



HAL
open science

Théorie asymptotique des représentations et applications à la théorie de Yang-Mills

Thibaut Lemoine

► **To cite this version:**

Thibaut Lemoine. Théorie asymptotique des représentations et applications à la théorie de Yang-Mills. Physique mathématique [math-ph]. Sorbonne Université, 2020. Français. NNT: . tel-03096870v1

HAL Id: tel-03096870

<https://theses.hal.science/tel-03096870v1>

Submitted on 5 Jan 2021 (v1), last revised 19 Nov 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



École Doctorale de Sciences Mathématiques de Paris Centre

THÈSE DE DOCTORAT

Discipline : Mathématiques

présentée par

Thibaut LEMOINE

Théorie asymptotique des représentations et applications à la théorie de Yang–Mills

dirigée par Thierry LÉVY

Rapportée par :

M. Pierre-Loïc MÉLIOT Université Paris-Saclay
M. James NORRIS Cambridge University

Soutenue le 29/10/2020 devant le jury composé de :

M. Philippe BIANE	Université Paris-Est	examineur
M. Cédric BOUTILLIER	Sorbonne Université	examineur
Mme Manon DEFOSSEUX	Université Paris Descartes	examinatrice
M. Cédric LECOUCVEY	Université de Tours	examineur
M. Thierry LÉVY	Sorbonne Université	directeur
Mme Mylène MAIDA	Université de Lille	examinatrice
M. Pierre-Loïc MÉLIOT	Université Paris-Saclay	rapporteur

REMERCIEMENTS

Mes premiers remerciements vont à Thierry, pour m'avoir fait confiance, conseillé et soutenu pendant ces quatre années de thèse. D'une part son exigence et sa rigueur m'auront été très utiles lors de la rédaction de mon premier article, mais aussi du présent manuscrit ; d'autre part ses nombreuses connaissances issues de domaines très variés auront été pour moi une source d'émerveillement et d'inspiration.

Je remercie chaleureusement Pierre-Loïc Méliot et James Norris pour avoir accepté d'être les rapporteurs de cette thèse. Je les remercie tout aussi bien pour leurs remarques positives que pour leurs commentaires constructifs et pertinents.

Merci à Philippe Biane, Cédric Boutillier, Manon Defosseux, Cédric Lecouvey et Mylène Maida de faire partie du jury ; c'est un honneur d'être évalué par des chercheurs et des chercheuses dont je respecte énormément les travaux.

Je ne remercierai sans doute jamais assez Damien Simon, qui m'a accompagné et encouragé depuis le début de mon master et qui m'a ensuite permis d'encadrer les TP de $\mathbb{C} + +$ pendant trois ans. Merci également de m'avoir fait découvrir le monde fascinant des probabilités non-commutatives et de la théorie des représentations ainsi que leurs interactions.

J'adresse un grand merci à Antoine Dahlqvist, dont les travaux avec James Norris sur le champ maître sur la sphère ont marqué un tournant dans ma compréhension de la théorie de Yang–Mills, et à plus forte raison, dans l'avancée de ma thèse. Je le remercie également pour la longue collaboration qu'il a entreprise avec moi depuis plus de deux ans, et qui commence à porter ses fruits.

Ma thèse a été marquée par une démarche de recherche relativement solitaire, en témoigne mon absence remarquée dans le bureau partagé du 201 ; je remercie toutefois tous ceux que j'ai pu y côtoyer, même ceux avec qui je n'ai pas eu le plaisir de faire davantage connaissance. Je partage par ailleurs ces remerciements avec tous les doctorants et chercheurs du laboratoire que j'ai croisés durant ces quatre années, et en particulier les organisateurs des différents groupes de travail auxquels j'ai pu participer et qui ont pu stimuler ma curiosité : je pense par exemple à Clément et Simon pour le groupe "combinatorics and random matrix theory", ou bien Carlo et Nicolas pour le groupe "noncommutative probability and rough paths".

D'autres remerciements s'imposent évidemment, envers ceux qui m'ont parfois taquiné, parfois supporté, souvent soutenu : mon entourage pour le moins éclectique. Bien qu'être exhaustif me sem-

ble pour le moins ambitieux, je pense notamment au personnel du Centre de Secours de Magny-en-Vexin que je côtoie depuis 5 ans, mes amis de la NS Team que j'ai vus s'éparpiller un peu partout en France (et en Europe), la Love Machine avec ses métaliflettes et ses week-ends YOLO, et ceux qui ne rentrent pas forcément dans des groupes mais qui comptent tout autant pour moi : Simon, Pierre, Paul, Antonin, Jonathan, Thomas.

Merci enfin à ma famille d'Île-de-France mais aussi celle d'Auvergne, tout particulièrement mes parents ; et puisqu'il convient de garder le(la) meilleur(e) pour la fin je remercie Alex pour être entrée (et restée !) dans ma vie, car sans elle je n'aurais jamais eu l'énergie d'entreprendre tout ce que j'ai accompli. Et ce n'est, je l'espère, que le début d'une belle aventure.

RÉSUMÉ

Cette thèse est consacrée à l'étude de la mesure de Yang–Mills euclidienne sur une surface compacte, avec pour groupe de structure le groupe unitaire $U(N)$ ou spécial unitaire $SU(N)$. Cette étude porte plus précisément sur le comportement asymptotique de cette mesure lorsque N tend vers l'infini, à l'aide des représentations asymptotiques du groupe unitaire.

Le premier chapitre, qui fait office d'introduction au sujet, explique en détails la construction de cette mesure en développant au préalable les diverses théories sur lesquelles elles s'appuie : la théorie de jauge, les probabilités non-commutatives et les représentations de groupes.

Nous montrons dans le chapitre 2, en nous reposant sur l'article [Lem19], que la fonction de partition de cette mesure de Yang–Mills converge, pour les surfaces compactes orientables de genre supérieur ou égal à 1 ou non orientables de genre supérieur ou égal à 2, vers une limite finie qui ne dépend que du genre, de l'orientabilité et de l'aire de la surface sous-jacente.

Dans le chapitre 3, nous construisons partiellement l'objet appelé *champ maître* sur les surfaces compactes orientables de genre 1 et plus, qui constitue la limite – au sens des probabilités non-commutatives – du champ aléatoire sur la surface sous-jacente dont la loi est donnée par la mesure de Yang–Mills.

Principaux résultats de cette thèse.

- Le théorème 1.4.10, page 81, concerne la convergence des fonctions de partition pour les surfaces compactes orientables de genre supérieur ou égal à 1 ;
- Le théorème 1.4.11, page 82, concerne la convergence des fonctions de partition pour les surfaces compactes non orientables de genre supérieur ou égal à 2 ;
- Le théorème 1.4.12, page 85, concerne la convergence en espérance et en variance des boucles de Wilson simples pour un lacet simple contractile sur une surface orientable de genre supérieur ou égal à 1 ;
- Le théorème 1.4.13, page 85, concerne la loi des holonomies le long de lacets d'homologie non nulle, et la concentration des boucles de Wilson pour des lacets simples non séparants, tous ces lacets étant sur une surface orientable de genre supérieur ou égal à 1.

ABSTRACT

This thesis is devoted to the study of Yang–Mills measure on a compact surface, with structure group the unitary group $U(N)$ or special unitary group $SU(N)$. This study is more precisely about this measure’s asymptotic behaviour in the large N limit, using asymptotic representations of the unitary group.

The first chapter, as an introduction to the subject, explains in details the construction of Yang–Mills measure after having developed the several theories it is based on: gauge theory, noncommutative probability and group representations.

We show in Chapter 2, based on the article [Lem19], that the partition function of this Yang–Mills measure converges, for orientable surfaces of genus greater or equal to 1 and non-orientable surfaces of genus greater or equal to 2, to a finite limit that only depends on the genus, the orientability and the area of the underlying surface.

In Chapter 3 we partially construct the so-called *master field* on orientable compact surfaces of genus greater or equal to 1, which is the limit – in a non-commutative probabilistic sense – to the random field on the underlying surface whose distribution is given by Yang–Mills measure.

Main results of this thesis.

- Theorem 1.4.10, page 81, concerns the convergence of the partition function for compact orientable surfaces of genus greater or equal to 1;
- Theorem 1.4.11, page 82, concerns the convergence of the partition function for compact non-orientable surfaces of genus greater or equal to 2;
- Theorem 1.4.12, page 85, concerns the convergence of Wilson loop expectation and variance for a contractible simple loop on an orientable surface of genus greater or equal to 1;
- Theorem 1.4.13, page 85, concerns the distribution of holonomies along loops with nonzero homology, and a concentration result about simple nonseparating loops, all these loops being on an orientable surface with genus $g \geq 1$.

INTRODUCTION (FR)

Les théories de jauge sont des théorie des champs sur des variétés qui jouent le rôle de l'espace-temps, et telles que les champs associés soient invariants par l'action d'un certain groupe de transformation, appelé le groupe de structure. La théorie de Yang–Mills euclidienne en deux dimensions est une version simplifiée de la théorie de jauge utilisée dans le Modèle Standard de la physique des particules ; elle peut être vue comme un modèle-jouet où tout peut être défini convenablement, tandis que la théorie de Yang–Mills quantique à quatre dimensions est loin d'être bien comprise d'un point de vue mathématique. En effet, le formalisme des intégrales de chemins utilisé en physique théorique peut être rendu rigoureux lorsque l'espace-temps est à deux dimensions. Cette théorie est devenu un champ d'application des probabilités depuis que Migdal [Mig75] a décrit la mesure de Yang–Mills sur le plan euclidien \mathbb{R}^2 à l'aide du noyau de la chaleur sur le groupe de structure. Après cette découverte, il est devenu possible de considérer la théorie de Yang–Mills à travers le prisme des matrices aléatoires, ce qui est le point de vue que nous adoptons dans cette thèse.

Le premier chapitre peut être considéré comme un cours sur la théorie de Yang–Mills en deux dimensions, et est écrit en français. Il couvre les prérequis en géométrie différentielle, en théorie des représentations et en probabilités non-commutatives nécessaires à la construction du processus d'holonomie de Yang–Mills, que nous esquissons ci-après. D'après les travaux de T. Lévy [Lévo3, Lévio], on considère une surface compacte connexe sans bord Σ , un groupe compact G , et un graphe orienté $\mathbb{G} = (\mathbb{V}, \mathbb{E}, \mathbb{E}^+, \mathbb{F})$ plongé dans Σ tel que toutes ses faces soient homéomorphes à des disques. Le *champ d'holonomie de Yang–Mills* est un processus stochastique $(H_\ell)_{\ell \in \mathcal{P}(\mathbb{G})}$ à valeurs dans G et indexé par l'ensemble des lacets sur \mathbb{G} obtenus par concaténation d'arêtes et de leurs inverses. La loi de ce processus peut être décrite en utilisant l'espace de configuration

$$\mathcal{C}_{\mathbb{G}}^G = G^{\mathbb{E}^+}.$$

Pour toute face F du graphe, on note $|F|$ son aire et ∂F son bord. On note également $(p_t)_{t \geq 0}$ le noyau de la chaleur sur G , et h_ℓ l'*application holonomie* définie pour un lacet $\ell = e_1^{\varepsilon_1} \cdots e_n^{\varepsilon_n}$ par

$$h_\ell : \begin{cases} \mathcal{C}_{\mathbb{G}}^G & \rightarrow G \\ g & \mapsto g_{e_n}^{\varepsilon_n} \cdots g_{e_1}^{\varepsilon_1} \end{cases}.$$

La loi du champ de Yang–Mills le long de ℓ , qui peut en quelque sorte être identifiée à la mesure de Yang–Mills¹, est alors donnée par la formule suivante :

$$\mathbb{E}[f(H_\ell)] = \frac{1}{Z_G(g, T)} \int_{G^{\mathbb{E}^+}} f(h_\ell) \prod_{F \in \mathbb{F}} p_{|F|}(h_{\partial F}(g)) dg, \quad (1)$$

¹Il s'agit plus précisément de la mesure image de la mesure de Yang–Mills par l'application holonomie.

où $Z_G(g, T)$ est une constante de normalisation, appelée *fonction de partition*, qui dépend du groupe de structure G , du genre g de la surface et de son aire totale T . En utilisant une sorte de théorème d'extension de Kolmogorov, il peut être montré que ce processus est en réalité une marginale finidimensionnelle d'un processus plus général $(H_\ell)_{\ell \in \mathcal{L}(\Sigma)}$ indexé par un ensemble $\mathcal{L}(\Sigma)$ de lacets sur la surface sous-jacente, mais nous allons rarement considérer le processus complet car tous les calculs reposent sur la formule (1) qui dépend du graphe considéré.

On considère le cas où le groupe de structure est le groupe unitaire $U(N)$ pour deux raisons principales :

- on s'attend à ce que le processus se comporte de manière agréable lorsque N tend vers l'infini, depuis l'article fondateur de 't Hooft [tH74] ;
- cela permet de décrire la loi du champ d'holonomie de Yang–Mills en termes de diagrammes de Young et de fonctions de Schur, qui ont des propriétés combinatoires particulièrement bonnes.

Plus concrètement, la décomposition de Fourier du noyau de la chaleur sur $U(N)$ repose sur les représentations irréductibles du groupe, qui sont caractérisées par des N -uplets décroissants d'entiers relatifs $\lambda = (\lambda_1 \geq \dots \geq \lambda_N)$ appelés *plus hauts poids*. On leur associe trois quantités :

(i) Le *caractère* de la représentation

$$\chi_\lambda(U) = \chi_\lambda(x_1, \dots, x_N), \quad \forall U \sim \text{diag}(x_1, \dots, x_N) \in U(N),$$

(ii) La *dimension* de la représentation, donnée par $d_\lambda = \chi_\lambda(1, \dots, 1)$,

(iii) Le *nombre de Casimir* de la représentation, donné par $c_2(\lambda)$ tel que $\Delta \chi_\lambda = -c_2(\lambda) \chi_\lambda$.

On a alors

$$p_t(U) = \sum_{\substack{\lambda_1 \geq \dots \geq \lambda_N \\ (\lambda_1, \dots, \lambda_N) \in \mathbb{Z}^N}} e^{-c_2(\lambda) \frac{t}{2}} d_\lambda \chi_\lambda(U), \quad \forall T > 0, \quad \forall U \in U(N).$$

Grâce à cette décomposition, (1) peut être réécrite de façon plus appropriée. Dans les deuxième et troisième chapitres, qui sont écrits en anglais car il s'agit de travaux de recherche destinés à être publiés, on utilise cette décomposition pour calculer les limites de différentes quantités reliées au processus d'holonomie de Yang–Mills.

Dans le chapitre 2, tiré de [Lem19], on donne une formule de la fonction de partition $Z_N(g, T) = Z_{U(N)}(g, T)$ pour une surface compacte orientable sans bord de genre $g \geq 1$ et d'aire totale T ; une étude asymptotique des représentations irréductibles de $U(N)$ nous mène à la limite suivante.

Théorème (Limites orientables, Thm. 2.1.2). *Soit Σ une surface orientable de genre g .*

(i) *Si $g \geq 2$, alors pour tout $T \in (0, +\infty)$, on a la convergence suivante :*

$$\lim_{N \rightarrow \infty} Z_N(g, T) = \sum_{n \in \mathbb{Z}} e^{-\frac{T}{2} n^2}. \quad (2)$$

(ii) Si $g = 1$, alors on considère $T \in (0, +\infty)$ et on pose $q = e^{-\frac{T}{2}}$. La convergence suivante est vérifiée :

$$\lim_{N \rightarrow \infty} Z_N(1, T) = \sum_{n \in \mathbb{Z}} e^{-\frac{T}{2}n^2} \prod_{m=1}^{\infty} (1 - q^m)^{-2}. \quad (3)$$

Un fait intéressant est que les limites impliquent deux formes modulaires connues pour avoir des liens avec la théorie des nombres : la fonction thêta de Jacobi et la fonction d'Euler. On établit aussi un résultat similaire pour des surface non orientables définies comme la somme connexe de g plans projectifs, avec $g \geq 2$: ce résultat est donné au théorème 2.1.3. Les preuves de ces théorèmes reposent principalement sur l'étude de plus hauts poids que l'on appellera *presque plats*, pour lesquels le nombre de Casimir possède des propriétés analytiques et combinatoires intéressantes.

Une autre raison pour laquelle il est intéressant de comprendre la limite du champ de Yang–Mills avec pour groupe de structure $U(N)$ lorsque $N \rightarrow \infty$ est l'existence supposée d'un *champ maître*. Il s'agit d'un champ déterministe défini comme la limite (au sens des probabilités non-commutatives) du champ d'holonomie de Yang–Mills, et son existence a déjà été prouvée sur le plan par Lévy [Lévi17] et sur la sphère par Dahlqvist et Norris [DN17]. Les deux preuves sont fondées sur deux outils complémentaires : l'espérance et la variance de boucles de Wilson pour des lacets simples (des lacets sans points d'auto-intersection), et les équations de Makeenko–Migdal.

L'espérance de la boucle de Wilson associée à un lacet ℓ est tout simplement $\mathbb{E}[\text{tr}(H_\ell)]$, et si sa limite existe, on s'attend à ce qu'elle corresponde à la valeur du champ maître pour le lacet ℓ . La variance associée est $\text{Var}[\text{tr}(H_\ell)]$, et le fait qu'elle tende vers 0 quand N tend vers l'infini garantit que le champ maître est déterministe. Un moyen de prouver la convergence de l'espérance et la variance de la boucle de Wilson pour un lacet simple est d'utiliser l'analyse harmonique sur $U(N)$, et c'est par exemple comme cela que Dahlqvist et Norris ont traité le cas de la sphère. Les équations de Makeenko–Migdal sont des équations différentielles satisfaites par l'espérance de la boucle de Wilson d'un lacet qui possède (au moins) un point d'auto-intersection. Grâce à ces équations, on peut déduire récursivement le calcul d'une espérance de boucle de Wilson pour un lacet à n points d'auto-intersection à partir d'espérances de boucles de Wilson pour des lacets à $n - 1$ points d'auto-intersection. Par induction, on peut alors construire (et calculer) le champ maître pour n'importe quel lacet avec un nombre fini de points d'auto-intersection.

Le but du chapitre 3 est de généraliser aux surfaces compactes orientables de genre 1 et plus les résultats obtenus dans le plan et la sphère en termes d'espérances et de variances de boucles de Wilson. Outre le fait que la mesure de Yang–Mills a une expression plus compliquée pour de telles surfaces, le principal obstacle est que les lacets simples peuvent être divisés en deux catégories :

- Les *lacs simples séparants*, qui ont une homologie nulle ;
- Les *lacs simples non séparants*, qui ont une homologie non nulle, et peuvent être identifiés à des lacets qui font le tour d'une anse, en considérant qu'une surface compacte orientable de genre g peut être continûment déformée en un tore à g anses.

On considère tout d'abord un lacet ℓ simple contractile. Il est en particulier séparant : il divise la surface en deux composantes connexes, dont l'une est homéomorphe à un disque. Un tel lacet peut donc être

complété en un graphe de sorte que l'on puisse appliquer (1) pour calculer l'espérance et la variance de sa boucle de Wilson. On montre dans les Thm. 3.2.5 et 3.2.6 que

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_\ell)] = e^{-\frac{t}{2}} \quad \text{and} \quad \lim_{N \rightarrow \infty} \text{Var}[\text{tr}(H_\ell)] = 0.$$

En particulier, l'espérance de la boucle de Wilson de ℓ possède la même limite que si ℓ était un lacet simple dans le plan, ce qui est assez inattendu. En fait, cela ne semble même pas avoir été conjecturé par les physiciens dans la littérature, comme c'est souvent le cas. Dans la Prop. 3.2.13, on donne également la vitesse de convergence vers zéro de la variance de la boucle de Wilson. Les preuves sont une fois de plus fondées sur la théorie des plus hauts poids presque plats, et en particulier sur la Prop. 3.2.10, qui est une généralisation que l'on a faite d'un résultat de Gross et Taylor [GT93]. On explique après cela comment ces résultats peuvent être appliqués à un théorème de Hall [Hal18] afin d'obtenir la limite de l'espérance et la variance de boucles de Wilson pour des lacets possédant éventuellement des points d'auto-intersection, qui sont contenus dans un petit disque topologique. Ensuite, on traite le cas d'un lacet simple non séparant : on montre dans le Thm. 3.2.16 que l'holonomie le long d'un tel lacet² est un unitaire de Haar, et dans la Prop. 3.2.17 que la variance de sa boucle de Wilson converge vers 0 quand N tend vers l'infini. Dans la fin du chapitre on explique ce qu'il reste à prouver afin d'obtenir une construction complète du champ maître.

²On prouve même cela pour tout lacet d'homologie non nulle, même s'il n'est pas simple.

INTRODUCTION (EN)

Gauge theories are field theories on manifolds, which play the role of spacetimes, and such that the associated fields are invariant with respect to a transformation group, named the structure group. The two-dimensional Euclidean Yang–Mills theory is a simplification of the gauge theory used in the Standard Model of particle physics; it can be viewed as a toy model where everything can be defined conveniently, whereas the four-dimensional quantum Yang–Mills theory is far from being well understood from a mathematical point of view. Indeed, the formalism of path integrals used in theoretical physics can be made rigorous when the underlying spacetime is two-dimensional. It has become a field of interest in probability theory since Migdal [Mig75] described the Yang–Mills measure on the Euclidean plane \mathbb{R}^2 in terms of the heat kernel of the structure group. After this discovery, it has been possible to consider the Yang–Mills theory through the prism of random matrix theory, which is the point of view we will adopt in this thesis.

The first chapter can be considered as a course on two-dimensional Yang–Mills theory, and is written in French. It covers the prerequisites in differential geometry, representation theory and noncommutative probability for the construction of Yang–Mills holonomy process, that we sketch hereafter. Following the work of T. Lévy [Lévo3, Lévio], we consider a compact connected closed surface Σ , a compact group G , and an oriented graph $\mathbb{G} = (\mathbb{V}, \mathbb{E}, \mathbb{E}^+, \mathbb{F})$ embedded in Σ such that all faces of \mathbb{G} are homeomorphic to disks. The *Yang–Mills holonomy field* is a G -valued stochastic process $(H_\ell)_{\ell \in \mathcal{P}(\mathbb{G})}$ indexed by the set of paths in \mathbb{G} obtained by concatenation of edges and their inverses. The distribution of this process can be described using the configuration space

$$\mathcal{C}_{\mathbb{G}}^G = G^{\mathbb{E}^+}.$$

For any face F of the graph, we denote by $|F|$ its area and ∂F its boundary. Let us also write $(p_t)_{t \geq 0}$ the heat kernel on G , and h_ℓ the *holonomy function* defined for a loop $\ell = e_1^{\varepsilon_1} \cdots e_n^{\varepsilon_n}$ as

$$h_\ell : \begin{cases} \mathcal{C}_{\mathbb{G}}^G & \rightarrow G \\ g & \mapsto g_{e_n^{\varepsilon_n}} \cdots g_{e_1^{\varepsilon_1}} \end{cases}.$$

The distribution of Yang–Mills field on ℓ , which can somehow be identified with Yang–Mills measure³, is then given by the following formula:

$$\mathbb{E}[f(H_\ell)] = \frac{1}{Z_G(g, T)} \int_{G^{\mathbb{E}^+}} f(h_\ell) \prod_{F \in \mathbb{F}} p_{|F|}(h_{\partial F}(g)) dg, \quad (4)$$

³It is more precisely the push-forward of Yang–Mills measure by the holonomy map.

where $Z_G(g, T)$ is a normalisation constant, called *partition function*, which depends on the structure group G , the genus g of the surface and its total area T . Using a kind of Kolmogorov extension theorem, it can be proved that this process is the finite-dimensional marginal of a more general process $(H_\ell)_{\ell \in \mathcal{L}(\Sigma)}$ indexed by some set of loops $\mathcal{L}(\Sigma)$ on the underlying surface, but we will rarely consider the whole process because all computations involve the formula (4) which depends on the graph we consider.

We consider the case when the structure group is the unitary group $U(N)$, for two main reasons:

- it is expected to behave nicely when N tends to infinity, since the seminal work of 't Hooft [tH74];
- it permits to describe the distribution of the Yang–Mills holonomy field in terms of Young diagrams and Schur functions, which have particularly good combinatorial properties.

In fact, the Fourier decomposition of the heat kernel on $U(N)$ relies on irreducible representations, which are labelled by nonincreasing N -tuples of integers $\lambda = (\lambda_1 \geq \dots \geq \lambda_N)$ called *highest weights*. We can associate to them three quantities:

- (i) The *character* of the representation

$$\chi_\lambda(U) = \chi_\lambda(x_1, \dots, x_N), \quad \forall U \sim \text{diag}(x_1, \dots, x_N) \in U(N),$$

- (ii) The *dimension* of the representation, which is $d_\lambda = \chi_\lambda(1, \dots, 1)$,

- (iii) The *Casimir number* of the representation, which is the number $c_2(\lambda)$ such that $\Delta \chi_\lambda = -c_2(\lambda) \chi_\lambda$.

We have then

$$p_t(U) = \sum_{\substack{\lambda_1 \geq \dots \geq \lambda_N \\ (\lambda_1, \dots, \lambda_N) \in \mathbb{Z}^N}} e^{-c_2(\lambda) \frac{t}{2}} d_\lambda \chi_\lambda(U), \quad \forall T > 0, \quad \forall U \in U(N).$$

Thanks to this decomposition, (4) can be rewritten in a more convenient way. In the second and third chapters, which are written in English because they are research works destined to be published, we use this decomposition to compute limits of different quantities related to the Yang–Mills holonomy process.

In Chapter 2, which is based on [Lem19], we give the formula of the partition function $Z_N(g, T) = Z_{U(N)}(g, T)$ for a closed compact orientable surface of genus $g \geq 1$ and total area T ; an asymptotic study of the irreducible representations of $U(N)$ yields the following limit.

Theorem (Orientable limits, Thm. 2.1.2). *Let Σ be an orientable surface of genus g .*

- (i) *If $g \geq 2$, then, for all $T \in (0, +\infty)$, the following convergences hold:*

$$\lim_{N \rightarrow \infty} Z_N(g, T) = \sum_{n \in \mathbb{Z}} e^{-\frac{T}{2} n^2}. \quad (5)$$

(ii) If $g = 1$, then consider $T \in (0, +\infty)$ and set $q = e^{-\frac{T}{2}}$. The following convergence holds:

$$\lim_{N \rightarrow \infty} Z_N(1, T) = \sum_{n \in \mathbb{Z}} e^{-\frac{T}{2}n^2} \prod_{m=1}^{\infty} (1 - q^m)^{-2}. \quad (6)$$

Interestingly, the limits involve two modular forms that are known to have links with number theory: the Jacobi theta function and the Euler function. We also state a similar result for nonorientable surfaces defined as the connected sum of g projective planes, with $g \geq 2$: it is given in Thm. 2.1.3. The proofs of these theorems mainly rely on the study of highest weights that we will call *almost flat*, for which the Casimir number has nice analytical and combinatorial properties.

Another reason why one would be interested in understanding the limit of Yang–Mills field with structure group $U(N)$ when $N \rightarrow \infty$ is the supposed existence of a *master field*. This is a deterministic field defined as the limit of the Yang–Mills holonomy field (as a noncommutative stochastic process), and it was already proven to exist in the plane by Lévy [Lév17] and in the sphere by Dahlqvist and Norris [DN17]. Both proofs are based on two complementary tools: the Wilson loop expectation and variance for simple loops (loops without self-intersections), and the Makeenko–Migdal equations.

The Wilson loop expectation associated to a given loop ℓ is simply $\mathbb{E}[\text{tr}(H_\ell)]$, and if its limit exists, it is expected to be the value of the master field for the loop ℓ . The associated Wilson loop variance is $\text{Var}[\text{tr}(H_\ell)]$, and the fact that it vanishes in the large N limit guarantees that the master field is deterministic. A way to prove the convergence of Wilson loop expectation and variance for a simple loop is to use harmonic analysis on $U(N)$, and that is for instance how Dahlqvist and Norris treated the case of the sphere. The Makeenko–Migdal equations are differential equations satisfied by the Wilson loop expectation of a loop with a self-intersection point. Thanks to these equations, the recursive computation of a Wilson loop expectation for a loop with n crossings can be deduced from the knowledge of Wilson loop expectation for loops with $n - 1$ crossings. By induction, one can then construct (and compute) the master field for any loop with a finite number of self-intersections.

The purpose of Chapter 3 is to generalize to compact orientable surfaces of genus 1 and higher the results obtained in the plane and the sphere in terms of Wilson loop expectations and variances. Beside the fact that the Yang–Mills measure has a more complicated expression for such surfaces, the main obstacle is that simple loops can be divided into two categories:

- The *separating simple loops*, which have zero homology,
- The *nonseparating simple loops*, which have nonzero homology and can be identified with loops that go around a handle, considering that a compact orientable surface of genus g can be continuously deformed into a torus with g handles.

We first consider a contractible simple loop ℓ . It is in particular separating: it splits the surface into two connected components, and one of these components is homeomorphic to a disk. We denote by t the area of this disk. Such a loop can then be completed into a graph and we can apply (4) to compute its Wilson loop expectation and variance. We show in Thm. 3.2.5 and 3.2.6 that

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_\ell)] = e^{-\frac{t}{2}} \quad \text{and} \quad \lim_{N \rightarrow \infty} \text{Var}[\text{tr}(H_\ell)] = 0.$$

In particular, the Wilson loop expectation has the same limit as if ℓ was a simple loop in the plane, which is quite unexpected. In fact, it does not even seem to have been conjectured by physicists in the literature, as it often is the case. In Prop. 3.2.13 we also give the rate of convergence to zero of the Wilson loop variance. The proofs are again based on the theory of almost flat highest weights, and in particular Prop. 3.2.10, which is generalization we did of a result by Gross and Taylor [GT93]. We explain afterwards how our results can be applied to a theorem by Hall [Hal18] in order to get the limit of Wilson loop expectation and variance for a loop with self-intersections that is contained in a small topological disk. Then, we treat the case of a nonseparating simple loop: we prove in Thm. 3.2.16 that the holonomy around such a loop⁴ is a Haar unitary, and in Prop. 3.2.17 that its Wilson loop variance converges to 0 in the large N limit. In the end of the chapter we explain what there remains to prove in order to get a complete construction of the master field.

⁴We even prove it for any loop with nonzero homology, even if it is not simple.

TABLE DES MATIÈRES

1	La théorie de Yang–Mills en deux dimensions	19
1.1	Prolégomènes géométriques	20
1.1.1	Variétés différentielles et espaces tangents	20
1.1.2	Fibrés vectoriels, champs de vecteurs et formes différentielles	24
1.1.3	Fibré principal et connexion	27
1.1.4	Courbure et action de Yang–Mills	32
1.1.5	Transformations de jauge	34
1.2	Probabilités non-commutatives	36
1.2.1	Espaces de probabilité non-commutatifs	37
1.2.2	Distributions non-commutatives	38
1.2.3	Indépendance et liberté	41
1.2.4	Convergence dans des espaces non-commutatifs	43
1.2.5	Mouvement brownien unitaire et limite d'échelle	45
1.3	Théorie des représentations	48
1.3.1	De la mesure de Haar sur un groupe	48
1.3.2	Définitions et constructions générales	50
1.3.3	Représentations des groupes compacts	55
1.3.4	Diagrammes de Young et fonctions de Schur	59
1.3.5	Représentations du groupe unitaire	63
1.4	La mesure de Yang–Mills sur une surface compacte	68
1.4.1	Mesure de Yang–Mills sur un graphe	69
1.4.2	Lois marginales du champ de Yang–Mills	74
1.4.3	Autour de la fonction de partition	77
1.4.4	Le champ maître sur une surface compacte	82
2	Yang–Mills measure and large N partition function	87
2.1	The Yang–Mills partition function on a compact surface	87
2.2	Algebraic tools	91
2.2.1	Witten zeta function	91
2.2.2	Almost flat highest weights	93
2.3	Large N limit of partition function on orientable surfaces	98
2.3.1	Genus $g \geq 2$	98
2.3.2	Genus $g = 1$	100
2.4	Large N limit of partition function on non-orientable surfaces	103

2.4.1	Genus $g \geq 3$	105
2.4.2	Genus $g = 2$	108
3	The master field on compact orientable surfaces	III
3.1	The master field, from Wilson to Makeenko–Migdal	III
3.1.1	Wilson loop expectations for simple loops	113
3.1.2	Makeenko–Migdal equations	117
3.2	Compact surfaces of higher genus	120
3.2.1	Contractible simple loops	120
3.2.2	Loops with self-intersections	147
3.2.3	Loops with nontrivial homology	149
3.2.4	Joint distribution for several loops	152
3.2.5	Remaining loops: the ‘missing link’	153

LA THÉORIE DE YANG–MILLS EN DEUX DIMENSIONS

Cette thèse porte sur une étude asymptotique de la mesure de Yang–Mills sur des surfaces compactes à l'aide de la théorie des représentations. Il s'agit, en d'autres termes, de l'utilisation d'outils *algébriques* pour étudier un objet *probabiliste* défini dans un cadre *géométrique*. C'est donc un sujet qui puise dans des théories a priori de natures très diverses, et qui pourtant sont intimement liées :

- La théorie de jauge en deux dimensions,
- Les probabilités non-commutatives,
- La théorie des représentations.

L'objectif de ce chapitre est de permettre à un lecteur non-initié de comprendre le contenu des chapitres suivants, qui constituent le véritable travail de recherche. Nous allons par conséquent rappeler – sans démonstration – les bases des théories mentionnées ci-dessus, ainsi que certains résultats nécessaires aux approfondissements développés dans les chapitres suivants. Ces rappels seront volontairement détaillés afin de rendre la thèse accessible à des lecteurs issus d'un parcours probabiliste traditionnel, et donc peu familiers de la géométrie différentielle, de la théorie des représentations, voire des probabilités libres qui demeurent une branche relativement récente des probabilités malgré leur popularité grandissante. Nous tenterons de notre mieux de mettre en exergue les intrications et les synergies entre ces branches, comme par exemple l'utilisation des représentations comme un analogue de la théorie de Fourier en probabilités non-commutatives, ou encore la réécriture d'intégrales de chemin sur le plan comme des moments de mouvements browniens sur des groupes.

La fin de ce chapitre s'articulera autour de la mesure de Yang–Mills sur des surfaces compactes, définie comme champ d'holonomie markovien en se fondant sur la construction de T. Lévy [Lévo3, Lévo] et reposant sur les théories rappelées préalablement. Nous expliquerons cette construction dans le cas des surfaces compactes connexes sans bord, avec pour groupe de structure le groupe unitaire $U(N)$ ou spécial unitaire $SU(N)$, puis développerons certains aspects de cette mesure en lien avec des mouvements browniens sur les groupes, l'analyse harmonique non-commutative, pour enfin évoquer les aspects asymptotiques que nous développerons dans les chapitres suivants. L'un de ces aspects, traité dans [Lem19], est la limite de sa fonction de partition ; l'autre, étudié conjointement avec Antoine Dahlqvist et qui va faire l'objet d'un futur article, est la limite d'une grande classe

d'observables liées à cette mesure, et qui constitue une avancée significative vers la construction du champ maître sur des surfaces compactes.

I.I PROLÉGOMÈNES GÉOMÉTRIQUES

La théorie de Yang–Mills, du nom de C. N. Yang et R. Mills qui sont les premiers à l'avoir conceptualisée dans l'article fondateur [YM54], est une théorie de jauge non-abélienne visant à unifier trois des quatre forces fondamentales (en l'occurrence les forces électromagnétique, faible et forte). En réalité, il serait judicieux de parler de théories de Yang–Mills au pluriel, car il s'agit de diverses théories décrivant des phénomènes différents, mais en utilisant le même cadre ; seuls diffèrent alors l'espace-temps et le groupe de structure.

Yang et Mills n'étaient toutefois pas les premiers à introduire la notion de jauge : dès 1865, Maxwell a décrit l'électrodynamique classique dans [Max65] en utilisant notamment un champ magnétique invariant par certaines transformations du potentiel agissant sur celui-ci, et ce furent alors les prémices des théories de jauge modernes. L'idée selon laquelle les interactions peuvent être décrites par des champs invariants par l'action de certains groupes de symétrie a accompagné le développement de toute la physique moderne jusqu'à ses modèles les plus récents. Dans le modèle standard de la physique des particules, on suppose que l'espace-temps est une variété pseudo-riemannienne de dimension 4 (trois dimensions spatiales et une dimension temporelle), et le groupe de structure dépend de l'interaction considérée. Par exemple, l'électromagnétisme peut être décrit à l'aide du groupe $U(1)$, l'interaction faible via le groupe $SU(2)$ ou encore l'interaction forte avec $SU(3)$. Plus globalement, la formulation actuelle des théories de jauge est la suivante : l'espace des configurations d'un système physique soumis à une interaction donnée par un groupe de symétrie G et évoluant dans un espace-temps M est modélisé par un G -fibré principal de base M . Nous allons ici considérer le “modèle-jouet” suivant : l'espace-temps M est de dimension 2 (donc une surface), et le groupe de symétrie G est un groupe de Lie compact, typiquement $U(N)$ ou $SU(N)$.

Afin d'aborder sans heurts cette théorie plus souvent connue des physiciens que des mathématiciens, nous allons introduire pas à pas les outils nécessaires à son appréhension. Après de brefs rappels de géométrie différentielle, nous nous pencherons sur la formalisation des théories de jauge à l'aide de connexions sur des fibrés principaux ; c'est en réalité cette dernière partie (paragraphe I.I.4) qui va servir dans la suite de la thèse, et le reste peut être survolé en première lecture.

