau la différence fondamentale existant entre un système évoluant dans un espace à deux dimensions ou dans un espace de dimension plus élevée.

Considérons à nouveau le cas d'un système de deux particules identiques.

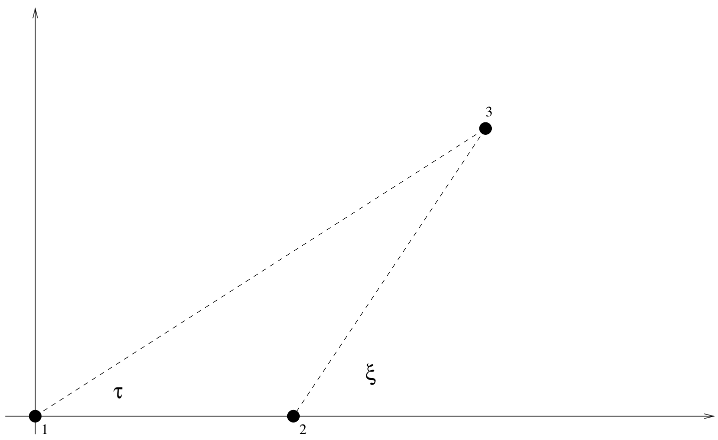
– Si $d = 2$, on a $B_2 = \mathbb{Z}$,

1. L'obtention du groupe fondamental d'une telle variété est un problème classique de topologie algébrique, résolu dans les années 60

– Si $d = 3$, on a $S_2 = \mathbb{Z}_2$.

Nous retrouvons le fait que le groupe de permutation S_2, isomorphe au groupe d'inversion, possède seulement deux représentations (il y a autant de classes que d'éléments, vu le caractère abélien) qui correspondent aux statistiques bosoniques et fermioniques ($\nu = 0, 1$). Le groupe B_2 admet quant à lui une infinité dénombrable de représentations, correspondant aux valeurs différentes que peut prendre le paramètre ν.

Avant de développer un peu plus avant le lien qui existe entre représentations de B_N et la dynamique du système, considérons l'exemple d'un ensemble de 3 particules identiques.



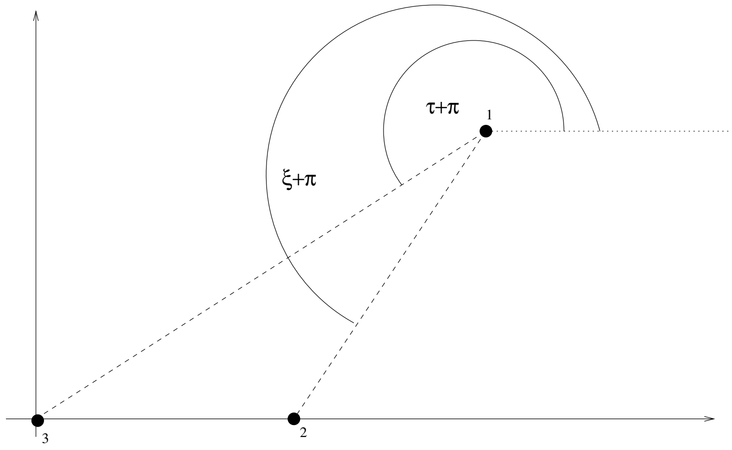
A l'instant t, une configuration du système est donnée par les valeurs des angles entre paires de particules. Ainsi, sur la figure précédente, on a

$$\varphi_{12}(t) = 0$$
$$\varphi_{13}(t) = \tau$$
$$\varphi_{23}(t) = \xi$$

Les angles $\varphi_{ij}(t)$ sont définis par

$$\varphi_{ij} = \tan^{-1}\left(\frac{x_j^2 - x_i^2}{x_j^1 - x_i^1}\right), \tag{5.2}$$

où (x_i^1, x_i^2) désignent les coordonnées de la particule i. Supposons qu'à l'instant t' les particules ont atteint la configuration illustrée dans la figure ci-dessous.



Cette situation correspond à l'échange des particules 1 et 3. La configuration finale est donnée par

$$\begin{aligned}\varphi_{12}(t') &= \xi + \pi \\ \varphi_{13}(t') &= \tau + \pi \\ \varphi_{23}(t') &= \pi\end{aligned}$$

Remarquons que l'on a

$$\sum_{i<j} \varphi_{ij}(t') - \sum_{i<j} \varphi_{ij}(t) = 3\pi.$$

Pour réaliser une boucle dans l'espace des configurations, un nombre entier d'échanges est nécessaire. Dans le cas qui nous occupe, ce nombre entier est 3. Remarquons que la configuration suivante

$$\begin{aligned}\varphi_{12}(t') &= \xi + 3\pi \\ \varphi_{13}(t') &= \tau + \pi \\ \varphi_{23}(t') &= \pi\end{aligned}$$

est identique à la précédente si l'on s'en tient à la position qu'occupe chaque particule. Elles correspondent à la même permutation des positions 1 et 3. Cependant, d'un point de vue dynamique, cinq permutations ont été effectuées :

$$\sum_{i<j} \varphi_{ij}(t') - \sum_{i<j} \varphi_{ij}(t) = 5\pi.$$

Sans développer plus avant la théorie du $braid - group$, nous dirons simplement que là où le groupe de permutation S_3 identifie les deux configurations finales, le groupe B_3 distingue les deux transformations différentes, au travers du nombre n de permutations. Ce groupe prend en quelque sorte en compte la dynamique de la permutation. Il existe également une représentation graphique. L'opération d'inversion correspondant à la première transformation est illustrée dans la figure suivante. Le groupe B_N est un groupe du type que nous avons rencontré lors de la détermination des groupes fondamentaux. Voir la note du bas de la page 24. Un générateur σ_i est associé à la permutation des particules i et $i+1$ et tout élément du groupe s'écrit comme un produit de générateurs. Ainsi, dans le cas de la première transformation, l'élément du groupe correspondant s'écrit $\sigma_1\sigma_2\sigma_1$. La seconde transformation amenant à la même configuration finale s'écrirait $\sigma_1^3\sigma_2\sigma_1$.
Puisque le groupe B_N tient compte de la dynamique amenant à une permutation de particules, il doit exister un lien avec la grandeur contenant la dynamique du système, le lagrangien. Revenons au problème des représentations du groupe. Dans le cas d'une simple permutation de la particule, disons k et $k+1$, nous avons

$$\chi(\sigma_k) = \exp\left(-i\nu\Delta\varphi_{k\,k+1}\right) = \exp\left(-i\nu \sum_{i<j} \Delta\varphi_{ij}^k\right),$$

où $\Delta\varphi_{ij}^k = \varphi_{ij}^k(t') - \varphi_{ij}^k(t) = \pi\delta_{ik}\delta_{j\,k+1}$. Cela peut être généralisé au cas d'une transformation quelconque

$$\chi(\alpha) = \exp\left(-i\nu \sum_{i<j} \int_t^{t'} d\tau \frac{d}{d\tau}\varphi_{ij}^{\alpha}(\tau)\right).$$

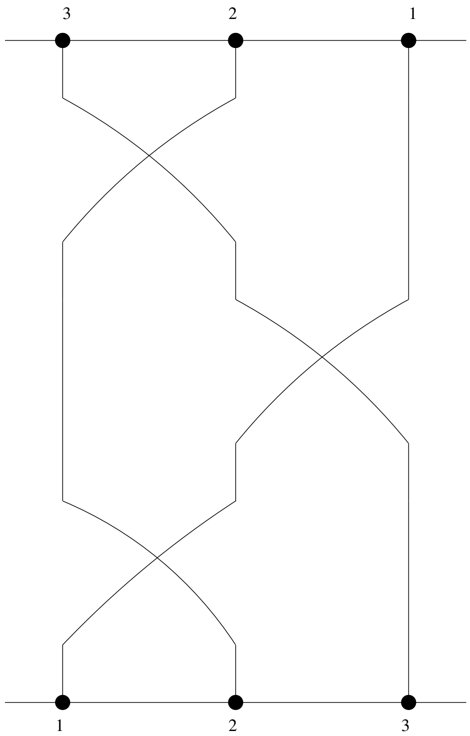


FIG. 5.1 – *Représentation graphique de l'opération d'inversion des particules 1 et 3. Les courbes représentent en quelque sorte les lignes d'univers des particules individuelles, le temps s'écoulant de bas en haut.*

Insérant cette relation dans (5.1), on obtient finalement

$$K(q,t';q,t) = \sum_{\alpha\in\pi_1(M_N^d)} \int [Dq_\alpha] e^{i\int_t^{t'} d\tau \left[L(q_\alpha(\tau),\dot{q}_\alpha(\tau)) - \nu \sum_{i<j} \frac{d}{d\tau}\varphi_{ij}^\alpha(\tau)\right]}.$$

Si nous redéfinissons le lagrangien

$$L' = L - \nu \sum_{i<j} \frac{d}{d\tau}\varphi_{ij}^\alpha(\tau), \tag{5.3}$$

nous obtenons une amplitude toute à fait classique, dans laquelle les différentes classes d'homotopie sont pondérées uniformément. En d'autres termes,

$$K(q,t';q,t) = \sum_{\alpha\in\pi_1(M_N^d)} \int [Dq_\alpha] e^{i\int_t^{t'} d\tau\, [L(q_\alpha(\tau),\dot{q}_\alpha(\tau))]}$$

décrit un système de bosons soumis à la nouvelle dynamique L'. Le terme additionnel ne modifie pas les équations de mouvement, puisqu'il s'agit d'une dérivée totale. Ce terme de surface modifie par contre drastiquement les propriétés statistiques des particules, reliée elle-même à la topologie de l'espace des configurations. Un terme de surface additionnel en liaison avec des propriétés topologiques sonne comme un air connu. Rappelons-nous simplement la transformation que nous avions opérée sur le lagrangien de Y.M. dans (4.39).

5.2.2 Cyons

Avant d'envisager l'application de la théorie de jauge de C.S. à la théorie des particules de statistique fractionnaire, nous allons considérer un cas très simple, constituant le prototype des

anyons .
Soit une particule de masse m chargée plongée dans un champ magnétique créé par un solénoïde infiniment long, de rayon nul, dirigé le long de l'axe des z. Nous ne nous intéressons qu'aux mouvements perpendiculaires à la direction du solénoïde, et considérons donc que la particule se meut dans un plan perpendiculaire à son axe.**?** Cette configuration de champ magnétique est décrite par le potentiel vecteur suivant

$$\mathbf{A}(\mathbf{r}) = \frac{\Phi}{2\pi}\left(\frac{-y}{r^2}\mathbf{e}_x + \frac{x}{r^2}\mathbf{e}_y\right),$$

donnant lieu au champ magnétique (pseudo-scalaire)

$$B = \nabla \wedge \mathbf{A} = \Phi\delta^{(2)}(\mathbf{r}),$$

qui, à cause de la fonction delta, porte souvent le nom de *vortex*. Le système à deux dimensions constitué d'un vortex localisé à l'origine et d'une particule chargée est appelé *cyon*. Le lagrangien non relativiste est

$$L = \frac{1}{2}mv^2 + \frac{e}{c}\mathbf{A}\cdot\mathbf{v}.$$

On a

$$\mathbf{p} = m\mathbf{v} + \frac{e}{c}\mathbf{A}$$

et

$$H = \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 = \frac{1}{2}mv^2.$$

La dynamique classique de ce système est donc identique à celle d'une particule libre. La présence du vortex donne simplement lieu à une redéfinition de l'impulsion canonique. Nous allons bien sûr examiner les effets quantiques de cette configuration particulière.
Le lagrangien étant invariant par rotation (il s'agit, rappelons-le, d'une rotation dans un plan), nous avons conservation du moment angulaire orbital "canonique" $\mathbf{J}_C$:

$$\begin{aligned}\mathbf{J}_C &= \mathbf{r}\wedge\mathbf{p}\\ &= \mathbf{r}\wedge m\mathbf{v} + e\mathbf{r}\wedge\mathbf{A}\\ &= \mathbf{J} + \frac{e\Phi}{2\pi c}\end{aligned}$$

On peut montrer que $\mathbf{J}_C$ possède le spectre conventionnel du moment angulaire du rotateur plan :

$$J_C = m\hbar, \quad m \in \mathbb{Z},$$

ce qui entraîne que le moment angulaire a

$$J = \hbar\left(m - \frac{e\Phi}{hc}\right)$$

pour valeurs propres. La valeur du moment angulaire pour $m = 0$

$$s = \frac{J(m=0)}{\hbar} = -\frac{e\Phi}{hc}, \tag{5.4}$$

est appelée *spin* du cyon. Il n'est, en toute généralité, ni entier, ni demi-entier. Le cyon constitue un exemple simple d'anyon.
Etablissons à présent les propriétés statistiques de telles particules. Considérons un sytème constitué de deux cyons identiques, et supposons que le flux Φ est localisé sur chacune des deux particules. Effectuons maintenant une rotation complète d'un cyon autour de l'autre, en

négligeant les interactions entre vortex et charges. Si, pour fixer les idées la particule 1 a effectué une révolution autour du vortex 2, nous savons que celle-ci acquiert une phase

$$\exp\left(-i\frac{e}{\hbar c}\int_C \mathbf{A}\cdot\mathbf{r}\right).$$

En utilisant le théorème de Stokes, il vient

$$\Delta\phi = \exp\left(-2i\pi\frac{e\Phi}{hc}\right).$$

La phase totale acquise par la fonction d'onde du système des deux particules est

$$\exp\left(-2i\pi\frac{2e\Phi}{hc}\right).$$

Il est alors immédiat d'en déduire la statistique du cyon

$$\nu = -\frac{2e\Phi}{hc},$$

et on a

$$\nu = 2s.$$

Il est important de remarquer que cette statistique n'est pas la statistique de la particule individuelle, mais celle du cyon : la particule associée au vortex de force Φ. Par exemple, une particule bosonique, associée à un flux de force $\Phi = \frac{1}{2}\frac{hc}{e}$ ($\frac{hc}{e}$ est l'unité fondamentale de flux) se comporte comme un fermion.

5.3 Chern-Simons et statistique fractionnaire

La procédure suivie dans le paragraphe précédent a l'avantage d'être très simple; nous avons vu qu'elle fait appel à des notions élémentaires de mécanique quantique. Cependant, l'opération de transformation d'une particule ordinaire en un anyon s'opère à l'aide de l'introduction un peu brutale d'un potentiel extérieur $\mathbf{A}$, par “collage” d'un vortex de champ magnétique sur la particule. De plus, elle s'inscrit dans un langage non relativiste. Une méthode un peu plus douce et peut-être un peu plus naturelle est de poser que $\mathbf{A}$ est la partie spatiale d'un champ de jauge $U(1)$ de dimension 2+1, dont l'action (invariante de Lorentz) est donnée par

$$S_{CS} = \int d^3x\mathcal{L}_{CS} = \int d^3x\frac{\kappa}{2}\,\varepsilon_{\mu\nu\rho}\,A^\mu\partial^\nu A^\rho.$$

Ce champ de jauge est couplé à un système matériel consistant en N particules ponctuelles non relativistes de masse m et de charge e. Le courant s'écrit

$$j^\mu(x) = \sum_{i=1}^{N} e\,v_i^\mu(t)\delta^{(2)}(\mathbf{x}-\mathbf{r}_i(t)) = (\rho,\mathbf{j}), \tag{5.5}$$

où $v_i^\mu(t) = (1,\mathbf{v}_i(t))(= (c,\mathbf{v}_i(t)))$. Les “vecteurs” j^μ et v^μ, ne sont bien sûr pas des trivecteurs au sens des transformations de Lorentz. Cette notation pseudo-relativiste est simplement un raccourci commode d'écriture qui permet, par exemple, d'exprimer l'équation de continuité

$$\partial_t\rho + \nabla\cdot\mathbf{j} = 0,$$

sous la forme

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

comme pour une théorie relativiste. Remarquons bien que l'on ne s'intéresse pas à la dynamique des particules matérielles. Celle-ci est supposée connue. Plus précisément, les fonctions $\mathbf{r}_i(t)$ sont des fonctions connues du temps et les $\mathbf{v}_i(t)$ sont leurs dérivées.
Toute la dynamique d'interaction entre les particules matérielles et le champ de jauge se trouve par contre contenue dans l'action

$$S = \int d^3x \left[\frac{\kappa}{2}\,\varepsilon_{\mu\nu\rho}\,A^\mu\partial^\nu A^\rho + A^\mu j_\mu\right].$$

Nous avons au chapitre deux étudié cette théorie de jauge en détail. Les équations de mouvement s'écrivent

$$^*F^\mu = \frac{-j^\mu}{\kappa},$$

ou encore

$$B = \frac{\rho}{\kappa}, \tag{5.6}$$

$$E_1 = \frac{j^2}{\kappa}, \tag{5.7}$$

$$E_2 = -\frac{j^1}{\kappa}. \tag{5.8}$$

Cette situation est fondamentalement différente du cas des cyons évoqués plus haut. Le champ magnétique B était imposé comme une contrainte extérieure, et nous avons du supposer, de manière très peu élégante, que le flux magnétique était localisé sur la particule matérielle. Dans le cas qui nous occupe, point n'est besoin d'introduire un champ extérieur “à la main” et nous voyons, au travers de (5.6), que le lien entre champ et particule découle naturellement des équations de champ. On peut cependant se demander s'il existe, sous forme d'une force de Lorentz, une self-interaction d'une particule individuelle. La réponse est négative car

$$\begin{aligned} E_i + \varepsilon_{ij}v^iB &= \frac{e}{\kappa}\varepsilon_{ij}(v^j - v^j)\delta^{(2)}(\mathbf{x}-\mathbf{r}(t)) \\ &= 0. \end{aligned} \tag{5.9}$$

Si nous intégrons (5.6) sur un petit disque centré sur une particule, nous obtenons le flux magnétique porté par celle-ci :

$$\Phi = \int_C d^2xB = \frac{e}{\kappa}.$$

Nous voyons une fois encore que la dynamique de C.S. réalise d'une manière naturelle ce que nous avions du imposer comme une contrainte externe pour un système de cyons.
Nous avons également vu lors de l'analyse de contraintes de cette théorie qu'il était possible au moyen des crochets de Dirac de faire apparaître une chaîne de contraintes de première classe tout à fait typique des théories de jauge habituelles, ainsi par exemple la condition de Gauss (2.102). Poser une condition de jauge a donc un sens, comme dans la théorie de Maxwell. Nous posons donc $A_0 = 0$. Nous savons que cette condition ne fixe pas totalement la jauge, car il nous est encore possible d'opérer une transformation sur les composantes spatiales du potentiel (cf. (4.35) dans un cas non abélien). Cette liberté résiduelle est éliminée [2]

$$\partial_i A^i = 0.$$

2. Afin d'alléger le discours et de dériver rapidement une solution analytique pour le potentiel, nous opérons ce raccourci, abusivement utilisé dans la littérature, en posant deux conditions de jauge. Une double condition de ce type ne peut se justifier dans le cadre de la théorie des systèmes contraints, à moins de modifier le lagrangien de départ.