I.I.I Variétés différentielles et espaces tangents

Dans un souci de pédagogie, nous commençons par les fondements de la géométrie différentielle, à savoir la définition des variétés et de leurs espaces tangents. Il s'agit plutôt d'un court formulaire visant à fixer les notations et permettre les calculs dans les paragraphes qui suivront, mais cela ne se substitue certainement pas à un cours introductif. Pour plus de rappels/approfondissements sur le sujet, le lecteur est invité à consulter les ouvrages [KN96, Lee13, Moroi] en anglais, ou bien [Laf12] en français.

Définition I.I.I. Une *variété différentielle réelle* de classe \mathcal{C}^∞ (ou *lisse*) et de dimension $n \in \mathbb{N}^*$ est un espace topologique M muni d'une famille $(U_i, \varphi_i)_{i \in I}$ telle que :

- $(U_i)_{i \in I}$ soit un recouvrement ouvert de M ;
- $\varphi_i : U_i \rightarrow \mathbb{R}^n$ soit un homéomorphisme sur son image ;
- $\varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \rightarrow \varphi_j(U_i \cap U_j)$ soit un \mathcal{C}^∞ -difféomorphisme.

Les couples (U_i, φ_i) sont appelés *cartes*, les ouverts U_i *ouverts de cartes*, φ_i les *applications coordonnées* et $\varphi_j \circ \varphi_i^{-1}$ les *changements de cartes* ou *changements de coordonnées*. L'ensemble des cartes de M s'appelle un *atlas*.

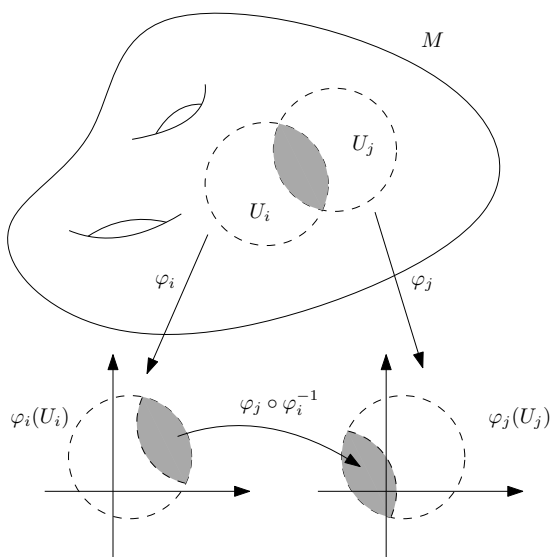


Figure 1.1: Une variété différentielle M de dimension 2, munie de deux cartes (U_i, φ_i) et (U_j, φ_j) , et de l'application de changement de coordonnées $\varphi_j \circ \varphi_i^{-1}$ dans l'intersection des ouverts de cartes U_i et U_j .

En d'autres termes, une variété est un espace topologique qui peut localement être assimilé à \mathbb{R}^n en se plaçant dans des cartes à l'aide des applications coordonnées. La Fig. 1.1 illustre le passage d'une carte à une autre dans une variété différentielle. Le calcul différentiel sur les variétés se résume donc à utiliser des coordonnées locales et utiliser les propriétés de \mathbb{R}^n .

Exemple. Soit M et N deux variétés de dimensions respectives n et m , munies respectivement d'un atlas $(U_i, \varphi_i)_{i \in I}$ et $(V_j, \psi_j)_{j \in J}$. Une fonction $f : M \rightarrow N$ est dite différentiable si pour tout $i \in I$ et $j \in J$ la fonction

$$\psi_j \circ f \circ \varphi_i^{-1} : \varphi_i(U_i) \rightarrow \psi_j(V_j)$$

est différentiable en tant que fonction de \mathbb{R}^n vers \mathbb{R}^m . En notant $\varphi_i = (x^1, \dots, x^n)^t$ et $\psi_j = (y^1, \dots, y^m)$, on peut donc écrire f à l'aide de fonctions $f^i \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$:

$$y^1(f) = f^1(x^1, \dots, x^n), \dots, y^m(f) = f^m(x^1, \dots, x^n).$$

Un des aspects importants de \mathbb{R}^n pour le calcul différentiel est sa structure d'espace vectoriel : en effet, la différentielle d'une fonction en un point est définie comme une application linéaire, ou encore

¹Il ne faut pas se laisser abuser par les notations : bien que x désigne plus traditionnellement des variables que des fonctions, x^k est pourtant une fonction, qui à un point de M associe sa k -ième coordonnée locale dans la carte (U_i, φ_i) .

les champs de vecteurs sont définis comme une famille de vecteurs indexés par l'espace sous-jacent. Ici, on ne suppose pas que la variété M possède une telle structure, et c'est pourquoi il est nécessaire d'introduire le concept d'espace tangent.

Définition 1.1.2. Soit M une variété différentielle de dimension n munie d'un atlas $(U_i, \varphi_i)_{i \in I}$ et $p \in M$ un point de cette variété. L'espace tangent à M en p est le quotient $T_p M$ du \mathbb{R} -espace vectoriel

$$\bigcup_{i: p \in U_i} \{i\} \times \mathbb{R}^n$$

par la relation d'équivalence suivante :

$$(i, v_1) \sim (j, v_2) \Leftrightarrow d(\varphi_j \circ \varphi_i^{-1})_{\varphi_i(p)} v_1 = v_2, \forall (i, j) \in I, \forall (v_1, v_2) \in (\mathbb{R}^n)^2.$$

Cette définition, bien que formellement très lourde, ne dit rien de plus que la chose suivante : l'espace tangent à M en p est un espace vectoriel assimilé à \mathbb{R}^n dans chaque carte, et pour lequel changer de carte revient à composer par la différentielle du changement de cartes dans la variété. Une façon plus commode de manipuler les vecteurs de $T_p M$ est de les voir comme des dérivations sur l'algèbre $\mathcal{F}(p)$ des fonctions de classe \mathcal{C}^1 au voisinage de p et à valeurs dans \mathbb{R} . Soit $t_0 \in]0, 1[$ un réel et $c : [0, 1] \rightarrow M$ une courbe de classe \mathcal{C}^1 sur M telle que $c(t_0) = p$. Alors le vecteur tangent à c en p est l'application linéaire

$$X : \begin{cases} \mathcal{F}(p) & \rightarrow \mathbb{R} \\ f & \mapsto X(f)_p := \left. \frac{df(c(t))}{dt} \right|_{t=t_0}. \end{cases} \quad (1.1)$$

Cette application X est bien une dérivation, au sens où l'on a

$$X(fg)_p = X(f)_p g(p) + f(p) X(g)_p, \forall (f, g) \in \mathcal{F}(p)^2. \quad (1.2)$$

Exemples.

- La sphère unité $S^n \subset \mathbb{R}^{n+1}$ est une variété différentielle, si on la munit des cartes (U_N, φ_N) et (U_S, φ_S) où :
 - $U_N = S^n \setminus \{N\}$ et $U_S = S^n \setminus \{S\}$ où $N = (0, \dots, 0, 1)$ et $S = (0, \dots, 0, -1)$ sont respectivement les pôles Nord et Sud ;
 - $\varphi_N : U_N \rightarrow \mathbb{R}^n$ est la projection stéréographique par rapport au pôle Nord, c'est-à-dire que pour tout $x \in U_N$, le point $(\varphi_N(x), 0)$ est l'unique point d'intersection de la droite (Nx) avec l'hyperplan $\mathbb{R}^n \times \{0\}$;
 - $\varphi_S : U_S \rightarrow \mathbb{R}^n$ est la projection stéréographique par rapport au pôle Sud, c'est-à-dire que pour tout $x \in U_S$, le point $(\varphi_S(x), 0)$ est l'unique point d'intersection de la droite (Sx) avec l'hyperplan $\mathbb{R}^n \times \{0\}$.

La Fig. 1.2 illustre la projection stéréographique du cercle S^1 par rapport au pôle nord.

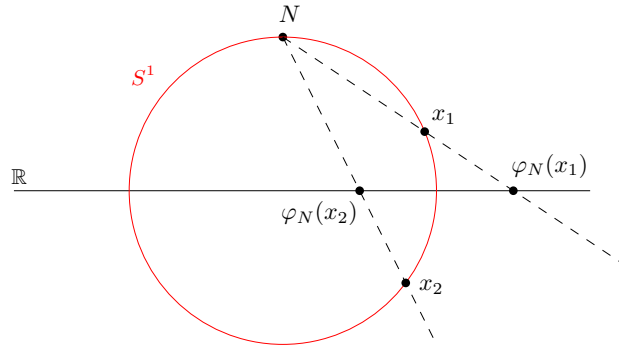


Figure 1.2: La projection stéréographique par rapport au pôle Nord dans la sphère $S^1 \subset \mathbb{R}^2$.

- Le 2-tore *algébrique* $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ est une variété différentielle quotient, et est difféomorphe à $(\mathbb{R}/\mathbb{Z})^2$ ainsi qu’au tore *géométrique*, défini comme la surface de révolution dans \mathbb{R}^3 obtenue en faisant tourner le cercle de centre $(2, 0, 0)$ de rayon 1 autour de l’axe (Oz) , et le difféomorphisme en question est induit par l’application

$$h : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R}^3 \\ (\theta, \phi) & \mapsto ((2 + \cos \theta) \cos \phi, (2 + \cos \theta) \sin \phi, \sin \theta) \end{cases} .$$

- Un *groupe de Lie* est un groupe G muni d’une structure de variété différentielle, et tel que l’application

$$\begin{cases} G \times G & \rightarrow G \\ (g, h) & \mapsto gh^{-1} \end{cases}$$

soit différentiable. Son espace tangent en l’élément neutre $\mathfrak{g} = T_e G$ est appelé *algèbre de Lie* et est muni d’une application bilinéaire antisymétrique $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$, appelée *crochet de Lie*, qui vérifie l’*identité de Jacobi*

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0. \quad (1.3)$$

Parmi les groupes de Lie les plus communs on trouve $GL_N(\mathbb{C})$, le groupe des matrices inversibles de taille $N \times N$ à coefficients complexes, ou encore $U(N)$, le sous-groupe de $GL_N(\mathbb{C})$ constitué des matrices unitaires, c’est-à-dire U telles que $UU^* = U^*U = I_N$, ou encore $SU(N)$, le sous-groupe de $U(N)$ constitué des matrices unitaires de déterminant 1. Ces groupes seront un des principaux objets d’attention tout au long de cette thèse. Leurs algèbres de Lie sont respectivement $\mathfrak{gl}_N(\mathbb{C}) = \mathcal{M}_N(\mathbb{C})$, $\mathfrak{u}_N = \{X \in \mathcal{M}_N(\mathbb{C}) : X^* = -X\}$, et $\mathfrak{su}_N = \{X \in \mathfrak{u}_N : \text{Tr}(X) = 0\}$.

Avant de plonger dans un niveau d’abstraction supplémentaire et d’introduire les fibrés, rassurons-nous un instant en montrant ce qui se passe dans \mathbb{R}^n pour les vecteurs tangents : en effet, cela permet de comprendre une bonne fois pour toutes les calculs en coordonnées locales. Il est à la fois commode et déroutant que \mathbb{R}^n soit son propre espace tangent en chacun de ses points ; pour lever cette ambiguïté nous allons considérer (e_1, \dots, e_n) la base canonique de \mathbb{R}^n , et $(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n})$ la base canonique de l’espace vectoriel des dérivations, isomorphe à \mathbb{R}^n . Soit $x = x^1 e_1 + \dots + x^n e_n \in \mathbb{R}^n$. Alors un vecteur tangent en x est donné par

$$X_x = X_1 \frac{\partial}{\partial x^1} + \dots + X_n \frac{\partial}{\partial x^n}, \quad (X_1, \dots, X_n) \in \mathbb{R}^n, \quad (1.4)$$

dans le sens où pour toute fonction $f \in \mathcal{F}(x)$

$$X_x(f) = X_1 \frac{\partial f(x)}{\partial x^1} + \cdots + X_n \frac{\partial f(x)}{\partial x^n} = (X_1, \dots, X_n) \cdot \nabla f(x) = (df)_x(X_1, \dots, X_n). \quad (1.5)$$

I.I.2 Fibrés vectoriels, champs de vecteurs et formes différentielles

Définition I.I.3. Soit $\pi : E \rightarrow M$ une application lisse entre deux variétés différentielles. Le triplet (E, π, M) est un *fibré vectoriel (réel) de rang k , d'espace total E et de base M* , si :

- Pour tout $x \in M$, la *fibre* $E_x = \pi^{-1}(\{x\})$ est un \mathbb{R} -espace vectoriel de dimension k ;
- Il existe un recouvrement de M par des ouverts $(U_i)_i$ tel que pour tout i , il existe un difféomorphisme $\varphi_i : \pi^{-1}(U_i) \rightarrow U_i \times \mathbb{R}^k$, appelé *trivialisat ion locale*, qui fait commuter le diagramme suivant :

$$\begin{array}{ccc} \pi^{-1}(U_i) & \xrightarrow{\varphi_i} & U_i \times \mathbb{R}^k \\ & \searrow \pi & \swarrow \text{pr}_1 \\ & U_i & \end{array} \quad (1.6)$$

- Pour tout $x \in M$, pour tout i tel que $x \in U_i$, l'application $\begin{cases} \mathbb{R}^k & \rightarrow \pi^{-1}(\{x\}) \\ v & \mapsto \varphi_i^{-1}(x, v) \end{cases}$ soit un isomorphisme d'espaces vectoriels.

En des termes plus élémentaires, construire un fibré vectoriel de base M et de rang k revient à placer de manière lisse, au-dessus de chaque point de M , une copie de \mathbb{R}^k , comme illustré par la Fig. 1.3. Cela se comprend dans la mesure où sur un ouvert U_i le fibré est identifié de manière canonique avec le produit cartésien $U_i \times \mathbb{R}^k$.

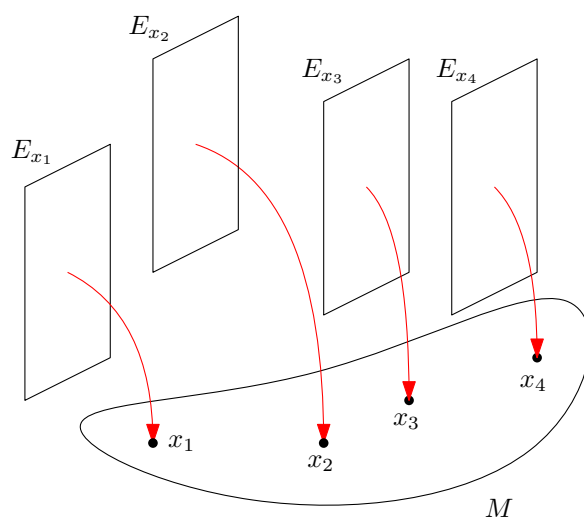


Figure 1.3: Un fibré vectoriel de rang 2 de base M : à chaque point $x \in M$ est associée une copie E_x du plan \mathbb{R}^2 .

On peut remarquer qu'il existe pour tout (i, j) tel que $U_i \cap U_j \neq \emptyset$ une application continue $\gamma_{ji} : U_i \cap U_j \rightarrow \text{GL}_k(\mathbb{R})$ qui vérifie :

$$\varphi_j \circ \varphi_i^{-1}(x, v) = (x, \gamma_{ji}(x)v), \quad \forall x \in U_i \cap U_j. \quad (1.7)$$

L'application $\varphi_j \circ \varphi_i^{-1}$ est appelée *application de transition* ou *changement de coordonnées*, et on peut vérifier que l'équation (1.7) équivaut à la commutativité du diagramme (1.6). Notons également que les applications γ_{ji} forment un *cocycle de Čech*, c'est-à-dire qu'elles vérifient

$$\gamma_{ji}(x)\gamma_{il}(x)\gamma_{lj}(x) = I_k, \quad \forall x \in U_i \cap U_j \cap U_l. \quad (1.8)$$

Exemples.

- Le fibré tautologique $(M \times \mathbb{R}^k, \text{pr}_1, M)$ est un fibré vectoriel, aussi appelé fibré trivial de rang k sur M .
- Le fibré tangent TM de la variété M est défini comme la somme disjointe des espaces tangents à M en chacun de ses points :

$$TM = \coprod_{x \in M} T_x M = \bigcup_{x \in M} \{x\} \times T_x M.$$

La projection associée est alors la projection sur la première coordonnée.

- Le fibré cotangent T^*M est le fibré dual de TM , c'est-à-dire qu'il est défini par la même projection mais que la fibre en x est l'espace vectoriel dual $(T_x M)^*$.

Définition 1.1.4. Soit (E, π, M) un fibré vectoriel et U un ouvert de M . Une *section* de π sur U est une application lisse $\sigma : U \rightarrow E$ telle que $\pi \circ \sigma = \text{id}_U$. On note $\Gamma(U, E)$ l'ensemble des sections de π sur U .

Notons que $\Gamma(U, E)$ est un $\mathcal{C}^\infty(U, \mathbb{R})$ -module. Une section est une sorte de “pseudo-inverse” de π : on associe de manière lisse, à un point $x \in M$, un élément de la fibre E_x (cf. Fig. 1.4), c'est-à-dire un vecteur. Par exemple on définit un *champ de vecteurs* sur une variété comme une section de son fibré tangent : cela consiste simplement à associer à chaque point p de la variété un élément $X = X_p$ de son espace tangent. De la même manière, une 1-forme différentielle sur une variété est définie comme une section du fibré cotangent. On note $\mathfrak{X}(M) = \Gamma(TM)$ l'ensemble des champs de vecteurs sur M et $\Omega^1(M) = \Gamma(T^*M)$ l'ensemble des 1-formes différentielles sur M .

La k -ème puissance extérieure du fibré cotangent définit un fibré $\wedge^k(T^*M)$, dont l'espace des sections $\Omega^k(M) = \Gamma(\wedge^k(T^*M))$ constitue ce que l'on appelle les *k -formes différentielles*. Si $\omega \in \Omega^k(M)$, $X^1, \dots, X^k \in \mathfrak{X}(M)$, alors on définit $\omega(X^1, \dots, X^k)$ par

$$(\omega(X^1, \dots, X^k))_x = \omega_x(X_x^1, \dots, X_x^k), \quad \forall x \in M.$$

En particulier, ω définit une application k -linéaire alternée du $\mathcal{C}^\infty(M)$ -module $\mathfrak{X}(M)$ sur $\mathcal{C}^\infty(M)$.

On peut enfin construire le fibré $\wedge^\bullet(T^*M) = \bigoplus_{k=0}^n \wedge^k(T^*M)$; l'ensemble de ses sections, noté $\Omega^\bullet(M)$, constitue l'ensemble des *formes différentielles* sur M , avec pour convention que les 0-formes sont les fonctions lisses sur M à valeurs dans \mathbb{R} . C'est une algèbre graduée qui vérifie $\Omega^\bullet(M) = \bigoplus_{k=0}^n \Omega^k(M)$.

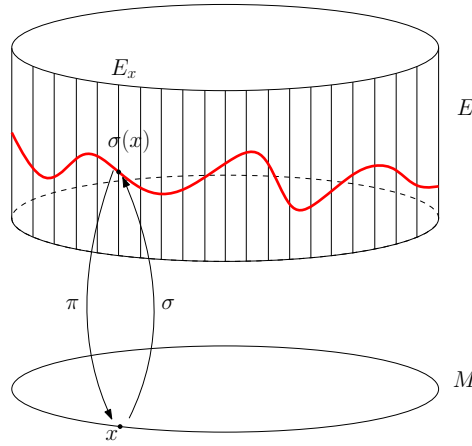


Figure 1.4: Une section (en rouge) d'un fibré en droites de base M .

Définition 1.1.5. Soit $\varphi : M \rightarrow N$ une application lisse entre deux variétés. Si $X \in \mathfrak{X}(M)$, on définit son *push-forward* $\varphi_*X \in \mathfrak{X}(\varphi(M))$ par

$$(\varphi_*X)_{\varphi(x)} = (d\varphi)_x X_x, \quad \forall x \in M$$

Si $\omega \in \Omega^k(N)$, on définit son *pull-back* $\varphi^*\omega \in \Omega^k(M)$ par

$$(\varphi^*\omega)_x(v_1, \dots, v_k) = \omega_{\varphi(x)}((d\varphi)_x v_1, \dots, (d\varphi)_x v_k), \quad \forall x \in M, \quad \forall v_1, \dots, v_k \in T_x M.$$

Définition 1.1.6. Soit M une variété compacte et $X \in \mathfrak{X}(M)$ un champ de vecteurs. Le *flot* de X est l'application

$$u^X : \begin{cases} \mathbb{R} \times M & \rightarrow M \\ (t, p) & \mapsto u_t^X(p) \end{cases}$$

solution de l'équation

$$\begin{cases} \frac{d}{dt} u_t^X(p) & = X(u_t^X)_p \\ u_0^X(p) & = p \end{cases} . \quad (1.9)$$

Le flot u^X d'un champ de vecteurs $X \in \mathfrak{X}(M)$ vérifie les propriétés suivantes :

- Pour tout $t \in \mathbb{R}$, l'application $u_t^X : M \rightarrow M$ est un \mathcal{C}^∞ -difféomorphisme d'inverse u_{-t}^X ;
- Pour tout $(s, t) \in \mathbb{R}^2$, on a la propriété de semi-groupe $u_t^X \circ u_s^X = u_{s+t}^X$.

La combinaison de ces assertions signifie que $(u_t^X)_{t \in \mathbb{R}}$ est un *groupe à un paramètre* de difféomorphismes, ou plus précisément que l'application

$$\begin{cases} \mathbb{R} & \rightarrow \text{Diff}(M) \\ t & \mapsto u_t^X \end{cases}$$

est un morphisme de groupes.

En pratique, l'équation (I.9) peut s'interpréter à l'aide de fonctions test, à l'instar des vecteurs tangents définis dans le paragraphe précédent : si $X \in \mathfrak{X}(M)$ est un champ de vecteurs et u^X son flot, alors pour tout $p \in M$ et pour toute fonction $f \in \mathcal{F}(p)$, on a

$$X(f)_p = \left. \frac{d}{dt} \right|_{t=0} f \circ u_t^X(p). \quad (\text{I.10})$$

Notons que cette formule est une simple réécriture de la règle de dérivation des fonctions composées.

Lorsque la variété est un groupe de Lie G , le flot d'un champ de vecteurs $X \in T_e G$ définit ce que l'on appelle l'*exponentielle de Lie*.

Définition I.1.7. Soit G un groupe de Lie muni de son action à gauche

$$\begin{cases} G \times G & \rightarrow G \\ (g, h) & \mapsto L_g(h) = gh \end{cases} .$$

- Un champ de vecteur $X \in \mathfrak{X}(G)$ est *invariant à gauche* si pour tout $g \in G$, $(L_g)_* X = X$.
- Une k -forme différentielle $\omega \in \Omega^k(G)$ est *invariante à gauche* si pour tout $g \in G$, $(L_g)^* \omega = \omega$.
- L'*application exponentielle* sur G est définie par

$$\exp : \begin{cases} T_e G & \rightarrow G \\ X & \mapsto u_1^{\tilde{X}}(e) \end{cases} , \quad (\text{I.11})$$

où \tilde{X} est l'unique champ de vecteurs invariant à gauche tel que $\tilde{X}_e = X$.

On note $\mathfrak{X}(G)^G$ (resp. $\Omega^k(G)^G$) l'ensemble des champs de vecteurs invariants à gauche (resp. des k -formes différentielles invariantes à gauche).

La dénomination "exponentielle" n'est pas fortuite puisqu'elle coïncide, lorsque $G = \text{GL}_N(\mathbb{C})$, avec l'exponentielle matricielle définie comme somme de la série exponentielle :

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k, \quad \forall A \in \mathcal{M}_N(\mathbb{C}) = \mathfrak{gl}_N(\mathbb{C}). \quad (\text{I.12})$$

La définition de l'exponentielle de Lie nécessite l'existence et l'unicité d'un champ de vecteurs invariant à gauche associé à un vecteur tangent en l'élément neutre de G : cela découle de la proposition suivante.

Proposition I.1.1. *On a les isomorphismes canoniques suivants entre espaces vectoriels : $\mathfrak{g} \cong \mathfrak{X}(G)^G$, et $\mathfrak{g}^* \cong \Omega^1(G)^G$.*

I.1.3 Fibré principal et connexion

Les fibrés vectoriels font partie de la famille extrêmement vaste des espaces fibrés, dont nous allons nous intéresser à un nouveau membre : les fibrés principaux. Au lieu d'être des copies d'un espace vectoriel, leurs fibres sont cette fois-ci des copies d'un groupe qui agit librement² sur l'espace total. Si G est un groupe de Lie, on considère les actions (à gauche) suivantes :

²Une action d'un groupe G sur un espace X est dite *libre* si le stabilisateur de chaque point de X est trivial, c'est-à-dire si pour tout $x \in X$ et pour tout $g \in G$, l'égalité $gx = x$ entraîne $g = e$.

– l'action de G sur lui-même par multiplication à gauche³

$$\begin{cases} G \times G & \rightarrow G \\ (g, h) & \mapsto L_g(h) = gh \end{cases} , \quad (I.I.3)$$

– L'action de G sur lui-même par multiplication de l'inverse à droite⁴

$$\begin{cases} G \times G & \rightarrow G \\ (g, h) & \mapsto R_g(h) = hg^{-1} \end{cases} , \quad (I.I.4)$$

– l'action de G sur lui-même par automorphisme intérieur

$$\begin{cases} G \times G & \rightarrow G \\ (g, h) & \mapsto \iota_g(h) = ghg^{-1} \end{cases} , \quad (I.I.5)$$

– l'action de G sur son algèbre de Lie \mathfrak{g}

$$\begin{cases} G \times \mathfrak{g} & \rightarrow \mathfrak{g} \\ (g, X) & \mapsto \text{Ad}_g(X) = (\iota_g)_*X = \frac{d}{dt}\Big|_{t=0} g \exp(tX) g^{-1} \end{cases} , \quad (I.I.6)$$

appelée *action adjointe*.

Définition I.I.8. Soit M une variété différentielle et G un groupe de Lie. Un G -*fibré principal* de base M est une variété différentielle P munie d'une action libre à droite

$$\begin{cases} P \times G & \rightarrow P \\ (p, g) & \mapsto p \cdot g \end{cases} ,$$

dont les fibres $(\pi^{-1}(x))_{x \in M}$ sont les orbites de P sous l'action de G ⁵, et d'une projection lisse $\pi : P \rightarrow M$ telle qu'il existe un recouvrement de M par des ouverts U_i et des difféomorphismes G -équivariants $\Phi_i : \pi^{-1}(U_i) \rightarrow U_i \times G$ qui font commuter le diagramme

$$\begin{array}{ccc} \pi^{-1}(U_i) & \xrightarrow{\Phi_i} & U_i \times G \\ & \searrow \pi & \swarrow \text{pr}_1 \\ & U_i & \end{array}$$

Tout comme pour les fibrés vectoriels, il existe également une définition équivalente des fibrés principaux à l'aide de cocycles de changements de coordonnées, que nous ne détaillons pas, mais qui figurent par exemple dans [Seno8a]. La notion de section définie pour un fibré vectoriel reste valable dans le cas d'un G -fibré principal, mais l'interprétation diffère. En effet, on peut considérer une section comme un moyen d'associer – toujours de manière lisse – à tout point x de M un représentant de la classe d'équivalence induite par P_x (en tant qu'orbite pour l'action de G). Contrairement aux fibrés vectoriels, les fibrés principaux voient leur trivialité caractérisée par l'existence d'une section globale.

³Nous avons déjà vu cette action au paragraphe précédent dans la Déf. I.I.7.

⁴Il est important que la multiplication se fasse par l'inverse, sinon on obtient une action à droite à la place.

⁵C'est-à-dire que pour tout $(p, q) \in P^2$ tel que $\pi(p) = \pi(q) = x$, il existe $g \in G$ tel que $q = p \cdot g$.

Proposition 1.1.2 ([Ste51], §8.3). *Soit (P, π, M) un G -fibré principal. Il est trivial si et seulement s'il existe une section globale $\sigma : M \rightarrow P$.*

L'action de G sur P définit par dérivation une application $\mathfrak{g} \rightarrow \mathfrak{X}(P)$ qui à $X \in \mathfrak{g}$ associe⁶ son *champ de vecteurs fondamental* \tilde{X} défini en $p \in P$ par

$$\tilde{X}_p = \left. \frac{d}{dt} \right|_{t=0} p \cdot \exp(tX). \quad (1.17)$$

Soit $P = M \times G$ un fibré principal trivial. On a, pour $p = (x, g) \in P$, $T_p P = T_x M \oplus T_g G$. Le sous-espace $T_x M$ est appelé *sous-espace vertical* de $T_p P$, et $T_g G$ le *sous-espace horizontal*. Pour un fibré principal quelconque, la décomposition est moins claire : si l'on peut facilement retrouver le sous-espace vertical de $T_p P$ pour un fibré (P, π, M) via

$$V_p = T_p P_{\pi(p)} \subset T_p P,$$

il n'y a en revanche pas de façon canonique de caractériser les vecteurs tangents horizontaux. C'est là que la notion de connexion entre en jeu. Il existe une multitude de définitions dans la littérature, nous choisissons la suivante, empruntée à [Moro1].

Définition 1.1.9. Soit (P, π, M) un G -fibré principal. Une *connexion principale* (ou simplement *connexion*) sur P est une application qui à $p \in P$ associe le sous-espace $H_p \subset T_p P$ qui vérifie les hypothèses suivantes :

- H_p est transversal à la fibre P_p , i.e.

$$T_p P = H_p \oplus V_p.$$

- L'ensemble $\{H_p, p \in P\}$ est invariant par l'action de G à droite ;
- L'application $p \mapsto H_p$ est lisse.

Le sous-espace $V_p = \ker d\pi_p$ est appelé *sous-espace vertical* et H_p le *sous-espace horizontal* de $T_p P$. Une définition conceptuellement différente mais qui nous sera utile est la suivante.

Définition 1.1.10. Soit (P, π, M) un G -fibré principal. Une *forme de connexion*⁷ sur P est une 1-forme différentielle $\omega \in \Omega^1(P) \otimes \mathfrak{g}$ sur P à valeurs dans \mathfrak{g} qui vérifie :

- $\omega(\tilde{X}_p) = X, \forall p \in P, \forall X \in \mathfrak{g}$,
- $(R_g)^* \omega = \text{Ad}(g^{-1})\omega, \forall g \in G$.

Le théorème suivant garantit que la notion de connexion et de forme de connexion sont parfaitement équivalentes.

Théorème 1.1.3 ([Moro1], Thm. 6.37). *Soit (P, π, M) un G -fibré principal et ω une forme de connexion sur P . L'application $p \mapsto H_p^\omega := \ker \omega_p$ est une connexion sur P . Réciproquement, si l'on a une connexion sur P , on peut lui associer canoniquement une forme de connexion ω .*

⁶Cela se fait de manière analogue à la Prop. 1.1.1.

⁷On parle aussi de *champ de jauge* en physique.

En conséquence de ce théorème, nous pouvons désormais identifier une connexion à sa forme de connexion, et nous nous autoriserons l'abus de langage suivant : le terme connexion désignera par la suite aussi bien l'application qui détermine un choix de sous-espace horizontal ou la forme différentielle à valeurs dans \mathfrak{g} qui lui est associée. Il nous est également possible d'étendre la notion de connexion à l'espace sous-jacent en tirant en arrière une connexion par une section.

Définition 1.1.11. Soit ω une connexion sur un G -fibré principal (P, π, M) et $\sigma : M \rightarrow P$ une section de ce fibré. On définit la *connexion A induite par ω* sur M , ou *champ de jauge* sur M , par $A = \sigma^*\omega$.

Étant donné une connexion ω sur un G -fibré principal (P, π, M) , et une courbe lisse $c : [a, b] \rightarrow M$, toute courbe $\tilde{c} : [a, b] \rightarrow P$ est appelée un *relèvement* de c si $\pi \circ \tilde{c} = c$ – en particulier, si σ est une section de P alors $\sigma \circ c$ est un relèvement de c . Un tel relèvement est dit *horizontal* (par rapport à la connexion ω) si pour tout t le vecteur tangent $\frac{d}{dt}\tilde{c}(t)$ est un vecteur horizontal, c'est-à-dire $\omega\left(\frac{d}{dt}\tilde{c}(t)\right) = 0$.

Proposition 1.1.4 ([Moroi], Prop. 6.36). Soit G un groupe de Lie compact et (P, π, M) un G -fibré principal, ω une connexion sur ce fibré, et $c : [a, b] \rightarrow M$ une courbe lisse. Pour tout $x_0 \in P_{c(a)}$, il existe un unique relèvement horizontal $\tilde{c} : [a, b] \rightarrow P$ de c tel que $\tilde{c}(a) = x_0$.

La Prop. 1.1.4 signifie que l'on peut *transporter* chaque point de la fibre au-dessus de $c(a)$ sur un point de la fibre au-dessus de $c(b)$ *parallèlement* à la direction définie par la connexion. Ce *transport parallèle* a donné naissance à la notion d'*holonomie*.

Définition 1.1.12. Soit (P, π, M) un G -fibré principal, σ une section de P , ω une connexion sur P et $c : [0, 1] \rightarrow M$ une courbe sur M . Alors l'*holonomie de ω le long de c* est la courbe $h : [0, 1] \rightarrow G$ solution de l'équation différentielle

$$\begin{cases} \frac{d}{dt}h(t) = -\sigma^*\omega\left(\frac{dc(t)}{dt}\right)h(t) \\ h(0) = e. \end{cases} \quad (1.18)$$

Le point terminal $h(1)$ de cette courbe calcule le transport parallèle le long de c déterminé par ω : si l'on note \tilde{c} le relèvement horizontal de c , celui-ci satisfait

$$\tilde{c}(1) = (\sigma(c))(0)h(1).$$

L'application h peut également s'interpréter comme une application $\text{hol}(\omega, c) : P_{c(0)} \rightarrow P_{c(1)}$ G -équivariante, que l'on nomme également *holonomie*, et qui à tout point $p \in P_{c(0)}$ associe $\tilde{c}(1)$, où \tilde{c} est l'unique relèvement horizontal de c tel que $\tilde{c}(0) = p$. On peut vérifier que les deux applications sont bien compatibles. La Fig. 1.5 illustre cette notion d'holonomie et de relèvement.

Mentionnons, à l'attention des géomètres les plus chevronnés, qu'il existe une notion plus générale de connexion, la *connexion d'Ehresmann*, qui permet de définir le transport parallèle sur des fibrés de natures plus diverses. N'étant pas nécessaires à la compréhension de cette thèse nous omettons cette notion, et renvoyons par exemple à [Moroi] ou [Pau14] pour en savoir plus.

Si les fibrés principaux sont au cœur de la théorie de jauge, il sera néanmoins commode de leur associer un fibré vectoriel, appelé fibré adjoint, qui permet de décrire les connexions en termes de formes différentielles sur la base du fibré.

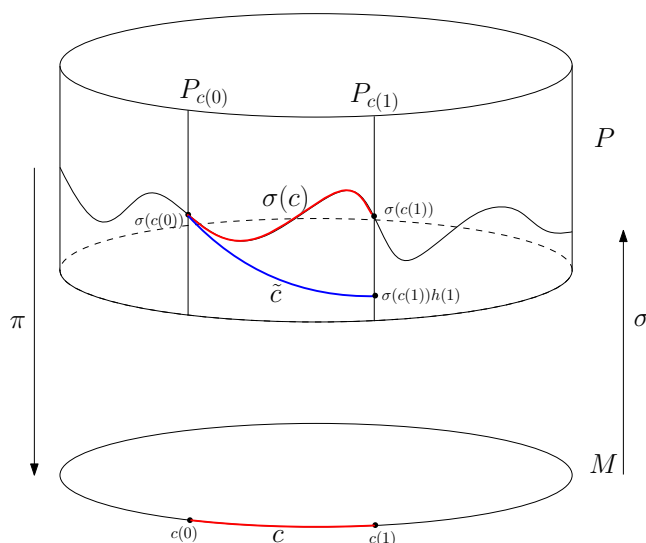


Figure 1.5: Le relèvement horizontal d'une courbe c sur M en une courbe \tilde{c} sur P , traduisant l'apparition de l'holonomie h par rapport à la section $\sigma(c)$.

Définition 1.1.13. Soit (P, π, M) un G -fibré principal, et V un espace vectoriel sur lequel G agit à gauche⁸. On construit un fibré vectoriel de base M et de fibre V de la façon suivante :

$$P \times_G V := (P \times V) / \sim,$$

où l'on définit la relation d'équivalence \sim sur $P \times V$ par

$$(p, \xi) \sim (q, \eta) \Leftrightarrow \exists g \in G, q = p \cdot g, \eta = g^{-1} \cdot \xi.$$

La projection correspondante est

$$\pi_V : \begin{cases} P \times_G V & \rightarrow M \\ [(p, \xi)] & \mapsto \pi(p) \end{cases}$$

$(P \times_G V, \pi_V, M)$ est appelé *fibré associé* à P pour l'action de G sur V .

La Déf. 1.1.13 peut se résumer par le diagramme suivant :

$$\begin{array}{ccc} P \times V & \xrightarrow{\pi \circ \text{pr}_1} & M \\ \downarrow & \nearrow \pi_V & \\ P \times_G V & & \end{array}$$

Ce diagramme définit π_V de manière non-ambiguë. En effet, pour tout $p \in P$, $P_{\pi(p)} = \{p \cdot g, g \in G\}$ par définition du fibré principal, ce qui signifie que $\pi(p) = \pi(p \cdot g)$ pour tout $g \in G$. Nous allons nous pencher sur un fibré associé particulier. Si (P, π, M) est un G -fibré principal, on définit son

⁸On verra dans le paragraphe 1.3 que cela définit une *représentation* de G .

fibré adjoint $\text{ad}(P)$ comme le fibré associé à P pour l'action adjointe de G sur \mathfrak{g} . Les fibres du fibré adjoint $\text{ad}(P)$ sont des copies de \mathfrak{g} , et on peut trouver des trivialisations locales (U_i, φ_i) telles que les applications de transition associées soient des applications continues à valeurs dans $\text{Aut}(\mathfrak{g})$. Si $s : U \subset M \rightarrow \text{ad}P$ est une section locale de P , on peut lui associer de manière unique une fonction $f : P \rightarrow \mathfrak{g}$ telle que

$$(R_g)^* f = \text{Ad}_{g^{-1}} \circ f, \forall g \in G.$$

On peut alors montrer la proposition suivante, qui donne leur forme finale aux connexions.

Proposition 1.1.5. *Soit (P, π, M) un G -fibré principal. On a l'isomorphisme suivant :*

$$\Omega^1(M) \otimes \text{ad}(P) \simeq \{\omega \in \Omega^1(P) \otimes \mathfrak{g} : \omega \text{ est équivariante et nulle sur les vecteurs verticaux}\}.$$

De plus l'espace \mathcal{A} des connexions sur (P, π, M) est un espace affine de direction $\Omega^1(M) \otimes \text{ad}(P)$.

La Prop. 1.1.5 peut s'interpréter de la façon suivante : la différence entre deux connexions sur P équivaut à une 1-forme différentielle sur M à valeur dans $\text{ad}(P)$.

I.I.4 Courbure et action de Yang–Mills

Dans cette section, nous allons définir l'action de Yang–Mills d'une connexion à l'aide de la courbure de celle-ci.

Définition 1.1.14. Soit ω une connexion sur un G -fibré principal (P, π, M) , on peut définir sa *courbure* Ω comme la 2-forme sur P à valeurs dans \mathfrak{g}

$$\Omega = d\omega + \frac{1}{2}[\omega \wedge \omega], \quad (1.19)$$

où $[\omega \wedge \omega]$ est une 2-forme définie par $[\omega \wedge \omega](X, Y) = 2[\omega(X), \omega(Y)]$, pour tous vecteurs X et Y tangents en un même point de P .

On peut montrer que la courbure d'une connexion sur P est G -équivariante et s'annule sur les vecteurs verticaux ; d'après la Prop. 1.1.5 on peut donc identifier Ω à une 2-forme sur M à valeurs dans $\text{ad}(P)$.

Afin de définir l'action de Yang–Mills d'une connexion, il nous est nécessaire d'introduire quelques notions de géométrie riemannienne.

Définition 1.1.15. Soit M une variété différentielle.

- (i) Une *métrique riemannienne* sur M est une *métrique* g sur le fibré tangent TM : plus précisément, g est une application qui à $p \in M$ associe de manière lisse une forme quadratique définie positive g_p sur $T_p M$. Le couple (M, g) est appelé *variété riemannienne*.
- (ii) Si (M, g) est une variété riemannienne orientée⁹, son *volume riemannien* $d\text{vol}$ est l'unique n -forme différentielle sur M positive et de norme 1.

⁹Une variété différentielle M de dimension n est *orientable* si elle admet un atlas dont les applications de transition ont un jacobien positif, ce qui équivaut au fait qu'elle possède une n -forme différentielle partout non nulle, appelée *forme volume*.

(iii) Si (M, g) est une variété riemannienne orientée de dimension n et (e_1, \dots, e_n) une base orthonormée de $T_p M$, où p est un point de M , alors on définit l'opérateur de Hodge

$$\star : \begin{cases} \bigwedge^k(T^*M) & \rightarrow \bigwedge^{n-k}(T^*M) \\ e^{i_1} \wedge \dots \wedge e^{i_k} & \mapsto e^{j_1} \wedge \dots \wedge e^{j_{n-k}} \end{cases},$$

où $e^{j_1} \wedge \dots \wedge e^{j_{n-k}}$ est tel que

$$e^{i_1} \wedge \dots \wedge e^{i_k} \wedge e^{j_1} \wedge \dots \wedge e^{j_{n-k}} = \text{dvol}.$$

En particulier, $\star 1 = \text{dvol}$.

Voyons ce que ces notions donnent en coordonnées locales. Si (M, g) est une variété riemannienne de dimension n , sa métrique s'écrit localement comme une matrice définie positive $(g_{ij}(x))_{1 \leq i, j \leq n}$ avec

$$g_{ij}(x) = g \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right).$$

On emploie en général la notation suivante¹⁰ : $g = g_{ij} dx^i dx^j$. Le volume riemannien de M s'écrit quant à lui

$$\text{dvol} = \sqrt{\det(g_{ij})} dx^1 \wedge \dots \wedge dx^n.$$

Étant donné que la courbure Ω d'une connexion ω s'interprète comme un élément de $\Omega^2(M) \otimes \text{ad}(P)$, dès lors que M est de dimension $n \geq 2$ on peut lui appliquer l'opérateur de Hodge, ce qui donne $\star \Omega \in \Omega^{n-2}(M) \otimes \text{ad}(P)$. Le produit extérieur $\Omega \wedge \star \Omega$ équivaut donc à un élément de $\Omega^n(M) \otimes \text{ad}(P) \otimes \text{ad}(P)$. De plus, le produit scalaire $\langle \cdot, \cdot \rangle$ invariant sur \mathfrak{g} définit une structure euclidienne sur $\text{ad}(P)$, qui se traduit par un produit scalaire que l'on va noter également $\langle \cdot, \cdot \rangle$. Cela nous permet de construire la n -forme différentielle $\langle \Omega \wedge \star \Omega \rangle \in \Omega^n(M)$.

Définition 1.1.16. Soit ω une connexion sur un G -fibré principal (P, π, M) où M est une variété riemannienne orientable, on définit son *action de Yang–Mills* par la quantité suivante :

$$S_{YM}(\omega) = \frac{1}{2} \int_M \langle \Omega \wedge \star \Omega \rangle. \quad (1.20)$$

Les définitions de la courbure et de l'action de Yang–Mills d'une connexion peuvent être rendues plus explicites en passant par une section du fibré : si (P, π, M) est un G -fibré principal, ω une connexion sur P et $\sigma : U \subset M \rightarrow P$ une section locale de ce fibré, alors on peut transporter ω sur M par pull-back en posant $A = \sigma^* \omega$. Cette connexion peut s'écrire en coordonnées locales sur U , en supposant que M soit de dimension n , $A = A_1 dx^1 + \dots + A_n dx^n$, où A_1, \dots, A_n sont des fonctions lisses de U sur \mathfrak{g} . De même, on peut transporter la courbure Ω en une 2-forme F^A sur M :

$$F^A := \sigma^* \Omega = \sum_{1 \leq i < j \leq n} \left(\frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j} + [A_i, A_j] \right) dx_i \wedge dx_j.$$

¹⁰Il s'agit de la *convention de sommation d'Einstein*, selon laquelle on effectue implicitement la somme sur les indices/exposants présents deux fois : c'est ici le cas de i et j .

La restriction sur U de l'action de Yang–Mills de ω devient finalement

$$S_{YM}(\omega)|_U = \frac{1}{2} \sum_{1 \leq i < j \leq n} \int_{\varphi(U)} \left\| \frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j} + [A_i, A_j] \right\|^2 dx_i \wedge dx_j. \quad (1.21)$$

Il est intéressant de remarquer que la dépendance de l'opérateur de Hodge et de l'intégrale sur M vis-à-vis de l'orientation de M se compensent, et donc que l'action S_{YM} est indépendante de l'orientation de M . Cela permet donc en particulier d'étendre sa définition à des variétés non-orientables.

La mesure de Yang–Mills *euclidienne*¹¹, telle qu'elle est introduite dans la littérature physique, est la suivante :

$$d\mu_{YM}(\omega) = \frac{1}{Z} e^{-\frac{1}{2T} S_{YM}(\omega)} d\omega. \quad (1.22)$$

Dans l'équation (1.22), Z désigne une constante de normalisation¹² appelée *fonction de partition*, $T > 0$ est une quantité appelée *constante de couplage*, et $d\omega$ est un équivalent de la mesure de Lebesgue sur l'espace des connexions. Le principal problème avec cette définition est que l'espace des connexions n'est pas localement compact, donc il n'existe pas de mesure de Radon invariante par translation sur cet espace, ce qui est l'idée que l'on se fait de $d\omega$.

I.I.5 Transformations de jauge

Une première étape vers une construction plus rigoureuse de la mesure de Yang–Mills serait d'utiliser les invariances de celle-ci par rapport à certaines transformations sur les connexions, que nous allons voir à présent.

Définition 1.1.17. Soit (P, π, M) un G -fibré principal. Une *transformation de jauge* sur P est un \mathcal{C}^∞ -difféomorphisme $j : P \rightarrow P$ qui préserve les fibres et dont l'action sur P commute avec l'action de G sur P . On note \mathcal{J} l'ensemble de ces transformations.

\mathcal{J} , muni de la loi de composition des fonctions, est un groupe de Lie de dimension infinie avec pour élément neutre la fonction identité sur P . On l'appelle *groupe de jauge*.

Exemple. Supposons que $P = M \times G$ est un fibré trivial. Alors l'action de G sur P est la suivante :

$$\begin{cases} (M \times G) \times G & \rightarrow M \times G \\ ((x, g), h) & \mapsto (x, g) \cdot h = (x, gh) \end{cases} .$$

Une transformation de jauge est donc un \mathcal{C}^∞ -difféomorphisme $j : (M \times G) \rightarrow (M \times G)$ tel que

$$j(x, g) \cdot h = j(x, gh), \quad \forall x \in M, \forall (g, h) \in G^2.$$

En effet, cette équation traduit le fait que les deux actions commutent, mais aussi que j préserve les fibres. L'application

$$\Phi : \begin{cases} \mathcal{J} & \rightarrow \mathcal{C}^\infty(M, G) \\ j & \mapsto J : (x \mapsto j(x, e)) \end{cases}$$

¹¹Cette mesure diffère de la mesure traditionnelle par une rotation de Wick, ce qui en fait une mesure positive au lieu d'une mesure complexe. En quelque sorte, on se place dans un cadre de physique statistique plutôt que de mécanique quantique, et la mesure de Yang–Mills euclidienne est une forme de mesure de Gibbs.

¹²La normalisation est faite de telle sorte que μ_{YM} soit une mesure de probabilité.

est une bijection, et l'action de \mathcal{J} sur $M \times G$ s'identifie à l'action de $\mathcal{C}^\infty(M, G)$ sur $M \times G$ de multiplication à gauche sur la seconde coordonnée.

Comme tous les fibrés que l'on va considérer sont au moins localement triviaux, sinon globalement, on peut désormais transporter la notion de transformation de jauge sur la variété sous-jacente : si M est une variété différentielle, on peut considérer localement une *transformation de jauge* sur M est une fonction lisse $f : U \rightarrow G$, où (U, φ) est une trivialisat on locale du fibr . Dans un souci d'all gement des notations, nous consid rons pour la fin de cette section que le fibr  (P, π, M) est trivial. L'ensemble $\mathcal{J} = \mathcal{C}^\infty(M, G)$ des transformations de jauge sur M est un groupe¹³ pour la multiplication point par point, et il agit sur l'espace \mathcal{A}_M des connexions sur M de la fa on suivante :

$$j \cdot A = \text{Ad}_j A + j^* \omega, \quad \forall A \in \mathcal{A}_M, \quad (1.23)$$

o  ω est la connexion, parfois appel e *forme de Maurer–Cartan*¹⁴, d finie pour $g \in G$ par

$$\omega_g : \begin{cases} T_g G & \rightarrow \mathfrak{g} \\ v & \mapsto (L_{g^{-1}})_* v \end{cases} .$$

Dans le cas o  G est un groupe lin aire on peut v rifier que l' quation (1.23) se r crit

$$(j \cdot A)v = j^{-1} A j v + j^{-1} d j v,$$

en tant que produit de vecteurs par des matrices.

Les transformations de jauge agissent sur la courbure par conjugaison :

$$F^{j \cdot A} = \text{Ad}_j F^A, \quad \forall A \in \mathcal{A}_M.$$

Ainsi, si le produit scalaire sur \mathfrak{g} est ad-invariant, on voit que l'action de Yang–Mills est invariante par transformation de jauge. C'est pourquoi il est possible de restreindre la d finition de $d\mu_{YM}$   l'espace $\mathcal{A}_M / \mathcal{J}$ dans l' quation (1.22).

N anmoins, l'espace $\mathcal{A}_M / \mathcal{J}$ n'est toujours pas localement compact, et donc la bonne d finition d'une mesure invariante par translation n'est pas assur e. Un moyen de le faire est de transporter la mesure sur $\mathcal{A}_M / \mathcal{J}$ en une mesure sur G par le biais d'holonomie le long de lacets.

D finition 1.1.18. Soit c une courbe sur M , et A une connexion sur M . L'*holonomie* de A le long de c est $h(1)$, o  $h : [0, 1] \rightarrow G$ est l'unique solution de l' quation diff rentielle suivante :

$$\begin{cases} \frac{d}{dt} h(t) & = -A \left(\frac{dc(t)}{dt} \right) h(t) \\ h(0) & = e. \end{cases} \quad (1.24)$$

Cette d finition de l'holonomie co cide avec la D f. 1.1.12, si l'on regarde la connexion A sur M comme le pull-back d'une connexion ω sur P par une section σ de P . Les diff rentes propri t s de l'holonomie sont r sum es dans la proposition suivante.

¹³Il s'agit ni plus ni moins que de l'image par Φ du groupe de jauge sur P , et de ce fait Φ devient un isomorphisme de groupes.

¹⁴cf. [DEF⁺99, Moroi]

Proposition 1.1.6. *Soit A une connexion sur M . Alors l'holonomie le long de A vérifie :*

(i) *Si c est une courbe sur M et φ un changement de paramétrage croissant, alors $\text{hol}(A, c \circ \varphi) = \text{hol}(A, c)$.*

(ii) *Si c_1 et c_2 sont des courbes sur M telles que $c_1(1) = c_2(0)$ alors*

$$\text{hol}(A, c_1 c_2) = \text{hol}(A, c_2) \text{hol}(A, c_1). \quad (1.25)$$

(iii) *Si c est une courbe sur M alors*

$$\text{hol}(A, c^{-1}) = \text{hol}(A, c)^{-1}. \quad (1.26)$$

(iv) *Si $j \in \mathcal{J}$ est une transformation de jauge et c une courbe sur M , alors*

$$\text{hol}(j \cdot A, c) = j(c(1))^{-1} \text{hol}(A, c) j(c(0)). \quad (1.27)$$

L'idée finale prend alors forme : la mesure image de la mesure de Yang–Mills sur $\mathcal{A}_M / \mathcal{J}$ par l'holonomie le long d'un lacet est une mesure sur G invariante par conjugaison. C'est cette mesure-là que nous pouvons définir de façon rigoureuse, et nous y reviendrons à la section 1.4.

1.2 PROBABILITÉS NON-COMMUTATIVES

Les probabilités non-commutatives, à l'instar de la sans doute plus célèbre *géométrie non-commutative*¹⁵ dont elles s'inspirent, portent un nom ambigu, sinon fallacieux. En effet, elles ne sont pas à opposer aux probabilités dites “classiques” où les variables aléatoires sont à valeurs dans un espace commutatif, mais il faut plutôt les voir comme une généralisation¹⁶, voire comme un changement de paradigme dans la manière d'aborder les probabilités. En effet, si les probabilités classiques et leur axiomatique de Kolmogorov se fondent sur la théorie de la mesure et l'intégrale de Lebesgue, comme en témoigne la présence de la tribu \mathcal{F} dans l'inéluctable $(\Omega, \mathcal{F}, \mathbb{P})$ introduit dans les cours de probabilités, les probabilités non-commutatives reposent quant à elles sur l'approche bourbakiste de l'intégrale, à savoir que l'on décrit l'intégrale par ses propriétés algébriques (forme linéaire sur un espace de fonctions) plutôt que géométriques (la mesure de l'aire sous la courbe d'une fonction, modulo le signe de cette dernière)¹⁷. Cette divergence de points de vues sur la bonne définition d'une intégrale (“doit-on définir l'intégrale à partir de la mesure ou inversement ?”) est élégamment expliquée en introduction du chapitre XI de [God03], et s'applique bien entendu à la théorie des probabilités.

Bien qu'elles semblent n'être qu'une (re)formalisation algébrique du cadre probabiliste classique, les probabilités non-commutatives ont vu leur utilité dans des applications assez diverses, souvent motivées par des modèles physiques, et l'on peut citer parmi ses branches les plus populaires la théorie des matrices aléatoires ou encore les probabilités libres. Nous allons présenter les bases de cette théorie nécessaires à la compréhension du cadre probabiliste dans lequel se place la mesure de Yang–Mills en deux dimensions ; pour des approfondissements le lecteur intéressé est invité à consulter les ouvrages [AGZ10, Mey86, NS06] qui apportent des points de vue complémentaires sur les probabilités non-commutatives.

¹⁵Développée par Alain Connes, cette théorie est résumée dans l'ouvrage de référence [Con90].

¹⁶On pourrait donc parler de “probabilités non nécessairement commutatives”, ou de “géométrie non nécessairement commutative”, par exemple, mais l'inélégance de la formulation a conduit à la formulation ambiguë que l'on connaît.

¹⁷Notons que dans la plupart des cas, le théorème de Riesz–Markov garantit que ces deux définitions coïncident.

1.2.1 Espaces de probabilité non-commutatifs

Définition 1.2.1. Un *espace de probabilité non-commutatif* est un couple (\mathcal{A}, τ) , où \mathcal{A} est une \mathbb{C} -algèbre unitaire, et τ est une forme linéaire sur \mathcal{A} qui vérifie $\tau(1_{\mathcal{A}}) = 1$, appelée *état*. τ est appelé *état tracial*, ou *trace*, s'il vérifie la propriété supplémentaire suivante :

$$\tau(ab) = \tau(ba), \quad \forall (a, b) \in \mathcal{A}^2.$$

On peut, selon les besoins de la théorie, enrichir la structure d'un espace de probabilité non-commutatif : en voici quelques exemples traditionnels.

Définition 1.2.2. Soit (\mathcal{A}, τ) un espace de probabilité non-commutatif.

- (i) Si \mathcal{A} est muni d'une involution $*$ antilinéaire telle que $(ab)^* = b^*a^*$ pour tout $(a, b) \in \mathcal{A}^2$ et telle que

$$\tau(a^*a) \geq 0, \quad \forall a \in \mathcal{A},$$

alors l'état τ est dit *positif* et (\mathcal{A}, τ) est appelé **-espace de probabilité*. L'état est dit *fidèle* s'il satisfait la propriété suivante pour tout $a \in \mathcal{A}$:

$$\tau(a^*a) = 0 \iff a = 0.$$

- (ii) Si (\mathcal{A}, τ) est un *-espace de probabilité muni d'une norme $\|\cdot\| : \mathcal{A} \rightarrow [0, \infty)$ pour laquelle τ est continu, telle que $(\mathcal{A}, \|\cdot\|)$ soit un espace vectoriel normé complet et qui vérifie

$$\|ab\| \leq \|a\|\|b\|, \quad \forall (a, b) \in \mathcal{A}^2,$$

et

$$\|a^*a\| = \|a\|^2, \quad \forall a \in \mathcal{A},$$

alors $(\mathcal{A}, \tau, \|\cdot\|)$ est appelé un *C*-espace de probabilité*¹⁸.

- (iii) Si $(\mathcal{A}, \tau, \|\cdot\|)$ est un C*-espace de probabilité tel que \mathcal{A} soit une algèbre de von Neumann¹⁹, alors $(\mathcal{A}, \tau, \|\cdot\|)$ est appelé *W*-espace de probabilité*.

Les C*-espaces de probabilités et les W*-espaces de probabilités bénéficient naturellement de la théorie des algèbres d'opérateurs développée entre autres par Gelfand, Naimark, von Neumann, ou encore Dixmier. Les *-espaces de probabilités englobent des espaces plus généraux, privilégiant parfois une approche plus combinatoire à défaut de bénéficier d'outils analytiques aussi puissants que ceux susmentionnés. Enfin, les espaces de probabilités non-commutatifs dépourvus d'involution sont un cas particulièrement abrupt à traiter, car ils ne possèdent même pas de notion de positivité.

¹⁸En particulier, \mathcal{A} est appelé C*-algèbre. Un théorème fondamental de la théorie des C*-algèbres, dû à Gelfand et Naimark [GN43], permet d'identifier une C*-algèbre commutative à l'algèbre des fonctions continues sur un espace localement compact et tendant vers 0 à l'infini, et une C*-algèbre quelconque à l'algèbre $\mathcal{B}(H)$ des opérateurs bornés sur un espace de Hilbert. Le livre de Dixmier [Dix64] fait office de référence sur le sujet.

¹⁹Une algèbre de von Neumann est une sous-algèbre de l'algèbre $\mathcal{B}(H)$ des opérateurs bornés sur un espace de Hilbert, qui est fermée pour la topologie faible sur $\mathcal{B}(H)$. On peut montrer que toute algèbre de von Neumann est une C*-algèbre.

Exemples. (i) Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. On pose

$$\mathcal{A} = L^{\infty-}(\Omega, \mathbb{P}) = \bigcap_{1 \leq p < \infty} L^p(\Omega, \mathbb{P}),$$

et $\tau : \mathcal{A} \rightarrow \mathbb{C}$ défini par

$$\tau(X) = \mathbb{E}[X] = \int_{\Omega} X(\omega) d\mathbb{P}(\omega).$$

Alors (\mathcal{A}, τ) est un espace de probabilité non-commutatif, correspondant à l'espace des variables aléatoires dont tous les moments sont finis. C'est même un $*$ -espace de probabilité, si l'on le munit de l'involution $X \mapsto \bar{X}$. Notons que les espaces L^p , pris individuellement, ne sont pas des espaces non-commutatifs : en effet, n'étant pas stables par la multiplication, en particulier ce ne sont pas des algèbres. En revanche, l'espace L^{∞} en est un, mais pas nécessairement intéressant car il ne contient pas certaines variables aléatoires fondamentales, telles que les gaussiennes.

- (ii) L'espace $(\mathcal{M}_N(\mathbb{C}), \text{tr}, *)$ où $\text{tr} = \frac{1}{N} \text{Tr}$ est la trace normalisée et $*$ est le passage à la matrice adjointe, est un $*$ -espace de probabilité.
- (iii) On a mentionné la théorie des matrices aléatoires comme application des probabilités non-commutatives, il est donc naturel de se demander dans quel espace ces fameuses matrices vivent ; il s'agit simplement du produit tensoriel des espaces des points précédents, à savoir

$$\mathcal{M}_N(\mathbb{C}) \otimes L^{\infty-}(\Omega, \mathbb{P}) \simeq \mathcal{M}_N(L^{\infty-}(\Omega, \mathbb{P})),$$

et on le munit de la trace $\tau = \text{tr} \otimes \mathbb{E}$, qui vérifie

$$\tau(M) = \int_{\Omega} \text{tr}(M(\omega)) d\mathbb{P}(\omega), \quad \forall M \in \mathcal{M}_N(L^{\infty-}(\Omega, \mathbb{P})).$$

Plus précisément, soit $A^{(N)}$ un élément de $\mathcal{M}_N(\mathbb{C}) \otimes (L^{\infty-}(\Omega, \mathbb{P}))$ où $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace de probabilités. $A^{(N)}$ n'est rien d'autre qu'une variable aléatoire sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans $\mathcal{M}_N(\mathbb{C})$, donc *stricto sensu* une matrice aléatoire.

1.2.2 Distributions non-commutatives

Si l'on veut suivre le chemin des probabilités classiques, une fois les espaces de probabilité et les variables aléatoires introduits, l'étape qui suit consiste à définir la *loi* d'une ou plusieurs variables aléatoires, et à trouver des moyens de la caractériser. Par exemple, en probabilités classiques, la loi d'une variable aléatoire réelle X sur $(\Omega, \mathcal{F}, \mathbb{P})$ est définie par la mesure image de \mathbb{P} par X , à savoir $\mathcal{P}_X = X_*\mathbb{P}$, et se caractérise notamment :

- (i) Par l'intégrale de fonctions mesurables bornées (ou positives), via la formule de transfert

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}} f(x) d\mathcal{P}_X(x),$$

ce qui donne une caractérisation *fonctionnelle* ;

- (ii) Par la fonction de répartition $f_X(x) = \mathbb{P}(X \leq x) = \mathbb{E}[\mathbf{1}_{]-\infty, x]}(X)$, qui équivaut à la caractérisation fonctionnelle du fait que les fonctions indicatrices d'intervalles $] - \infty, x]$ engendrent les fonctions mesurables positives/bornées sur \mathbb{R} ;
- (iii) Par la fonction caractéristique $\Phi_X : t \mapsto \mathbb{E}[e^{itX}]$, ce qui correspond à la transformée de Fourier inverse de \mathcal{P}_X , et on peut alors parler de caractérisation *spectrale*.

Dans le cas non-commutatif, les définitions peuvent différer selon les propriétés de l'espace considéré. La plus générale est la suivante.

Définition 1.2.3. Soit (\mathcal{A}, τ) un $*$ -espace de probabilité, et a un élément de \mathcal{A} . La loi de a est la forme linéaire $\mu : \mathbb{C}\langle X, X^* \rangle \rightarrow \mathbb{C}$ définie par

$$\mu(P(X, X^*)) = \tau(P(a, a^*)), \quad \forall P \in \mathbb{C}\langle X, X^* \rangle.$$

On peut identifier l'espace des polynômes non-commutatifs à deux indéterminées et à coefficients complexes, en tant qu'algèbre, à l'algèbre engendrée par les mots (ou monômes non-commutatifs) en les variables a et a^* , *i.e.*

$$w(a, a^*) = a^{\alpha_1} a^{*\beta_1} \dots a^{\alpha_k} a^{*\beta_k}, \quad (\alpha_1, \beta_1, \dots, \alpha_k, \beta_k) \in \mathbb{N}^{2k}, \quad k \in \mathbb{N},$$

ou encore

$$w(a, a^*) = a^{\varepsilon_1} \dots a^{\varepsilon_n}, \quad (\varepsilon_1, \dots, \varepsilon_n) \in \{1, *\}^n, \quad n \in \mathbb{N}.$$

Le produit des monômes se traduit explicitement par la concaténation des mots. Les images de mots en a et a^* par τ constituent ce que l'on appelle les $*$ -moments de a .

Heuristiquement, la Déf. 1.2.3 revient à prendre la caractérisation fonctionnelle des variables aléatoires classiques et à remplacer les fonctions mesurables bornées par les fonctions polynomiales, ôtant de ce fait toute notion d'analyse et se focalisant sur la partie algébrique. Des interprétations plus analytiques peuvent être obtenues sous des hypothèses plus fortes : par exemple, si $a \in \mathcal{A}$ vérifie $aa^* = a^*a$ (on dit que a est un élément *normal* de \mathcal{A}), et s'il existe une mesure μ sur \mathbb{C} à support compact telle que

$$\int_{\mathbb{C}} z^k \bar{z}^\ell d\mu(z) = \tau(a^k a^{*\ell}), \quad \forall (k, \ell) \in \mathbb{N}^2,$$

alors μ caractérise également la distribution de a et s'apparente à la mesure image \mathcal{P}_X du cas classique. Si de plus a est autoadjoint, alors $\text{supp}(\mu) \subset \mathbb{R}$ et les $*$ -moments de a correspondent aux moments de μ au sens classique. Par ailleurs, si \mathcal{A} est un C^* -espace de probabilité et si $a \in \mathcal{A}$ est normal, alors la mesure μ décrite précédemment existe, et est appelée *mesure spectrale* de a .

Exemples.

- (i) Soit (\mathcal{A}, τ) un $*$ -espace de probabilité. Un élément $u \in \mathcal{A}$ est un *unitaire de Haar* s'il vérifie $uu^* = u^*u = 1$ et si

$$\tau(u^n) = 0, \quad \forall n \in \mathbb{Z} \setminus \{0\}.$$

- (ii) Soit (\mathcal{A}, τ) un $*$ -espace de probabilité et $r > 0$ un réel positif. Si $x \in \mathcal{A}$ est autoadjoint de mesure spectrale $\frac{2}{r^2\pi} \sqrt{r^2 - t^2} \mathbf{1}_{[-r,r]}(t) dt$, c'est-à-dire que

$$\tau(x^k x^{*\ell}) = \int_{-r}^r t^{k+\ell} \sqrt{r^2 - t^2} \frac{2dt}{r^2\pi},$$

alors x est appelé *élément semi-circulaire de rayon r* (ou de variance r^2). La loi associée est la *loi du demi-cercle*, et il s'agit de l'analogue de la loi normale en probabilités non-commutatives. Notons que la loi du demi-cercle est caractérisée par ses moments, et que ceux-ci ont une formulation explicite à l'aide des *nombres de Catalan* $C_n = \frac{1}{n+1} \binom{2n}{n}$, $n \in \mathbb{N}$: si x est un élément semi-circulaire de rayon r , alors pour tout $k \in \mathbb{N}$ on a

$$\tau(x^{2k}) = \left(\frac{r}{2}\right)^{2k} C_k, \quad \tau(x^{2k+1}) = 0. \quad (1.28)$$

- (iii) Soit (\mathcal{A}, τ) un espace de probabilité non-commutatif. Les éléments appartenant à $\mathbb{C}1_{\mathcal{A}}$ sont, par définition, des constantes, et correspondent exactement aux variables aléatoires constantes dans le cadre classique où $\mathcal{A} = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ puisqu'elles vérifient bien $\tau(z1_{\mathcal{A}}) = z\tau(1_{\mathcal{A}}) = z$ pour tout $z \in \mathbb{C}$.
- (iv) Soit $A^{(N)}$ un élément de $\mathcal{M}_N(\mathbb{C}) \otimes (L^\infty(\Omega, \mathbb{P}))$ où $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace de probabilités. Si $A^{(N)}$ est une matrice normale, alors la mesure spectrale de $A^{(N)}$ existe et est donnée par l'intensité de la *mesure empirique spectrale*

$$\mu_A = \mathbb{E}[\widehat{\mu}_{A^{(N)}}] = \mathbb{E}\left[\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i(A^{(N)})}\right]$$

où $\lambda_i : \mathcal{M}_N(\mathbb{C}) \rightarrow \mathbb{C}$ est la fonction qui à une matrice complexe de taille $N \times N$ associe sa i -ème valeur propre²⁰.

L'exemple (iii) ci-dessus justifie *a posteriori* la condition, qui aurait pu sembler superfétatoire, selon laquelle un espace de probabilité non-commutatif soit muni d'une unité : en effet, il est naturel de demander que les constantes soient mesurables quel que soit l'espace sur lequel on se place.

Étant donné des éléments $(a_1, \dots, a_n) \in \mathcal{A}^n$, on peut également définir la distribution jointe de ces variables aléatoires, en généralisant la Déf. 1.2.3 à un nombre quelconque de variables.

Définition 1.2.4. Soit (\mathcal{A}, τ) un $*$ -espace de probabilité, et (a_1, \dots, a_n) des éléments de \mathcal{A} . La *distribution jointe* de a_1, \dots, a_n est la forme linéaire μ sur $\mathbb{C}\langle X_1, X_1^*, \dots, X_n, X_n^* \rangle$ définie par

$$\mu(P(X_1, X_1^*, \dots, X_n, X_n^*)) = \tau(P(a_1, a_1^*, \dots, a_n, a_n^*)), \quad \forall P \in \mathbb{C}\langle X_1, X_1^*, \dots, X_n, X_n^* \rangle.$$

Notons que la distribution jointe au sens des probabilités non-commutatives n'admet pas nécessairement d'analogue en probabilités classiques, notamment si les variables aléatoires ne commutent pas. Plus précisément, si deux variables aléatoires a et b hermitiennes ne commutent pas, il n'existe pas toujours de mesure sur \mathbb{R}^2 dont les moments sont ceux du couple (a, b) .

²⁰Le choix de l'ordre parmi les valeurs propres n'a pas d'importance puisque la mesure empirique spectrale est une fonction symétrique des valeurs propres de $A^{(N)}$.

Exemple. Soit $(\mathcal{M}_N(\mathbb{C}), \text{tr})$ l'espace de probabilité non-commutatif défini par les matrices de taille $N \times N$. On se donne $(A, B) \in \mathcal{M}_N(\mathbb{C})^2$ un couple de matrices hermitiennes. Alors en particulier A et B admettent des mesures spectrales en tant que matrices normales. En revanche, si elles ne commutent pas, il n'existe pas nécessairement de mesure sur \mathbb{R}^2 admettant les mêmes moments que leur distribution jointe

$$\mu(P(X, Y)) = \text{tr}(P(A, B)), \quad \forall P \in \mathbb{C}\langle X, Y \rangle.$$

En effet, si l'on prend A et B les matrices de $\mathcal{M}_2(\mathbb{C})$ définies par

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

alors on peut notamment vérifier que

$$ABAB = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \quad ABBA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Or s'il existait une mesure μ sur \mathbb{R}^2 admettant les mêmes moments que la loi jointe de A et B , en particulier on aurait

$$\text{tr}(ABAB) = \int_{\mathbb{R}^2} x^2 y^2 d\mu(x, y) = \text{tr}(ABBA),$$

ce qui est impossible puisque $\text{tr}(ABAB) = 1$ tandis que $\text{tr}(ABBA) = \frac{3}{2}$.

1.2.3 Indépendance et liberté

En probabilités classiques, l'indépendance se traduit tout d'abord à l'échelle des tribus : la famille de tribus $(\mathcal{G}_1, \dots, \mathcal{G}_k)$ est indépendante si pour tout $A_i \in \mathcal{G}_i$ on a la factorisation $\mathbb{P}(\bigcap_i A_i) = \prod_i \mathbb{P}(A_i)$. Des variables aléatoires X_1, \dots, X_n respectivement à valeurs dans $(E_1, \mathcal{E}_1), \dots, (E_n, \mathcal{E}_n)$ sont indépendantes si leurs tribus engendrées sont indépendantes, et on a l'équivalence entre les points suivants :

- (i) X_1, \dots, X_n sont indépendantes ;
- (ii) $\mathcal{P}_{(X_1, \dots, X_n)} = \mathcal{P}_{X_1} \otimes \dots \otimes \mathcal{P}_{X_n}$;
- (iii) $\mathbb{E}[\prod_i f_i(X_i)] = \prod_i \mathbb{E}[f_i(X_i)]$, pour toutes $f_i : E_i \rightarrow \mathbb{R}$ mesurables bornées.
- (iv) (pour des X_i à valeurs dans \mathbb{R}) $\Phi_{(X_1, \dots, X_n)}(t_1, \dots, t_n) = \Phi_{X_1}(t_1) \cdots \Phi_{X_n}(t_n)$, pour tout $(t_1, \dots, t_n) \in \mathbb{R}^n$.

Les point (ii) et (iii) sont parfaitement explicites quant à la nature algébrique de l'indépendance : en effet, la mesure produit $\mathcal{P}_{X_1} \otimes \dots \otimes \mathcal{P}_{X_n}$ du point (ii) est le produit tensoriel des mesures, vues comme formes linéaires sur l'espace des fonctions mesurables bornées à valeurs dans \mathbb{R} , et le point (iii) caractérise justement ce produit tensoriel du point de vue des formes linéaires. C'est pourquoi on parle d'indépendance *tensorielle*. Il est possible de redéfinir cette forme d'indépendance dans un cadre non-commutatif, mais nous allons voir que ce n'est pas spécialement naturel – ou adapté.

Définition 1.2.5. Soit $(\mathcal{A}_i)_{i \in I}$ des sous-algèbres unitaires d'un espace de probabilité non-commutatif (\mathcal{A}, τ) . Elles sont dites *tensoriellement indépendantes* (ou simplement *indépendantes*) si les \mathcal{A}_i commutent deux à deux et si τ se factorise comme suit :

$$\tau\left(\prod_j a_j\right) = \prod_j \tau(a_j), \quad \forall a_j \in \mathcal{A}_j, \quad \forall j \in J, \quad \forall J \subset I \text{ fini.} \quad (1.29)$$

Des variables aléatoires $(a_i)_{i \in I}$ sont dites *indépendantes* si les algèbres qu'elles engendrent respectivement sont indépendantes.

Si l'on garde en tête l'idée que τ joue le rôle de l'espérance par rapport à la mesure de probabilité dans le cas non-commutatif, et \mathcal{A} l'algèbre des fonctions mesurables bornées, il est clair que l'équation (1.29) correspond à la caractérisation fonctionnelle de l'indépendance. Le problème est que cette notion n'est utile que pour des variables aléatoires qui commutent : par exemple, si a et b ne commutent pas, l'équation (1.29) ne permet pas de calculer (sous forme de factorisation) toutes les quantités de la forme

$$\tau(a^{k_1} b^{\ell_1} \dots a^{k_n} b^{\ell_n}),$$

et en particulier on voit qu'elle ne permet pas de décrire intégralement la *-loi jointe de variables aléatoires qui ne commutent pas à partir de leurs *-lois respectives, comme c'est le cas en probabilités classiques. C'est pourquoi l'indépendance tensorielle est insuffisante, du moins pour décrire les interactions de variables aléatoires qui ne commutent pas.