Celle-ci permet d'obtenir une expression pour le potentiel vecteur. En effet, (5.6) s'écrit

$$\partial_1 A^2 - \partial_2 A^1 = \frac{\rho}{\kappa},$$

et en dérivant par rapport à x^1, puis en utilisant la condition de jauge secondaire, il vient

$$\nabla^2 A^2 = \frac{1}{\kappa}\partial_1 \rho,$$

et, d'une manière analogue,

$$\nabla^2 A^1 = -\frac{1}{\kappa}\partial_2 \rho.$$

La fonction de Green du laplacien à deux dimensions est donnée par

$$\nabla^2 \left(\frac{1}{2\pi}\log|\mathbf{x}|\right) = \delta^{(2)}(\mathbf{x}).$$

En effectuant la convolution

$$A^i(x) = \frac{1}{2\pi\kappa}\int d^2y\, \varepsilon^{ij}[\partial_j \rho](y)\log|\mathbf{x}-\mathbf{y}|,$$

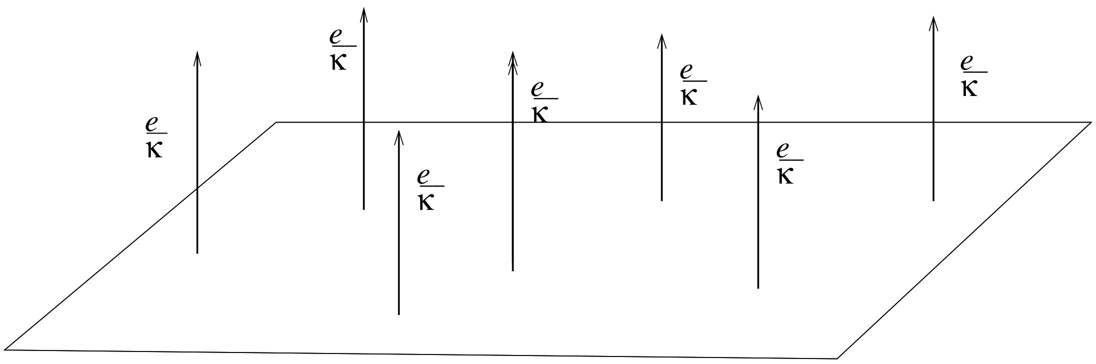
il vient sans difficulté, en utilisant (5.5),

$$A^i(x) = \frac{e}{2\pi\kappa}\varepsilon^{ij}\sum_{k=1}^{N}\frac{x^j - r_k^j(t)}{|\mathbf{x}-\mathbf{r}_k(t)|}.$$

Avec cette solution, nous pouvons dériver le champ magnétique agissant sur une particule individuelle, dont la position est $\mathbf{r}_a(t)$.

$$B(\mathbf{r}_a) = \frac{e}{\kappa}\sum_{k\neq a}^{N}\delta^{(2)}(\mathbf{r}_a - \mathbf{r}_k), \tag{5.10}$$

ce qui montre que chaque particule voit les $N-1$ autres comme sources d'un flux magnétique de force $\frac{e}{\kappa}$.



Exmaminons à présent la statistique à laquelle répondent de telles particules. Considérons à nouveau un système de deux particules matérielles soumises à un potentiel de jauge de C.S. . Si une d'entre elles effectue une rotation complète autour de l'autre, la fonction d'onde acquiert une phase

$$\exp\left(ie\oint \mathbf{A}\cdot d\mathbf{x}\right) = \exp\left(i\frac{e^2}{\kappa}\right).$$

Le spin est alors

$$\nu = \frac{e^2}{2\pi\kappa}, \tag{5.11}$$

qui à nouveau peut prendre des valeurs non entières ou non demi-entières. Des particules chargées soumises à un potentiel de jauge de C.S. sont donc des anyons. Il est ainsi possible de transformer des bosons en fermions, ou vice-versa, ou encore en particules de spin quelconque en ajustant la constante de couplage κ.
Remarquons que le potentiel (5.3) peut s'écrire

$$\begin{aligned} A^i(\mathbf{r}_1, \cdots, \mathbf{r}_N) &= \frac{e}{2\pi\kappa}\varepsilon^{ij}\frac{\partial}{\partial r_i^j}\sum_{j\neq i}^{N}\ln|\mathbf{r}_i - \mathbf{r}_j| \\ &= -\frac{e}{2\pi\kappa}\frac{\partial}{\partial r_i^j}\sum_{j\neq i}^{N}\varphi_{ij}, \end{aligned}$$

où φ_{ij} est donné par (5.2). Le lagrangien (non-relativiste) de la dynamique des particules soumises au champs de jauge de C.S. est

$$\begin{aligned} L &= \sum_{i=1}^{N}\left(\frac{1}{2}mv_i^2 + e\mathbf{v}_i\cdot\mathbf{A}_i\right) \\ &= \sum_{i=1}^{N}\left(\frac{1}{2}mv_i^2\right) - \frac{e^2}{2\pi\kappa}\sum_{i=1}^{N}\sum_{j\neq i}v_i^k\frac{\partial}{\partial r_i^k}\varphi_{ij} \\ &= \sum_{i=1}^{N}\left(\frac{1}{2}mv_i^2\right) - \frac{e^2}{2\pi\kappa}\sum_{i<j}(v_i^k - v_j^k)\frac{\partial}{\partial r_i^k}\varphi_{ij}, \end{aligned}$$

en utilisant la propriété $\frac{\partial}{\partial r_j^k}\varphi_{ij} = -\frac{\partial}{\partial r_i^k}\varphi_{ij}$. La somme du deuxième terme de la dernière relation n'étant rien d'autre que la dérivée totale $\frac{d}{dt}\varphi_{ij}$, il vient finalement

$$L = \sum_{i=1}^{N}\left(\frac{1}{2}mv_i^2\right) - \frac{e^2}{2\pi\kappa}\left(\sum_{i<j}\frac{d}{dt}\varphi_{ij}\right).$$

En se rappelant la relation (5.3), nous obtenons directement la valeur du paramètre de statistique

$$\nu = \frac{e^2}{2\pi\kappa},$$

en accord avec (5.11).

Application Les anyons ne sont pas seulement des fantaisies topologiques amusantes et d'un intérêt purement mathématique. Ils jouent un rôle important dans certains phénomènes physiques. Il est évident qu'il ne s'agit pas de particules réelles, au sens habituel des particules matérielles, puisque le monde dans lequel nous vivons possède trois dimensions spatiales. Il existe cependant des systèmes que l'on rencontre dans l'étude de la matière condensée dont l'évolution peut être décrite dans un espace à deux dimensions spatiales. Un exemple important est celui de l'effet Hall quantique. Des électrons sont capturés dans une très fine couche à l'interface de deux semi-conducteurs différents. Le système est porté à une température basse (de l'ordre de quelques Kelvin) et soumis à un champ magnétique de forte intensité (plus de 10 Tesla) perpendiculaire à la couche. Température et champ magnétique confinent le mouvement des électrons au plan de l'interface en gelant tout déplacement dans la direction perpendiculaire à celle-ci. L'effet Hall quantique est caractérisé par le fait que la conductance de Hall, σ_H est quantifiée en unité de e^2/h :

$$\sigma_H = \nu \frac{e^2}{h}.$$

Le nombre ν peut être un entier (Effet Hall entier) ou une fraction (Effet Hall fractionnaire). Le problème quantique est apparenté au problème de Landa, ou problème d'une particule chargée dans un champ magnétique. Dans le cas d'école étudié dans tout cours de mécanique quantique, on montre que le problème se ramène à l'étude d'une particule dans un puits de potentiel harmonique. Bien que le problème que nous décrivons soit infiniment plus complexes, cette caractéristique fondamentale subsiste en toile de fonds. Le problème se ramène en fait à l'étude de la population des niveaux de Landau. Dans le cas de l'effet Hall entier, la quantification de σ_H est une manifestation directe de la quantification d'un système d'électrons sans interaction dans un champ magnétique extérieur; les électrons répondent à la statistique de Fermi et la populatin des niveaux de Landau est essentiellement dictée par le principe d'exclusion. Si l'interaction coulombienne entre électrons est prise en compte, on assiste à une modification de l'état fondamental et à une répartition plus complexe des électrons sur les niveaux. La tendance est à la condensation des électrons sur ce nouvel état fondamental, et seule une fraction des niveaux de Landau sont occupés, ce qui est un signe d'une modification des propriétés statistiques du système. Un grand atout de la théorie est qu'elle possède des solutions exactes. Il existe par exemple une expression analytique pour la fonction d'onde de l'état fondamental. Il est alors relativement simple d'établir les propriétés statistiques relatives à ces solutions. Le caractère anyonique apparaît lorsqu'on examine les excitations de l' état fondamental. Comme pour l'oscillateur harmonique, on associe une pseudo-particule au quantum d'excitation (le phonon en est un exemple). Ces quasi-particules manifestent des propriétés typiques de particules de statistique fractionnaire.

5.4 Quelques caractéristiques

5.4.1 Propagateur

Dans l'étude que nous avons réalisée au cours du chapitre 2, nous avons mentionné brièvement une propriété tout à fait caractéristique de cette théorie : l'apparition d'une masse dans la théorie de jauge. Nous insistons ici sur le fait que la génération de cette masse est tout à fait indépendante d'un quelconque procédé apparenté au mécanisme de Higgs. Nous allons le montrer explicitement en calculant le propagateur du photon. Nous repartons donc du lagrangien $\mathcal{L}_{MCS}$ que nous écrivons sous une forme légèrement différente

$$\mathcal{L}_{MCS} = -\frac{1}{4e^2} F_{\mu\nu} F^{\mu\nu} + \frac{\kappa}{2} \varepsilon_{\mu\nu\rho} A^\mu \partial^\nu A^\rho,$$

avec la constante de couplage e^2, de dimension 1, introduite ici au dénominateur, afin de garder un potentiel de jauge de dimension 1, comme dans le cas quadri-dimensionnel. La constante de couplage κ est donc sans dimension. Nous introduisons une condition de jauge relativiste, en ajoutant un terme $\mathcal{L}_{gf}$.

$$\mathcal{L}_{MCS} = -\frac{1}{4e^2} F_{\mu\nu} F^{\mu\nu} + \frac{\kappa}{2} \varepsilon_{\mu\nu\rho} A^\mu \partial^\nu A^\rho - \frac{1}{2ae^2} (\partial_\mu A^\mu)^2.$$

En développant le terme cinétique et “intégrant par parties”, il vient

$$\mathcal{L}_{MCS} = \frac{1}{2e^2} A_\nu \left[\partial^2 g^{\mu\nu} - \frac{a-1}{a} \partial^\mu \partial^\nu + \kappa e^2 \varepsilon^{\mu\nu\rho} \partial_\rho \right] A_\mu.$$

Dans l'espace des impulsions, le propagateur Δ est défini par

$$\frac{1}{2e^2}\left[-p^2 g^{\mu\nu} + \frac{a-1}{a}p^\mu p^\nu + i\kappa e^2 \varepsilon^{\mu\nu\rho}p_\sigma\right]\Delta_{\nu\rho}(p) = \delta^\mu_\rho.$$

En posant

$$\Delta_{\nu\rho}(p) = A(p)g_{\nu\rho} + B(p)p_\nu p_\rho + C(p)\varepsilon_{\lambda\nu\rho}p^\lambda,$$

on obtient le système d'équation suivant

$$\begin{aligned} -A(p) - i\kappa e^2 C(p) &= 1 \\ -\frac{1}{a}p^2 B(p) + \frac{a-1}{a}A(p) + i\kappa e^2 C(p) &= 0 \\ -p^2 C(p) + i\kappa e^2 A(p) &= 0 \end{aligned}$$

Après un peu d'algèbre élémentaire, il vient

$$\Delta_{\mu\nu}(p) = 2\,e^2\left[\frac{-p^2 g_{\mu\nu} - p_\mu p_\nu - i\kappa e^2 \varepsilon_{\mu\nu\rho}p^\rho}{p^2(p^2-\kappa^2 e^4)} + a\frac{p_\mu p_\nu}{p^4}\right],$$

montrant clairement la présence d'un pôle en $p^2 = \kappa^2 e^4$.