Définition 1.2.6. Soit $(\mathcal{A}_i)_{i \in I}$ des sous-algèbres unitaires d'un espace de probabilité non-commutatif (\mathcal{A}, τ) . Elles sont dites *librement indépendantes* (ou simplement *libres*) si, pour tout $n \in \mathbb{N}$, pour tout $(i_1, \dots, i_n) \in I^n$ tel que $i_k \neq i_{k+1}$ pour tout $1 \leq k \leq n$ ²¹ et pour tout $a_k \in \mathcal{A}_{i_k}$, $\tau(a_k) = 0$ implique

$$\tau(a_1 \dots a_n) = 0.$$

Des variables aléatoires $(a_i)_{i \in I}$ sont *libres* si les algèbres unitaires qu'elles engendrent respectivement sont libres.

Cette définition paraît souvent contre-intuitive et inutilement compliquée de prime abord pour quiconque est habitué aux probabilités classiques ; elle l'est moins pour les algébristes qui pourront reconnaître une analogie avec la notion de liberté dans les groupes.

Définition 1.2.7. Soit G un groupe et $(G_i)_{i \in I}$ des sous-groupes de G . Les (G_i) sont *libres* si pour tout $n \in \mathbb{N}$, pour tous $i_1 \neq \dots \neq i_n$ et pour tout $g_k \in G_{i_k} \setminus \{e\}$, on a

$$g_1 \dots g_n \neq e.$$

Cette notion de liberté au sein de la théorie des groupes induit par extension celle de *produit libre* de groupes, et même d'algèbres. Il est rassurant de voir que cette analogie n'est pas infondée, comme nous allons le voir. En effet, pour un groupe donné G , on peut munir son algèbre de groupe $\mathbb{C}G$ d'une structure d'espace de probabilité non-commutatif, en posant $\tau\left(\sum_{g \in G} \alpha_g g\right) = \alpha_e$, et on a alors le résultat suivant.

²¹On utilisera par la suite la notation $i_1 \neq \dots \neq i_n$ pour signifier cette condition, en gardant à l'esprit que \neq n'est pas une relation transitive.

Proposition 1.2.1. *Soit G un groupe et $(G_i)_{i \in I}$ des sous-groupes de G . Les points suivants sont équivalents :*

- (i) *Les groupes $(G_i)_i$ sont libres ;*
- (ii) *Les algèbres de groupes $\mathbb{C}G_i$, en tant que sous-espaces de probabilité de $(\mathbb{C}G, \tau)$, sont libres.*

Il n'est pas difficile de s'assurer que la relation de liberté est symétrique. Outre le lien entre la liberté en probabilités non-commutatives et la liberté en théorie des groupes, nous allons voir dans le lemme suivant que la liberté de variables aléatoires permet de s'assurer que leur loi jointe soit entièrement caractérisée par leurs lois individuelles respectives, tout comme pour des variables aléatoires indépendantes en probabilités classiques.

Lemme 1.2.2. *Soit (\mathcal{A}, τ) un espace de probabilité non-commutatif et $(\mathcal{A}_i)_{i \in I}$ une famille de sous-espaces libres. Soit \mathcal{B} l'algèbre engendrée par les \mathcal{A}_i , alors $\tau|_{\mathcal{B}}$ est entièrement déterminé par $(\tau|_{\mathcal{A}_i})_{i \in I}$.*

1.2.4 Convergence dans des espaces non-commutatifs

Contrairement aux probabilités classiques pour lesquelles de nombreuses notions de convergence coexistent et possèdent des caractéristiques propres, les espaces de probabilités non-commutatifs, sans qu'on les pourvoie de structure analytique supplémentaire, ne possèdent qu'un seul type de convergence : la convergence en distribution non-commutative.

Définition 1.2.8. *Soit $(\mathcal{A}_N, \tau_N)_{N \in \mathbb{N}}$ une suite d'espaces de probabilité non-commutatifs et (\mathcal{A}, τ) un espace de probabilité non-commutatif. Une suite de variables aléatoires $(a_N)_{N \in \mathbb{N}}$, où $a_N \in \mathcal{A}_N$ pour tout N , converge en distribution non-commutative vers $a \in \mathcal{A}$ quand N tend vers l'infini si*

$$\lim_{N \rightarrow \infty} \tau_N(a_N^n) = \tau(a^n), \quad \forall n \in \mathbb{N}. \quad (1.30)$$

Si l'on se rapporte au cas classique, cette convergence correspond à la fois à la convergence en loi et à la convergence dans L^p pour tout $p \in \mathbb{N}$. En guise d'illustration, voici deux formes du théorème central limite, l'une fondée sur l'indépendance tensorielle, l'autre sur la liberté.

Théorème 1.2.3. *Soit (\mathcal{A}, τ) un $*$ -espace de probabilité et $(a_N)_{N \in \mathbb{N}^*}$ une suite d'éléments de \mathcal{A} indépendants et identiquement distribués tels que pour tout $N \in \mathbb{N}^*$:*

- a_N est autoadjoint ;
- $\tau(a_N) = 0$ et $\tau(a_N^2) = \sigma^2 < \infty$.

Alors on a la convergence en distribution non-commutative suivante :

$$\frac{a_1 + \cdots + a_N}{\sqrt{N}} \rightarrow x, \quad (1.31)$$

où x est une variable aléatoire normale centrée de variance σ^2 .

Théorème 1.2.4. *Soit (\mathcal{A}, τ) un $*$ -espace de probabilité et $(a_N)_{N \in \mathbb{N}^*}$ une famille libre d'éléments de \mathcal{A} identiquement distribués tels que pour tout $N \in \mathbb{N}^*$:*

- a_N est autoadjoint ;
- $\tau(a_N) = 0$ et $\tau(a_N^2) = \sigma^2 < \infty$.

Alors on a la convergence en distribution non-commutative suivante :

$$\frac{a_1 + \cdots + a_N}{\sqrt{N}} \rightarrow s, \quad (1.32)$$

où s est un élément semi-circulaire de variance σ^2 .

La Déf. 1.2.8 admet une généralisation naturelle aux familles quelconques de variables aléatoires, en particulier les processus.

Définition 1.2.9. Soit $(\mathcal{A}_N, \tau_N)_{N \in \mathbb{N}}$ une suite d'espaces de probabilité non-commutatifs et (\mathcal{A}, τ) un espace de probabilité non-commutatif. Soit I un ensemble d'indices et on considère pour tout $i \in I$ la suite de v.a. $(a_N^{(i)})_{N \in \mathbb{N}}$ avec $a_N^{(i)} \in \mathcal{A}_N$ et $a^{(i)} \in \mathcal{A}$. Alors :

- (i) $(a_N^{(i)})_{i \in I, N \in \mathbb{N}}$ converge en distribution non-commutative vers $(a^{(i)})_{i \in I}$ si on a

$$\lim_{N \rightarrow \infty} \tau_N(a_N^{(i_1)} \cdots a_N^{(i_n)}) = \tau(a^{(i_1)} \cdots a^{(i_n)}), \quad \forall n \in \mathbb{N}, \forall (i_1, \dots, i_n) \in I^n. \quad (1.33)$$

- (ii) $(a_N^{(i)})_{i \in I, N \in \mathbb{N}}$ converge en $*$ -distribution vers $(a^{(i)})_{i \in I}$ si $(a_N^{(i)}, (a_N^{(i)})^*)_{N \in \mathbb{N}, i \in I}$ converge en distribution vers $(a^{(i)}, (a^{(i)})^*)_{i \in I}$.

Dans le contexte spécifique des matrices aléatoires, la structure analytique de l'espace auquel appartient l'objet limite permet de définir des notions de convergences plus fines et plus variées, s'approchant de celles que l'on retrouve en probabilités classiques.

Définition 1.2.10. Soit $(A_1^{(N)}, \dots, A_n^{(N)})_{N \in \mathbb{N}}$ une suite de n -uplets de matrices aléatoires, c'est-à-dire qu'il existe un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ tel que pour tout n et pour tout N , $A_n^{(N)} : \Omega \rightarrow \mathcal{M}_N(\mathbb{C})$ soit une variable aléatoire, ou encore que $A_n^{(N)}$ soit un élément de $\mathcal{M}_N(L^\infty(\Omega, \mathbb{P}))$. Soit $\|\cdot\|_{\mathcal{M}_N(\mathbb{C})}$ la norme de Frobenius donnée par $\|M\|_{\mathcal{M}_N(\mathbb{C})} = \sqrt{\text{Tr}(MM^*)}$, ainsi que $(a_1, \dots, a_n) \in \mathcal{A}^n$ un n -uplet d'éléments d'un W^* -espace de probabilité $(\mathcal{A}, \tau, \|\cdot\|)$.

- (i) La suite de n -uplets $(A_1^{(N)}, \dots, A_n^{(N)})$ converge en distribution non-commutative presque-sûrement vers (a_1, \dots, a_n) si pour presque-tout ω , la famille $(A_1^{(N)}(\omega), \dots, A_n^{(N)}(\omega)) \in \mathcal{M}_N(\mathbb{C})$ converge en distribution non-commutative vers (a_1, \dots, a_n) au sens de la Déf. 1.2.9.
- (ii) La suite de n -uplets $(A_1^{(N)}, \dots, A_n^{(N)})$ converge en distribution non-commutative fortement vers (a_1, \dots, a_n) si la convergence a lieu en distribution non-commutative presque-sûrement et si pour tout $P \in \mathbb{C}\langle X_1, \dots, X_n \rangle$,

$$\lim_{N \rightarrow \infty} \|P(A_1^{(N)}, \dots, A_n^{(N)})\|_{\mathcal{M}_N(\mathbb{C})} = \|P(a_1, \dots, a_n)\|_{\mathcal{A}}, \text{ p.s.} \quad (1.34)$$

1.2.5 Mouvement brownien unitaire et limite d'échelle

Dans cette section, on met en pratique la théorie développée précédemment dans le but d'étudier un processus stochastique matriciel qui sera présent (de manière plus ou moins explicite) dans la théorie de Yang–Mills en deux dimensions, à savoir le mouvement brownien sur le groupe de structure $U(N)$ dans notre cas. Il est construit comme solution d'une équation différentielle stochastique par rapport au mouvement brownien sur \mathfrak{u}_N ; ce dernier peut, comme nous allons le voir, s'interpréter comme une matrice aléatoire dont les coefficients sont des combinaisons linéaires (à coefficients complexes) de mouvements browniens réels. Il est également, à une rotation près, égal au mouvement brownien hermitien introduit par Dyson [Dys62] qui a été étudié plus en détail que son analogue unitaire : par exemple le processus de ses valeurs propres, appelé *mouvement brownien de Dyson*²², est le sujet de nombreuses recherches dont certaines sont liées aux questions d'universalité mentionnées au paragraphe 1.2.1.

Définition 1.2.11. Soit \mathfrak{u}_N l'algèbre de Lie de $U(N)$, munie d'un produit scalaire $\langle \cdot, \cdot \rangle$ Ad-invariant. Le *mouvement brownien* sur \mathfrak{u}_N est le processus gaussien $K = (K_t)_{t \geq 0}$ à valeurs dans \mathfrak{u}_N qui vérifie :

$$\mathbb{E}[\langle X, K_t \rangle \langle Y, K_s \rangle] = (s \wedge t) \langle X, Y \rangle, \quad \forall (X, Y) \in \mathfrak{u}_N^2, \quad \forall (s, t) \in (\mathbb{R}_+)^2. \quad (1.35)$$

Il est possible de se ramener à l'étude de mouvements browniens réels en passant en coordonnées : étant donné une base orthonormée (X_1, \dots, X_d) de \mathfrak{u}_N et $d = \dim \mathfrak{u}_N = N^2$ mouvements browniens réels standards $((B_t^{(i)})_{t \geq 0}, 1 \leq i \leq d)$, on peut construire explicitement K par

$$K_t = \sum_i B_t^{(i)} X_i.$$

Munissons en l'occurrence \mathfrak{u}_N du produit scalaire suivant :

$$\langle X, Y \rangle = N \operatorname{Tr}(X^* Y) = N^2 \operatorname{tr}(X^* Y). \quad (1.36)$$

On peut se servir du plongement de \mathfrak{u}_N dans $\mathcal{M}_N(\mathbb{C})$ et utiliser la base canonique (E_{ij}) de $\mathcal{M}_N(\mathbb{C})$ pour décrire une base orthonormale de \mathfrak{u}_N . On pose, pour $1 \leq k < \ell \leq N$,

$$X_{k\ell} = \frac{1}{\sqrt{2N}}(E_{k\ell} - E_{\ell k}), \quad Y_{k\ell} = \frac{i}{\sqrt{2N}}(E_{k\ell} + E_{\ell k}),$$

et pour $1 \leq k \leq N$

$$Z_k = \frac{i}{\sqrt{N}} E_{kk}.$$

Cette famille est bien orthonormale pour le produit scalaire défini plus haut, et comme \mathfrak{u}_N est de dimension $d = N^2$ on en déduit bien que c'est une base orthonormale. En conclusion, on obtient la forme matricielle explicite suivante du mouvement brownien sur \mathfrak{u}_N :

$$K_t = \begin{pmatrix} \frac{iB_t^1}{\sqrt{N}} & \frac{B_t^{1,2} + iB_t'^{1,2}}{\sqrt{2N}} & \frac{B_t^{1,3} + iB_t'^{1,3}}{\sqrt{2N}} & \dots & \frac{B_t^{1,N} + iB_t'^{1,N}}{\sqrt{2N}} \\ \frac{B_t^{1,2} - iB_t'^{1,2}}{\sqrt{2N}} & \frac{iB_t^2}{\sqrt{N}} & \frac{B_t^{2,3} + iB_t'^{2,3}}{\sqrt{2N}} & \dots & \frac{B_t^{2,N} + iB_t'^{2,N}}{\sqrt{2N}} \\ \frac{B_t^{1,3} - iB_t'^{1,3}}{\sqrt{2N}} & \frac{B_t^{2,3} - iB_t'^{2,3}}{\sqrt{2N}} & \frac{iB_t^3}{\sqrt{N}} & \ddots & \frac{B_t^{3,N} + iB_t'^{3,N}}{\sqrt{2N}} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \frac{B_t^{1,N} - iB_t'^{1,N}}{\sqrt{2N}} & \frac{B_t^{2,N} - iB_t'^{2,N}}{\sqrt{2N}} & \frac{B_t^{3,N} - iB_t'^{3,N}}{\sqrt{2N}} & \dots & \frac{iB_t^N}{\sqrt{N}} \end{pmatrix} \quad (1.37)$$

²²On trouvera sa définition et ses principales propriétés au paragraphe 4.3.1 de [AGZ10].

La variation quadratique $d[K]_t$ correspond en fait au produit matriciel $(dK_t dK_t) = -I_N dt$. Une façon de voir cela est de faire apparaître l'élément de Casimir $C_{\mathfrak{u}_N}$, que l'on va décrire au paragraphe 1.3.3 :

$$dK_t \otimes dK_t = \sum_{k,\ell=1}^d dB_t^{(k)} dB_t^{(\ell)} x_k \otimes x_\ell = C_{\mathfrak{u}_N} dt.$$

Or l'image de $C_{\mathfrak{u}_N}$ par l'application $x \otimes y \mapsto xy$ est $c_{\mathfrak{u}_N} I_N$, où $c_{\mathfrak{u}_N} = -1$, ce qui donne la formule annoncée pour $d[K]_t$.

Définition 1.2.12. Soit $U \in U(N)$. Le mouvement brownien sur $U(N)$ issu de U est l'unique processus $(U_t)_{t \geq 0}$ solution forte de l'équation différentielle stochastique matricielle suivante :

$$dU_t = U_t \circ dK_t = dK_t U_t + \frac{c_{\mathfrak{u}_N}}{2} U_t dt, \quad (1.38)$$

avec pour condition initiale $U_0 = U$.

Le passage en coordonnées nous permet de voir K comme une transformation linéaire (complexe) d'un mouvement brownien dans \mathbb{R}^d ; la bonne nouvelle est alors que les outils habituels du calcul d'Itô multidimensionnel s'appliquent au processus K . Si l'on interprète (1.38) comme un système d'EDS par rapport à des mouvements browniens réels, on se convainc sans peine que cette équation est bien définie ; on peut également montrer que pour tout t , U_t appartient bien à $U(N)$: en effet, si l'on applique la formule d'Itô multidimensionnelle à $U_t^* U_t$ on obtient $d(U_t U_t^*) = 0$. On en déduit que $U_t U_t^* = I_N$ pour tout t et donc $U_t \in U(N)$.

Exemple. Dans le cas du groupe $U(1)$, son algèbre de Lie est $\mathfrak{u}_1 = i\mathbb{R}$, que l'on peut munir du produit scalaire usuel. Ainsi, un mouvement brownien sur \mathfrak{u}_1 est donné par $(iB_t)_t$, où (B_t) est un mouvement brownien standard sur \mathbb{R} , et donc le mouvement brownien sur $U(1)$ est la solution forte de l'équation différentielle stochastique

$$dU_t = iU_t dB_t - \frac{1}{2} U_t dt.$$

On peut vérifier sans peine que $(e^{iB_t})_t$ convient ; on remarque que le mouvement brownien sur le cercle correspond bien à un "enroulement" du mouvement brownien réel, au sens où son angle, modulo 2π , est égal à (B_t) .

L'exemple ci-dessus est important, à notre avis, pour deux raisons. Premièrement, il apparaît, de façon renormalisée, dans une factorisation du mouvement brownien sur $U(N)$, que l'on trouve par exemple dans [Dah17].

Lemme 1.2.5. Soit (U_t) un mouvement brownien sur $U(N)$. Alors $(U_t)_t$ a la même loi que $(e^{iB_t/N} S_t)_t$ où (B_t) est un mouvement brownien standard et (S_t) est un mouvement brownien sur $SU(N)$ ²³ indépendant de (B_t) .

²³Ce processus est construit exactement comme le mouvement brownien sur \mathfrak{su}_N : c'est la solution forte de l'EDS

$$dS_t = dK_t^{\mathfrak{su}_N} S_t + \frac{c_{\mathfrak{su}_N}}{2} S_t dt,$$

où $c_{\mathfrak{su}_N} = -1 + \frac{1}{N^2}$ est associé à l'élément de Casimir $C_{\mathfrak{su}_N}$ et $K^{\mathfrak{su}_N}$ est un mouvement brownien sur \mathfrak{su}_N .

Deuxièmement, une approche – que nous n'utiliserons pas dans cette thèse mais qu'il apparaît intéressant de mentionner – pour analyser le mouvement brownien sur $U(N)$ est de regarder le processus de ses valeurs propres ; à l'instar du mouvement brownien de Dyson qui caractérise les valeurs propres du mouvement brownien hermitien on peut décrire le processus $(e^{i\theta_1(t)}, \dots, e^{i\theta_N(t)})$ des valeurs propres de (U_t) comme un système de particules sur $U(1)$, et plus particulièrement une famille de mouvements browniens sur $U(1)$ conditionnés à ne pas se rencontrer.

Jusqu'à présent on a introduit le mouvement brownien unitaire en utilisant une représentation matricielle de celui-ci pour se ramener au calcul stochastique multidimensionnel réel. Nous allons désormais l'étudier de façon plus intrinsèque, à travers le prisme des probabilités non-commutatives introduites aux paragraphes précédents. On se donne un espace de probabilités $(\Omega, \mathcal{F}, \mathbb{P})$ et on considère le sous-ensemble de $\mathcal{M}_N(\mathbb{C}) \otimes L^\infty(\Omega, \mathbb{P})$ constitué des matrices aléatoires unitaires. C'est un sous-espace de l'espace des matrices aléatoires sur Ω ; il est muni de la structure de groupe induite par $U(N)$. En particulier, dans ce cadre on peut calculer les $*$ -moments du mouvement brownien unitaire comme suit :

$$\tau(P(U_t, U_t^*)) = \mathbb{E}[\text{tr}(P(U_t, U_t^*))] = \int_{U(N)} \text{tr}(P(U, U^{-1})) p_t(U) dU, \quad (1.39)$$

où p_t est le noyau de la chaleur sur $U(N)$ que nous étudierons dans la section 1.3.5.

Lorsque N tend vers l'infini, ces moments admettent une limite finie qui déterminent un processus appelé *mouvement brownien unitaire libre*, ou *mouvement brownien multiplicatif libre* ; cela explique en partie l'utilisation de la trace normalisée dans le calcul des moments, et en ce sens la limite peut s'interpréter comme une limite d'échelle. Ce processus limite a été défini par Biane [Bia97] dans le langage des probabilités libres, mais des résultats similaires ont été obtenus par Singer [Sin95] et Rains [Rai97] de façon concomitante, le premier dans un contexte proche de cette thèse, à savoir l'étude asymptotique de la mesure de Yang–Mills planeaire, et le second dans un contexte d'étude asymptotique de mouvements browniens sur des groupes matriciels.

Théorème 1.2.6. *Lorsque N tend vers l'infini, le mouvement brownien $(U_t)_{t \geq 0}$ sur $U(N)$ converge presque-sûrement en distribution non-commutative vers un processus $(u_t)_{t \geq 0}$ unitaire tel que ses incréments multiplicatifs $u_t u_s^*$ soient stationnaires et libres²⁴, et tel que pour tout $n \in \mathbb{N}^*$,*

$$\tau(u_t^n) = \tau(u_t^{-n}) = e^{-\frac{nt}{2}} \sum_{k=0}^{n-1} \frac{(-t)^k}{k!} n^{k-1} \binom{n}{k+1}. \quad (1.40)$$

Cette convergence se traduit de la façon suivante : pour tout $k \in \mathbb{N}$, $(t_1, \dots, t_k) \in (\mathbb{R}_+)^k$, et $P \in \mathbb{C}\langle X_1, \dots, X_{2k} \rangle$,

$$\lim_{N \rightarrow \infty} \text{tr}(P(U_{t_1}^{(N)}, (U_{t_1}^{(N)})^*, \dots, U_{t_k}^{(N)}, (U_{t_k}^{(N)})^*)) = \tau(P(u_{t_1}, u_{t_1}^*, \dots, u_{t_k}, u_{t_k}^*)) \text{ p.s.} \quad (1.41)$$

Une version “forte” de ce théorème a plus récemment été démontrée par Collins, Dahlqvist et Kemp [CDK18] : ils ont en effet montré que la convergence de la formule 1.41 reste valable si l'on remplace la trace par la norme d'opérateurs, ce qui permet de remplacer la convergence du Thm. 1.2.6 en une convergence forte au sens de la Déf. 1.2.10.

²⁴Il s'agit là de l'analogie, en probabilités libres, d'un processus de Lévy.

1.3 THÉORIE DES REPRÉSENTATIONS

La théorie des représentations consiste à identifier un espace relativement général à un sous-espace de $\text{End}(V)$ pour un espace vectoriel V donné, et ainsi bénéficier de la théorie de la réduction des endomorphismes – entre autres – afin de décomposer les objets étudiés dans une base bien choisie. En d’autres termes, on peut voir la théorie des représentations comme une version algébrique de la théorie spectrale, ou de la théorie de Fourier.

Bien que nous la nommions “théorie” au singulier, elle est multiple et foisonnante, et surtout se rapporte à des objets qui peuvent être de natures fondamentalement différentes, et de structures plus ou moins riches ; nous allons cependant tenter d’en donner une définition informelle la plus générale possible avant d’en étudier des aspects spécifiques : étant donné une catégorie \mathcal{C} et un objet A de celle-ci, une représentation de A dans un espace vectoriel V est la donnée d’un morphisme de A vers $\text{End}(V)$ ou $\text{GL}(V)$ ²⁵, ce qui se traduit de manière synthétique par :

$$(\rho, V) \text{ représentation de } A \in \text{Obj}(\mathcal{C}) \Leftrightarrow (V \in \text{Obj}(\text{Vec}), \rho \in \text{Hom}_{\mathcal{C}}(A, \text{GL}(V))).$$

Exemples.

- (i) Dans la catégorie Grp des groupes, une représentation d’un groupe G est la donnée d’un espace vectoriel V et d’un morphisme de groupes $G \rightarrow \text{GL}(V)$.
- (ii) Dans la catégorie TopGrp des groupes topologiques, une représentation de G est donnée par un morphisme de groupes topologiques $G \rightarrow \text{GL}(V)$, soit un morphisme de groupes qui soit continu.
- (iii) Dans la catégorie $k\text{-Alg}$ des k -algèbres associatives, où k est un corps, une représentation de degré n d’une algèbre A est la donnée d’un morphisme de k -algèbres $A \rightarrow M_n(k)$.
- (iv) Dans la catégorie $C^*\text{Alg}$ des C^* -algèbres, une représentation d’une C^* -algèbre A est la donnée d’un espace de Hilbert H et d’un morphisme de C^* -algèbres $A \rightarrow \mathcal{B}(H)$.

Notre objectif ici est de développer la théorie de Fourier, ou tout du moins son analogue, dans le groupe $U(N)$. Nous allons donc, dans un premier temps, énoncer quelques résultats généraux sur les représentations des groupes, en particulier des groupes compacts, qui vont nous mener au théorème fondamental de Peter-Weyl et à la décomposition spectrale du Laplacien sur un groupe compact. Dans un second temps, nous verrons comment ces résultats s’appliquent au cas du groupe unitaire.

1.3.1 De la mesure de Haar sur un groupe

Aussi sûr que la théorie de Fourier et l’analyse harmonique “classiques” présupposent une connaissance de la théorie de l’intégration sur \mathbb{R}^n , la théorie des représentations, notamment appliquée à l’analyse harmonique non-commutative, nécessite une connaissance de la théorie de l’intégration sur d’autres groupes. Loin de nous l’idée de détailler cette théorie, mais il semble indispensable qu’avant

²⁵Cela présuppose, bien entendu, que $\text{End}(V)$ ou $\text{GL}(V)$ appartienne à la catégorie \mathcal{C} , et cela ne fonctionne donc pas pour toute catégorie.

d'aller plus loin nous introduisons au moins la notion de *mesure de Haar* et quelques-unes de ses propriétés. Le chapitre 5 de [Faro8] constitue une introduction élémentaire au sujet, et un développement plus complet pourra être trouvé au chapitre 15 de [God03], voire pour les plus hardis au chapitre VII de [Bou63] – la monographie [Wei40] pourra aussi être consultée avec profit.

Definition 1.3.1. Soit G un groupe topologique localement compact. Une mesure de Radon positive μ sur G est *invariante à gauche*²⁶ si elle vérifie

$$\int_G f(gx) d\mu(x) = \int_G f(x) d\mu(x), \quad \forall g \in G, \forall f \in \mathcal{C}_c(G),$$

et *invariante à droite* si elle vérifie

$$\int_G f(xg) d\mu(x) = \int_G f(x) d\mu(x), \quad \forall g \in G, \forall f \in \mathcal{C}_c(G).$$

Une mesure invariante à gauche (resp. à droite) non nulle, si elle existe, est appelée *mesure de Haar* à gauche (resp. à droite).

L'existence de cette mesure a été démontrée pour un groupe localement compact métrisable et séparable par Haar [Haa33], puis par Weil [Wei40] en toute généralité. Lorsqu'un groupe admet une mesure de Haar finie, elle est unique à un facteur près ; c'est un résultat que l'on doit à von Neumann [vN36] et Kakutani [Kak38]. En particulier pour tout groupe compact G il existe une mesure de Haar à gauche μ sur G telle que $\mu(G) = 1$: c'est de cette mesure qu'il s'agira lorsque nous parlerons de *mesure de Haar* par la suite, et on la notera dg .

Proposition 1.3.1. Soit G un groupe localement compact.

(i) Si μ est une mesure de Haar à gauche sur G et g est un élément de G , la forme linéaire

$$\begin{cases} \mathcal{C}_c(G) & \rightarrow \mathbb{R} \\ f & \mapsto \int_G f(gxg^{-1}) d\mu(x) \end{cases}$$

définit une mesure de Haar à gauche. Il existe donc $\Delta(g) > 0$ tel que

$$\int_G f(gxg^{-1}) d\mu(x) = \Delta(g) \int_G f(x) d\mu(x),$$

et on appelle module l'application $\Delta : G \rightarrow \mathbb{R}_+$.

(ii) Le module Δ est un morphisme continu de groupes de G vers \mathbb{R}_+ .

(iii) Si le module de G est constant et égal à 1, on dit que G est unimodulaire. Dans ce cas, toute mesure de Haar à gauche est une mesure de Haar à droite, et elle est invariante par passage à l'inverse.

(iv) Si G est compact ou abélien, alors il est unimodulaire.

²⁶On peut faire le rapprochement avec la définition 1.1.7 lorsque G est un groupe de Lie, et en utilisant la correspondance entre mesures et 1-formes différentielles.

1.3.2 Définitions et constructions générales

Définition 1.3.2. Soit G un groupe (resp. un groupe topologique, un groupe de Lie), et $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . Une *représentation* de G est la donnée d'un \mathbb{K} -espace vectoriel V et d'un morphisme de groupes $\rho : G \rightarrow \text{GL}(V)$.

Par défaut, dans ce qui suit, on supposera toute représentation continue si G est un groupe topologique et lisse si G est un groupe de Lie. Notons qu'on peut également définir une représentation de G par une action de groupe $(g, v) \mapsto g \cdot v$ sur un espace vectoriel V . Il arrive alors que V soit appelé un G -module car l'action de g sur V ressemble formellement à la multiplication externe d'un anneau A sur un A -module. L'expression " G -module" est toutefois un abus de langage, ou tout au moins une métonymie : en réalité on peut effectivement munir V d'une structure de module sur l'algèbre du groupe G . Cette correspondance mérite d'être développée.

Définition 1.3.3. Soit G un groupe, et $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . Son *algèbre de groupe* $\mathbb{K}[G]$ est la \mathbb{K} -algèbre constituée des combinaisons linéaires d'éléments de G à coefficients dans \mathbb{K} .

Notons que lorsque G est engendré par des générateurs $\{g_1, \dots, g_k\}$, on peut voir $\mathbb{K}[G]$ comme l'image de l'algèbre $\mathbb{K}\langle X_1, \dots, X_k \rangle$ des polynômes non-commutatifs à k variables et à coefficients dans \mathbb{K} par la spécialisation $(X_1, \dots, X_k) \mapsto (g_1, \dots, g_k)$. La terminologie " G -module" peut alors s'expliquer par la proposition suivante, dont la démonstration constitue une simple vérification.

Proposition 1.3.2. Soit (ρ, V) une représentation de G . Alors V peut être muni d'une structure de $\mathbb{K}(G)$ -module en étendant \mathbb{K} -linéairement l'action de G sur V définie par

$$g \cdot v = \rho(g)v, \quad \forall g \in G, \quad \forall v \in V.$$

Les deux définitions d'une représentation que nous avons introduites possèdent chacune ses avantages : la première rend explicite le morphisme ρ tandis que la seconde est plus concise et permet d'alléger les notations, particulièrement lorsqu'on raisonne de manière diagrammatique. Dans ce qui suit, nous les utiliserons indistinctement selon ce qui semblera le plus pertinent.

Passons à présent au produit tensoriel de représentations, qui permet de considérer l'action simultanée du groupe sur plusieurs espaces vectoriels comme une action sur leur produit tensoriel.

Définition 1.3.4. Soit (ρ_1, V_1) et (ρ_2, V_2) deux représentations de G . Le *produit tensoriel* de ces représentations est défini comme la représentation $\rho_1 \otimes \rho_2$ sur l'espace $V_1 \otimes V_2$ vérifiant les deux propriétés (équivalentes) suivantes :

$$g \cdot (v_1 \otimes v_2) = (g \cdot v_1) \otimes (g \cdot v_2), \quad \forall v_1 \otimes v_2 \in V_1 \otimes V_2, \quad \forall g \in G;$$

$$(\rho_1 \otimes \rho_2)(g) = \rho_1(g) \otimes \rho_2(g), \quad \forall g \in G.$$

Note : on peut étendre cette définition au produit tensoriel d'un nombre fini de représentations $\rho_1 \otimes \dots \otimes \rho_n$, et ce, de manière associative. Cela permet de voir l'algèbre tensorielle

$$T(V) = \bigoplus_{n=0}^{\infty} V^{\otimes n}$$

comme une représentation de G . Notons que G agit sur $T(V)$ par automorphismes d'algèbre graduée²⁷.

Dans les applications de la théorie des représentations aux probabilités non-commutatives, le produit tensoriel de représentations occupe une place de choix, comme l'illustrent la dualité de Schur–Weyl ou bien la décomposition d'un produit de caractères – nous aurons le loisir de présenter ces résultats au chapitre 3. Aussi semble-t-il utile de rappeler quelques propriétés plus ou moins connues du produit tensoriel d'espaces vectoriels ; pour plus de détails, on renvoie à [Proo7, Chap.5].

Pour commencer, l'application

$$\Phi_{U,V} : \begin{cases} V \otimes U^* & \xrightarrow{\cong} & \text{Hom}(U, V) \\ v \otimes \varphi & \mapsto & u \mapsto \langle \varphi, u \rangle v \end{cases} \quad (1.42)$$

définit un isomorphisme canonique. Toute application $A \in \text{Hom}(U, V)$, représentée matriciellement par (a_{ij}) dans des bases $\{e_1, \dots, e_n\}$ et $\{f_1, \dots, f_n\}$ respectivement de U et V , correspond au tenseur $\sum_{i,j} a_{ij} f_i \otimes e^j$, en notant $\{e^1, \dots, e^n\}$ la base duale sur U^* . En effet, pour tout k on constate que

$$\Phi_{U,V}(f_i \otimes e^j)(e_k) = \langle e^j, e_k \rangle f_i = \delta_{jk} f_i$$

et comme pour tout k on a $Ae_k = \sum_i a_{ik} f_i$, on conclut par linéarité.

Soit (ρ_1, U) et (ρ_2, V) deux représentations de G . L'action de G sur $\text{Hom}(U, V)$ donnée par $(g \cdot A)(u) = g \cdot A(g^{-1} \cdot u)$ se traduit par le diagramme

$$\begin{array}{ccc} U & \xrightarrow{A} & V \\ g \downarrow & & \downarrow g \\ U & \xrightarrow{g \cdot A} & V \end{array}$$

Un cas particulier de l'isomorphisme Φ est celui où $V = \mathbb{C}$, et ρ_2 est la représentation triviale, i.e. $\rho_2(g)z = z, \forall z \in \mathbb{C}$; la représentation définie par le procédé précédent n'est autre que la représentation contragrédiente ρ^\vee sur $U^* = \text{Hom}(U, \mathbb{C})$. Aussi, l'action de G sur $\text{Hom}(U, V)$ peut se réinterpréter en termes de produit tensoriel comme

$$g \cdot A = (\rho_2 \otimes \rho_1^\vee)(g)(\Phi_{U,\mathbb{C}}^{-1}(A)).$$

Un autre cas particulier d'isomorphisme est celui où $U = V$, et qui permet de définir la trace d'un tenseur.

Définition 1.3.5. Soit $u \otimes \varphi$ un tenseur de $U \otimes U^*$. Sa *trace* est le scalaire défini par

$$\text{Tr}_U(u \otimes \varphi) = \langle \varphi, u \rangle.$$

Il n'est pas difficile de vérifier que cette définition coïncide avec celle de la trace d'un endomorphisme représenté matriciellement :

$$\text{Tr}_U(u \otimes \varphi) = \text{Tr}(\Phi_U(u \otimes \varphi)).$$

La trace satisfait la propriété de factorisation suivante, que l'on peut rapprocher de la factorisation qui caractérise l'indépendance tensorielle en probabilités non-commutatives.

²⁷Cette action s'opère diagonalement sur les k -tenseurs : $g \cdot (v_1 \otimes \dots \otimes v_k) = (g \cdot v_1) \otimes \dots \otimes (g \cdot v_k)$.

Proposition 1.3.3. Soit $A \in \text{End}(U)$, $B \in \text{End}(V)$. La trace de $A \otimes B \in \text{End}(U \otimes V)$ vérifie

$$\text{Tr}(A \otimes B) = \text{Tr}(A)\text{Tr}(B).$$

De la même façon qu'on peut comprendre un espace vectoriel plus simplement lorsqu'on le décompose en sous-espaces vectoriel, nous allons voir qu'il est souvent utile de décomposer une représentation en d'autres représentations plus simples.

Définition 1.3.6. Soit (ρ, V) une représentation d'un groupe G . Si W est un sous-espace de V stable par $\rho(g)$ pour tout $g \in G$, on dit que (ρ, W) est une *sous-représentation* de (ρ, V) , ou encore que W est un *sous- G -module* de V .

Il apparaît important de noter que l'on ne parle pas de *restriction* pour autant. En effet, la représentation *restreinte* est la restriction d'une représentation d'un groupe G à la représentation d'un sous-groupe H .

Une représentation ρ est dite *irréductible* si elle n'admet pas de sous-représentation non-triviale. On peut voir les représentations irréductibles comme les atomes de la théorie des représentations, et un des principaux enjeux de cette branche est de caractériser les représentations irréductibles, puis de déterminer des règles de décomposition de certaines représentations en représentations irréductibles d'autre part. Commençons par introduire la notion d'entrelacement, qui comme on va le voir par la suite, permet d'établir des isomorphismes entre deux représentations irréductibles.

Définition 1.3.7. Si (ρ_1, V_1) et (ρ_2, V_2) sont deux représentations de G et $A \in \text{Hom}(V_1, V_2)$ vérifie

$$A\rho_1(g) = \rho_2(g)A, \quad \forall g \in G, \tag{1.43}$$

alors on dit que A *entrelace* ρ_1 et ρ_2 .