5.4.2 Algèbre de Poincaré à 2+1 dimensions

A trois dimensions, les générateurs de l'algèbre de Poincaré, $L_{\mu\nu}$ pour les transformations de Lorentz et P_μ pour les translations, satisfont les relations de commutation tout à fait classique

$$\begin{aligned} [J^\mu, J^\nu] &= i\varepsilon^{\mu\nu\rho}J_\rho, \\ [J^\mu, P^\nu] &= i\varepsilon^{\mu\nu\rho}P_\rho, \\ [P^\mu, P^\nu] &= 0. \end{aligned}$$

Le pseudo-vecteur J_μ est associé au tenseur antisymétrique $L_{\mu\nu}$ par la relation $J^\mu = \frac{1}{2}\varepsilon^{\mu\nu\rho}L_{\nu\rho}$. Les représentations irréductibles de cette algèbre sont caractérisées par les valeurs propres des deux opérateurs de Casimir :

$$P^2 = P^\mu P_\mu, \quad W = P_\mu J^\mu,$$

où W est l'opérateur (pseudo-scalaire) de Pauli-Lubanski. Les représentations Φ sont définies par

$$P^2\Phi = m^2\Phi, \quad W\Phi = -sm\Psi,$$

définissant la masse m et le spin s.

1. Dans le cas d'un champ $\phi(p)$ dans l'espace des impulsions sur lequel P^μ agit par simple multiplication et pour lequel J^μ est simplement un opérateur de moment angulaire orbital :

   $$P^\mu\phi = p^\mu\phi, \quad J^\mu\phi = -i\varepsilon^{\mu\nu\rho}p_\nu\frac{\partial}{\partial p^\rho}\phi,$$

   les équations aux valeurs propres se réduisent simplement à l'équation de Klein-Gordon, et $P\cdot J = 0$. Il s'agit, bien sûr, d'un champ scalaire de spin 0.
2. Pour un champ à deux composantes ψ, considérons

   $$J^\mu = -i\varepsilon^{\mu\nu\rho}p_\nu\frac{\partial_\nu}{\partial^\rho}\mathbf{1} - \frac{1}{2}\gamma^\mu,$$

   pour lequel la seconde équation de mouvement donne l'équation de Dirac $(i\gamma^\mu\partial_\mu - m)\Psi = 0$, correspondant à la valeur propre de spin $s = \frac{1}{2}$.

3. Le troisième cas est celui des particules vectorielles. Prenons

$$(J^\mu)_{\alpha\beta} = -i\varepsilon^{\mu\nu\rho} p_n u \frac{\partial}{\partial p^\rho} \delta_{\alpha\beta} + i\varepsilon^\mu_{\alpha\beta}.$$

La seconde équation aux valeurs propres nous donne

$$(P \cdot J)_{\alpha\beta}\ {}^*F^\beta = i\varepsilon^\mu_{\alpha\beta} p_\mu\ {}^*F^\beta = -sm\ *F_\alpha$$

et nous retrouvons la forme des équations de mouvement (2.72), avec $m = \kappa e^2$ et $s = \pm 1$.

D'une façon encore différente, nous avons fait apparaître la masse κe^2. Ceci nous permet une nouvelle fois de remarquer que celle-ci est vraiment intrinsèque à la théorie de jauge et totalement indépendante d'un procédé de génération. Le terme de Chern-Simons donne lieu à des équations de mouvement pour des champs de jauge appartenant à une représentation irréductible de l'algèbre de Poincaré et n'est nullement un terme de brisure de symétrie, comme dans le cas du procédé de Higgs.

5.4.3 Algèbre de Dirac à 2+1 dimensions

A trois dimensions, il est possible de construire une algèbre de Clifford, à l'aide de matrices 2×2. Les champs fermioniques sont donc représentés par des spineurs à deux composantes. Une représentation naturelle est la représentation de Dirac, contruite sur les matrices de Pauli :

$$\begin{aligned}
\gamma^0 = \sigma^3 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \\
\gamma^1 = i\sigma^1 &= \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \\
\gamma^2 = i\sigma^2 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\end{aligned}$$

Elles vérifient, bien sûr, les relations d'anticommutation de l'algèbre

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu},$$

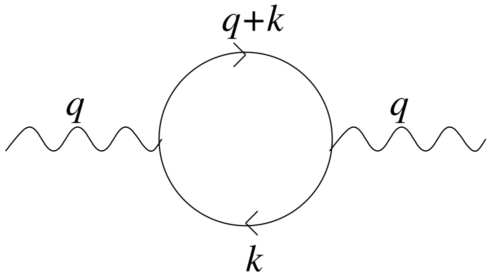
et la relation

$$\mathrm{tr}(\gamma^\mu \gamma^\nu \gamma^\rho) = -2i\varepsilon^{\mu\nu\rho}. \tag{5.12}$$

Cette dernière diffère de l'espace à 4 dimensions où nous savons que la trace du produit d'un nombre impair de matrices de Dirac est nulle. Il est aussi à noter qu'il n'est pas possible de construire une matrice du type γ^5 qui anticommute avec toutes les matrices de Dirac. La notion de chiralité est donc inexistante dans une théorie à trois dimensions.

5.4.4 Induction du terme de Chern-Simons

Une caractéristique également typique de la théorie est que le terme de C.S. peut être induit par le calcul de corrections radiatives dans le cadre de l'électrodynamique quantique à 3 dimensions, en l'absence de ce terme dans le lagrangien.
Considérons la correction à une boucle au propagateur du photon



Les règles de Feynman sont analogues aux règles habituelles de la QED à quatre dimensions. On est amené à calculer l'intégrale

$$\Pi_2^{\mu\nu}(q) = \int \frac{d^3k}{(2\pi)^3} \mathrm{tr}\left[\gamma^\mu \frac{\not q + \not k + m}{(q+k)^2 - m^2} \gamma^\nu \frac{\not k + m}{k^2 - m^2}\right].$$

Dans le calcul habituel de la correction à quatre dimensions le numérateur se développe en éliminant les traces d'un produit de trois matrices de Dirac. Dans le cas présent, ces termes doivent être pris en compte, vu (5.12). La partie classique se calcule de la manière habituelle, avec régularisation. Le terme additionnel s'écrit

$$\Pi_{2,3d}^{\mu\nu}(q) = \varepsilon^{\mu\nu\rho} p_\rho \Pi_{3d}(q),$$

où

$$\Pi_{3d}(q) = 2im \int \frac{d^3k}{(2\pi)^3} \frac{1}{[(p+k)^2 - m^2][k^2 - m^2]}.$$

Cette intégrale est convergente. Après passage dans l'espace euclidien et aux coordonnées polaires, on obtient

$$\Pi_{3d}(q) = \frac{1}{2\pi} \frac{m}{|p|} \arcsin\left(\frac{|p|}{\sqrt{p^2 + 4m^2}}\right).$$

La valeur de cette intégrale n'est pas fondamentale pour notre discussion. Par contre, la forme de $\Pi_{2,3d}^{\mu\nu}(q)$ est très intéressante. En effet, si nous représentons formellement les deux photons externes par A^μ et A^ν, il vient en contractant

$$A_\mu(q)\Pi_{2,3d}^{\mu\nu}(q)A_\nu(q) \propto A_\mu(q)\varepsilon^{\mu\nu\rho} p_\rho A_\nu(q),$$

et en passant à l'espace des coordonnées

$$A_\mu(x)\varepsilon^{\mu\nu\rho}\partial_\rho A_\nu(x),$$

soit le terme bien connu de Chern-Simons.
Une discussion plus détaillée et plus rigoureuse nous amènerait à développer perturbativement l'action effective. Retenons simplement qu'à trois dimensions, les propriétés spéciales des matrices de Dirac, en particulier (5.12), conduisent directement à l'induction du terme de C.S. Ceci résulte du calcul d'une correction à l'ordre d'une boucle. Etant donné l'origine topologique du terme de C.S., on peut s'attendre à ce que celui-ci ne subisse pas de correction ultérieure aux ordres plus élevés. Coleman et Hill ont en effet démontré quen dans la théorie abélienne, il n'y a plus de contribution au terme de C.S. au-delà de la correction à l'ordre d'une boucle. Ce résultat est connu sous le nom de "théorème de Coleman et Hill". Le rôle particulier joué par le terme de C.S. dans le cadre de la renormalisation de l'électrodynamique quantique dans le plan est un autre caractéristique de la théorie.