L'entrelacement peut se traduire par le diagramme commutatif suivant :

$$\begin{array}{ccc} V_1 & \xrightarrow{A} & V_2 \\ g \downarrow & & \downarrow g \\ V_1 & \xrightarrow{A} & V_2 \end{array}$$

et on dit que A est une application G -équivariante. L'ensemble des applications G -équivariantes de V_1 vers V_2 est noté $\text{Hom}_G(V_1, V_2)$. C'est l'ensemble des points fixes de l'action de G sur $\text{Hom}(V_1, V_2)$. En effet, soit $A \in \text{Hom}_G(V_1, V_2)$. On a

$$(g \cdot A)(u) = g \cdot A(g^{-1} \cdot u) = A(u)$$

si et seulement si

$$g \cdot A(v) = A(g \cdot v),$$

en posant $v = g^{-1} \cdot u$.

Le lemme suivant est fondamental, à tel point qu'il nous semble utile d'en donner la (courte) démonstration.

Lemme 1.3.4 (Schur).

- (i) Soit (ρ_1, V_1) et (ρ_2, V_2) deux représentations irréductibles de dimension finie d'un groupe G . Si $A \in \text{Hom}(V_1, V_2)$ est G -équivariante et non nulle, alors c'est un isomorphisme.
- (ii) Soit (ρ, V) une représentation complexe irréductible de G , et $A \in \text{End}(V)$ une application qui commute avec ρ . Alors il existe $\lambda \in \mathbb{C}$ tel que

$$A = \lambda I_V.$$

Preuve. (i) Le sous-espace $\ker A$ est stable. En effet, soit $x \in \ker A$, $g \in G$, alors $\rho_2(g)Ax = 0 = A\rho_1(g)x$ donc $\rho_1(g)x \in \ker A$. Or ρ_1 est irréductible donc soit $A = 0$ soit A est injective. Par un raisonnement similaire on montre la surjectivité de A .

(ii) Par le théorème fondamental de l'algèbre, il existe $\lambda \in \mathbb{C}$ tel que $A - \lambda I_V$ ne soit pas inversible. En appliquant le point (i), on en déduit que $A - \lambda I_V = 0$. \square

Le lemme de Schur permet notamment de définir des classes d'équivalence sur les représentations irréductibles, en considérant que deux représentations irréductibles sont équivalentes si et seulement si elles sont entrelacées par une application linéaire non nulle. On note \widehat{G} l'ensemble des classes d'équivalence de représentations irréductibles.

Définition 1.3.8. Soit (ρ, V) une représentation complexe irréductible de dimension finie d'un groupe G . Son *caractère* est la fonction

$$\chi_\rho : \begin{cases} G & \rightarrow \mathbb{C} \\ g & \mapsto \text{Tr}(\rho(g)) \end{cases} . \quad (1.44)$$

Un corollaire immédiat du lemme de Schur est que deux représentations irréductibles équivalentes possèdent le même caractère, donc que \widehat{G} est en bijection avec l'ensemble des caractères irréductibles. Aussi emploierons-nous la même notation pour cet ensemble. Un autre corollaire immédiat du lemme de Schur, utilisant cette fois-ci le point (ii), est que les représentations irréductibles d'un groupe abélien sont toutes de dimension 1, et qu'elles sont égales à leur caractère. Lorsque G est abélien et localement compact, \widehat{G} est lui-même un groupe abélien localement compact, appelé *groupe dual*. Ce cadre permet une généralisation du développement en série de Fourier des fonctions périodiques, vues comme des fonction de $G = \text{U}(1)$ dont le groupe dual est $\widehat{G} = \mathbb{Z}$. Nous développerons cet exemple à la lumière du théorème de Peter–Weyl dans la section 1.3.3.

Définition 1.3.9. Soit (ρ, V) une représentation complexe de dimension finie d'un groupe compact G . La représentation ρ est dite :

- (i) *réelle* si V est muni d'une structure réelle invariante, *i.e.* une application antilinéaire G -équivariante $c : V \rightarrow V$ telle que $c^2 = \text{Id}_V$;
- (ii) *quaternionique* si V est muni d'une structure quaternionique invariante, *i.e.* une application antilinéaire G -équivariante $j : V \rightarrow V$ telle que $j^2 = -\text{Id}_V$;

(iii) *complexe* si elle n'est ni réelle, ni quaternionique.

La classification ci-dessus admet une caractérisation intéressante, donnée par la proposition suivante.

Proposition 1.3.5 ([BtD95], Prop.6.8). *Soit (ρ, V) une représentation irréductible d'un groupe compact G , de caractère χ . Alors*

$$\int_G \chi(g^2) dg = \begin{cases} 1 & \text{si } \rho \text{ est une représentation réelle,} \\ 0 & \text{si } \rho \text{ est une représentation complexe,} \\ -1 & \text{si } \rho \text{ est une représentation quaternionique.} \end{cases} \quad (1.45)$$

Le membre de gauche de l'équation (1.45) est appelé *indicateur de Frobenius–Schur* de λ et se note parfois ι_λ . Il jouera un rôle déterminant dans le chapitre 2 pour calculer la limite de la fonction de partition de Yang–Mills sur une surface non-orientable.

Nous terminons cette section avec une brève revue des représentations d'algèbres de Lie.

Définition 1.3.10. Soit \mathfrak{g} une algèbre de Lie. une *représentation* de \mathfrak{g} est la donnée d'un \mathbb{C} -espace vectoriel V et d'une application linéaire $\rho : \mathfrak{g} \rightarrow \text{End}(V)$ qui est un morphisme d'algèbres de Lie, i.e.

$$\rho([X, Y]) = [\rho(X), \rho(Y)], \quad \forall (X, Y) \in \mathfrak{g}^2. \quad (1.46)$$

Exemple. Si \mathfrak{g} est une algèbre de Lie, on définit sa *représentation adjointe* ad sur elle-même par

$$\text{ad}_X(Y) = [X, Y], \quad \forall (X, Y) \in \mathfrak{g}^2. \quad (1.47)$$

La théorie des représentations des algèbres de Lie est similaire à celle des représentations des groupes de Lie, notamment en raison du troisième théorème de Lie (cf. [Sero6, II.V.8, Thm.3]) qui établit que pour toute algèbre de Lie \mathfrak{g} de dimension finie il existe un groupe de Lie G connexe et simplement connexe tel que \mathfrak{g} soit l'algèbre de Lie de G . Forts de cette information, nous pouvons alors transporter les représentations des groupes de Lie vers celles des algèbres, *via* un procédé de dérivation en l'élément neutre.

Proposition 1.3.6. *Soit G un groupe de Lie linéaire, \mathfrak{g} son algèbre de Lie, et (ρ, V) une représentation continue de G dans un espace vectoriel de dimension finie. On pose*

$$d\rho : \mathfrak{g} \rightarrow V, \quad X \mapsto \frac{d}{dt} \rho(\exp tX)|_{t=0}. \quad (1.48)$$

Cela donne une représentation de \mathfrak{g} dans V , appelée représentation dérivée de ρ .

La représentation adjointe Ad de G sur \mathfrak{g} , définie par l'action adjointe (1.16), admet pour représentation dérivée la représentation adjointe ad de \mathfrak{g} dans elle-même :

$$\text{ad}_X(Y) = \frac{d}{dt} \text{Ad}(\exp(tX))|_{t=0} Y = [X, Y]. \quad (1.49)$$

Ces deux représentations sont reliées par l'application exponentielle, puisqu'on a la relation

$$\text{Ad}_{\exp(X)} = \exp(\text{ad}_X), \quad \forall X \in \mathfrak{g}. \quad (1.50)$$

En réalité on a même un résultat encore plus fort, qui découle notamment du troisième théorème de Lie.

Theorem 1.3.7 ([Sero6], II.V.8, Thm. 2). *La catégorie des groupes de Lie complexes connexes et simplement connexes est équivalente à la catégorie des algèbres de Lie complexes de dimension finie.*

Cette équivalence est fondamentale, dans le sens où les représentations d'algèbres de Lie (qui sont en général plus faciles à étudier) déterminent exactement les représentations de groupes de Lie lorsque ces derniers sont connexes et simplement connexes. Nous verrons que l'on peut également déduire les représentations d'autres groupes, non nécessairement connexes ou simplement connexes, à partir des représentations de leur algèbre de Lie : nous traiterons en détail le cas de $U(N)$ dans le paragraphe 1.3.5.

1.3.3 Représentations des groupes compacts

Lorsque le groupe G est compact, on peut apparenter la décomposition de ses représentations à la réduction d'endomorphismes, à l'aide de représentations unitaires. On va également voir qu'il est possible de définir un opérateur qui commute avec toutes les représentations irréductibles, et qui jouera le rôle du laplacien.

Définition 1.3.II. Soit (ρ, V) une représentation de G . Supposons V muni d'une structure d'espace de Hilbert. La représentation ρ est dite *unitaire* si elle est à valeurs dans $U(V) = \{U \in \text{End}(V) : U^*U = UU^* = \text{Id}_V\}$.

Comme la proposition suivante le montre, l'utilisation de représentations unitaires permet de décomposer toute représentation d'un groupe compact en sous-représentations irréductibles.

Proposition 1.3.8. *Soit (ρ, V) une représentation de dimension finie d'un groupe compact G .*

- (i) *Il existe un produit scalaire $\langle \cdot, \cdot \rangle$ sur V pour lequel ρ est une représentation unitaire.*
- (ii) *Tout sous-espace de V stable par ρ admet un supplémentaire stable.*
- (iii) *L'espace V peut être décomposé en somme directe de sous-espaces stables irréductibles.*

À l'aide de cette proposition et du lemme de Schur, on obtient les *relations d'orthogonalité de Schur*.

Théorème 1.3.9. *Soit (ρ, V) une représentation unitaire irréductible de G . Alors*

$$\int_G \langle \rho(g)u, v \rangle \overline{\langle \rho(g)u', v' \rangle} dg = \frac{1}{d_\rho} \langle u, u' \rangle \langle v, v' \rangle, \quad \forall (u, v, u', v') \in V^4. \quad (\text{I.51})$$

Soit (ρ_1, V_1) et (ρ_2, V_2) deux représentations irréductibles de G non équivalentes. Alors

$$\int_G \langle \rho_1(g)u, v \rangle \overline{\langle \rho_2(g)u', v' \rangle} dg = 0, \quad \forall (u, v) \in V_1^2, \quad \forall (u', v') \in V_2^2. \quad (\text{I.52})$$

Ces relations sont un pas vers le théorème central de la théorie des représentations des groupes compacts, et qui a posé les jalons de l'analyse harmonique non-commutative²⁸.

²⁸En particulier, la théorie des fonctions sphériques sur une paire de Gelfand s'apparente à la théorie qui découle du théorème de Peter–Weyl, et permet par exemple de traiter le cas d'un groupe localement compact. On pourra consulter [Far80] pour un aperçu de cette théorie.

Théorème 1.3.10 (Peter–Weyl). *Soit G un groupe compact, \widehat{G} l'ensemble des classes d'équivalences de ses représentations unitaires irréductibles.*

(i) *Pour tout $\lambda \in \widehat{G}$ et $(\rho_\lambda, V_\lambda)$ une représentation de classe λ , on note \mathcal{M}_λ l'espace vectoriel engendré par les fonctions²⁹ $(\rho_{ij} : g \mapsto \langle \rho_\lambda(g)e_i, e_j \rangle)_{i,j}$, (e_i) étant une base de V_λ . Alors*

$$L^2(G) = \widehat{\bigoplus_{\lambda \in \widehat{G}} \mathcal{M}_\lambda}, \quad (1.53)$$

où ' $\widehat{}$ ' désigne la fermeture dans $L^2(G)$.

(ii) *Une fonction $f \in L^2(G)$ est dite centrale si pour tout $g, h \in G$, $f(hgh^{-1}) = f(g)$. On note $Z^2(G)$ le sous-espace de $L^2(G)$ constitué des fonctions centrales³⁰. Alors l'ensemble des caractères χ_λ , où λ parcourt \widehat{G} , est une base hilbertienne de $Z^2(G)$.*

Pour comprendre le fonctionnement de ce théorème, illustrons-le dans le cas de $U(1)$. L'isomorphisme de groupes topologiques

$$\begin{cases} U(1) & \rightarrow \mathbb{R}/2\pi\mathbb{Z} \\ z = e^{i\theta} & \mapsto \theta \end{cases} \quad (1.54)$$

permet d'établir la correspondance biunivoque entre les fonctions centrales³¹ sur $U(1)$ et les fonctions 2π -périodiques :

$$f \in Z(U(1)) \leftrightarrow f^0 : \begin{cases} \mathbb{R}/2\pi\mathbb{Z} & \rightarrow \mathbb{C} \\ \theta & \mapsto f(e^{i\theta}) \end{cases}$$

D'après le théorème de Peter–Weyl, de telles fonctions sont donc décomposables dans la base hilbertienne des caractères irréductibles de $U(1)$. Or, comme il s'agit ici d'un groupe abélien, ses représentations irréductibles sont toutes de dimension 1 et égales à leur caractère. On peut montrer que les caractères en question sont les fonctions $(\chi_n)_{n \in \mathbb{Z}}$ définies par

$$\chi_n : \begin{cases} U(1) & \rightarrow \mathbb{C} \\ x & \mapsto x^n \end{cases}.$$

Par conséquent, toute fonction $f \in Z(U(1))$ admet la décomposition

$$f(z) = \sum_{n \in \mathbb{Z}} c_n(f) z^n, \quad (1.55)$$

²⁹De telles fonctions sont appelées *coefficients matriciels*, car elles représentent les coefficients de la représentation $\rho_\lambda(g)$ vue comme une matrice dans la base (e_i) . Notons que l'on peut réinterpréter la relation d'orthogonalité de Schur (1.51) comme suit :

$$\int_G \rho_{ij}(g) \overline{\rho_{kl}(g)} dg = \frac{1}{d_\rho} \delta_{ik} \delta_{jl}.$$

³⁰Le Z est là pour signifier que c'est précisément le *centre* (*Zentrum* en allemand) de l'algèbre $L^2(G)$ munie du produit de convolution, ce qui se vérifie immédiatement.

³¹Comme $U(1)$ est abélien, toute fonction $f: U(1) \rightarrow \mathbb{C}$ est centrale.

où $(c_n(f))_{n \in \mathbb{Z}}$ est une famille de nombres complexes telle que la série entière du membre de droite soit de rayon de convergence au moins 1. En transposant l'équation (1.55) pour les fonctions 2π -périodiques, on obtient alors la formule classique du développement en série de Fourier

$$f^0(\theta) = \sum_{n \in \mathbb{Z}} c_n(f^0)(\cos(n\theta) + i \sin(n\theta)). \quad (1.56)$$

La décomposition de fonctions sur $U(1)$ en série de Fourier admet en réalité une généralisation, qui est un corollaire direct du théorème de Peter–Weyl et des relations d'orthogonalité de Schur : la *formule de Plancherel*. Si $\lambda \in \widehat{G}$ est une classe d'équivalence de représentations irréductibles de G et $(\rho_\lambda, V_\lambda)$ un représentant de cette classe, on associe à toute fonction f intégrable sur G son *coefficient de Fourier*

$$\widehat{f}(\lambda) = \int_G f(g) \rho_\lambda(g^{-1}) dg.$$

On a alors le théorème suivant.

Théorème 1.3.II (Théorème de Plancherel). *Soit G un groupe compact, $f \in L^2(G)$ une fonction de carré intégrable sur G . Alors f est égale³² à sa série de Fourier*

$$f(g) = \sum_{\lambda \in \widehat{G}} d_\lambda \text{Tr}(\widehat{f}(\lambda) \rho_\lambda(g)), \quad \forall g \in G. \quad (1.57)$$

Si l'on utilise la définition des coefficients de Fourier et les propriétés de ρ_λ on peut remarquer que l'équation (1.57) peut se récrire à l'aide des caractères irréductibles :

$$f(g) = \sum_{\lambda \in \widehat{G}} d_\lambda (f * \chi_\lambda)(g), \quad \forall g \in G. \quad (1.58)$$

Il est intéressant de mentionner que tous les résultats abordés dans cette section s'appliquent en particulier aux groupes finis, en notant que pour un tel groupe, une mesure de Haar est un multiple de la mesure de comptage.

Nous allons à présent nous tourner vers un second aspect de la théorie des représentations des groupes compacts : la résolution d'équations aux dérivées partielles. En effet, un des exemples canoniques de l'utilisation de l'analyse harmonique dans \mathbb{R}^n est la caractérisation du *noyau de la chaleur*, c'est-à-dire la solution fondamentale de l'équation de la chaleur $\frac{1}{2} \Delta_x f(x, t) = \frac{\partial}{\partial t} f(x, t)$. Celle-ci se fonde sur une étude spectrale de l'opérateur laplacien sur \mathbb{R}^n en utilisant la transformée de Fourier ; nous allons voir comment cela fonctionne en remplaçant \mathbb{R}^n par un groupe compact G muni de son algèbre de Lie \mathfrak{g} . On munit \mathfrak{g} d'une structure euclidienne, en choisissant un produit scalaire $\langle \cdot, \cdot \rangle$ invariant par la représentation adjointe de G :

$$\langle \text{Ad}(g)X, \text{Ad}(g)Y \rangle = \langle X, Y \rangle, \quad \forall g \in G, \quad \forall (X, Y) \in \mathfrak{g}^2.$$

Notons qu'à partir d'un produit scalaire $\langle \cdot | \cdot \rangle$ quelconque sur \mathfrak{g} , on peut construire un tel produit scalaire invariant en posant

$$\langle X, Y \rangle = \int_G (\text{Ad}(g)X | \text{Ad}(g)Y) dg.$$

³²Il s'agit là d'une égalité dans $L^2(G)$, c'est-à-dire que la série du terme de droite de (1.57) converge dans $L^2(G)$.

Définition 1.3.12. Soit $X = (X_1, \dots, X_d)$ une base orthonormée de \mathfrak{g} . On appelle *élément de Casimir* le tenseur

$$C_{\mathfrak{g}} = \sum_{i=1}^d X_i \otimes X_i.$$

Une propriété importante pour garantir la bonne définition de ce tenseur est qu'il ne dépend pas de la base orthonormée choisie. En effet, soit $Y = (Y_1, \dots, Y_d)$ une autre base orthonormée, posons $Q = (Q_{ij})_{1 \leq i, j \leq d}$ la matrice de changement de base de X vers Y , qui est une matrice orthogonale, c'est-à-dire que $(Q^{-1})_{ij} = Q_{ji}$ pour tout $(i, j) \in \{1, \dots, d\}$. On a

$$\begin{aligned} \sum_{j=1}^d Y_j \otimes Y_j &= \sum_{j=1}^d \sum_{i=1}^d Q_{ij} X_i \otimes Q_{ij} X_i \\ &= \sum_{i=1}^d \sum_{j=1}^d Q_{ij} Q_{ij} X_i \otimes X_i \\ &= \sum_{i=1}^d \left(\sum_{j=1}^d Q_{ij} (Q^{-1})_{ji} \right) X_i \otimes X_i \\ &= \sum_{i=1}^d X_i \otimes X_i, \end{aligned}$$

en utilisant l'orthogonalité de Q .

Étant donné une représentation (ρ, V) de \mathfrak{g} , on peut associer à l'élément de Casimir un opérateur $\Omega_{\rho} \in \text{End}(V)$, appelé *opérateur de Casimir* de la représentation ρ , défini par

$$\Omega_{\rho} = m \circ (\rho \otimes \rho)(C_{\mathfrak{g}}) = \sum_{i=1}^d \rho(X_i)^2, \quad (1.59)$$

où $m : x \otimes y \mapsto xy$ est l'opération de multiplication dans $\mathcal{M}_N(\mathbb{C})$. On peut vérifier que Ω_{ρ} commute à la représentation ρ , et si ρ est irréductible le lemme de Schur implique qu'il existe $c_{\rho} \in \mathbb{C}$ tel que

$$\Omega_{\rho} = -c_{\rho} \text{Id}_V. \quad (1.60)$$

Le signe '−' provient de la remarque suivante : si $\rho = d\pi$ où π est une représentation de G telle que ρ n'est pas la représentation triviale, alors en prenant un produit scalaire (\cdot, \cdot) sur V pour lequel π est unitaire, on a

$$d\pi(X)^* = -d\pi(X), \quad \forall X \in \mathfrak{g},$$

et pour tout $v \in V$ non nul,

$$(\Omega_{\rho} v, v) = - \sum_i \|\rho(X_i)v\|^2 < 0.$$

Ainsi, on obtient que $c_{\rho} > 0$.

L'élément de Casimir $C_{\mathfrak{g}}$ est en réalité intimement lié au laplacien sur le groupe G .

Définition 1.3.13. L'opérateur de Laplace–Beltrami (ou *laplacien*) sur G est l'opérateur défini par

$$\Delta_G : \begin{cases} \mathcal{C}^2(G) & \rightarrow \mathcal{C}^0(G) \\ f & \mapsto \Delta_G f : g \mapsto \sum_{i=1}^d \frac{d^2}{dt^2} \Big|_{t=0} f(g \exp(tX_i)) \end{cases}, \quad (1.61)$$

où (X_1, \dots, X_d) est une base orthonormée de \mathfrak{g} .

La dérivée de Lie, qui correspond à l'analogie, dans un groupe de Lie, de la dérivée directionnelle selon un vecteur tangent, définit une représentation ρ de \mathfrak{g} sur $\mathcal{C}^1(G)$:

$$\rho(X)f : g \mapsto \frac{d}{dt} \Big|_{t=0} f(g \exp(tX)), \quad (1.62)$$

et il vient naturellement que

$$\Delta_G f = \Omega_\rho f, \quad \forall f \in \mathcal{C}^2(G). \quad (1.63)$$

Par ailleurs, si $\lambda \in \widehat{G}$ est une classe d'équivalence de représentations irréductibles et $(\rho_\lambda, V_\lambda)$ est une représentation irréductible de classe λ , alors toute fonction représentative $f \in \mathcal{M}_\lambda$ peut s'écrire

$$f(g) = \text{Tr}(A_f \rho_\lambda(g)),$$

avec $A_f \in \text{End}(V)$. Il vient que

$$\Delta_G f(g) = \text{Tr}(\Omega_{\rho_\lambda} A_f \rho_\lambda(g)),$$

et d'après (1.60), en notant $c_\lambda = c_{\rho_\lambda}$, on en déduit que

$$\Delta_G f = -c_\lambda f. \quad (1.64)$$

La décomposition du laplacien en fonctions propres permet de décrire la décomposition du noyau de la chaleur sur G à l'aide des caractères irréductibles. Nous allons montrer cela dans la section 1.3.5, dans le cas où G est le groupe unitaire. Nous verrons également plus tard que l'élément de Casimir dont on s'est servi pour décrire le laplacien apparaît dans la construction du mouvement brownien sur le groupe – ce qui n'a rien d'étonnant lorsqu'on le considère comme un processus de Markov de générateur $\frac{1}{2}\Delta_G$ et dont le semi-groupe de convolution est justement donné par le noyau de la chaleur.

1.3.4 Diagrammes de Young et fonctions de Schur

Avant de décrire les représentations de $U(N)$ nous devons introduire quelques objets combinatoires que nous retrouverons dans nos calculs, et qui apparaissent dans de nombreux modèles de physique statistique³³.

Une *partition* $\lambda \vdash n$ d'un entier naturel n est un k -uplet d'entiers naturels $\lambda = (\lambda_1, \dots, \lambda_k)$, que l'on peut toujours supposer ordonné : $\lambda_1 \geq \dots \geq \lambda_k \geq 0$. Les coefficients de la partition sont appelés des *parts*. Un *diagramme de Young* est la représentation graphique d'une telle partition

³³On en trouvera une liste non exhaustive d'exemples dans [BP13] et [GP15].

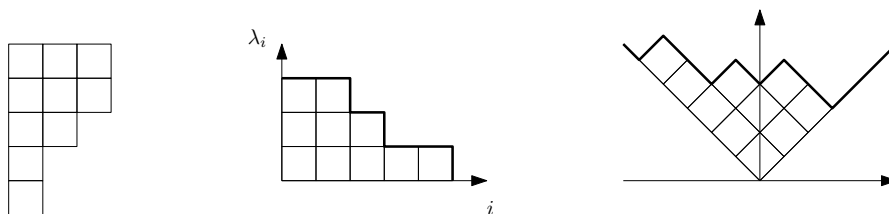


Figure 1.6: Le diagramme de Young associé à la partition $(3, 3, 2, 1, 1) \vdash 9$, en notation anglo-saxonne (à gauche), cartésienne (au milieu) et russe (à droite).

sous forme d'empilement de carrés (ou de boîtes). En notation anglo-saxonne, c'est-à-dire orientée dans le sens lexicographique, un diagramme de Young se lit de gauche à droite et de haut en bas – à la manière des coordonnées d'une matrice. La i -ème ligne du diagramme associé à λ contient exactement λ_i cases, et toutes les lignes sont alignées à gauche. Il existe d'autres conventions que l'on retrouve fréquemment : la notation française, qui correspond à une symétrie de la notation anglo-saxonne par rapport à l'axe des abscisses, et la notation russe, présentée par exemple dans [NO06, §4.2], qui correspond à une rotation de 45 degrés de la notation française dans le sens antihoraire. Quant à nous, nous emploierons dans cette thèse une notation moins canonique : la notation cartésienne, où l'on visualise le diagramme comme le graphe d'une fonction en escalier. Cette notation se prêterait bien, par la suite, à la caractérisation de ce que l'on appellera *plus hauts poids presque plats*. Trois des notations précédentes sont illustrées à la Fig. 1.6.

On introduit deux types de relations d'ordre partiel pour les diagrammes de Young :

- Si $m \leq n$, $\lambda \vdash n$ et $\mu \vdash m$, on note $\lambda \subset \mu$ si $\mu_i \leq \lambda_i$ pour tout i (ce qui revient à dire que le diagramme associé à μ est contenu dans celui associé à λ) ;
- Si $\lambda = (\lambda_1, \dots, \lambda_n)$ et $\mu = (\mu_1, \dots, \mu_{n-1})$, on note $\mu \prec \lambda$ et on dit que λ et μ sont *entrelacés* si $\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \dots \geq \mu_{n-1} \geq \lambda_n$.

Lorsque $\mu \subset \lambda$, on peut définir un autre objet combinatoire appelé *diagramme de Young gauche*³⁴ $\lambda - \mu$, comme l'ensemble des cases qui restent lorsqu'on retire celles de μ à celles de λ , comme l'illustre la Fig. 1.7. On peut montrer (cf. [Mac15]) que si $\mu \subset \lambda$ et $\mu \prec \lambda$ alors $\lambda - \mu$ est une bande horizontale, c'est-à-dire que le diagramme de Young gauche possède au plus une seule case par colonne.

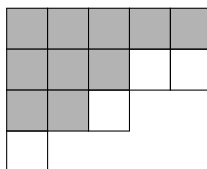


Figure 1.7: Décomposition du diagramme $(5, 5, 3, 1)$ en le diagramme $(5, 3, 2, 0)$ (en gris) et le diagramme gauche $(5, 5, 3, 1) - (5, 3, 2, 0)$ (en blanc). Notons que dernier est une bande horizontale.

Toutes les définitions qui précèdent se généralisent sans aucun souci à des n -uplets décroissants d'entiers relatifs $\lambda = (\lambda_1 \geq \dots \geq \lambda_n) \in \mathbb{Z}^n$, aussi appelés *signatures* ; on peut toujours utiliser

³⁴ou *skew-diagram* en anglais.

une représentation sous forme de diagrammes, en généralisant la notation française : le diagramme de $\lambda = (\lambda_1 \geq \dots \geq \lambda_N)$ est représenté dans le réseau \mathbb{Z}^2 par l'ensemble $\{(i, \lambda_i), 1 \leq i \leq N\}$, comme illustré à la Fig. 1.8. Nous verrons au chapitre 2, plus précisément au paragraphe 2.2.2, que l'on peut décomposer des signatures en partitions de manière à obtenir des informations sur leur forme globale, ce qui nous permettra de définir la notion de *plus haut poids presque plat*.

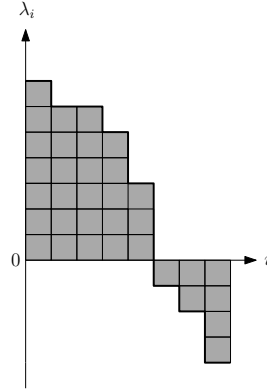


Figure 1.8: Le diagramme associé à la signature $(7, 6, 6, 5, 3, -1, -2, -4)$.

Soit $\lambda = (\lambda_1 \geq \dots \geq \lambda_N) \in \mathbb{Z}^N$ une signature. On lui associe le polynôme de Laurent suivant :

$$a_\lambda(x_1, \dots, x_N) := \det \left[x_i^{\lambda_j + N - j} \right]_{1 \leq i, j \leq N}$$

Il est alterné/antisymétrique puisque le fait de permuter x_i et x_j revient à permuter les lignes i et j de la matrice dont a_λ est le déterminant.

Proposition 1.3.12. a_λ est divisible dans $\mathbb{Z}[x_1^{\pm 1}, \dots, x_N^{\pm 1}]$ par le polynôme de Vandermonde

$$V_N(x_1, \dots, x_N) := \prod_{1 \leq i < j \leq N} (x_i - x_j) = \det \left[x_i^{N-j} \right]_{1 \leq i, j \leq N},$$

et le polynôme de Laurent résultant de cette division est une fonction symétrique notée s_λ .

Remarque : parfois dans la littérature on trouve la définition suivante du polynôme de Vandermonde : $V_N(x_1, \dots, x_N) = \det [x_i^{j-1}] = \prod_{i < j} (x_j - x_i)$, qui diffère de la première par un facteur ± 1 en fonction de N .

Définition 1.3.14. soit $\lambda = (\lambda_1 \geq \dots \geq \lambda_N) \in \mathbb{Z}^N$ une signature. Le polynôme de Laurent s_λ associé, défini par

$$s_\lambda(x_1, \dots, x_N) = \frac{\det \left[x_i^{\lambda_j + N - j} \right]}{V_N(x_1, \dots, x_N)}, \quad \forall (x_1, \dots, x_N) \in \mathbb{C}^N \quad (1.65)$$

et introduit à la proposition précédente, s'appelle la *fonction de Schur* associée à λ .

Exemples. Considérons des cas particuliers de signatures à coefficients positifs : si $N \in \mathbb{N}^*$ et $r \in \mathbb{N}^*$ sont deux entiers on pose $\lambda = (r) \in \mathbb{N}^N$ lorsque $\lambda = (r, 0, \dots, 0)$, ce qui correspond à un diagramme d'une seule ligne, et si $r \leq N$ on pose $\lambda = (1^r) \in \mathbb{N}^N$ lorsque $\lambda = (\underbrace{1, \dots, 1}_r, 0, \dots, 0)$,

ce qui correspond à un diagramme d'une seule colonne. On peut vérifier que

$$s_{(r)}(x_1, \dots, x_N) = h_r(x_1, \dots, x_N),$$

où h_r est la r -ième fonction symétrique complète, définie comme la somme de tous les monômes de degré r . De même on peut vérifier que

$$s_{(1^r)}(x_1, \dots, x_N) = e_r(x_1, \dots, x_N),$$

où e_r est la r -ième fonction symétrique élémentaire, définie comme la somme de tous les produits de r éléments distincts parmi les N variables. Voici des exemples plus précis de telles fonctions lorsque $N = 3$.

$$\begin{aligned} s_{(1,0,0)} &: (x, y, z) \mapsto x + y + z, \\ s_{(1,1,0)} &: (x, y, z) \mapsto xy + xz + yz, \\ s_{(2,0,0)} &: (x, y, z) \mapsto x^2 + xy + xz + y^2 + yz + z^2, \\ s_{(1,1,1)} &: (x, y, z) \mapsto xyz, \\ s_{(3,0,0)} &: (x, y, z) \mapsto x^3 + x^2y + x^2z + xy^2 + xyz + xz^2 + y^3 + y^2z + yz^2 + z^3. \end{aligned}$$

Si l'on note $|\lambda| = \lambda_1 + \dots + \lambda_N$ on remarque que dans les exemples précédents s_λ est un polynôme homogène de degré $|\lambda|$; c'est en réalité valable pour tout λ tel que $\lambda_N \geq 0$, et c'est une conséquence de la formule de *Jacobi–Trudi*³⁵

$$s_\lambda = \det(h_{\lambda_i + j - i})_{1 \leq i, j \leq N}. \quad (1.66)$$

Si l'on autorise les λ_i à être négatifs, l'homogénéité persiste bien que s_λ devienne un polynôme de Laurent.

La formule suivante, appelée *règle de branchement*, permet de calculer récursivement n'importe quelle fonction de Schur à N variables à partir de celles à $N - 1$ variables. Elle possède par ailleurs une interprétation en termes de représentations du groupe unitaire.

Proposition 1.3.13. Soit $\lambda = (\lambda_1 \geq \dots \geq \lambda_N) \in \mathbb{Z}^N$. Pour tout $(x_1, \dots, x_N) \in \mathbb{C}^N$, on a

$$s_\lambda(x_1, \dots, x_N) = \sum_{\substack{\mu = (\mu_1 \geq \dots \geq \mu_{N-1}) \in \mathbb{Z}^{N-1} \\ \mu \prec \lambda}} s_\mu(x_1, \dots, x_{N-1}) x_N^{|\lambda| - |\mu|}. \quad (1.67)$$

Une dernière formule qui va nous intéresser est la *formule de Pieri*.

Proposition 1.3.14. Soit $\lambda = (\lambda_1 \geq \dots \geq \lambda_N) \in \mathbb{N}^N$ une partition et r un entier naturel positif. On a l'égalité suivante :

$$h_r s_\lambda = \sum_{\mu} s_\mu, \quad (1.68)$$

où la somme s'effectue sur toutes les signatures μ telles que $\lambda \subset \mu$ et telles que $\mu - \lambda$ est une bande horizontale de taille r .

³⁵On en trouvera une démonstration dans [FH91, §A.2].

Les fonctions de Schur possèdent une multitude d'autres propriétés combinatoires que l'on pourra notamment retrouver dans [Mac15] et [Sta99], mais dans le cadre de cette thèse elles auront avant tout la vocation d'être les caractères des représentations irréductibles de $U(N)$, comme nous allons le voir au prochain paragraphe.

1.3.5 Représentations du groupe unitaire

Dans cette section, nous allons nous concentrer sur le cas où le groupe G est le groupe unitaire $U(N)$. Notre objectif final est de démontrer la décomposition suivante du noyau de la chaleur p_T sur $U(N)$, dont nous définirons les différentes composantes :

$$p_T(U) = \sum_{\substack{\lambda_1 \geq \dots \geq \lambda_N \\ (\lambda_1, \dots, \lambda_N) \in \mathbb{Z}^N}} e^{-c_2(\lambda) \frac{T}{2}} d_{\lambda} s_{\lambda}(U), \quad \forall T > 0, \forall U \in U(N). \quad (1.69)$$

En premier lieu, nous allons donner les représentations irréductibles de $U(N)$ et leurs caractères ; nous allons pour ce faire utiliser la théorie des plus hauts poids. Nous exhiberons ensuite une expression de l'opérateur de Casimir et de ses valeurs propres associées aux fonctions de Schur, pour enfin exprimer le noyau de la chaleur à l'aide des fonctions de Schur, ce qui donnera l'équation (1.69).

On rappelle que $U(N)$ est un sous-groupe compact et connexe de $GL_N(\mathbb{C})$. Son algèbre de Lie \mathfrak{u}_N est un espace vectoriel réel mais pas complexe, et sa complexifiée n'est autre que $\mathfrak{gl}_N(\mathbb{C}) = \mathcal{M}_N(\mathbb{C})$. Le sous-groupe

$$T_N = \{t = \text{diag}(t_1, \dots, t_N), t_1, \dots, t_N \in \mathbb{C}, |t_j| = 1\}$$

du groupe $U(N)$ en est un tore maximal. Son algèbre de Lie est donnée par

$$\mathfrak{t}_N = \{t = \text{diag}(t_1, \dots, t_N), t_1, \dots, t_N \in i\mathbb{R}\}.$$

La complexifiée de \mathfrak{t}_N , notée \mathfrak{h}_N , correspond à la restriction de $\mathfrak{gl}_N(\mathbb{C})$ aux matrices diagonales, et la sous-algèbre nilpotente de $\mathfrak{gl}_N(\mathbb{C})$ constituée des matrices triangulaires supérieures strictes est notée \mathfrak{n}_N .

Théorème 1.3.15. *L'algèbre de Lie $\mathfrak{gl}_N(\mathbb{C})$ admet la décomposition en somme directe*

$$\mathfrak{gl}_N(\mathbb{C}) = \mathfrak{h}_N \oplus \mathfrak{n}_N \oplus \mathfrak{n}_N^* \quad (1.70)$$

Soit (π, V) une représentation de dimension finie de $U(N)$. On étend π par différentiation en une représentation $d\pi$ de l'algèbre de Lie \mathfrak{u}_N , puis, par complexification, en une représentation de $\mathfrak{gl}_N(\mathbb{C})$. Une forme linéaire μ sur \mathfrak{h}_N est un *poids* de la représentation π s'il existe un vecteur $v \in V$ non nul tel que pour tout $H \in \mathfrak{h}_N$, on ait $d\pi(H)(v) = \mu(H)v$. Tout poids de la représentation π est de la forme

$$\mu(H) = \sum_{i=1}^N \mu_i h_i, \quad \mu_i \in \mathbb{Z}, \quad \forall H = \text{diag}(h_1, \dots, h_N) \in \mathfrak{h}_N, \quad (1.71)$$

et on appelle *poids* de \mathfrak{h}_N toutes les formes linéaires sur \mathfrak{h}_N de cette forme.

L'ensemble P des poids de \mathfrak{h}_N est donc en bijection avec \mathbb{Z}^N . On note $P(\pi)$ le sous-ensemble de P constitué des poids de la représentation π . Parmi les poids de \mathfrak{h}_N , on distingue les poids *dominants* et les poids *fortement dominants*.

Définition 1.3.15. On note P^+ (resp. P^{++}) l'ensemble des poids μ dominants (resp. fortement dominants), c'est-à-dire tels que $\mu_1 \geq \dots \geq \mu_N$ (resp. $\mu_1 > \dots > \mu_N$).

Proposition 1.3.16. On a la décomposition suivante :

$$V = \bigoplus_{\mu \in P(\pi)} V_{\mu}, \quad (1.72)$$

où V_{μ} est le sous-espace propre généralisé associé à μ , c'est-à-dire

$$V_{\mu} = \{v \in V : \forall H \in \mathfrak{h}_N, d\pi(H)v = \mu(H)v\}.$$

Définition 1.3.16. Un vecteur non nul $v \in V$ est un *vecteur de plus haut poids* s'il existe un poids $\lambda \in P(\pi)$ tel que

$$d\pi(H)v = \lambda(H)v, \forall H \in \mathfrak{h}_N$$

et

$$d\pi(X)v = 0, \forall X \in \mathfrak{n}_N.$$

Par la décomposition (1.70) on voit que si v est un vecteur de plus haut poids et $M = H + X_+ \in \mathfrak{gl}_N(\mathbb{C})$, alors $d\pi(M)v = \lambda(H)v$.

Théorème 1.3.17 (Théorème du plus haut poids). Soit (π, V) une représentation de dimension finie de $U(N)$.

(i) π possède un vecteur de plus haut poids.

(ii) π est irréductible si et seulement si tous les vecteurs de plus haut poids sont proportionnels entre eux. En particulier, ils sont associés à un même poids dominant λ que l'on appellera plus haut poids de π .

Ce théorème définit une application

$$\varphi : \begin{cases} \widehat{U}(N) & \rightarrow P^+ \\ \chi_{\pi} & \mapsto \lambda \end{cases},$$

et il se trouve qu'elle est bijective, comme l'indique le théorème suivant.

Théorème 1.3.18. Pour tout $\lambda \in P^+$ il existe une représentation irréductible de $U(N)$ de plus haut poids λ . De plus, si π est une représentation irréductible de $U(N)$ de plus haut poids λ et de caractère χ_{π} , alors pour tout $t \in T_N$,

$$\chi_{\pi}(t) = s_{\lambda}(t), \quad (1.73)$$

où s_{λ} est la fonction de Schur associée à λ ³⁶. En particulier, si deux représentations irréductibles ont le même plus haut poids, alors elles sont équivalentes.

³⁶On utilise implicitement la bijection $T_N \simeq U(1)^N \subset \mathbb{C}^N$ pour étendre les fonctions de Schur au tore maximal T_N .

Cela suffit à décrire les caractères des représentations irréductibles car toute fonction centrale sur $U(N)$ est entièrement caractérisée par sa restriction à T_N . Terminons par le calcul de la dimension d'une représentation irréductible à l'aide de son plus haut poids. Si (π, V) est une représentation irréductible de plus haut poids λ , sa dimension $d_\lambda = s_\lambda(1, \dots, 1)$ n'est pas donnée directement par la formule (1.65), qui lui donne une forme indéterminée. Pour contourner cela, on peut par exemple appliquer la fonction de Schur à l'ensemble des puissances d'un paramètre $q \neq 1$, ce qui donne

$$s_\lambda(1, q, \dots, q^{N-1}) = \prod_{i < j} \frac{q^{\lambda_i + N - i} - q^{\lambda_j + N - j}}{q^{j-1} - q^{i-1}},$$

et de faire tendre q vers 1. On obtient la *formule de la dimension de Weyl* :

$$d_\lambda = s_\lambda(1, \dots, 1) = \prod_{1 \leq i, j \leq N} \frac{\lambda_i - i - \lambda_j + j}{j - i}. \quad (1.74)$$

On a par ailleurs la *règle de branchement* suivante, comme corollaire direct de la Prop. 1.3.13, qui permet de calculer par induction les dimensions de représentations irréductibles de $U(N)$ à partir de celles de $U(N-1)$.

Proposition 1.3.19 (Règle de branchement sur $U(N)$). *Soit $\lambda \in \widehat{U}(N)$ un plus haut poids de $U(N)$. Alors on a*

$$d_\lambda = \sum_{\mu < \lambda} d_\mu, \quad (1.75)$$

où la somme est effectuée sur tous les plus hauts poids $\mu \in \widehat{U}(N-1)$ entrelacés avec λ , c'est-à-dire tels que l'on obtient le diagramme de λ à partir de celui de μ en ajoutant une case.

Si λ et μ sont deux plus hauts poids de $U(N)$ tels que le diagramme de λ est obtenu à partir de celui de μ en ajoutant une case, on utilisera plutôt la notation $\lambda \searrow \mu$ ou $\mu \nearrow \lambda$. On retrouve une relation de ce type dans la proposition suivante, qui est un corollaire immédiat de la formule de Pieri (Prop. 1.3.14).

Proposition 1.3.20. *Soit $\lambda \in \widehat{SU}(N)$ un plus haut poids de $SU(N)$. Alors on a*

$$\mathrm{Tr}(x) s_\lambda(x) = \sum_{\substack{\mu \in \widehat{SU}(N) \\ \mu \searrow \lambda}} s_\mu(x), \quad \forall x \in SU(N). \quad (1.76)$$

Tournons-nous à présent vers l'élément de Casimir C_{u_N} . Bien que sa définition soit indépendante du choix de la base de \mathfrak{u}_N , il nous sera nécessaire d'en exhiber une afin d'effectuer certains calculs. On rappelle que \mathfrak{u}_N peut être munie du produit scalaire défini en (1.36), et qu'on obtient une base orthonormale pour ce produit scalaire en posant, pour $1 \leq k < \ell \leq N$,

$$X_{k\ell} = \frac{1}{\sqrt{2N}}(E_{k\ell} - E_{\ell k}), \quad Y_{k\ell} = \frac{i}{\sqrt{2N}}(E_{k\ell} + E_{\ell k}),$$

et pour $1 \leq k \leq N$

$$Z_k = \frac{i}{\sqrt{N}}E_{kk}.$$

Proposition 1.3.21. *L'élément de Casimir $C_{\mathfrak{u}_N}$ s'écrit³⁷ :*

$$C_{\mathfrak{u}_N} = -\frac{1}{N} \sum_{k,\ell=1}^N E_{k\ell} \otimes E_{\ell k}. \quad (1.77)$$

Le tenseur $\sum_{k,\ell=1}^N E_{k\ell} \otimes E_{\ell k}$ qui apparaît dans le terme de droite de l'équation (1.77) n'est autre que l'opérateur de transposition

$$T : \begin{cases} \mathbb{C}^N \otimes \mathbb{C}^N & \rightarrow \mathbb{C}^N \otimes \mathbb{C}^N \\ x \otimes y & \mapsto y \otimes x \end{cases}.$$

Cette expression de l'élément de Casimir sert entre autres à déduire les formules suivantes, que l'on retrouve notamment dans [Sen08a], ou plus récemment dans [DHK13] sous le nom de "formules magiques".

Proposition 1.3.22. *Soit (X_1, \dots, X_d) une base orthonormée de \mathfrak{u}_N pour le produit scalaire (1.36). Alors pour tout $(A, B) \in \mathcal{M}_N(\mathbb{C})^2$, on a les formules suivantes :*

$$\sum_{i=1}^d X_i^2 = -I_N, \quad (1.78)$$

$$\sum_{i=1}^d X_i A X_i = -\operatorname{tr}(A) I_N, \quad (1.79)$$

$$\sum_{i=1}^d \operatorname{tr}(X_i A) X_i = -\frac{1}{N^2} A, \quad (1.80)$$

$$\sum_{i=1}^d \operatorname{tr}(X_i A) \operatorname{tr}(X_i B) = -\frac{1}{N^2} \operatorname{tr}(AB). \quad (1.81)$$

Remarque. L'équation (1.78) justifie le calcul $m(C_{\mathfrak{u}_N}) = -I_N$, où $m : x \otimes y \mapsto xy$ est l'opération de multiplication sur les matrices, que l'on a utilisée au paragraphe 1.2.5 pour calculer la variation quadratique du mouvement brownien unitaire.

Nous avons vu à la section 1.3.3 que si l'on prend une représentation irréductible (π, V) de $U(N)$, disons par exemple une représentation de plus haut poids λ , alors l'opérateur de Casimir associé vérifie

$$\Omega_\lambda := \sum_{i=1}^d d\pi(X_i)^2 = -c_2(\lambda) \operatorname{Id}_V \quad (1.82)$$

pour un certain $c_2(\lambda) \geq 0$. On peut désormais donner l'expression explicite de $c_2(\lambda)$.

Proposition 1.3.23. *Soit $\lambda \in \widehat{U}(N)$ un plus haut poids. Alors on a*

$$c_2(\lambda) = \frac{1}{N} \left(\sum_{i=1}^N \lambda_i^2 + \sum_{1 \leq i < j \leq N} (\lambda_i - \lambda_j) \right) = \frac{1}{N} \left(\sum_{i=1}^N \lambda_i^2 + \sum_{i=1}^N \lambda_i (N + 1 - 2i) \right). \quad (1.83)$$

³⁷Attention, cette égalité est valable dans $\mathcal{M}_N(\mathbb{C}) \otimes_{\mathbb{C}} \mathcal{M}_N(\mathbb{C})$ mais pas dans $\mathcal{M}_N(\mathbb{C}) \otimes_{\mathbb{R}} \mathcal{M}_N(\mathbb{C})$!

Preuve. Soit π une représentation de $U(N)$ de plus haut poids λ . Dans la base orthonormée $(X_{k\ell}, Y_{k\ell}, Z_k)$, on a :

$$\begin{aligned}\pi(X_{k\ell})^2 &= \frac{1}{2N} (\pi(E_{k\ell})^2 + \pi(E_{\ell k})^2 - \pi(E_{k\ell})\pi(E_{\ell k}) - \pi(E_{\ell k})\pi(E_{k\ell})), \\ \pi(Y_{k\ell})^2 &= \frac{1}{2N} (-\pi(E_{k\ell})^2 - \pi(E_{\ell k})^2 - \pi(E_{k\ell})\pi(E_{\ell k}) - \pi(E_{\ell k})\pi(E_{k\ell})).\end{aligned}$$

On en déduit que

$$\begin{aligned}-\Omega_\lambda &= \sum_{k=1}^N \pi(E_{kk})^2 + \sum_{k \neq \ell} \pi(E_{k\ell})\pi(E_{\ell k}) \\ &= \sum_{k=1}^N \pi(E_{kk})^2 + \sum_{1 \leq k < \ell \leq N} [\pi(E_{k\ell}), \pi(E_{\ell k})].\end{aligned}$$

On a en outre, pour tout $k < \ell$, $[E_{k\ell}, E_{\ell k}] = E_{kk} - E_{\ell\ell}$, ce qui implique

$$-\Omega_\lambda = \sum_{k=1}^N \pi(E_{kk})^2 + \sum_{1 \leq k < \ell \leq N} (\pi(E_{kk}) - \pi(E_{\ell\ell})).$$

Enfin, soit $v \in V$ un vecteur de plus haut poids. On a $\pi(E_{kk})v = \lambda_k v$ pour tout k , et $\pi(E_{k\ell})v = 0$ pour tout $k < \ell$. On en déduit donc la première égalité de (1.83) ; la seconde se déduit de la première par réarrangement des sommes. \square

Nous pouvons enfin énoncer et démontrer le résultat annoncé au début de cette section. On appelle *noyau de la chaleur* sur $U(N)$ la solution $p : (t, U) \in [0, \infty[\times U(N) \mapsto p_t(U)$ de l'équation de la chaleur

$$\begin{cases} \frac{d}{dt} p_t(U) &= \Delta_{U(N)} p_t(U), \quad \forall t > 0, \quad \forall U \in U(N), \\ p_0(U) &= \delta_{I_N}(U) \end{cases} \quad (1.84)$$

Théorème 1.3.24. *Le noyau de la chaleur sur $U(N)$ admet la décomposition suivante dans $L^2(U(N))$:*

$$p_t(U) = \sum_{\lambda \in \widehat{U}(N)} e^{-c_2(\lambda) \frac{t}{2}} d_\lambda s_\lambda(U), \quad \forall t > 0, \quad \forall U \in U(N). \quad (1.85)$$

Preuve. Pour alléger les notations, et parce qu'il n'y a aucune ambiguïté sur le groupe sous-jacent, on notera Δ le laplacien sur $U(N)$. D'après la formule de Plancherel (1.58) le noyau de la chaleur admet la décomposition de Fourier suivante dans $L^2(U(N))$:

$$p_t(U) = \sum_{\lambda \in \widehat{U}(N)} d_\lambda (p_t * s_\lambda)(U).$$

Il nous faut alors calculer, pour tout $\lambda \in \widehat{U}(N)$ et tout $U \in U(N)$, le produit de convolution $(p_t * s_\lambda)(U)$. On a, en utilisant la linéarité de l'intégrale et l'équation de la chaleur,

$$\frac{d}{dt} (p_t * s_\lambda)(U) = \left(\left(\frac{d}{dt} p_t \right) * s_\lambda \right) (U) = \frac{1}{2} ((\Delta p_t) * s_\lambda)(U).$$

Par ailleurs, on peut vérifier³⁸ que $((\Delta p_t) * s_\lambda)(U) = (p_t * (\Delta s_\lambda))(U)$, et il est clair que $s_\lambda \in \mathcal{M}_\lambda$, donc l'équation (1.64) implique que

$$((\Delta p_t) * s_\lambda)(U) = -c_2(\lambda)(p_t * s_\lambda)(U),$$

où $c_2(\lambda)$ est donné dans la Prop. 1.3.23. On en déduit que, pour tout $\lambda \in \widehat{U}(N)$ et $U \in U(N)$ fixés, la fonction $u_{\lambda,U} \mapsto (p_t * s_\lambda)(U)$ est solution de l'équation différentielle

$$u'_{\lambda,U}(t) = -\frac{c_2(\lambda)}{2}u_{\lambda,U}(t),$$

et il existe alors une constante $C_{\lambda,U} \in \mathbb{R}$ telle que

$$(p_t * s_\lambda)(U) = C_{\lambda,U}e^{-c_2(\lambda)\frac{t}{2}}.$$

La condition initiale $p_0(U) = \delta_{I_N}(U)$ permet de déterminer $C_{\lambda,U}$:

$$(p_0 * s_\lambda)(U) = s_\lambda(U) = C_{\lambda,U},$$

ce qui permet de conclure la preuve. □

I.4 LA MESURE DE YANG–MILLS SUR UNE SURFACE COMPACTE

Suite aux travaux de Yang et Mills – notamment [YM54] – qui ont donné leur nom à la théorie, une autre avancée cruciale dans le développement de la mesure de Yang–Mills a été [Mig75], dans lequel Migdal décrit ce que devrait être cette mesure sur une surface, ou tout du moins une version discrète ; on y voit notamment la première apparition du développement du noyau de la chaleur sur le groupe de structure sur la base des caractères irréductibles. Ont suivi alors les travaux de Gross [Gro88] puis de Driver [Dri89] qui ont mené à la construction rigoureuse de la mesure de Yang–Mills planeaire. Il a fallu ensuite attendre les années 90 pour que Sengupta la construise sur la sphère [Sen92] puis sur toute surface compacte [Sen97]. Parallèlement à cela, Witten [Wit91] a étudié plus en détail la mesure de Yang–Mills discrète, dans des perspectives différentes – son but était de calculer le volume symplectique de l'espace des modules de connexions plates sur une surface de Riemann.

Dans cette section, nous décrivons la construction de la mesure de Yang–Mills telle qu'elle a été effectuée par T. Lévy [Lévo3, Lévio], qui est complémentaire de celle de Sengupta, et repose sur une formalisation rigoureuse de la théorie discrète de Witten, adjointe d'un passage à la limite continue par des arguments topologiques et probabilistes. À partir de maintenant, le terme *surface* désignera, sauf mention explicite³⁹, une *variété différentielle réelle de dimension 2 compacte connexe sans bord*, et on notera Σ l'ensemble de ces surfaces. Le choix de restreindre notre définition de surface provient de la classification suivante, dont on peut trouver par exemple la démonstration dans [Mas91].

³⁸Cela repose essentiellement sur la formule d'intégration par parties suivante : en reprenant la représentation ρ de u_N sur $\mathcal{C}^1(U(N))$ définie par (1.62), on a pour toutes fonctions $f_1, f_2 \in \mathcal{C}^1(G)$,

$$\langle \rho(X)f_1, f_2 \rangle_{L^2(U(N))} = -\langle f_1, \rho(X)f_2 \rangle_{L^2(U(N))}.$$

³⁹La seule exception à laquelle nous accorderons de l'importance sera le plan euclidien \mathbb{R}^2 , pour lequel une grande partie des résultats ont précédé ceux sur les surfaces compactes.

Théorème 1.4.1 (Classification des surfaces). *Toute surface $M \in \Sigma$ est homéomorphe à l'une des surfaces suivantes :*

- (i) *La somme connexe de g tores⁴⁰,*
- (ii) *La somme connexe de g plans projectifs, avec $g \geq 1$.*

Dans le premier cas, la surface est dite *orientable*, et g est appelé son *genre*. On peut montrer que cette surface est homéomorphe à un polygone à $4g$ côtés⁴¹, appelé *domaine fondamental*, dont le bord orienté dans le sens antihoraire s'écrit (à un choix du côté de départ près) $[a_1, b_1] \cdots [a_g, b_g]$, où

- $\{a_1, b_1, \dots, a_g, b_g\}$ sont des générateurs du groupe fondamental de la surface,
- $[x, y] = xyx^{-1}y^{-1}$ désigne le commutateur de deux éléments d'un groupe,
- Le polygone est recollé le long des côtés identiques en préservant leur orientation

Autrement dit, cela revient à quotienter le polygone par la relation $[a_1, b_1] \cdots [a_g, b_g] = 1$.

Dans le second cas, on appellera également g son *genre*, bien que cette dénomination soit moins canonique, et la surface est dite *non orientable*. On peut montrer que cette surface est homéomorphe à un polygone à $2g$ côtés, dont le bord orienté dans le sens antihoraire s'écrit $a_1^2 \cdots a_g^2$.

Le domaine fondamental du tore est illustré à la Fig. 1.9, et ceux de deux surfaces plus exotiques à la Fig. 1.10.

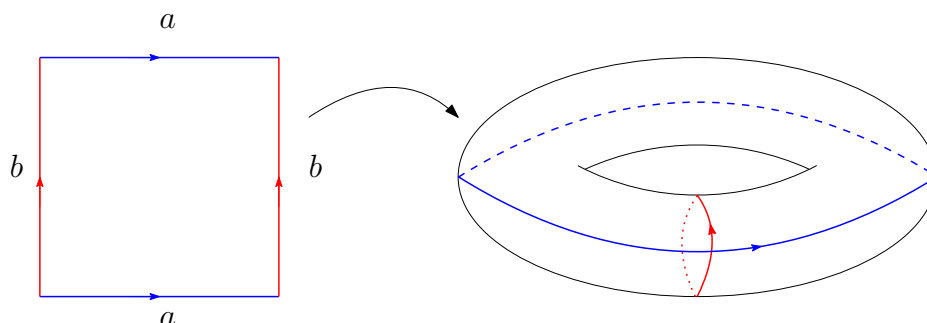


Figure 1.9: Le domaine fondamental d'un 2-tore (à gauche), et ce même tore (à droite) obtenu par recollement du domaine le long des générateurs a et b de son groupe fondamental.

1.4.1 Mesure de Yang–Mills sur un graphe

Afin de définir la mesure de Yang–Mills sur une surface, il nous faut tout d'abord la définir sur un graphe plongé dans cette surface. Aussi commençons-nous par quelques définitions sur les graphes plongés.

⁴⁰Si $g = 0$ par convention cela donne une sphère ; autrement on peut voir une telle surface comme un tore à g anses.

⁴¹On notera que, pour le genre zéro, cela n'a pas de sens de parler d'un polygone à zéro côté ; dans ce cas précis cela revient à prendre un disque et à le recoller le long de son bord.

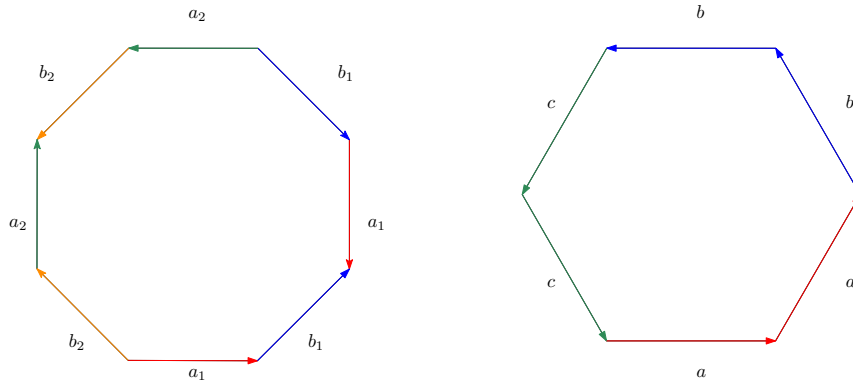


Figure 1.10: Le domaine fondamental d'une surface orientable de genre 2 (à gauche), et d'une surface non-orientable de genre 3 (à droite).

Definition 1.4.1. Soit $M \in \Sigma \cup \{\mathbb{R}^2\}$ une surface.

- (i) Un *chemin paramétré* sur M est une application lipschitzienne $c : [0, 1] \rightarrow M$ qui est soit constante sur $[0, 1]$, soit constante sur aucun sous-intervalle de $[0, 1]$.
- (ii) Deux chemins paramétrés sur M sont *équivalents* s'ils ne diffèrent que par un changement de paramétrage bi-lipschitzien croissant.
- (iii) L'ensemble des *chemins* sur M est le quotient de l'ensemble des chemins paramétrés par la relation d'équivalence ci-dessus⁴². On note $\mathcal{P}(M)$ l'ensemble des chemins sur M .
- (iv) Si c est un chemin, on définit son *inverse* c^{-1} comme $c^{-1} = c \circ \varphi$, où φ est un changement de paramétrage décroissant.
- (v) Un chemin c tel que $c(0) = c(1)$ est appelé un *lacet* (ou une *boucle*), et est dit *simple* si le chemin paramétré c est injectif sur $[0, 1[$. On note $\mathcal{L}(M)$ l'ensemble des lacets sur M , et $\mathcal{L}_m(M)$ l'ensemble des lacets sur M dont le point de départ (et donc d'arrivée) est $m \in M$.
- (vi) Une *arête* est un chemin injectif.
- (vii) Un *graphe* $(\mathbb{E}, \mathbb{V}, \mathbb{F})$ sur M est la donnée :
 - d'un ensemble \mathbb{E} d'arêtes tel que pour tout $e \in \mathbb{E}$, on ait $e^{-1} \in \mathbb{E}$, et tel que deux arêtes distinctes qui ne sont pas l'inverse l'une de l'autre ne se rencontrent qu'à leurs extrémités,
 - de l'ensemble \mathbb{V} des extrémités des arêtes, appelées *sommets*,
 - de l'ensemble \mathbb{F} des composantes connexes de $M \setminus \{e([0, 1]), e \in \mathbb{E}\}$, appelées *faces*.
- (viii) Un graphe est dit *admissible* si toutes ses faces sont homéomorphes à des disques.
- (ix) Soit $(\mathbb{E}, \mathbb{V}, \mathbb{F})$ un graphe sur M . Une *orientation* de ce graphe est la donnée d'un sous-ensemble $\mathbb{E}^+ \subset \mathbb{E}$ tel que pour tout $e \in \mathbb{E}$,

$$|\mathbb{E}^+ \cap \{e, e^{-1}\}| = 1.$$

⁴²Pour alléger les notations, comme nous n'allons considérer que des chemins au sens défini ici, nous noterons c aussi bien la classe d'équivalence à paramétrage croissant près ou bien un de ses représentants, sans ambiguïté.

On peut voir un exemple de graphe admissible orienté sur la Fig. I.II.



Figure I.II: À gauche : un graphe admissible non orienté plongé dans le plan. À droite : une orientation possible de ce même graphe.

L'ensemble $\mathcal{L}_m(M)$ des lacets issus du même point $m \in M$, muni de la concaténation, est un monoïde. En revanche, il est moins clair que l'on puisse le munir d'une structure de groupe ; Hambly et Lyons [HL10] ont montré qu'en définissant une relation d'équivalence sur ces lacets on peut obtenir un groupe, appelé *groupe des lacets réduits*. Nous ne détaillons pas cette relation d'équivalence, par souci de simplicité, et renvoyons à l'article original ou bien à [Lévi17, § 5.7] qui replace ce résultat dans le contexte de la théorie de Yang–Mills en deux dimensions.

Théorème I.4.2 ([LZ04], Thm.I.3.10). *Soit $M \in \Sigma$ une surface orientable de genre g et un graphe $(\mathbb{E}, \mathbb{V}, \mathbb{F})$ sur cette surface. Alors le graphe vérifie*

$$|\mathbb{V}| - \frac{1}{2}|\mathbb{E}| + |\mathbb{F}| = 2 - 2g.$$

La quantité $\chi(g) = 2 - 2g$ qui apparaît dans le théorème précédent s'appelle la *caractéristique d'Euler*, et c'est un invariant topologique : elle est la même pour deux surfaces homéomorphes.

Nous pouvons désormais introduire la mesure tant attendue : la mesure de Yang–Mills sur une surface.

Définition I.4.2. Soit $M \in \Sigma$ une surface, G un groupe compact, $\mathbb{G} = (\mathbb{E}, \mathbb{V}, \mathbb{F})$ un graphe sur M et \mathbb{E}^+ une orientation de celui-ci. L'espace de configuration associé au graphe $(\mathbb{E}, \mathbb{V}, \mathbb{F})$ est l'espace

$$\mathcal{C}_{\mathbb{G}}^G = G^{\mathbb{E}^+}$$

des fonctions de \mathbb{E}^+ à valeurs dans G . On notera $g : e \in \mathbb{E}^+ \mapsto g_e$ un élément générique de cet espace de configuration, et dg sa mesure de Haar⁴³.

En somme, une configuration revient à associer à chaque arête orientée de \mathbb{G} un élément du groupe G . La théorie de Yang–Mills décrite dans [Lévo3, Lévio] revient à étudier des configurations aléatoires sur cet espace, c'est-à-dire des familles de matrices aléatoires (parce qu'en pratique G sera pour nous un groupe matriciel) indexées par les lacets tracés dans un graphe. Nous allons voir que ces matrices sont les images des lacets en question par une fonction multiplicative sur les lacets.

⁴³Notons que cette mesure ne dépend pas du choix de l'orientation du graphe, puisque la mesure de Haar d'un groupe compact est invariante par l'application $x \mapsto x^{-1}$.

Définition 1.4.3. Soit G un groupe et $(\mathbb{E}, \mathbb{V}, \mathbb{F})$ un graphe admissible sur une surface M , muni d'une orientation \mathbb{E}^+ . Si $c = e_1^{\varepsilon_1} \cdots e_n^{\varepsilon_n}$ est un chemin passant par les arêtes e_1, \dots, e_n de \mathbb{E}^+ avec $\varepsilon_i \in \{-1, 1\}$ pour tout $1 \leq i \leq n$, alors on lui associe l'*holonomie*

$$h_c : \begin{cases} \mathcal{C}_{\mathbb{G}}^G & \rightarrow G \\ g & \mapsto g_{e_n}^{\varepsilon_n} \cdots g_{e_1}^{\varepsilon_1} \end{cases} . \quad (1.86)$$

L'application $c \mapsto h_c$ satisfait la même propriété de multiplicativité que celle décrite dans la Prop. 1.1.6 pour l'holonomie d'une connexion le long d'un lacet.

Il reste désormais à munir notre espace de configuration d'une mesure de probabilités, comme annoncé plus tôt. Le dernier ingrédient nécessaire à la définition de cette mesure est la donnée d'une mesure d'aire sur M , c'est-à-dire d'une mesure borélienne absolument continue par rapport à la mesure de Lebesgue dans toute carte. On l'utilisera pour mesurer l'aire des faces du graphe et on notera $|F|$ l'aire de la face $F \in \mathbb{F}$.

Définition 1.4.4. La *mesure de Yang–Mills* sur \mathbb{G} de groupe de structure G est la mesure de probabilité $\mu_{T,G,\mathbb{G}}$ sur $\mathcal{C}_{\mathbb{G}}^G$ définie par

$$d\mu_{T,G,\mathbb{G}}(g) = \frac{1}{Z_{T,G,\mathbb{G}}} \prod_{F \in \mathbb{F}} p_{|F|}(h_{\partial F}(g)) dg, \quad (1.87)$$

où $Z_{T,G,\mathbb{G}}$ est la constante de normalisation de $\mu_{T,G,\mathbb{G}}$, appelée *fonction de partition*, et vérifie

$$Z_{T,G,\mathbb{G}} = \int_{\mathcal{C}_{\mathbb{G}}^G} \prod_{F \in \mathbb{F}} p_{|F|}(h_{\partial F}(g)) dg. \quad (1.88)$$

L'équation (1.87) est parfois nommée dans la littérature *formule de Driver–Sengupta*, car elle a été démontrée par Driver [Dri89] dans le cas où M est le plan euclidien, et par Sengupta [Sen92, Sen97] pour la sphère S^2 puis pour toute surface compacte. Il est intéressant de remarquer que l'on n'a pas spécifié d'orientation pour le bord des faces dans la formule (1.87); la propriété de multiplicativité de l'application h et l'invariance du noyau de la chaleur $p_{|F|}$ par inversion permettent de remarquer que

$$p_{|F|}(h_{(\partial F)^{-1}}) = p_{|F|}(h_{\partial F}), \quad \forall F \in \mathbb{F}.$$

Par conséquent, la définition de la mesure de Yang–Mills est indépendante du choix de l'orientation du graphe, et, *a fortiori*, de l'orientation de la surface. La mesure de Yang–Mills peut donc être tout aussi bien définie sur une surface non orientable. Notons en outre que la définition de la mesure de Yang–Mills ne dépend pas du point de base lorsqu'on considère le bord d'une face F , puisque le noyau de la chaleur $p_{|F|}$ est également invariant par conjugaison, et changer le point de base de ∂F revient à conjuguer par des arêtes qui constituent ce bord.

La formule de Driver–Sengupta permet de définir un processus stochastique $(h_c)_{c \in \mathcal{P}(\mathbb{G})}$ à valeurs dans le groupe G et indexé par l'ensemble $\mathcal{P}(\mathbb{G})$ des chemins possibles dans le graphe \mathbb{G} ⁴⁴. Le principal intérêt de ce processus est qu'il est stable par subdivision du graphe. En effet, si \mathbb{G}_1 et \mathbb{G}_2 sont

⁴⁴Autrement dit, tous les chemins que l'on peut former en concaténant des arêtes du graphe.

des graphes admissibles orientés tels que \mathbb{G}_2 est *plus fin* que \mathbb{G}_1 , *i.e.* si toute arête de \mathbb{G}_1 est un chemin dans \mathbb{G}_2 , on peut construire une inclusion $\mathcal{C}_{\mathbb{G}_2} \rightarrow \mathcal{C}_{\mathbb{G}_1}$ par

$$\iota_{2 \rightarrow 1} \left\{ \begin{array}{l} \mathcal{C}_{\mathbb{G}_2} \rightarrow \mathcal{C}_{\mathbb{G}_1} \\ g \mapsto (e \in \mathbb{E}_1 \mapsto h_e(g)) \end{array} \right. ,$$

où chaque arête e de \mathbb{G}_1 est vue comme un chemin de \mathbb{G}_2 . On peut montrer, en utilisant la propriété de semi-groupe du noyau de la chaleur $(p_t)_{t \geq 0}$ sur G , que $\mu_{T,G,\mathbb{G}_1} = (\iota_{2 \rightarrow 1})_* \mu_{T,G,\mathbb{G}_2}$, et que les fonctions de partition Z_{T,G,\mathbb{G}_1} et Z_{T,G,\mathbb{G}_2} sont égales. En particulier on peut définir $Z_{T,G}$ comme la valeur de la fonction de partition de la mesure de Yang–Mills, qui ne dépend donc pas du choix d’un graphe sur la surface. On peut également construire une limite inductive des mesures de Yang–Mills sur des graphes plongés dans une même surface, mais cette construction est nettement plus difficile à montrer. En effet, étant donné deux graphes \mathbb{G}_1 et \mathbb{G}_2 , il n’existe pas toujours un troisième graphe qui soit à la fois plus fin que \mathbb{G}_1 et \mathbb{G}_2 . De plus, il n’est pas toujours garanti qu’un lacet, même très régulier, puisse s’écrire comme concaténation d’arêtes d’un graphe – son complémentaire peut par exemple posséder une infinité de composantes connexes (par exemple on ne peut pas trouver de graphe qui contienne le lacet issu de la concaténation du segment $[0, 1/\pi]$ et du graphe de la fonction $x \mapsto x \sin(1/x)$ sur $[0, 1/\pi]$, comme illustré à la Fig. 1.12). Un choix possible pour résoudre ces difficultés est de munir M d’une structure riemannienne telle que le volume riemannien coïncide avec la mesure d’aire, et de restreindre l’ensemble des chemins aux chemins géodésiques par morceaux, comme montré dans [Lévo3, Lévio], et il est ensuite envisageable d’étendre cette construction à des lacets moins réguliers par des raisonnements de densité. On obtient alors le théorème suivant.

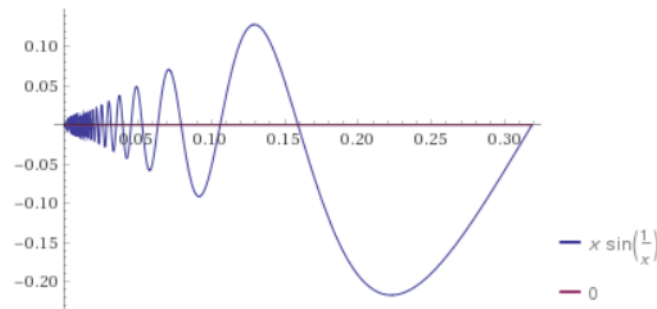


Figure 1.12: Un lacet de \mathbb{R}^2 que l’on ne peut pas écrire comme un chemin tracé dans un graphe.

Théorème 1.4.3. *Soit $M \in \Sigma$ une surface munie d’une mesure d’aire d’aire totale T , G un groupe de Lie compact, \mathfrak{g} son algèbre de Lie munie d’un produit scalaire invariant. Alors il existe un processus $(H_c)_{c \in \mathcal{P}(M)}$ à valeurs dans G et indexé par les chemins sur M , qui vérifie les hypothèses suivantes :*

- (i) *Pour tout graphe $\mathbb{G} = (\mathbb{E}, \mathbb{V}, \mathbb{F})$, la famille $(H_e)_{e \in \mathbb{E}}$ a pour loi $\mu_{T,G,\mathbb{G}}$.*
- (ii) *Pour toute suite $(c_n)_{n \geq 1}$ de chemins partageant les mêmes points de départ et d’arrivée, si $c_n \rightarrow c$ uniformément et la longueur de c_n converge vers la longueur de c , alors on a la convergence en probabilité $H_{c_n} \rightarrow H_c$.*

Ces hypothèses caractérisent la loi, dans le sens où deux processus qui les vérifient ont nécessairement même loi.

Contrairement à ce que l'on pourrait croire, la surface du théorème 1.4.3 n'a pas besoin d'être munie d'une métrique riemannienne particulière. En effet, bien que dans le point (ii) la longueur des chemins considérés dépende du choix d'une métrique, le fait que la longueur de c_n converge vers celle de c en est indépendant. Le processus $(H_c)_{c \in \mathcal{P}(M)}$ est appelé *champ d'holonomie de Yang–Mills*, et satisfait l'invariance de jauge déjà rencontrée à l'équation (1.27), que l'on réécrit ici :

$$\left((j(c(1)))^{-1} H_c j(c(0)) \right)_{c \in \mathcal{P}(M)} \stackrel{(d)}{=} (H_c)_{c \in \mathcal{P}(M)},$$

pour toute transformation de jauge $j : M \rightarrow G$. Cette invariance possède deux implications distinctes :

- Si $c \in \mathcal{P}(M)$ est un chemin qui n'est pas un lacet, c'est-à-dire que $c(1) \neq c(0)$, alors H_c suit une loi uniforme sur G ;
- Si $c \in \mathcal{P}(M)$ est un lacet, alors la loi de H_c est invariante par conjugaison sur G .

Le champ d'holonomie possède par ailleurs une autre forme d'invariance, à savoir l'invariance par difféomorphismes qui préservent l'aire. Cette caractéristique est démontrée par exemple dans [Lévo3, Prop.I.7.1].