5.4.5 Théorie non abélienne - quantification de la constante de couplage

Le lagrangien de la théorie de C.S. non abélienne s'écrit

$$\mathcal{L}_{CS} = \kappa\varepsilon^{\mu\nu\rho}\mathrm{tr}\left(\mathcal{A}_\mu\partial_\nu\mathcal{A}_\rho + \frac{2}{3}\mathcal{A}_\mu\mathcal{A}_\nu\mathcal{A}_\rho\right). \tag{5.13}$$

On aura reconnu l'argument de la dérivée du second membre de (4.6), écrit à trois dimensions. Les équations de champ sont tout à fait analogues au cas abélien, la courbure $F^{\mu\nu}$ étant comme toujours donnée par

$$F_{\mu\nu} = \partial_\mu\mathcal{A}_\nu - \partial_\nu\mathcal{A}_\mu + [\mathcal{A}_\mu, \mathcal{A}_\nu].$$

La différence avec le cas abélien se remarque dans le comportement du lagrangien non abélien sous une transformation de jauge

$$\mathcal{A}_\mu \to g^{-1}A_\mu g + g^{-1}\partial_\mu g.$$

Il vient en effet

$$\mathcal{L}_{CS} \to \mathcal{L}_{CS} - \kappa\varepsilon^{\mu\nu\rho}\partial_\mu\mathrm{tr}\left(\partial_\nu g g^{-1}\mathcal{A}_\rho\right) - \frac{\kappa}{3}\varepsilon^{\mu\nu\rho}\mathrm{tr}\left(g^{-1}\partial_\mu g\, g^{-1}\partial_\nu g\, g^{-1}\partial_\rho g\right).$$

Nous voyons que l'action, contrairement au cas abélien n'est pas invariante de jauge, mais change d'une quantité égale au nombre d'enroulement (*winding number*) de la transformation de jauge construite sur g. On a donc

$$\mathcal{S}_{CS} \to \mathcal{S}_{CS} - 8\pi^2\kappa N.$$

Ceci a une conséquence importante dans le cadre de la quantification de la théorie non abélienne. En effet, lors de la quantification, il s'agit de s'assurer que

$$\exp\left(iS\right)$$

demeure invariante sous une transformation de jauge (quantification dans le formalisme fonctionnel). Cela est réalisé pour des valeurs particulières de la constante de couplage κ :

$$\kappa = \frac{k}{4\pi}, \quad k \in \mathbb{Z}. \tag{5.14}$$

Notons qu'il s'agit ici de la contante de couplage non renormalisée, désignée avec une notation standard "κ_{bare}". Il faut également s'assurer qu'une constante de couplage renormalisée satisfasse à la condition de quantification précédente. Il a été montré que dans le cadre d'une théorie de Yang-Mills $SU(n)$ couplée au terme (5.13), sans fermion, que la correction à l'ordre d'une boucle entraîne une renormalisation de la constante de couplage, sous la forme d'un très beau résultat

$$4\pi\kappa_{\mathrm{ren}} = 4\pi\kappa_{\mathrm{bare}} + n.$$

Non seulement la condition de quantification (5.14) est satisfaite, mais le nombre entier qui apparaît grâce à la correction est l'entier n du groupe de jauge $SU(n)$. A l'heure actuelle, il n'a pas encore été démontré que les corrections aux ordres ultérieurs ne modifient pas la relation précédente. Des calculs à l'ordre de deux boucles montrent qu'à cet ordre de perturbation, il n'y a aucune contribution. Il est à noter que la manipulation des intégrales divergentes dans le cadre de la renormalisation des théories de C.S. est loin d'être évidente; la présence du tenseur antisymétrique complique en effet les schémas de régularisation.

Chapitre 6

Conclusion

Revenons d'abord quelques instants sur les interactions entre mathématique et physique que nous avions évoquées dans l'introduction. Nous avons pu constater au cours de ce travail la solidité du formalisme mathématique dans lequel s'expriment les théories de jauge rencontrées en physique. Ce formalisme s'articule en effet sur des notions géométriques et topologiques bien connues. Il est des domaines de la physique où le langage formel s'appuie sur des bases beaucoup moins robustes voire inexistantes. Pensons simplement à la notion d'espace de Fock, autour duquel s'organise le formalisme de la seconde quantification de la théorie des champs, qui n'est pas encore défini correctement de manière mathématique. Plus parlant encore est la méthode des intégrales de chemin ou intégrales fonctionnelles, employées à nouveau dans le cadre de la seconde quantification. Les méthodes de calcul développées dans ce langage s'obtiennent par simple analogie avec le calcul intégral classique sans aucune justification. La quantification des théories de jauge, que nous n'avons pas abordée directement ici, est par contre un domaine à part entière de la physique mathématique, ces deux mots juxtaposés évoquant des interactions nombreuses et rigoureuses.

Comme annoncé dans l'introduction, ce travail se voulait avant tout un aperçu de plusieurs situations où ces interactions géométriques et topologiques peuvent être mises en évidence. Ainsi, nous n'avons par exemple pas abordé le problème de la quantification de la théorie de Chern-Simons, nous limitant à une description de la théorie classique et à son origine topologique. Une étude plus poussée de la théorie quantique nous amènerait à discuter par exemple de la renormalisation de cette théorie de jauge, que nous n'avons fait qu'effleurer dans les derniers paragraphes du chapitre précédent, en examinant par exemple le comportement infra-rouge et ultra-violet, la nature des contre-termes, etc. Il serait également intéresssant de développer plus avant le formalisme induit par le terme de Chern-Simons dans le cadre de la description théorique de l'effet Hall quantique. Le lecteur souhaitant obtenir plus d'informations sur la théorie de C.S. peut consulter l'article de V.Dunne ([12]), dans lequel l'auteur dresse un tableau plus complet et plus précis. En ce qui concerne la partie mathématique, nous n'avons, là non plus, pas étudié de manière générale le problème épineux de la quantification d'une théorie de jauge. L'angle θ de la théorie de Yang-Mills, n'en constitue qu'un aspect particulier. La vraie question est : quelle traduction en termes géométriques peut-on donner à la fixation d'une jauge indispensable à la quantification ? L'opérateur BRS que nous avons très succinctement discuté à la fin du chapitre 4 est un début de réponse. Pour le reste, on peut consulter par exemple les ouvrages de M.Henneaux de l'université de Mons. L'étude des anomalies, dans lequel nous avons rapidement introduit la théorie des index pourrait également être poursuivie, par exemple par la lecture des chapitres consacrés aux théorèmes d'index dans ([3]) ou à la description algébrique des anomalies et de la théorie des lagrangiens effectifs dans ([1]).

Ce travail a été écrit dans l'esprit d'une synthèse, d'une clarification des notions qui y ont été exposées. Si celui-ci permet ultérieurement à une personne intéressée par ces aspects encore souvent méconnus des théories de jauge une introduction dans un formalisme qui, il faut le dire, est souvent rebutant, ce but aura été doublement atteint.

Remerciement Je tiens à remercier Monsieur A.Burnel qui a accepté de patronner ce mémoire de fin d'études pour sa disponibilité et ses encouragements.

Annexe A

Espace topologique

A.1 Définition

Pour un ensemble donné E, choisissons de désigner par E_α une partie de E. Soit Y un ensemble, fini ou infini, de parties de E : $Y = \{E_\alpha\}$. E est un espace topologique si on peut trouver un ensemble Y dont les éléments vérifient les conditions suivantes :

1. $\emptyset \in Y,\ E \in Y$,
2. $\bigcup_\alpha E_\alpha \in Y$,
3. $\bigcap_\alpha E_\alpha \in Y$, avec $\{E_\alpha\}$ fini.

L'ensemble Y est appelé une topologie sur E et ses éléments sont appelés des **ouverts**. L'espace topologique ainsi construit se note (E, Y). On écrit très souvent les axiomes 1 et 2 ci-dessus en français sous la forme : "*Toute union d'ouverts est ouverte*" et "*Toute intersection finie d'ouverts est ouverte*".

A.2 Exemples

- *Topologie triviale* $Y = \{\emptyset, E\}$.
- *Topologie "naturelle"* Si E est un espace métrique (E est muni d'une distance d), on peut introduire la notion de *boule ouverte* : $\{y \in E \mid d(x,y) < \varepsilon\}$, étant donné un point x de E et un réel positif ε. Cela étant, une partie P de E est ouverte si tout point de P est le centre d'une boule ouverte incluse dans P. On peut vérifier que l'ensemble des parties ouvertes de E satisfait aux trois axiomes, ce qui entraîne que les ouverts définis à l'aide de la notion de distance sont bien des ouverts au sens de la topologie.

Il existe encore d'autres topologies, plus exotiques, que l'on peut attacher à E, mais nous ne retiendrons que la topologie naturelle, sur laquelle repose une grande partie de l'analyse fonctionnelle, comme nous allons le constater immédiatement.

A.3 Continuité

Un des grands intérêts de la topologie est d'introduire la notion de continuité d'une application à l'aide des ouverts.

$f\ :\ (E, X) \to (F, Y)$ *est* **continu** *sur E si, pour tout ouvert Ω de F, $f^{-1}(\Omega)$ est ouvert.*

Il peut paraître curieux de baser la définition sur l'application inverse. Il suffit de considérer une simple application discontinue sur un intervalle de la droite pour se convaincre que les problèmes liés à la discontinuité interviennent au retour de l'application et non pas à l'aller.

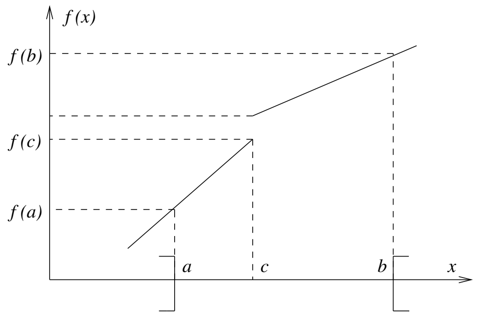


FIG. A.1 –

On remarque que f, ici une fonction réelle, manifestement discontinue, envoie l'intervalle ouvert $]a, b[$ sur la réunion de deux intervalles ouverts que nous pourrions écrire $]f(a), f(c)[$ et $]f(c+\varepsilon), f(b)[$.

Une autre définition, basée sur la notion de voisinage est plus parlante :

L'application $f : (E, X) \to (F, Y)$ est continue en $P \in E$ si pour tout voisinage de $f(P)$ dans F, il existe un voisinage de P telle que son image soit entièrement incluse dans $f(P)$.

L'application de la figure ci-dessus est discontinue car il est aisé de trouver un intervalle ouvert I contenant $f(c)$ tel qu'il n'existe aucun intervalle ouvert contenant c dont l'image est entièrement incluse dans $I\cdot$

A.4 Homéomorphisme

Cette notion, fondamentale en topologie, permet de comparer des espaces topologiques.

(E, X) et (F, Y) sont **homéomorphes** *s'il existe une application $f : (E, X) \to (F, Y)$ telle que*

- *f soit bijective,*
- *f et f^{-1} soient continues.*

On dit que les espaces E et F sont topologiquement équivalents, ou encore qu'il est impossible de les distinguer par des moyens topologiques. D'une manière plus intuitive, on imagine que l'on peut passer d'un espace à l'autre par "déformation continue" [1]. Le prototype de deux espaces non équivalents est donné par la figure ci-dessous, que l'on retrouve dans tous les livres de topologie, lorsqu'on y aborde la notion d'homéomorphisme.

1. Il convient cependant d'être vigilant avec cette image intuitive : nous verrons que l'on peut rétracter continûment une droite sur un de ces points. Il est évident que ces deux objets ne sont pas en correspondance biunivoque

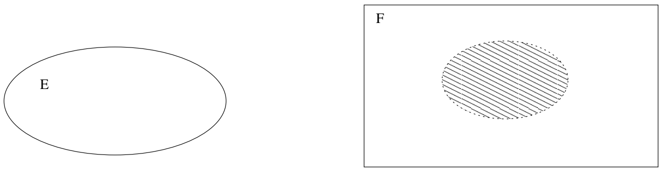


FIG. A.2 –

Du point de vue géométrique, le but de la topologie est d'étudier des classes d'ensembles de points topologiquements équivalents et de décrire leurs représentants à l'aide de propriétés qui restent inchangées par homéomorphisme (*invariants topologiques*). Un des grands avantages à répartir les espaces en classe d'équivalence est que l'on peut choisir de représenter les éléments d'une classe par celui possédant la géométrie la plus simple. On préférera dans l'exemple suivant la figure de droite à celle de gauche :

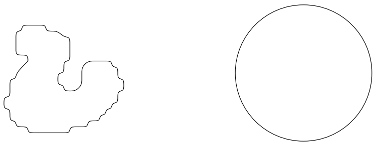


FIG. A.3 –

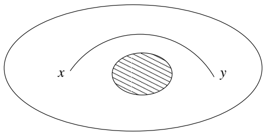
Exemples

- Tout intervalle ouvert de $\mathbb{R}$ est homéomorphe à $\mathbb{R}$. Par exemple, soit
$$E =]-\frac{\pi}{2}, \frac{\pi}{2}[, \qquad F = \mathbb{R}.$$
On a bien sûr $f : E \to F : f(x) = \tan(x)$, continue et d'inverse continue.
- Dans le même ordre d'idée, on pourrait montrer que tout boule ouverte de $\mathbb{R}^2$ est homéomorphe à $\mathbb{R}^2$.
- On peut montrer que la dimension d'un espace est un invariant topologique.

Nous donnerons dans la suite des exemples d'espaces non équivalents.

A.5 Connexité - connexité par arcs

Un espace (E, X) est **connexe** *s'il n'existe pas de partition de E en deux parties non vides de E. Il est* **connexe par arcs** *si pour tous $x, y \in E$, il existe un arc $C : [0,1] \to (E, X)$ tel que $C(0) = x$ et $C(1) = y$.*



- La connexité par arcs entraîne la connexité, la réciproque est fausse.
- La connexité et la connexité par arcs sont conservées par continuité.
- Les deux formes de connexité sont des invariants topologiques.

A.6 Compacité

Nous introduirons cette notion par celle de recouvrement. Une famille $F = \{F_\alpha\}$ d'ensembles est un recouvrement d'un ensemble E si $\bigcup_\alpha F_\alpha \supset E$.

Un espace topologique (E, X) est **compact** *si de tout recouvrement ouvert de E, on peut extraire un recouvrement fini.*

La compacité est une notion plus subtile que l'idée de finitude que sa définition évoque. Ainsi, l'intervalle ouvert $]0,1[$ bien que borné n'est pas compact. En effet, considérons le recouvrement ouvert suivant : $\{]\frac{1}{n}, 1-\frac{1}{n}[,\ n \in \mathbb{N},\ n > 2\}$. Il est impossible d'en extraire un recouvrement fini. Rappelons le cours d'analyse de première candidature : une partie de $\mathbb{R}^n$ est compacte ssi elle est bornée et fermée.

- Toute image continue d'un compact est compacte.
- La compacité est un invariant topologique.

Nous pouvons à présent donner des exemples d'espaces non équivalents.

- $[0,1]$ et $]0,1[$ ne sont pas topologiquement équivalents car le premier est compact et l'autre non. Un intervalle compact n'est donc pas équivalent à $\mathbb{R}$.
- Une parabole n'est pas homéomorphe à une hyperbole car la première est connexe et la seconde non.
- Le cercle S^1 n'est pas homéomorphe à $\mathbb{R}$ car l'un est compact et l'autre ne l'est pas.
- Anticipant sur la notion d'espace simplement connexe, le cercle S^1 n'est pas équivalent à un intervalle compact, car celui-ci est simplement connexe ("rétractable" en un point) et le cercle non.

Annexe B

Notions élémentaires de géométrie différentielle

B.1 Variétés différentiables

Soit M un espace topologique. Une **carte** pour M est la donnée de U, ouvert de M, et d'un homéomorphisme de U sur un ouvert Ω de $\mathbb{R}^n$:

$$\alpha : U \to \Omega$$

Il est ainsi possible d'identifier un point x de M à un point de $\mathbb{R}^n$, en d'autres termes ce point x possède les coordonnées :

$$x^1 = \alpha^1(x), \ \ldots \ , \ x^n = \alpha^n(x).$$

Soient à présent deux cartes (U_1, α_1) et (U_2, α_2) pour M telles que $U_1 \cap U_2 \neq \emptyset$. Tout point x de l'intersection possède deux ensembles de coordonnées :

$$x^i = \alpha_1^i(x), \quad y^i = \alpha_2^i(x).$$

La relation entre les deux est donnée par

$$y^i(x) = \alpha_2^i(\alpha_1^{-1}(x)),$$

c-à-d un homéomorphisme entre deux ouverts de $\mathbb{R}^n$:

$$\alpha_2 \circ \alpha_1^{-1} : \alpha_1(U_1 \cap U_2) \to \alpha_2(U_1 \cap U_2).$$

Une famille de cartes $\{(U_r, \alpha_r)\}$ telles que $\cup_r U_r = M$ est un **atlas** de M. L'espace M équipé d'un atlas est une **variété**. Si en outre les cartes de l'atlas sont telles que les homéomorphismes $\alpha_s \circ \alpha_r^{-1}$ sont C^∞, alors M est une **variété différentiable**.

B.2 Exemple

Traitons une variété que nous retrouverons à plusieurs reprises dans la suite :

$$S^1 = \{(x, y) \in \mathbb{R}^2 \text{ t.q. } x^2 + y^2 = 1\}.$$

Nous définissons α_1^{-1} et α_2^{-1} par

$$\alpha_1^{-1} :]0, 2\pi[\to S^1, \qquad \theta \mapsto (\cos(\theta), \sin(\theta)),$$

$$\alpha_2^{-1} :]-\pi, \pi[\to S^1, \quad \theta \mapsto (\cos(\theta), \sin(\theta)).$$

On vérifie aisément que ces deux applications sont bijectives et continues. En tant qu'homéomorphismes, elles réalisent, avec leur ouvert correspondant un atlas, et S^1 est une variété. Bien sûr, elles sont également C^∞, donc notre cercle est une variété différentiable.

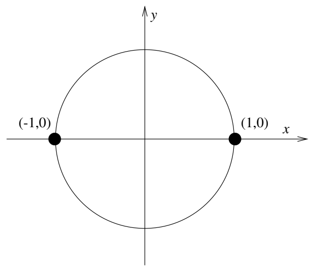


FIG. B.1 – S^1, *variété différentiable*

B.3 Variété Riemannienne

Dans la définition de la variété, nous avons considéré un espace topologique M, sans préciser de topologie particulière. Si nous voulons à présent utiliser la topologie naturelle, nous devons envisager M comme un espace métrique. Un raisonnement très naïf pourrait nous amener à écrire pour l'exemple précédent $d(x,y) = |x-y|$ pour deux points x et y de S^1, c-à-d utiliser la métrique de l'espace $\mathbb{R}^2$, dans lequel la variété est immergée. Or nous voulons bien entendu définir la distance en “restant” sur la variété. Ceci est réalisé grâce à la notion de métrique riemannienne.

Une variété différentiable est **riemannienne** *si on peut lui associer en tout point une forme bilinéaire et symétrique g que l'on appelle* **tenseur métrique**.

La distance entre deux points x et $x + dx$ dans M est donnée par

$$ds^2 = g_{ij} dx^i dx^j \quad (i, j = 1, \ldots, \dim(M)).$$

Exemples:
- $\mathbb{R}^n$ est une variété riemannienne pour $g_{ij} = \delta_{ij}$.
- L'espace de Minkowski pour $g_{\mu\nu} = [\mathrm{diag}(1, -1, -1, -1)]_{\mu\nu}$. C'est un exemple de métrique non définie positive.

Pour des exemple plus évolués, voir bien sûr le cours de relativité.

B.3.1 Applications entre variétés

Dans ce paragraphe, M et N sont des variétés différentiables.