1.4.2 Lois marginales du champ de Yang–Mills

Le champ d'holonomie de Yang–Mills, en tant que processus stochastique, n'est pas simple à étudier comme un tout. Aussi, à l'instar du mouvement brownien réel $(B_t)_{t \geq 0}$ qui est plus facile à manipuler par le truchement de ses marginales fini-dimensionnelles $(B_{t_1}, \dots, B_{t_n})_{t_1, \dots, t_n \geq 0}$, nous allons décrire $(H_c)_{c \in \mathcal{P}(M)}$ à l'aide de marginales adéquates. Pour cela il nous faut fixer un graphe admissible orienté $\mathbb{G} = (\mathbb{E}, \mathbb{V}, \mathbb{F})$ sur M , et considérer le groupe des lacets réduits $\mathcal{L}_m^{\text{red}}(\mathbb{G})$ sur le graphe \mathbb{G} , avec pour point de base $m \in \mathbb{V}$. Ce groupe est défini comme le quotient du monoïde $\mathcal{L}_m(\mathbb{G})$ des lacets sur \mathbb{G} avec pour point de base le sommet m , pour la relation d'équivalence suivante : $\ell \sim \ell'$ si et seulement si on peut passer de ℓ à ℓ' en ajoutant ou en supprimant des sous-lacets de la forme ee^{-1} , comme illustré à la Fig. 1.13. Le *lacet réduit* associé à une telle classe d'équivalence est le seul lacet qui ne contient pas de sous-lacet de la forme ee^{-1} .

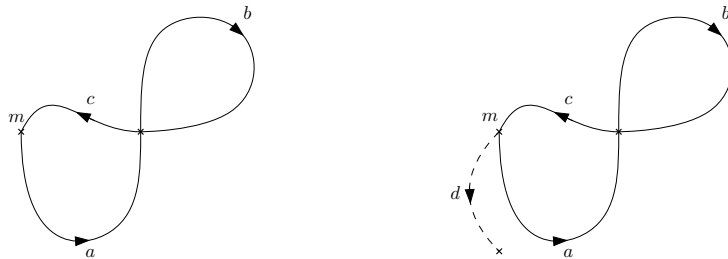


Figure 1.13: Deux lacets $\ell = abc$ et $\ell' = abcdd^{-1}$ appartenant à la même classe d'équivalence dans $\mathcal{L}_m^{\text{red}}(\mathbb{G})$.

La structure algébrique de $\mathcal{L}_m^{\text{red}}(\mathbb{G})$ va permettre de définir les marginales recherchées.

Proposition 1.4.4. *Le groupe $\mathcal{L}_m^{\text{red}}(\mathbb{G})$ est un groupe libre de rang $\frac{1}{2}|\mathbb{E}| - |\mathbb{V}| + 1$.*

Notons que d'après la proposition précédente et le Thm. 1.4.2, le rang de $\mathcal{L}_m^{\text{red}}(\mathbb{G})$ est égal à $|\mathbb{F}| + 2g - 1$ lorsque M est une surface orientable de genre g .

Proposition 1.4.5. *Supposons que M est une surface orientable de genre g , et $\mathbb{G} = (\mathbb{E}, \mathbb{V}, \mathbb{F})$ un graphe admissible sur M . Alors le groupe $\mathcal{L}_m^{\text{red}}(\mathbb{G})$ admet la présentation suivante⁴⁵ :*

$$\mathcal{L}_m^{\text{red}}(\mathbb{G}) = \langle a_1, b_1, \dots, a_g, b_g, \ell_1, \dots, \ell_f | [a_1, b_1] \cdots [a_g, b_g] = \ell_1 \cdots \ell_f \rangle, \quad (1.89)$$

où

- $f = |\mathbb{F}|$ est le nombre de faces du graphe,
- $(a_1, b_1, \dots, a_g, b_g)$ sont des générateurs du groupe fondamental de M ,
- pour tout $i \in \{1, \dots, f\}$, ℓ_i est un lasso qui entoure la face F_i , c'est-à-dire un lacet de la forme $c\partial F_i c^{-1}$ avec c un chemin qui part de m et rejoint le bord de F_i ,
- pour toute fonction intégrable $f : G^{2g+f} \rightarrow \mathbb{C}$, on a

$$\int_{\mathcal{C}_G^G} f(H_{a_1}, H_{b_1}, \dots, H_{a_g}, H_{b_g}, H_{\ell_1}, \dots, H_{\ell_f}) d\mu_{T, G, \mathbb{G}} = \frac{1}{Z_{T, G}} \int_{G^{2g+f-1}} f(x_1, y_1, \dots, x_g, y_g, z_1, \dots, z_f) \prod_{1 \leq i \leq f} p_{|F_i|}(z_i) \prod_{\substack{1 \leq i \leq g \\ 1 \leq j \leq f-1}} dx_i dy_i dz_j, \quad (1.90)$$

où l'on a posé $z_f = z_{f-1}^{-1} \cdots z_1^{-1} [x_1, y_1] \cdots [x_g, y_g]$.

Exemple. On considère le graphe \mathbb{G} de la Fig. 1.14 sur un tore d'aire totale T . Si l'on pose $t = |F_1|$ et si l'on se donne $f : G \rightarrow \mathbb{C}$, on peut calculer $\mathbb{E}[f(H_{\ell_1})] = \int_{\mathcal{C}_G^G} f(H_{\ell_1}) d\mu_{T, G, \mathbb{G}}$ en utilisant la Prop. 1.4.5 et cela donne

$$\mathbb{E}[f(H_{\ell_1})] = \frac{1}{Z_{T, G}} \int_{G^3} f(z) p_t(z) p_{T-t}(z^{-1}[x, y]) dx dy dz. \quad (1.91)$$

Nous verrons au chapitre 3 vers quelle quantité converge ce genre d'intégrale, lorsque $G = \text{U}(N)$ et que N tend vers l'infini, en prenant pour fonctions test les fonctions puissances $f : g \mapsto \text{tr}(g^n)$, $n \in \mathbb{N}$.

La Prop. 1.4.5 admet une version simplifiée lorsque M est la sphère ou un disque, à savoir que le groupe $\mathcal{L}_m^{\text{red}}(M)$ admet pour base (en tant que groupe libre) des lassos autour de chaque face sauf une (dans le cas du plan, on exclut en règle générale la face non bornée). Sous la mesure de Yang–Mills, les holonomies le long de ces lassos sont des variables aléatoires à valeurs dans G , indépendantes et dont la densité est le noyau de la chaleur $p_{|F|}$ pris au temps correspondant à l'aire de la face associée au lasso. Cela revient, en quelque sorte, à considérer le cas d'un graphe admissible sur une surface compacte dont on ferait tendre l'aire d'une face vers l'infini. En effet, dans l'intégrale (1.90) la variable z_f exprime la contrainte de bord de la face "extérieure", et si on l'exclut de la fonction test tout en faisant tendre $|F_f|$ vers l'infini on obtient la même factorisation qui caractérise l'indépendance des variables

⁴⁵Les générateurs relatifs à cette présentation sont appelés *tame generators* dans [Lévio].

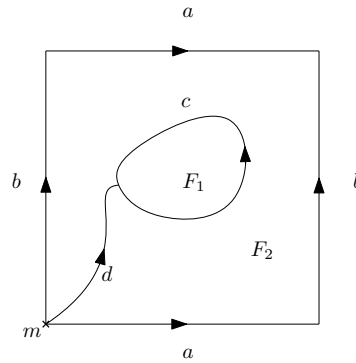


Figure 1.14: Un exemple de graphe admissible sur le tore : le graphe possède deux faces F_1 et F_2 , et $\mathcal{L}_m^{\text{red}}(\mathbb{G})$ admet pour présentation $\langle a, b, \ell_1, \ell_2 \mid \ell_2 = \ell_1^{-1}[a, b] \rangle$, où $\ell_1 = dcd^{-1}$.

de lassos dans le cas plan. Pour illustrer cette idée de “face infinie”, reprenons l’exemple de la Fig. 1.14 en faisant tendre $|F_2|$ vers l’infini. On obtient alors un graphe plongé dans le plan, donné par la Fig. 1.15, et l’équation (1.91) devient⁴⁶

$$\mathbb{E}[f(H_{\ell_1})] = \int_G f(z)p_t(z)dz. \quad (1.92)$$

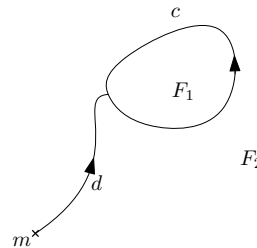


Figure 1.15: Le graphe obtenu en faisant tendre l’aire de F_2 vers l’infini dans le graphe de la Fig. 1.14.

On reconnaît alors la loi du mouvement brownien sur G , pris au temps correspondant à l’aire t de la face F_1 . Cela donne peut-être une vision plus claire de ce que représente le champ d’holonomie de Yang–Mills sur le plan : c’est le processus qui, pour tout graphe \mathbb{G} dans le plan, attribue à chaque face F du graphe un mouvement brownien sur G pris au temps $|F|$ de sorte que les mouvements browniens associés à toutes les faces soient indépendants.

En revanche, la description du champ de Yang–Mills en termes de mouvement brownien est moins claire pour une surface compacte de genre g , de par la relation qui unit les lassos de chaque face et les générateurs du groupe fondamental. Reprenons par exemple le cas du lacet simple dans un tore, selon la configuration de la Fig. 1.14. On note a et b les deux générateurs du groupe fondamental du tore, et on peut montrer que les variables associées H_a et H_b sont des unitaires de Haar (on va même en montrer une version généralisée au Thm. 3.2.16). On note $\ell = dcd^{-1}$ le lasso associé à la face F_1 . Pour toute fonction test $f : G \rightarrow \mathbb{C}$, on a vu l’expression de $\mathbb{E}[f(H_\ell)]$ à l’équation (1.91). Sa loi correspond à celle au temps $t = |F_1|$ d’un mouvement brownien sur G issu de e , conditionné à avoir

⁴⁶Notons qu’on admet ici que la limite de $Z_{T,G}$, quand T tend vers l’infini, est de 1 ; cela se démontre par exemple en utilisant le développement de $Z_{T,G}$ le long des caractères de G .

la même loi au temps T que $[H_a, H_b]$. En d'autres termes, H_ℓ a la loi d'un "pont brownien sur G entre l'élément neutre et le commutateur de deux unitaires de Haar". Dans le chapitre 3 nous verrons dans quelle mesure la limite des moments du champ d'holonomie de Yang–Mills avec pour groupe de structure $U(N)$ (resp. $SU(N)$) s'approche de celle des moments du mouvement brownien unitaire (resp. spécial unitaire).

1.4.3 Autour de la fonction de partition

Nous avons vu au paragraphe précédent que pour une surface compacte, la fonction de partition de Yang–Mills ne dépendait pas du graphe choisi. Nous allons en donner une expression compatible avec l'étude asymptotique que nous effectuerons au chapitre 2. Cette expression peut notamment être trouvée dans [LM15] pour la sphère et dans [Wit91] pour les surfaces de genre supérieur, mais nous allons la redémontrer en détail, d'une part car cela illustre les résultats évoqués dans la section sur la théorie des représentations, et d'autre part car certains éléments de démonstration seront réutilisés dans le chapitre 3 pour calculer des boucles de Wilson.

Puisque la fonction de partition de Yang–Mills sur une surface ne dépend pas du graphe, il est raisonnable de considérer un graphe le plus simple possible. Nous allons distinguer deux situations :

- Dans le cas de la sphère, on peut tracer un lacet simple qui la sépare en deux faces d'un graphe, et le lacet constitue alors l'unique arête du graphe ;
- Dans le cas d'une surface orientable ou non orientable de genre $g \geq 1$, on peut prendre pour graphe le domaine fondamental, constitué alors d'une seule face polygonale à $4g$ (resp. $2g$) côtés dans le cas orientable (resp. non orientable).

On notera $Z_N(g, T)$ (resp. $Z_N^-(g, T)$) la fonction de partition de Yang–Mills sur une surface compacte orientable (resp. non orientable) de genre g . Il est possible d'exprimer toutes ces fonctions de partitions à l'aide des représentations irréductibles de $U(N)$, plus précisément la décomposition de Fourier du noyau de la chaleur et l'indicateur de Frobenius–Schur ι_λ défini en (1.45).

Avant de donner les formules des fonctions de partition selon l'orientabilité et le genre de la surface, mentionnons un résultat utile sur les représentations des groupes compacts.

Proposition 1.4.6 ([Faro8], Prop.5.2). *Soit (ρ, V) une représentation irréductible d'un groupe compact G . Alors*

$$\int_G \chi_\rho(xgyg^{-1})dg = \frac{1}{d_\rho} \chi_\rho(x)\chi_\rho(y). \quad (1.93)$$

Cette proposition implique le résultat suivant, qui va nous servir à calculer les fonctions de partition de surfaces de genre non nul.

Corollaire 1.4.7. *Soit (ρ, H) une représentation irréductible unitaire d'un groupe compact G de caractère χ . Alors pour tout $x \in G$ on a les équations*

$$\int_{G^2} \chi(x[y, z])dydz = \frac{\chi(x)}{d_\rho^2}, \quad (1.94)$$

et

$$\int_{G^2} \chi(xy^2z^2)dydz = \frac{\iota_\rho}{d_\rho} \chi(x). \quad (1.95)$$

Preuve. L'équation (1.94) découle directement de la proposition précédente et des relations d'orthogonalité de Schur, en utilisant la décomposition de χ dans une base orthonormale de H . Pour montrer (1.95), on commence par effectuer un changement de variable $y \leftarrow yz^{-1}$, ce qui donne à l'aide de la Prop. 1.4.6

$$\int_{G^2} \chi(xy^2z^2) dydz = \int_{G^2} \chi(xyz^{-1}yz) dydz = \frac{1}{d_\rho} \int_G \chi(xy) \chi(y) dy.$$

On utilise ensuite le fait que χ soit centrale pour écrire $\chi(y) = \int_G \chi(z^{-1}yz) dz$, ce qui donne

$$\frac{1}{d_\rho} \int_G \chi(xy) \chi(y) dy = \frac{1}{d_\rho} \int_{G^2} \chi(xy) \chi(z^{-1}yz) dydz.$$

En utilisant la Prop. 1.4.6 à nouveau, il vient

$$\frac{1}{d_\rho} \int_{G^2} \chi(xy) \chi(z^{-1}yz) dydz = \int_{G^3} \chi(xyz^{-1}yzg^{-1}) dydzdg.$$

On effectue alors le changement de variable $z \leftarrow zg$, ce qui donne

$$\int_{G^3} \chi(xyz^{-1}yzg^{-1}) dydzdg = \int_{G^3} \chi(xyz^{-1}yz) dydzdg.$$

On a alors montré l'équation suivante :

$$\int_{G^2} \chi(xy^2z^2) dydz = \int_{G^2} \chi(xyz^{-1}yz) dydz.$$

En utilisant la Prop. 1.4.6 et la définition de ι_ρ , on en déduit bien l'équation (1.95). \square

Bien que nous ne rentrions pas en détails dans ces considérations topologiques chères aux physiciens, mentionnons tout de même que les équations (1.94) et (1.95) possèdent une interprétation en termes de chirurgie des surfaces, et permettent de relier la fonction de partition des sommes connexes de certaines surfaces à partir de celle des surfaces de départ. C'est notamment à l'aide de cette chirurgie que Witten démontre les formules qui vont suivre dans [Wit91], et on peut également retrouver des raisonnements similaires dans [Lévio, Lévo3], voire dans [LéVII] où cette correspondance entre chirurgie des surfaces et fonctions de partitions est reliée aux théories quantiques des champs topologiques (ou TQFT, pour *topological quantum field theories* en anglais). Quant à nous, nous nous contenterons d'un exemple informel qui illustre l'analogie entre l'équation (1.95) avec $x = e$ et la chirurgie de la bouteille de Klein, et pour l'illustrer nous emprunterons les illustrations très éclairantes de [Sti93].

La bouteille de Klein est une surface non-orientable de genre 2 obtenue par recollement d'un cylindre comme indiqué à la Fig. 1.16. Cependant, son domaine fondamental ne correspond pas au polygone à l'origine de cette construction ; pour cela il faut ajouter une arête au graphe de départ, puis découper ce graphe le long de cette arête et le recoller le long d'une autre arête. Ces étapes sont illustrées à la Fig. 1.17. On peut vérifier que la démonstration de l'équation (1.95) repose sur les mêmes étapes.

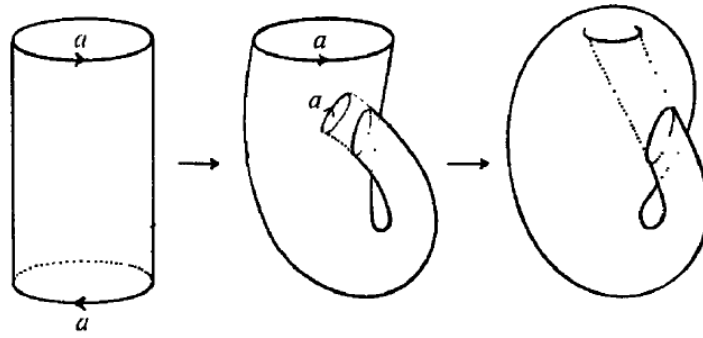


Figure 1.16: En recollant les bords du cylindre de gauche, on obtient la bouteille de Klein de droite. L'étape du milieu montre que pour préserver l'orientation du bord recollé la surface doit se traverser elle-même.

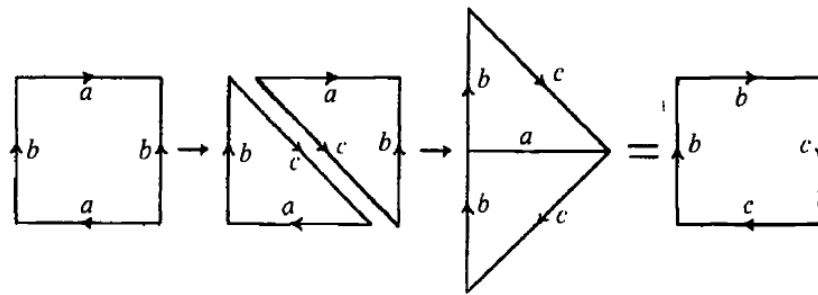


Figure 1.17: Le polygone (à gauche) correspondant à la construction de la bouteille de Klein de la Fig. 1.16 peut être transformé en le domaine fondamental de la bouteille de Klein (à droite) par découpage et recollement.

Proposition 1.4.8. Soit $M \in \Sigma$ une surface d'aire totale T . La fonction de partition de Yang-Mills sur M avec pour groupe de structure $U(N)$ est donnée par :

(i) Si M est une sphère,

$$Z_N(0, T) = \sum_{\lambda \in \widehat{U}(N)} e^{-c_2(\lambda) \frac{T}{2}} d_\lambda^2; \tag{1.96}$$

(ii) Si M est une surface orientable de genre $g \geq 1$,

$$Z_N(g, T) = \sum_{\lambda \in \widehat{U}(N)} e^{-c_2(\lambda) \frac{T}{2}} d_\lambda^{2-2g}; \tag{1.97}$$

(iii) Si M est une surface non-orientable de genre $g \geq 1$,

$$Z_N^-(g, T) = \sum_{\lambda \in \widehat{U}(N)} e^{-c_2(\lambda) \frac{T}{2}} d_\lambda^{2-g} (\iota_\lambda)^g. \tag{1.98}$$

Preuve. (i) On considère un lacet simple ℓ sur la sphère, qui sépare celle-ci en deux faces d'aires respectives t et $T - t$. On suppose sans perdre de généralités que ℓ est orienté positivement, en tant que

bord de la face d'aire t , et négativement, en tant que bord de la face d'aire $T - t$. Pour toute fonction $f : U(N) \rightarrow \mathbb{C}$, on a

$$\mathbb{E}[f(H_\ell)] = \frac{1}{Z_N(0, T)} \int_{U(N)} f(x) p_t(x) p_{T-t}(x^{-1}) dx,$$

et comme pour la fonction constante $f = 1$ il est clair que $\mathbb{E}[f(H_\ell)] = 1$, on obtient que

$$Z_N(0, T) = \int_{U(N)} p_t(x) p_{T-t}(x^{-1}) dx,$$

ce qui se simplifie en

$$Z_N(0, T) = p_T(I_N)$$

d'après la propriété de semi-groupe du noyau de la chaleur sur $U(N)$. On applique alors le Thm. 1.3.24 et cela donne la formule voulue, en utilisant le fait que $s_\lambda(I_N) = d_\lambda$.

(ii) On se donne des générateurs $(a_1, b_1, \dots, a_g, b_g)$ du groupe fondamental de M , et le graphe constitué d'une face F d'aire T , dont le bord orienté positivement s'écrit $\partial F = [a_1, b_1] \cdots [a_g, b_g]$. On a vu que cette face correspond au domaine fondamental de M . Il vient alors, d'après la Prop. 1.4.5, que la fonction de partition s'écrit

$$Z_N(g, T) = \int_{U(N)^{2g}} p_T([x_1, y_1] \cdots [x_g, y_g]) dx_1 dy_1 \cdots dx_g dy_g.$$

Cela se réécrit, à l'aide du Thm. 1.3.24 :

$$Z_N(g, T) = \sum_{\lambda \in \widehat{U}(N)} e^{-c_2(\lambda) \frac{T}{2}} d_\lambda \int_{U(N)^{2g}} s_\lambda([x_1, y_1] \cdots [x_g, y_g]) dx_1 dy_1 \cdots dx_g dy_g.$$

En appliquant l'équation (1.94) g fois, on obtient que

$$\int_{U(N)^{2g}} s_\lambda([x_1, y_1] \cdots [x_g, y_g]) dx_1 dy_1 \cdots dx_g dy_g = \frac{1}{d_\lambda^{2g-1}}.$$

On en déduit bien la formule voulue.

(iii) On se donne des générateurs (a_1, \dots, a_g) du groupe fondamental de M , et le graphe constitué d'une face F d'aire T , dont le bord orienté positivement s'écrit $\partial F = a_1^2 \cdots a_g^2$. Comme pour le point (ii) la fonction de partition s'écrit alors

$$Z_N^-(g, T) = \int_{U(N)^g} p_T(x_1^2 \cdots x_g^2) dx_1 \cdots dx_g.$$

En utilisant l'équation (1.95) g fois, on obtient

$$\int_{U(N)^g} s_\lambda(x_1^2 \cdots x_g^2) dx_1 \cdots dx_g = \frac{(\iota_\lambda)^g}{d_\lambda^{g-1}}.$$

On en déduit, comme pour le point (ii), la formule voulue. □

L'étude asymptotique de la fonction de partition sur la sphère a été l'objet de recherches des physiciens Douglas et Kazakov [DK93], qui ont constaté que l'énergie libre de la mesure de Yang–Mills sur la sphère d'aire T , c'est-à-dire la quantité

$$F(T) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \log Z_N(0, T)$$

admettait une *transition de phase* à la valeur $T = \pi^2$: sa dérivée troisième possédait un saut en cette valeur. Ce résultat a été démontré rigoureusement par Lévy et Maïda [LM15] sous la forme suivante, à l'aide de la théorie des grandes déviations et de la théorie de la minimisation sous contrainte.

Théorème 1.4.9 (Transition de phase de Douglas–Kazakov). *Pour tout $T \geq 0$, la quantité $F(T)$ existe. Cela définit une fonction $F : \mathbb{R}_+^* \rightarrow \mathbb{R}$ qui est de classe \mathcal{C}^2 sur \mathbb{R}_+^* et de classe \mathcal{C}^∞ sur $\mathbb{R}_+^* \setminus \{\pi^2\}$. Par ailleurs, la dérivée troisième de F admet un saut de $-\frac{2}{\pi^6}$ en π^2 :*

$$\lim_{\substack{T \rightarrow \pi^2 \\ T < \pi^2}} F^{(3)}(T) = -\frac{1}{\pi^6} \text{ et } \lim_{\substack{T \rightarrow \pi^2 \\ T > \pi^2}} F^{(3)}(T) = -\frac{3}{\pi^6}. \quad (1.99)$$

En ce qui concerne les fonctions de partition associées aux autres surfaces compactes, leur étude asymptotique est bien plus récente, et suit un régime totalement différent puisqu'elles admettent une limite finie. Ce résultat a été démontré dans [Lem19], et fait l'objet du chapitre 2. Les limites en question font apparaître deux formes modulaires classiques en théorie des nombres. La première est la *fonction d'Euler* définie sur le disque unité ouvert D par

$$\phi(q) = \prod_{m=1}^{\infty} (1 - q^m), \quad \forall q \in D,$$

et la seconde est *fonction thêta de Jacobi* définie sur $\mathbb{C} \times \mathbb{H}^+$, où \mathbb{H}^+ est le demi-plan de Poincaré, par

$$\vartheta(z, \tau) = \sum_{n \in \mathbb{Z}} e^{i\pi n^2 \tau + 2i\pi n z}, \quad \forall (z, \tau) \in \mathbb{C} \times \mathbb{H}^+.$$

Plus précisément, nous démontrerons dans ce chapitre les deux théorèmes suivants. Dans ceux-ci, on considère aussi bien les fonctions de partition $Z_N(g, T)$ et $Z_N^-(g, T)$ avec pour groupe de structure $U(N)$ que les fonctions de partition $Z_N'(g, T)$ et $Z_N'^-(g, T)$ avec pour groupe de structure $SU(N)$, qui admettent une formulation similaire à celle de la Prop. 1.4.8.

Théorème 1.4.10 (Chap. 2, Thm. 2.1.2). *Soit Σ une surface compacte connexe sans bord orientable de genre g .*

(i) *Si $g \geq 2$, alors pour tout $T \in (0, +\infty)$, on a la convergence suivante :*

$$\lim_{N \rightarrow \infty} Z_N(g, T) = \theta(T/2) \text{ et } \lim_{N \rightarrow \infty} Z_N'(g, T) = 1. \quad (1.100)$$

De plus,

$$\lim_{N \rightarrow \infty} Z_N'(g, 0) = 1. \quad (1.101)$$

(ii) Si $g = 1$, alors pour tout $T \in (0, +\infty)$ on pose $q = e^{-\frac{T}{2}}$, et on a la convergence suivante :

$$\lim_{N \rightarrow \infty} Z_N(1, T) = \frac{\theta(\frac{T}{2})}{\phi(q)^2} \text{ et } \lim_{N \rightarrow \infty} Z'_N(1, T) = \frac{1}{\phi(q)^2}. \quad (\text{I.102})$$

Théorème I.4.II (Chap. 2, Thm. 2.I.3). Soit Σ une surface compacte connexe sans bord non orientable de genre g .

(i) Si $g \geq 3$, alors pour tout $T \in (0, +\infty)$, on a la convergence suivante :

$$\lim_{N \rightarrow \infty} Z_N^-(g, T) = \theta(T/2) \text{ et } \lim_{N \rightarrow \infty} Z'_N{}^-(g, T) = 1. \quad (\text{I.103})$$

De plus,

$$\lim_{N \rightarrow \infty} Z'_N{}^-(g, 0) = 1. \quad (\text{I.104})$$

(ii) Si $g = 2$, alors pour tout $T \in (0, +\infty)$ on pose $q = e^{-\frac{T}{2}}$, et on a la convergence suivante :

$$\lim_{N \rightarrow \infty} Z_N^-(2, T) = \frac{\theta(\frac{T}{2})}{\phi(q^2)} \text{ et } \lim_{N \rightarrow \infty} Z'_N{}^-(2, T) = \frac{1}{\phi(q^2)}. \quad (\text{I.105})$$

I.4.4 Le champ maître sur une surface compacte

L'idée d'étudier les aspects asymptotiques de la théorie de Yang–Mills sur $U(N)$ avec $N \rightarrow \infty$ remonte aux années 70, notamment avec l'article [tH74] dans lequel 't Hooft a émis la conjecture que la théorie de jauge associée à $SU(N)$ devait considérablement se simplifier, d'un point de vue combinatoire, lorsque N tend vers l'infini⁴⁷. Cela a petit à petit conduit à la définition d'un objet limite de la mesure de Yang–Mills sur une surface, appelé *champ maître*. Celui-ci a fait l'objet de nombreuses recherches dans les années 90 dans le cas du plan [GG95, Gop96, Sin95, Xu97], puis dans les années 2000-2010 [Sen08b, AS12, Lévi17, Dahi16, CDG17] dans un cadre plus général. Il a par ailleurs été étudié sur la sphère, par exemple dans [FMS11, LM15], avant d'être construit explicitement dans [DN17], mais son existence dans le cadre d'autres surfaces compactes est encore un problème ouvert⁴⁸. En guise de première étape, Hall [Hal18] a permis, sous d'importantes hypothèses de convergence de traces d'holonomies sur des laces simples, d'en déduire l'existence du champ maître sur des surfaces plus générales, mais restreint à des lacets contenus dans un disque topologique.

Une première approche du champ maître, suggérée par Singer [Sin95], est de le définir comme une connexion A_∞ sur un G_∞ -fibré principal P_∞ que pour tout lacet c avec un nombre fini de points d'auto-intersection,

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_c)] = \text{tr}_\infty(\text{hol}(A_\infty, c)), \quad (\text{I.106})$$

⁴⁷Pour être plus précis, bien que nous ne développerons pas les notions citées, les diagrammes planaires dominants, lorsque N tend vers l'infini, sont ceux qui possèdent des quarks sur les arêtes. 't Hooft établit également un lien entre cet ensemble de diagrammes planaires dominants et certains types de cordes. Ces considérations bien plus physiques que mathématiques dépassent néanmoins largement le cadre de cette thèse.

⁴⁸Il semble que son existence soit pourtant admise par les physiciens : elle est décrite comme un théorème par Singer [Sin95] mais celui-ci a annoncé que les détails d'un développement mathématique de cette construction étaient en cours d'élaboration – sans donner suite jusqu'à présent.

où tr_∞ est un caractère du groupe G_∞ . Cette définition est reprise dans le cas de la sphère dans [DN17], et elle est autant intuitive qu'elle est peu commode car il est difficile de rendre explicites P_∞ et G_∞ ; par exemple lorsqu'on considère le champ de Yang–Mills avec pour groupe de structure $U(N)$, rien ne garantit que le groupe G_∞ associé soit le groupe $U(\infty)$ défini comme limite inductive, et même si c'était le cas, il s'avère que $U(\infty)$ est un groupe pour lequel l'analyse harmonique est difficile à effectuer : le lecteur intéressé pourra consulter les travaux d'Olshanski [Olso3, BO05, GO16] sur le sujet. Toutefois, il est possible de contourner cette approche en utilisant le langage des probabilités non-commutatives.

Définition 1.4.5. Soit $M \in \Sigma \cup \{\mathbb{R}^2\}$ une surface et m un point de M . Un *champ maître* sur M est la donnée d'un $*$ -espace de probabilité (\mathcal{A}, τ) et d'un processus stochastique non-commutatif $(h_\ell)_{\ell \in \mathcal{L}_m(M)}$ sur \mathcal{A} qui vérifie les hypothèses suivantes :

(i) Pour tout $(\ell_1, \ell_2) \in \mathcal{L}_m(M)^2$,

$$h_{\ell^{-1}} = h_\ell^* = h_\ell^{-1} \text{ et } h_{\ell_1 \ell_2} = h_{\ell_2} h_{\ell_1} ; \quad (1.107)$$

(ii) Si $(\ell_n)_{n \in \mathbb{N}}$ est une suite de lacets qui converge uniformément vers un lacet ℓ , alors (h_{ℓ_n}) converge vers h_ℓ dans $L^2(\mathcal{A}, \tau)$.

(iii) Soit $(H_c)_{c \in \mathcal{P}(M)}$ le champ d'holonomie de Yang–Mills défini au Thm. 1.4.3 avec pour groupe de structure $U(N)$ ou $SU(N)$. Pour tout lacet $\ell \in \mathcal{L}_m(M)$ on a

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_\ell)] = \tau(h_\ell). \quad (1.108)$$

Cette définition, empruntée à [Lévi17], dit en réalité probablement les mêmes choses que celle de Singer vue plus haut, mais son avantage est de ne pas nécessiter une explicitation du $*$ -espace de probabilité, lequel devrait, toujours d'après [Sin95], être lié à un "anneau de Murray–von Neumann d'opérateurs"⁴⁹ de type II_1 pour une représentation donnée de type II_1 de G_∞ .

Le point (iii) de la Déf. 1.4.5 signifie que le champ maître peut être vu comme le processus limite, au sens des probabilités libres, du champ d'holonomie de Yang–Mills. La démonstration de convergences de la forme (1.108) sera au cœur du chapitre 3. Pour des lacets simples notamment, c'est-à-dire des lacets qui ne possèdent pas de point d'auto-intersection, on peut montrer que pour une surface compacte de genre $g \geq 1$ la décomposition de Fourier de $\mathbb{E}[\text{tr}(H_\ell)]$ suit le même régime que celle de la fonction de partition de Yang–Mills, à savoir que seule une catégorie notable de plus hauts poids contribue à la limite – les *plus hauts poids presque-plats* que nous introduirons dès le chapitre 2.

En dépit du postulat actuel en physique théorique selon lequel le champ maître existe pour tous types de surface (voir note 48), seules deux surfaces ont reçu un traitement complet et rigoureux d'un point de vue mathématique, et ce, dans les années 2010 :

- Le plan, qui a été notamment traité d'un côté par Lévy [Lévi17] et de l'autre par Anshelevich et Sengupta [AS12] ;
- La sphère, qui a été traitée dans son intégralité par Dahlqvist et Norris [DN17].

⁴⁹Cette dénomination désigne en réalité une algèbre de von Neumann.

Pour ces deux surfaces, la construction du champ maître s'est opérée de façon similaire, selon un procédé que nous détaillerons au début du chapitre 3. Pour l'instant, nous nous contentons d'en dévoiler un ingrédient dont nous avons encore omis l'existence : les *équations de Makeenko–Migdal*. Elles permettent de relier l'holonomie le long d'un lacet possédant n points d'auto-intersection aux holonomies le long de deux lacets possédant $n - 1$ points d'auto-intersection. La construction du champ maître repose alors sur une sorte de récurrence dont l'initialisation est l'étude des lacets simples.

Pour introduire les équations de Makeenko–Migdal, il nous faut présenter le cadre formel dans lequel elles sont satisfaites. On choisit une surface $M \in \Sigma \cup \{\mathbb{R}^2\}$ et un graphe \mathbb{G} sur cette surface.

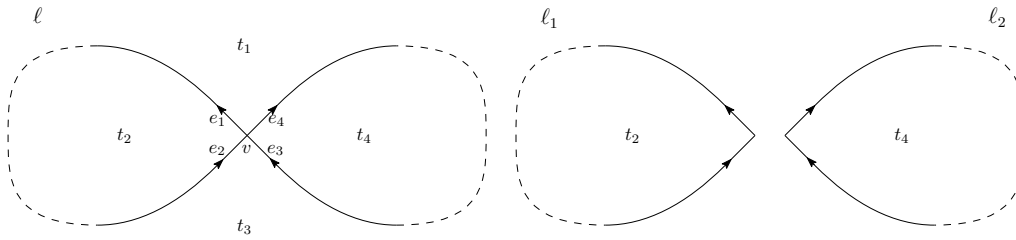


Figure 1.18: Le lacet ℓ (à gauche) peut être séparé, au niveau de son point d'auto-intersection v , en deux lacets ℓ_1 et ℓ_2 (à droite).

Si ℓ est un lacet sur M obtenu en concaténant des arêtes de \mathbb{G} et leurs inverses, et s'il possède un point d'auto-intersection en $v \in S$, alors on peut le séparer en deux lacets ℓ_1 et ℓ_2 comme dans la Fig. 1.18, et ceux-ci possèdent un point d'auto-intersection de moins que ℓ . Les faces F_1, F_2, F_3 et F_4 dont les bords communs sont traversés par ℓ et qui se rejoignent en v peuvent être toutes distinctes, ou bien certaines peuvent être éventuellement confondues. On note respectivement t_1, t_2, t_3 et t_4 leurs aires, vues comme des variables réelles positives. On a vu que $\mathbb{E}[\text{tr}(H_\ell)]$ dépend notamment de ces variables ; l'équation de Makeenko–Migdal associée au découpage de ℓ en ℓ_1 et ℓ_2 s'écrit

$$\left(\frac{\partial}{\partial t_1} - \frac{\partial}{\partial t_2} + \frac{\partial}{\partial t_3} - \frac{\partial}{\partial t_4} \right) \mathbb{E}[\text{tr}(H_\ell)] = \mathbb{E}[\text{tr}(H_{\ell_1})\text{tr}(H_{\ell_2})]. \quad (\text{I.109})$$

Une des conjectures concernant le champ maître est la suivante : pour tous lacets ℓ_1 et ℓ_2 , on a la factorisation asymptotique

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_{\ell_1})\text{tr}(H_{\ell_2})] = \lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_{\ell_1})] \lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_{\ell_2})]. \quad (\text{I.110})$$

Cette factorisation est le fruit d'un phénomène de concentration des traces d'holonomies autour de valeurs déterministes. Par conséquent, les moments du champ maître, s'ils existent, définissent une fonction Φ sur les lacets qui satisfait la version asymptotique suivante des équations de Makeenko–Migdal pour un lacet $\ell(t_1, t_2, t_3, t_4)$ qui se sépare en $\ell_1(t_2)$ et $\ell_2(t_4)$ comme dans la Fig. 1.18 :

$$\left. \frac{d}{dt} \right|_{t=0} \Phi_{\ell(t_1+t, t_2-t, t_3+t, t_4-t)} = \Phi_{\ell_1(t_2)} \Phi_{\ell_2(t_4)}. \quad (\text{I.111})$$

Cette fonction permet de décrire intégralement le champ maître. En effet, soit ℓ un lacet sur M et n un entier naturel, alors $\tau(h_\ell^n) = \tau(h_{\ell^n}) = \Phi_{\ell^n}$ où ℓ^n est simplement le lacet qui parcourt n fois

ℓ . Notons que si le champ maître h est une fonction multiplicative, ce n'est toutefois pas le cas de Φ *a priori* : en effet, rien ne dit par exemple que $\tau(h_\ell^2) = \tau(h_\ell)^2$ – la multiplicativité est remplacée par la version asymptotique des équations de Makeenko–Migdal. La propriété de factorisation asymptotique des traces d'holonomies (1.110) implique par ailleurs un fait remarquable : la convergence (1.108) peut être renforcée en une convergence en probabilité quand N tend vers l'infini

$$\mathrm{tr}(H_\ell) \xrightarrow{\mathbb{P}} \Phi_\ell, \quad (1.112)$$

où l'on a alors la convergence de variables aléatoires complexes vers une constante réelle. Il nous est donc possible de donner une nouvelle description du champ maître, peut-être plus simple que celle de la Déf. 1.4.5.