Une fonction $f : M \to \mathbb{R}$ est **différentiable en** *$x \in M$ si, parmi les cartes de l'atlas, il en existe une (U, α), $x \in U$, telle que $f \circ \alpha^{-1} : \alpha(U) \to \mathbb{R}$ soit différentiable au sens des fonctions ordinaires en $\alpha(x)$.*

Une application $F : M \to N$ est **différentiable en** *$x \in M$ s'il existe une carte (U, α) pour M, $x \in U$, et une carte (V, β) pour N, $F(x) \in V$ telles que $\beta \circ F \circ \alpha^{-1}$ soit différentiable en $\alpha(x)$.*
F est un **difféomorphisme** *s'il est bijectif, différentiable et d'inverse différentiable. M et N*

sont donc dans ce cas homéomorphes et $\dim(M) = \dim(N)$.
Les difféomorphismes de M sur lui-même sont des automorphismes et forment un groupe [1] noté $\mathrm{diff}(M)$.

B.3.2 Produit cartésien

Soient M et N deux variétés différentiables, de dimensions respectives n et p, dont les atlas sont respectivement$\{(U_r, \alpha_r)\}$ et $\{(V_s, \beta_s)\}$, respectivement. On définit des cartes pour $M \times N$ par

$$\alpha_r \times \beta_s : U_r \times V_s \to \mathbb{R}^{n+p}, \quad (x, y) \mapsto (\alpha_r(x),\, \beta_s(y)).$$

L'atlas est alors donné par $\{(U_r \times V_s,\, \alpha_r \times \beta_s)\}$. On peut montrer que, par ces constructions, le produit cartésien de variétés différentiables est une variété différentiable. Ainsi,

$$T^n = S^1 \times \ldots S^1,$$

c-à-d le tore de dimension n est différentiable.

B.4 Vecteurs - Espace tangent

Considérons une variété différentiable M de dimension n et soit $\{(U, \alpha)\}$ une carte pour M. Soit

$$f : U \to \mathbb{R}$$

une fonction différentiable au point $x \in U$. On définit les opérateurs $\frac{\partial}{\partial x^i}\big|_x$ agissant sur f par

$$\left.\frac{\partial}{\partial x^i}\right|_x f = \left.\frac{\partial}{\partial x^i}(f \circ \alpha^{-1})\right|_{\alpha(x)} = \frac{\partial}{\partial x^i} f(x^1, \cdots, x^n).$$

L'espace vectoriel réel sous-tendu par

$$\left.\frac{\partial}{\partial x^1}\right|_x, \ldots, \left.\frac{\partial}{\partial x^n}\right|_x \tag{B.1}$$

est appelé *espace tangent à la variété* M *au point* x et est noté $T_x M$. Ainsi, tout élément de $T_x M$, appelé *vecteur* au point x s'écrit

$$v = \sum_{i=1}^{n} v^i \left.\frac{\partial}{\partial x^i}\right|_x, \tag{B.2}$$

où v_i, nombres réels, sont les composantes de v dans la base (B.1). Un vecteur tangent au point x est donc un opérateur de dérivation qui agit sur des fonctions différentiables en x. On peut retrouver l'intuition familière du vecteur en considérant quelques exemples.

- Les vecteurs fondamentaux définis en (B.1) sont les vecteurs de base habituels car on peut écrire

$$e_j(x) = \sum_{i=1}^{n} \delta_{ij} \left.\frac{\partial}{\partial x^i}\right|_x .$$

1. Pour la propriété de fermeture, il suffit d'écrire $\beta \circ f \circ g \circ \alpha^{-1} = \beta \circ f \circ \alpha^{-1} \circ \alpha \circ \beta^{-1} \circ \beta \circ g \circ \alpha^{-1}$ et d'utiliser le fait que M est une variété différentiable

– Dans le cas d'une courbe $x(t)$ dans U de paramètre t, l'opérateur de dérivée

$$\left.\frac{d}{dt}\right|_{t_0} = \sum_{i=1}^{n} \frac{dx^i}{dt}(t_0) \left.\frac{\partial}{\partial x^i}\right|_{x(t_0)}$$

est un vecteur tangent à la courbe au point $x(t_0)$, ses composantes sont $\frac{dx^i}{dt}(t_0)$.

Si, dans la définition (B.2) on laisse varier le point x dans U, on obtient une application

$$v : U \to \bigcup_{x\in U} T_xM, \quad x \mapsto v(x) = \sum_{i=1}^{n} v^i(x) \left.\frac{\partial}{\partial x^i}\right|_{x},$$

appelé *champ vectoriel* v sur U. Un champ vectoriel sur M est la donnée de champs vectoriels v_r sur U_r tels que $v_r = v_s$ sur $U_r \cap U_s$, où $\{(U_r, \alpha_r)\}$ est un atlas pour M.

B.4.1 Crochet de Lie

Il est possible de définir un commutateur, ou crochet de Lie de deux champs vectoriels v et w sur M. C'est un champ vectoriel donné par

$$[v,w] = \sum_{i,j} \left(v^i \frac{\partial}{\partial x^i} w^j - w^i \frac{\partial}{\partial x^i} v^j \right) \frac{\partial}{\partial x^j}.$$

Lorsqu'il agit sur une fonction différentiable f, ce champ vectoriel s'écrit

$$[v,w](f) = v(w(f)) - w(v(f)).$$

On peut vérifier qu'il satisfait aux propriétés de bilinéarité, d'antisymétrie, ainsi qu'à l'identité de Jacobi, ce qui fait de l'ensemble des champs vectoriels sur M une algèbre de Lie. Il s'agit en fait de l'algèbre de Lie du groupe des difféomorphismes évoqué ci-dessus.

B.4.2 Application tangente - "Push forward"

A une application

$$F : M \to N,$$

entre deux variétés M et N, on associe une application du type

$$F_* : T_xM \to T_{F(x)}N,$$

entre les espaces tangents correspondants. Celle-ci est définie de la manière suivante. Pour

$$v = \sum_{i=1}^{n} v^i \left.\frac{\partial}{\partial x^i}\right|_{x} \in T_xM,$$

on définit

$$F_*(v) = \sum_{j=1}^{p}\sum_{i=1}^{n} v^i \frac{\partial F^j}{\partial x^i} \left.\frac{\partial}{\partial y^j}\right|_{F(x)},$$

où $y^j = F^j(x^1, \ldots, x^n)$ avec $j = 1, \ldots, p$.

B.5 Formes différentielles

Une $p-$forme en un point x d'une variété M est une application

$$\phi_x : \overbrace{T_xM \times \ldots \times T_xM}^{p} \to V, \quad (v_1, \ldots, v_p) \mapsto b$$

où V est un espace vectoriel, linéaire en ses p arguments, et telle que

$$\phi(\ldots, v_i, \ldots, v_j, \ldots) = -\phi(\ldots, v_j, \ldots, v_i, \ldots).$$

L'ensemble des p-formes au point x est noté $\Lambda^p(T_xM, V)$.
Une $p-$forme sur M est une application

$$\varphi : M \to \Lambda^p(T_xM, V), \quad x \mapsto \phi_x.$$

L'espace des $p-$formes sur M est noté $\Lambda^p(M, V)$.

B.5.1 Produit extérieur

Le produit extérieur de deux formes sur M est l'application

$$\wedge : \Lambda^p(M, V) \times \Lambda^q(M, V) \to \Lambda^{p+q}(M, V)$$

définie par

$$(\varphi \wedge \psi)(v_1, \ldots, v_{p+q}) = \frac{1}{p!q!} \sum_{\text{perm}} \text{sign}(\pi) \phi(v_{\pi(1)}, \ldots, v_{\pi(p)}) \psi(v_{\pi(p+1)}, \ldots, v_{\pi(p+q)}).$$

Propriétés

1. Bilinéarité
2. Associativité
3. $\varphi \wedge \psi = (-1)^{pq} \psi \wedge \phi, \quad \varphi \in \Lambda^p(M, V), \; \psi \in \Lambda^q(M, V)$. Ainsi, $\varphi \wedge \varphi = 0$, si φ est de degré impair.

Grâce au produit extérieur, il est possible de décomposer une $p-$forme sur une base de p 1-formes

$$\beta^{i_1} \wedge \ldots \wedge \beta^{i_p}, \quad 1 \leq i_1 < \cdots < i_p \leq n.$$

On a

$$\varphi = \frac{1}{p!} \sum_{i_1, \cdots, i_p} \varphi_{i_1 \cdots i_p} \beta^{i_1} \wedge \ldots \wedge \beta^{i_p},$$

où la matrice $\varphi_{i_1 \cdots i_p}$ des composantes est antisymétrique. La dimension de l'espace des $p-$ formes sur une variété de dimension n est

$$\text{dim}\Lambda^p(M, V) = C_n^p.$$

Les 1-formes de base "naturelles" sont définies à l'aide des champs vectoriels de base (B.1). La 1-forme dx^i est définie par

$$dx^i \left(\frac{\partial}{\partial x^j} \right) = \delta_j^i.$$

B.5.2 Retour - "Pull back" d'une forme différentielle

Considérons à nouveau une application entre deux variétés M et N

$$F : M \to N.$$

On lui associe une application du type

$$F^* : \Lambda^p(N, V) \to \Lambda^p(M, V)$$

entre les espaces de formes qui leur sont associés. Pour $\varphi \in \Lambda^p(N, V)$, celle-ci est définie pour tout $x \in M$ et tous $v_i \in T_x M$ par

$$(F^*\varphi)_x(v_1, \ldots, v_p) = \varphi_{F(x)}(F_*(v_1), \ldots, F_*(v_p)).$$

Propriétés

1. Linéarité
2. $F^*(\varphi \wedge \psi) = (F^*\varphi) \wedge (F^*\psi)$

B.5.3 Dérivée extérieure

La dérivée extérieure d'une $p-$forme sur M est l'application

$$d : \Lambda^p(M, V) \to \Lambda^{p+1}(M, V),$$

définie par

$$d\varphi = \frac{1}{p!} \sum_{i_1, \cdots, i_p} \frac{\partial}{\partial x^i} \varphi_{i_1 \cdots i_p} dx^i \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_p}.$$

Propriétés

1. Linéarité
2. Règle de Leibniz : $d(\varphi \wedge \psi) = (d\varphi) \wedge \psi + (-1)^p \varphi \wedge (d\psi), \quad \varphi \in \Lambda^p(M, V)$.
3. $d^2 = 0$
4. Pour l'application F définie ci-dessus $F^* d\varphi = d(F^*\varphi)$
5. Relation avec le crochet de Lie :

$$d\varphi(u, v) = u(\varphi(v)) - v(\varphi(u)) - \varphi([u, v]) \tag{B.3}$$

Exemples

1. Pour une 0-forme, c-à-d une fonction f, on a

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} dx^i.$$

   Il s'agit de la 1-forme gradient.
2. Pour une 2-forme φ sur $\mathbb{R}^3$ $\varphi = \frac{1}{2} \sum_{ijk=1}^{3} \varepsilon_{ijk} \varphi_i dx^j \wedge dx^k$, on a

$$d\varphi = \sum_{i=1}^{3} \frac{\partial \varphi_i}{\partial x^i} dx^1 \wedge dx^2 \wedge dx^3.$$

   Si les x_i sont les coordonnées cartésiennes, on retrouve la divergence du champ vectoriel dont les φ_i sont les composantes.
3. Pour une 1-forme $\varphi = \sum_{i=1}^{3} \varphi_i dx^i$, sur $\mathbb{R}^3$, on montre de même que la dérivée extérieure est le rotationnel du champ vectoriel associé à φ si elle exprimée dans une base cartésienne.

B.6 Lemme de Poincaré - Cohomologie de de Rham

Le lemme de Poincaré est la généralisation dans le langage des formes du théorème de primitivation de fonctions de plusieurs variables.
Une forme $\varphi \in \Lambda^p(M,V)$ est *fermée* si $d\varphi = 0$. Elle est *exacte* s'il existe $\psi \in \Lambda^{p-1}(M,V)$ telle que $\varphi = d\psi$.
Le lemme de Poincaré fait appel à la notion de variété contractile, abordée dans le chapitre 2.

Sur une variété contractile, une forme différentielle est fermée si et seulement si elle est exacte.

L'ensemble des p-formes fermées sur M est noté $Z^p(M)$, et $B^p(M)$ est l'ensemble des $p-$formes exactes sur M. On a bien sûr $B^p(M) \subset Z^p(M)$. Le groupe de cohomologie de de Rham est défini par

$$H^p(M) = Z^p(M)/B^p(M).$$

Ce groupe est susceptible de nous donner des informations sur la topologie de la variété M. Ainsi par exemple, pour une variété contractile, $H^p(M) = e$, le groupe trivial, vu le lemme de Poincaré.

B.7 Intégration sur les variétés

Soit M une variété de dimension n, de frontière ∂M, $\{(U_r, \alpha_r)\}$ un atlas pour M et une n-forme $\varphi = f dx^1 \wedge dx^n$. L'intégrale de φ sur un domaine K contenu dans voisinage U_r de M est

$$\int_K \varphi = \int_{\alpha_r(K)} (\alpha_r^{-1})^* \varphi = \int_{\alpha_r(K)} f(x) d^n x,$$

soit l'intégrale classique d'une fonction sur un ouvert de $\mathbb{R}^n$. Si K n'est pas entièrement contenu dans U_r, il faut utiliser une partition de l'unité h_r. On a alors

$$\int_K \varphi = \sum_r \int_{K \cap U_r} h_r \varphi.$$

Il faut également veiller à définir correctement l'orientation des variétés (normales extérieures, sens de parcours d'une courbe ...) Si F est une application entre deux variétés M et N, de même dimension n, on a, pour $K \subset M$

$$\int_K F^* \varphi = \int_{F(K)} \varphi.$$

Enfin, on a l'important et très beau théorème de Stokes

$$\int_M d\varphi = \int_{\partial M} \varphi,$$

qui reprend sous une forme condensée tous les théorèmes célèbres d'intégration : la formule de Stokes, bien sûr, tous les théorèmes de Green, dont le fameux théorème d'Ostrogradski, et dans le cas de la variété $\mathbb{R}$, le calcul de l'intégrale par variation de la primitive.

B.8 Etoile de Hodge

L'étoile de Hodge est l'application

$$* : \Lambda^p(M,V) \to \Lambda^{n-p}(M,V)$$

définie pour une $p-$ forme $\varphi = \frac{1}{p!}\sum_{i_1,\dots,i_p}\varphi_{i_1\dots i_p}dx^{i_1}\wedge\cdots\wedge dx^{i_p}$ sur une variété riemannienne M par

$$^*\varphi = \frac{1}{(n-p)!}\sum_{i_{p+1},\dots,i_n}\left[\frac{1}{p!}\sum_{i_1,\cdots,i_p}\varepsilon_{i_1\dots i_n}\sqrt{|\det g|}\sum_{j_1,\cdots,j_p}\varphi_{j_1\dots j_p}g^{i_1j_1}\dots g^{i_pj_p}\right]dx^{i_{p+1}}\wedge\cdots\wedge dx^{i_n}.$$

Insistons sur le fait que cette application nécessite une métrique pour la variété M. Le cas le plus connu d'application de l'étoile de Hodge est celui du vecteur polaire que l'on associe à trois dimensions à un tenseur antisymétrique : le vecteur de Poisson, par exemple, ou le champ magnétique.

L'étoile de Hodge permet de définir un nouvel opérateur de dérivation, l'opérateur de codérivée

$$\delta : \Lambda^p(M) \to \Lambda^{p-1}, \quad \varphi \mapsto (-1)^{np+n+1+s}\ {}^*d^*\varphi,$$

où n est la dimension de M, et la métrique est de signature $r-s$. Nous introduisons cet opérateur un peu particulier car celui-ci permet d'introduire une généralisation du laplacien aux formes :

$$\Delta : \Lambda^p(M) \to \Lambda^p, \quad \Delta = -(d\delta + \delta d).$$

Δ se réduit au laplacien dans le cas de la métrique euclidienne, et au d'alembertien $\Box$ dans le cas de la métrique de Minkowski.

Terminons par une propriété de l'étoile de Hodge vis-à-vis de l'intégration. Avec des notations claires, on a

$$\int \varphi \wedge\ {}^*\psi = \int \psi \wedge\ {}^*\varphi.$$

B.9 Variété parallélisable

Un référentiel d'une variété M de dimension n est un ensemble de n champs vectoriels défini sur un sous-ensemble de M, linéairement indépendants en chaque point de leur ensemble de définition. Par exemple, un système de coordonnées (U,α) de coordonnées $x^1,\dots,x^n$ définit un référentiel $\frac{\partial}{\partial x^1},\dots,\frac{\partial}{\partial x^n}$ sur U. Mais il n'y a en général pas de référentiel défini sur toute la variété M. Si tel est le cas, on dit que M est *parallélisable.* Les variétés parallélisables sont l'exception plutôt que la règle. Par exemple, S^n est parallélisable pour n=1,3 et 7.

Bibliographie

[1] M.Göckeler and T.Schücker. *Differential Geometry, gauge theories and gravity*, Cambridge University Press, 1987.

[2] C.Nash and S.Sen. *Topology and geometry for physicists*, Academic Press, 1983.

[3] M.Nakahara. *Geometry, topology and physics.* Adam Hilger, 1990.

[4] A.A. Belavin et al. "Pseudoparticle solutions of the Yang-Mills Equations", *Phys. Lett.*, **59 B**(1) (1975) 85.

[5] A.Trautman and M.E.Mayer. "A brief introduction to the geometry of gauge fields" *in* "Lectures given at the XX. Internationale Universitätswochen für Kernphysik, Schladming, Austria", (1981) 433. Springer-Verlag, 1981.

[6] M.E.Peskin and D.V.Schroeder. *An Introduction to Quantum Field Theory*, Addison-Wesley Publishing Company, 1995.

[7] K.K. Uhlenbeck. "Removable Singularities in Yang-Mills Fields", *Bull. of the Am. Math. Soc.*, **1**(3) (1979) 579.

[8] S.S. Chern and J.Simons. "Characteristic Forms and Geometric Invariants", *Ann. Math.*, **99** (1974) 48.

[9] S.Deser. "Chern-Simons Terms as an Example of the Relations Between Mathematics and Physics" *preprint* math-ph/9805020.

[10] S.Deser, R.Jackiw and S.Templeton. "Topologically Massive Gauge Theories", *Ann. Phys.*, **140** (1982) 372.

[11] A.Lerda. *Anyons, Quantum Mechanics of Particles with Fractional Statistics*, Springer-Verlag, Lecture Notes in Physics, m 14, 1992.

[12] G.V. Dunne. "Aspects of Chern-Simons Theory", Lectures at the 1998 Les Houches Summer School, *preprint* hep-th/9902115.

[13] P.Sharan, A.Mehra, K.Dasgupta and M.Sami. "Gauge conditions for an Abelian Chern-Simons system consistent with equations of motion", *preprint* hep-th/9903098.

[14] R.Jackiw and C.Rebbi. "Vacuum Periodicity in a Yang-Mills Quantum Theory", *Phys. Rev. Lett.*, **37**(3) (1976) 172.

[15] S.Deser, R.Jackiw and S.Templeton. "Three-Dimensional Massive Gauge Theories", *Phys. Rev. Lett.*, **48**(15) (1982) 975.

[16] A.Heil, A.Kersch, B.Reifenhäuser and H.Vogel. "Application of topological ideas to the Aharonov-Bohm effect", *Eur. J. Phys*, **9** (1988) 200.

[17] R.Jackiw. "Introduction to the Yang-Mills quantum theory", *Rev. of Mod. Phys.*, **52**(4) (1980) 661.

[18] R.Jackiw. "Gauge theories in three dimensions" *in* "Gauge Theories of the Eighties", Proceedings, Äkäslompolo, Finland 1982, Springer Verlag, Lectures Notes in Physics 181, 1983.

[19] K.Fujikawa, *Phys. Rev. Lett.*, **42** (1979) 1195; *Phys. Rev.*, **D21** (1980) 2848.

Table des matières