Définition 1.4.6. Soit $M \in \Sigma \cup \{R^2\}$ une surface et m un point de M . Un *champ maître* sur M est une fonctionnelle $\Phi : \mathcal{L}_m(M) \rightarrow \mathbb{R}$ qui vérifie l'équation (1.112) pour tout $\ell \in \mathcal{L}_m(M)$.

Ainsi, l'accent est mis sur la convergence en probabilité des traces d'holonomies, aussi appelées *boucles de Wilson*, vers les valeurs du champ maître associée.

Comme annoncé un peu plus tôt, le chapitre 3 sera dédié à l'étude du champ maître. Nous y développerons brièvement la théorie utilisée pour le construire sur le plan et sur la sphère, puis nous l'étudierons sur les surfaces compactes connexes orientables de genre supérieur ou égal à 1. En particulier, nous démontrerons le théorème suivant, qui constitue un résultat inédit.

Théorème 1.4.12 (Chap. 3, Thm. 3.2.5 et 3.2.6, Prop. 3.2.13). *Soit $\Sigma_{g,T}$ une surface compacte connexe sans bord orientable de genre $g \geq 1$ et d'aire T . Soit ℓ un lacet simple contractile entourant un domaine d'aire t . Lorsque le groupe de structure est $U(N)$ ou $SU(N)$ avec N qui tend vers l'infini, la trace du champ d'holonomie de Yang–Mills sur $\Sigma_{g,T}$ associé au lacet ℓ vérifie*

$$\lim_{N \rightarrow \infty} \mathbb{E}[\mathrm{tr}(H_\ell)] = e^{-t/2}, \quad (1.113)$$

$$\lim_{N \rightarrow \infty} \mathrm{Var}[\mathrm{tr}(H_\ell)] = 0. \quad (1.114)$$

De plus, la vitesse de convergence de la variance est de l'ordre de $N^{\varepsilon-1}$, pour $\varepsilon > 0$ arbitrairement petit.

Nous verrons par ailleurs que ce théorème permet de vérifier une conjecture de Hall [Hal8] que nous rappellerons. Nous discuterons ensuite d'autres types de lacets que l'on rencontre dans ces nouvelles surfaces, à savoir des lacets d'homologie non nulle. Nous montrerons des résultats partiels relatifs à ceux-ci, notamment le théorème suivant.

Théorème 1.4.13 (Chap. 3, Thm. 3.2.16 et Prop. 3.2.17). *Soit $\Sigma_{g,T}$ une surface compacte connexe sans bord orientable de genre $g \geq 1$ et d'aire T .*

(i) *Si ℓ est un lacet d'homologie non nulle sur $\Sigma_{g,T}$ alors pour tout $n \in \mathbb{Z}^*$, $\mathbb{E}[\mathrm{tr}(H_\ell^n)] = 0$.*

(ii) *Si ℓ est un générateur du groupe fondamental de $\Sigma_{g,T}$, alors $\lim \mathrm{Var}[\mathrm{tr}(H_\ell)] = 0$.*

Nous terminerons par une description des lacets qui échappent encore au champ maître, ainsi que quelques pistes de réflexion que nous envisageons.

YANG–MILLS MEASURE AND LARGE N PARTITION FUNCTION

This chapter describes the results obtained in [Lem19], which are essentially the computations of several limits associated to the Yang–Mills partition function when the gauge group is $U(N)$ or $SU(N)$. After giving the definition of the Yang–Mills partition function, we will introduce two major tools needed to study its asymptotics: the *Witten zeta function* and the *almost flat highest weights*. The rest of the chapter will be devoted to the computation of the limits of partition functions.

2.1 THE YANG–MILLS PARTITION FUNCTION ON A COMPACT SURFACE

First and foremost, let us mention that, throughout this chapter, a *surface* corresponds to *compact connected surface without boundary*. This choice is motivated by Thm. 1.4.1, that we recall here.

Theorem 2.1.1 (Classification of compact surfaces). *Let Σ be a surface. Then, it is homeomorphic to either one of the following:*

- (i) *The connected sum of g 2-tori¹,*
- (ii) *The connected sum of g projective planes.*

In the first case, the surface is said to be *orientable*. The integer g denotes the *genus* of the surface, and appears for example in the computation of the Euler characteristic. In the second one, the surface is said to be *non-orientable*, and we will also call g the *genus*; however it will not contribute the same way to the Euler characteristic. Indeed, this number is equal to $2 - 2g$ when the surface is orientable and $2 - g$ when it is not. Let us recall a few definitions from the previous chapter, in order to get a quick overview of the objects of interest. We want to define the Yang–Mills partition function in a way that we can compute its limit: for that, we will need two quantities related to the irreducible representations of $U(N)$ or $SU(N)$.

Definition 2.1.1. Let $\lambda \in \widehat{U}(N)$ be a non-increasing N -tuple of integers and $\mu \in \widehat{SU}(N)$ be a non-increasing N -tuple of integers such that $\mu_N = 0$.

¹If $g = 0$ then by convention it is a sphere; otherwise it can also be seen as a 2-torus with g handles.

(i) The *dimension* d_λ of a representation of highest weight λ is equal to²

$$d_\lambda = \prod_{1 \leq i < j \leq N} \frac{\lambda_i - \lambda_j + j - i}{j - i} = \prod_{1 \leq i < j \leq N} \left(1 + \frac{\lambda_i - \lambda_j}{j - i} \right). \quad (2.1)$$

In the case of a representation of highest weight μ it is equal to

$$d_\mu = \prod_{1 \leq i < j \leq N} \frac{\mu_i - \mu_j + j - i}{j - i} = \prod_{1 \leq i < j \leq N} \left(1 + \frac{\mu_i - \mu_j}{j - i} \right). \quad (2.2)$$

(ii) The (quadratic) *Casimir number* is a non-negative real number $c_2(\lambda)$ that satisfies $\Delta s_\lambda = -c_2(\lambda)s_\lambda$, and is equal to³

$$c_2(\lambda) = \frac{1}{N} \left(\sum_{i=1}^N \lambda_i^2 + \sum_{1 \leq i < j \leq N} (\lambda_i - \lambda_j) \right). \quad (2.3)$$

in the case of $SU(N)$, it is denoted $c'_2(\mu)$ and is equal to

$$c'_2(\mu) = \frac{1}{N} \left(\sum_{i=1}^N \mu_i^2 - \frac{1}{N} \left(\sum_{i=1}^N \mu_i \right)^2 + \sum_{1 \leq i < j \leq N} (\mu_i - \mu_j) \right). \quad (2.4)$$

Given an orientable surface Σ with genus g and area T , the *Yang–Mills partition function* on Σ with structure group $U(N)$ is defined as the following sum⁴:

$$Z_N(g, T) = \sum_{\lambda \in \widehat{U}(N)} e^{-c_2(\lambda) \frac{T}{2}} d_\lambda^{2-2g}. \quad (2.5)$$

The same kind of formula also holds when one replaces $U(N)$ by $SU(N)$ as a structure group:

$$Z'_N(g, T) = \sum_{\lambda \in \widehat{SU}(N)} e^{-c'_2(\lambda) \frac{T}{2}} d_\lambda^{2-2g}. \quad (2.6)$$

Now, if the surface Σ is non-orientable, defined as the connected sum of g projective planes, we can still define its Yang–Mills partition function using a slightly different formula. If we defined ι_λ is the so-called *Frobenius–Schur indicator* of λ by

$$\iota_\lambda = \int_G s_\lambda(g^2) dg, \quad (2.7)$$

then we have

$$Z_N^-(g, T) = \sum_{\lambda \in \widehat{U}(N)} e^{-\frac{T}{2} c_2(\lambda)} d_\lambda^{2-g} (\iota_\lambda)^g, \text{ if the structure group is } U(N), \quad (2.8)$$

²Cf. Equation (1.74).

³Cf. Prop. 1.3.23.

⁴Cf. Prop. 1.4.8 for this formula and the following ones.

$$Z'_N(g, T) = \sum_{\lambda \in \widehat{\text{SU}}(N)} e^{-\frac{T}{2} c'_2(\lambda)} d_\lambda^{2-g} (\iota_\lambda)^g, \text{ if the structure group is } \text{SU}(N). \quad (2.9)$$

The purpose of this chapter is to compute the limits, as N tends to infinity, of the partition function on Σ depending on its orientability, genus and structure group. Let us introduce two functions that naturally appear in the limit, and that come from number theory – or more specifically, from the theory of modular forms.

Definition 2.1.2. Let $\mathbb{H}^+ = \{\tau \in \mathbb{C} : \Im(\tau) > 0\}$ be the complex upper-half plane and $D = \{q \in \mathbb{C} : |q| < 1\}$ the open unit disk.

(i) The *Jacobi theta function* is the function $\vartheta : \mathbb{C} \times \mathbb{H}^+ \rightarrow \mathbb{C}$ defined by

$$\vartheta(z; \tau) = \sum_{n \in \mathbb{Z}} e^{i\pi n^2 \tau + 2i\pi n z}. \quad (2.10)$$

(ii) The *Euler function* is the function $\phi : D = \{q \in \mathbb{C} : |q| < 1\} \rightarrow \mathbb{C}$ defined by

$$\phi(q) = \prod_{m=1}^{\infty} (1 - q^m). \quad (2.11)$$

As we only need special values of Jacobi theta function, let us also define an intermediate function

$$\theta : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ t & \mapsto \vartheta(0; \frac{it}{2\pi}) = \sum_{n \in \mathbb{Z}} e^{-tn^2} \end{cases}. \quad (2.12)$$

We can now state the two main theorems from [Lem19], which give the limit of the $\text{U}(N)$ and $\text{SU}(N)$ Yang–Mills partition function on an orientable surface of genus $g \geq 1$ and on a non-orientable surface of genus $g \geq 2$.

Theorem 2.1.2 (Orientable limits). *Let Σ be an orientable surface of genus g .*

(i) *If $g \geq 2$, then, for all $T \in (0, +\infty)$, the following convergences hold:*

$$\lim_{N \rightarrow \infty} Z_N(g, T) = \theta(T/2) \text{ and } \lim_{N \rightarrow \infty} Z'_N(g, T) = 1. \quad (2.13)$$

Moreover,

$$\lim_{N \rightarrow \infty} Z'_N(g, 0) = 1. \quad (2.14)$$

(ii) *If $g = 1$, then consider $T \in (0, +\infty)$ and set $q = e^{-\frac{T}{2}}$. The following convergences hold:*

$$\lim_{N \rightarrow \infty} Z_N(1, T) = \frac{\theta(\frac{T}{2})}{\phi(q)^2} \text{ and } \lim_{N \rightarrow \infty} Z'_N(1, T) = \frac{1}{\phi(q)^2}. \quad (2.15)$$

Theorem 2.1.3 (Non-orientable limits). *Let Σ be a non-orientable surface homeomorphic to the sum of g projective planes.*

(i) If $g \geq 3$, then, for all $T \in (0, +\infty)$, the following convergences hold:

$$\lim_{N \rightarrow \infty} Z_N^-(g, T) = \theta(T/2) \text{ and } \lim_{N \rightarrow \infty} Z_N'^-(g, T) = 1. \quad (2.16)$$

Moreover,

$$\lim_{N \rightarrow \infty} Z_N'^-(g, 0) = 1. \quad (2.17)$$

(ii) If $g = 2$, then consider $T \in (0, +\infty)$ and set $q = e^{-\frac{T}{2}}$. The following convergences hold:

$$\lim_{N \rightarrow \infty} Z_N^-(2, T) = \frac{\theta(\frac{T}{2})}{\phi(q^2)} \text{ and } \lim_{N \rightarrow \infty} Z_N'^-(2, T) = \frac{1}{\phi(q^2)}. \quad (2.18)$$

Before we prove these theorems we can highlight the following fact: there seems to be an asymptotic factorization

$$\lim_{N \rightarrow \infty} Z_N(g, T) = Z_1(g, T) \lim_{N \rightarrow \infty} Z_N'(g, T). \quad (2.19)$$

Indeed, it is easy to remark that $Z_1(g, T) = \theta(T/2)$ from Equation (2.5), and the factorization now seems clear. Our proofs will always rely on the fact that it is easier to compute the limits in the $SU(N)$ case and that we can deduce the $U(N)$ limit from it, but this factorization will not be a part of the reasoning. The closest fact we will use is the existence of a bijection between ‘dual sets’, understood as the sets of irreducible representations:

$$\widehat{SU}(N) \times \widehat{U}(1) \simeq \widehat{U}(N). \quad (2.20)$$

However, there is no such relationship between the underlying groups: all we know is that $SU(N) \times U(1)$ is a group covering of $U(N)$ with the homomorphism given by $\varphi_N : (M, z) \mapsto zM$. φ_N is actually a surjective Lie group morphism, with

$$\ker \varphi_N = \{(I, 1), (e^{2i\pi/N} I, e^{-2i\pi/N}), \dots, (e^{2i\pi(N-1)/N} I, e^{-2i\pi(N-1)/N})\} \simeq \mathbb{Z}^N.$$

In particular, it appears that φ_N factorizes into a Lie group isomorphism

$$\begin{array}{ccc} SU(N) \times U(1) & \xrightarrow{\varphi_N} & U(N) \\ \downarrow & \nearrow \Phi_N & \\ (SU(N) \times U(1)) / \mathbb{Z}^N & & \end{array} \quad (2.21)$$

Furthermore, another well-known Lie group homomorphism is given by $\det : U(N) \rightarrow U(1)$ with kernel $SU(N)$, which leads to the following short exact sequence

$$1 \longrightarrow SU(N) \xrightarrow{\iota} U(N) \xrightarrow{\det} U(1) \longrightarrow 1. \quad (2.22)$$

This short sequence means that $U(N)$ is a *group extension* of $U(1)$ by $SU(N)$. This is even a *split* extension because there exists a global section $s : U(1) \rightarrow U(N)$: we can take for example $s(z) = \text{diag}(z, 1, \dots, 1) \in U(N)$. There is a result on group theory that says that a split extension $A \rightarrow H \rightarrow G$ is equivalent to the fact that H is isomorphic to the semidirect product $A \rtimes_{\rho} G$ for some

action $\rho : G \times A \rightarrow A$. One can find a proof of this fact in [Ser78, §8.2].

Regarding these relations between $U(N)$, $SU(N)$ and $U(1)$, there might be a deeper explanation, or at least interpretation, of the factorization (2.19). It might also be probable that this has a link with the splitting of unitary Brownian motion stated in Lemma 1.2.5. Unfortunately, we didn't manage yet to relate the asymptotic factorization of partition functions to the splitting of Brownian motion or to a limiting behaviour of the diagrams (2.21) or (2.22).

2.2 ALGEBRAIC TOOLS

Despite the fact that Theorems 2.1.2 and 2.1.3 give the limit of a normalization constant of a probability measure, the core of our proof will rely more on algebraic tools rather than probabilistic ones. Let us introduce in this section those tools, namely the Witten zeta function, which is an infinite series associated with a Lie algebra, and the almost flat highest weights, which constitute a subset of highest weights that plays a fundamental role in the large N limit of partition functions.

2.2.1 Witten zeta function

Let \mathfrak{su}_N be the Lie algebra of the Lie group $SU(N)$. One defines its *Witten zeta function* $\zeta_{\mathfrak{su}_N} : \mathbb{C} \rightarrow \mathbb{C}$ as the meromorphic continuation of the series

$$\zeta_{\mathfrak{su}_N}(s) = \sum_{\lambda \in \widehat{SU}(N)} d_{\lambda}^{-s}. \quad (2.23)$$

If we consider the partition function $Z'_N(g, 0)$ as the limit of the partition function of an orientable surface with genus g and area T when T tends to 0, we see that it becomes a special value of Witten zeta function:

$$Z'_N(g, 0) = \zeta_{\mathfrak{su}_N}(2g - 2). \quad (2.24)$$

This zeta function, as we will see later, controls the convergence of all partition functions associated with orientable surfaces with genus $g \geq 2$ or non-orientable surfaces with genus $g \geq 3$; it was actually studied for different reasons by Witten [Wit91], and later by Zagier [Zag94] who presumably coined its name 'Witten zeta function'.

The following Proposition summarizes the asymptotic properties of Witten zeta functions that we will need later on.

Proposition 2.2.1. *For all real $s > 1$, one has*

$$\sup_{N \geq 1} \zeta_{\mathfrak{su}_N}(s) = \sup_{N \geq 1} \sum_{\lambda \in \widehat{SU}(N)} d_{\lambda}^{-s} < \infty.$$

More precisely,

$$\lim_{N \rightarrow \infty} \zeta_{\mathfrak{su}_N}(s) = 1 \quad \text{and} \quad \lim_{N \rightarrow \infty} \sum_{\substack{\lambda \in \widehat{SU}(N) \\ \lambda \neq (0, \dots, 0)}} d_{\lambda}^{-s} = 0.$$

The proof of this proposition relies on three lemmas.

Lemma 2.2.2. *For all $s > 1$ and all $N \geq 1$, one has*

$$\sum_{\lambda \in \widehat{\text{SU}}(N)} d_{\lambda}^{-s} \leq \prod_{k=1}^{N-1} \sum_{n \geq k} \binom{n}{k}^{-s}. \quad (2.25)$$

Proof. Let us choose $s > 1$ and $N \geq 1$. In the left-hand side of (2.25), which is a sum over $\lambda_1 \geq \dots \geq \lambda_N \geq 0$, let us make the change of variables

$$m_1 = \lambda_1 - \lambda_2 + 1, \dots, m_{N-1} = \lambda_{N-1} - \lambda_N + 1.$$

The new variables m_1, \dots, m_{N-1} are now independent, and positive. Using (2.2), we find

$$d_{\lambda} = \prod_{1 \leq i < j \leq N} \frac{m_i + \dots + m_{j-1}}{j - i}, \quad (2.26)$$

so that

$$\begin{aligned} \sum_{\lambda_1 \geq \dots \geq \lambda_N = 0} d_{\lambda}^{-s} &= \sum_{m_1, \dots, m_{N-1} \geq 1} \prod_{1 \leq i < j \leq N} \frac{(j - i)^s}{(m_i + \dots + m_{j-1})^s} \\ &= \sum_{m_1, \dots, m_{N-1} \geq 1} \prod_{k=1}^{N-1} \prod_{i=1}^k \frac{(k - i + 1)^s}{(m_i + \dots + m_k)^s} \quad (k = j - 1) \end{aligned}$$

Since $m_i + \dots + m_{k-1} \geq k - i$, we obtain

$$\begin{aligned} \sum_{\lambda_1 \geq \dots \geq \lambda_N = 0} d_{\lambda}^{-s} &\leq \sum_{m_1, \dots, m_{N-1} \geq 1} \prod_{k=1}^{N-1} \prod_{i=1}^k \frac{(k - i + 1)^s}{(m_k + k - i)^s} \\ &= \sum_{m_1, \dots, m_{N-1} \geq 1} \prod_{k=1}^{N-1} \binom{k + m_k - 1}{k}^{-s} \\ &= \prod_{k=1}^{N-1} \sum_{n \geq k} \binom{n}{k}^{-s}, \end{aligned}$$

which is the announced upper bound. □

Lemma 2.2.3. *For all real $s > 1$,*

$$\sum_{k \geq 1} \sum_{n > k} \binom{n}{k}^{-s} < \infty.$$

Proof. We use the fact that for k between 2 and $n - 2$, the inequality $\binom{n}{k} \geq \binom{n}{2}$ holds. Thus,

$$\sum_{k \geq 1} \sum_{n > k} \binom{n}{k}^{-s} \leq 2^{-s} + \sum_{n=3}^{\infty} \left(\frac{2}{n^s} + (n-3) \frac{2^s}{n^s (n-1)^s} \right)$$

which is indeed finite for $s > 1$. □

Lemma 2.2.4. *Let λ be an element of $\widehat{\text{SU}}(N)$. If $\lambda = (0, \dots, 0)$ then $d_\lambda = 1$, otherwise $d_\lambda \geq N$.*

proof. Let us use again the variables m_1, \dots, m_{N-1} introduced in the proof of Lemma 2.2.2. It is manifest on the expression (2.26) of d_λ that this dimension is increasing in each of the variables m_1, \dots, m_r . The case where each of these variables is equal to 1 is the case where $\lambda = (0, \dots, 0)$ and $d_\lambda = 1$. Any other irreducible representation has a dimension that is at least equal to the dimension of one of the representations

$$\lambda_1 = (1, 0, \dots, 0), \lambda_2 = (1, 1, 0, \dots, 0), \dots, \lambda_{N-1} = (1, \dots, 1, 0).$$

These representations, which are the exterior powers of the standard representation of $\text{SU}(N)$, have dimensions

$$d_{\lambda_k} = \binom{N}{k} \geq N, \quad k \in \{1, \dots, N-1\}.$$

Thus, $d_\lambda \geq N$, as expected. \square

We can now prove Proposition 2.2.1.

Proof of Proposition 2.2.1. The bound obtained in Lemma 2.2.2 can be rewritten as

$$\sum_{\lambda_1 \geq \dots \geq \lambda_N = 0} d_\lambda^{-s} \leq \prod_{k=1}^{N-1} \left[1 + \sum_{n>k} \binom{n}{k}^{-s} \right] \leq \exp \sum_{k=1}^{\infty} \sum_{n>k} \binom{n}{k}^{-s}$$

and this last bound, independent of N , is finite by Lemma 2.2.3. This proves the first assertion.

For the second, let us introduce a real $s' \in (1, s)$ and use Lemma 2.2.4. We find

$$\sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \neq (0, \dots, 0)}} d_\lambda^{-s} \leq N^{-(s-s')} \sum_{\lambda \in \widehat{\text{SU}}(N)} d_\lambda^{-s'},$$

which tends to 0 as N tends to infinity. \square

2.2.2 Almost flat highest weights

From two integer partitions $\alpha = (\alpha_1 \geq \dots \geq \alpha_r > 0)$ and $\beta = (\beta_1 \geq \dots \geq \beta_s > 0)$ of respective lengths r and s , and an integer $n \in \mathbb{Z}$, we can form, for all $N \geq r + s + 1$, the highest weight

$$\lambda_N(\alpha, \beta, n) = (\alpha_1 + n, \dots, \alpha_r + n, \underbrace{n, \dots, n}_{N-r-s}, n - \beta_s, \dots, n - \beta_1) \in \widehat{\text{U}}(N), \quad (2.27)$$

which we also denote by $\lambda(\alpha, \beta, n)$ when there is no doubt on the value of N . We extend this definition in the obvious way to the cases where one or both of the partitions α and β are the empty partition.

We can also form the highest weight

$$\lambda_N(\alpha, \beta) = \lambda_N(\alpha, \beta, \beta_1) \in \widehat{\text{SU}}(N),$$

with the convention that $\lambda_N(\alpha, \emptyset) = \lambda_N(\alpha, \emptyset, 0) = (\alpha_1, \dots, \alpha_r, 0)$.

These constructions are illustrated in Fig. 2.1 below. The reader may have noticed that the definition of $\lambda_N(\alpha, \beta, n)$ still makes sense when $N = r + s$ and wonder why we exclude this case. The reason is that under the stronger assumption $N \geq r + s + 1$, it is possible to recover α and β unambiguously from the data of $\lambda_N(\alpha, \beta, n)$, r and s . A counterexample with $r = s = 1$ and $N = 2$ is given in Fig. 2.2. Without the data of r and s , there are usually multiple ways of writing a highest weight in the form $\lambda_N(\alpha, \beta, n)$, see also Fig. 2.2. Finally, it should be emphasized that every highest weight of $U(N)$ or $SU(N)$ can be written as $\lambda_N(\alpha, \beta, n)$ or $\lambda_N(\alpha, \beta)$.

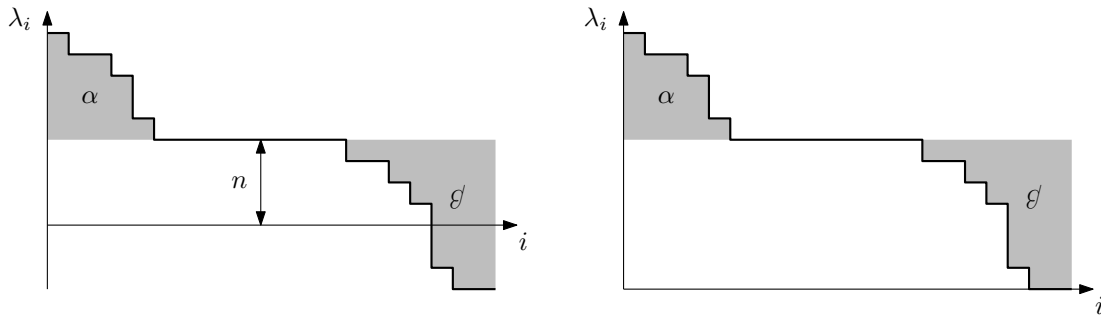


Figure 2.1: From two partitions α and β and an integer $n \in \mathbb{Z}$, we can form the highest weights $\lambda(\alpha, \beta, n) \in \widehat{U}(N)$ (on the left) and $\lambda(\alpha, \beta) \in \widehat{SU}(N)$ (on the right).

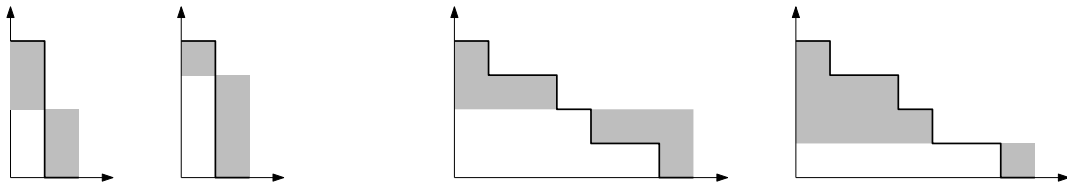


Figure 2.2: On the left: the highest weight $(4, 0)$ can be written in several ways as $\lambda_2(\alpha, \beta)$ with α and β of length 1. On the right: the highest weight $(4, 3, 3, 2, 1, 1, 0)$ is equal to $\lambda_7((2, 1, 1), (2, 1, 1))$ as well as to $\lambda_7((3, 2, 2, 1), (1))$.

The construction above has a kind of ‘dual process’, that is, the decomposition of any highest weight of $U(N)$ or $SU(N)$ as $\lambda(\alpha, \beta, n)$ or $\lambda(\alpha, \beta)$. In the case of $SU(N)$, this process is based on the observation that for all $M_1, M_2 \geq 0$, the map

$$\begin{aligned} \widehat{SU}(M_1 + 1) \times \widehat{SU}(M_2 + 1) &\xrightarrow{\sim} \widehat{SU}(M_1 + M_2 + 1) \\ (\alpha, \beta) &\longmapsto \lambda_{M_1+M_2+1}(\alpha, \beta) \end{aligned} \tag{2.28}$$

is a bijection. We have to make a slightly different construction depending on the parity of N : let $\lambda \in \widehat{SU}(N)$.

- If $N = 2M + 1$, then there is a unique $\alpha_\lambda \in \widehat{SU}(M + 1)$ and a unique $\beta_\lambda \in \widehat{SU}(M + 1)$ such that $\lambda = \lambda_N(\alpha_\lambda, \beta_\lambda)$, therefore we take $M_1 = M_2 = M$ in (2.28).

- If $N = 2M$, then there is a unique $\alpha_\lambda \in \widehat{\text{SU}}(M)$ and a unique $\beta_\lambda \in \widehat{\text{SU}}(M + 1)$ such that $\lambda = \lambda_N(\alpha_\lambda, \beta_\lambda)$, therefore we take $M_1 = M - 1$ and $M_2 = M$ in (2.28).

In the case of $\text{U}(N)$ we have a bijection

$$\Phi : \begin{cases} \widehat{\text{SU}}(N) \times \mathbb{Z} & \xrightarrow{\sim} \widehat{\text{U}}(N) \\ (\lambda, n) & \mapsto \lambda + n = (\lambda_1 + n, \dots, \lambda_N + n), \end{cases} \quad (2.29)$$

which guarantees that any highest weight of $\text{U}(N)$ can be written as $\lambda + n$ with $\lambda \in \widehat{\text{SU}}(N)$ and $n \in \mathbb{Z}$. Hence, we have the decomposition $\lambda + n = \lambda_N(\alpha_\lambda, \beta_\lambda) + n$ for any highest weight of $\text{U}(N)$.

The Casimir number of a highest weight can be expressed conveniently through this decomposition, as we will show below. First, let us recall the definition of the content of a box of a diagram, which is mentioned in particular in [Sta99, VO04].

Definition 2.2.1. Let $\alpha = (\alpha_1 \geq \dots \geq \alpha_r \geq 0)$ be a non-increasing sequence of integers, seen as a Young diagram. For any box (i, j) of this diagram, that is, any (i, j) such that $j \leq \alpha_i$, we call *content* of the box (i, j) the quantity $c(i, j) = j - i$. We also define the *total content* $K(\alpha)$ of α as the sum of the contents of the boxes of α .

An example is given on Fig. 2.3.

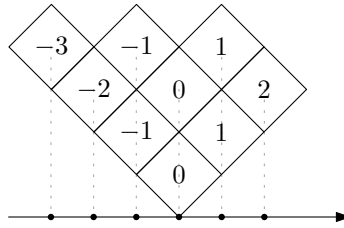


Figure 2.3: Filling of the boxes of $(3, 3, 2, 1)$ with their respective contents. The Young diagram is represented here in the so-called Russian way, where the content of a box is its abscissa.

The main result of this section is the following.

Proposition 2.2.5. Let α and β be two partitions of respective lengths r and s . Let n be an integer. Then, provided $N \geq r + s$, we have

$$c_2(\lambda_N(\alpha, \beta, n)) = |\alpha| + |\beta| + n^2 + \frac{2}{N}(K(\alpha) + K(\beta) + n(|\alpha| - |\beta|)) \quad (2.30)$$

in the unitary case, and

$$c'_2(\lambda_N(\alpha, \beta)) = |\alpha| + |\beta| + \frac{2}{N}(K(\alpha) + K(\beta)) + \frac{1}{N^2}(|\alpha| - |\beta|)^2 \quad (2.31)$$

in the special unitary case.

Proof. Let us start with the unitary case. Using the definition of Casimir number and the definition of $\lambda(\alpha, \beta, n)$, we obtain

$$\begin{aligned} Nc_2(\lambda_N(\alpha, \beta, n)) &= \sum_{i=1}^r \alpha_i^2 + \sum_{1 \leq i < j \leq r} (\alpha_i - \alpha_j) + 2n|\alpha| + \sum_{i=1}^s \beta_i^2 + \sum_{1 \leq i < j \leq s} (\beta_i - \beta_j) - 2n|\beta| \\ &\quad + |\alpha|(N - r - s) + |\beta|(N - r - s) + \sum_{\substack{1 \leq i \leq r \\ 1 \leq j \leq s}} (\alpha_i + \beta_j) + Nn^2, \end{aligned}$$

which can be rearranged into

$$\begin{aligned} Nc_2(\lambda_N(\alpha, \beta, n)) &= N(|\alpha| + |\beta| + n^2) + 2n(|\alpha| - |\beta|) + \sum_{i=1}^r \alpha_i^2 + \sum_{1 \leq i < j \leq r} (\alpha_i - \alpha_j) - r|\alpha| \\ &\quad + \sum_{i=1}^s \beta_i^2 + \sum_{1 \leq i < j \leq s} (\beta_i - \beta_j) - s|\beta|. \end{aligned}$$

On the other hand,

$$K(\alpha) = \sum_{i=1}^r \frac{\alpha_i(\alpha_i + 1)}{2} - i\alpha_i = \frac{1}{2} \left(\sum_{i=1}^r \alpha_i^2 + \sum_{1 \leq i < j \leq r} (\alpha_i - \alpha_j) - r|\alpha| \right)$$

and we find (2.30) as announced.

Concerning the special unitary case, we need to subtract from $c_2(\lambda)$ the quantity $\frac{1}{N^2} (\sum \lambda_i)^2$, which leads to

$$c'_2(\lambda_N(\alpha, \beta)) = c_2(\lambda_N(\alpha, \beta, \beta_1)) - \frac{1}{N^2} (|\alpha| - |\beta| + Nn)^2$$

from which (2.31) follows easily. \square

The proof of Theorem 2.1.2.(ii) will rely on two estimates of the Casimir number: one that helps proving the convergence of the sum of $q^{c'_2(\lambda)}$ over almost flat highest weights λ to the expected limit, and one that helps controlling the sum over remaining highest weights.

Lemma 2.2.6. *Let $\lambda \in \widehat{\text{SU}}(N)$. Set $k = |\alpha_\lambda| + |\beta_\lambda|$. Then the following inequalities hold:*

$$k - \frac{k^2}{N} \leq c'_2(\lambda) \leq k + \frac{k^2}{N} + \frac{k^2}{N^2}, \quad (2.32)$$

$$\frac{k}{2} \leq c'_2(\lambda). \quad (2.33)$$

Proof. We start from the expression of $c'_2(\lambda) = c'_2(\lambda_N(\alpha_\lambda, \beta_\lambda))$ given by (2.31). The point is to bound $K(\alpha_\lambda)$ and $K(\beta_\lambda)$.

- The partition function can be decomposed as

$$Z_N = S_{N,1} + S_{N,2} + S_{N,3} + S_{N,4},$$

with $S_{N,i} = \sum_{\lambda \in \Lambda_{N,i}^+} q^{c_2(\lambda)}$ for $1 \leq i \leq 4$.

- $S_{N,1}$ converges, when $N \rightarrow \infty$, to the limit stated in Theorem 2.1.2.(ii).
- $S_{N,2}$, $S_{N,3}$ and $S_{N,4}$ all converge to 0.

This scheme of proof highlights the importance of almost flat highest weights in the large N asymptotics of Yang–Mills measure on the torus: they are somehow the only weights contributing to the limit of its partition function.

2.3 LARGE N LIMIT OF PARTITION FUNCTION ON ORIENTABLE SURFACES

2.3.1 Genus $g \geq 2$

The special unitary case

We will start by proving Theorem 2.1.2.(i) in the special unitary case. Let us first reduce the problem to the case where $T = 0$ and $g = 2$.

Lemma 2.3.1. *For all $g \geq 0$, all $T \geq 0$, and all $N \geq 1$, we have*

$$1 \leq Z'_N(g, T) \leq Z'_N(2, 0).$$

It follows from this lemma that the special unitary case of Theorem 2.1.2.(i) is implied by the assertion

$$\lim_{N \rightarrow \infty} Z'_N(2, 0) = 1, \tag{2.36}$$

which we will prove in this section.

Proof of Lemma 2.3.1. The N -tuple $(0, \dots, 0)$ has dimension 1 and Casimir number 0. Thus, it contributes 1 to the partition function $Z'_N(g, T)$, which explains the first inequality. The second inequality is an immediate consequence of the fact that all Casimir numbers are non-negative, and that all dimensions are positive integers. \square

We can finally turn to the proof of (2.36), and therefore of the special unitary variant of Theorem 2.1.2.(i). This will entirely rely on Proposition 2.2.1 about asymptotic properties of Witten zeta function.

Proof of Theorem 2.1.2.(i) in the special unitary case. On one hand, Lemma 2.3.1 states that

$$Z'_N(2, 0) \geq 1.$$

On the other hand,

$$Z'_N(2, 0) = \sum_{\lambda \in \widehat{\text{SU}}(N)} d_\lambda^{-2} = 1 + \sum_{\lambda \neq (0, \dots, 0)} d_\lambda^{-2}.$$

Using Lemma 2.2.4, we find

$$Z'_N(2, 0) \leq 1 + N^{-\frac{1}{2}} \sum_{\lambda \neq (0, \dots, 0)} d_\lambda^{-\frac{3}{2}} \leq 1 + N^{-\frac{1}{2}} \sup_{N \geq 1} \sum_{\lambda \in \widehat{\text{SU}}(N)} d_\lambda^{-\frac{3}{2}}.$$

Thanks to Proposition 2.2.1, this implies

$$\limsup_{N \rightarrow \infty} Z'_N(2, 0) \leq 1$$

and this concludes the proof of (2.36), hence of Theorem 2.1.2.(i) in the special unitary case. \square

The unitary case

We treat the unitary case of Theorem 2.1.2.(i) using our understanding of the special unitary case, and the bijection $\Phi : (\lambda, n) \mapsto \lambda + n$ given in (2.29). We will keep throughout this section the notation λ for an element of $\widehat{\text{SU}}(N)$, n for an element of \mathbb{Z} and $\lambda + n$ for the corresponding element of $\widehat{\text{U}}(N)$ for the sake of consistency.

The first observation is the following. We use the notation $|\lambda| = \lambda_1 + \dots + \lambda_N$.

Lemma 2.3.2. *We have the equality*

$$c_2(\lambda + n) = c'_2(\lambda) + \left(n + \frac{|\lambda|}{N} \right)^2. \quad (2.37)$$

Proof. The proof is a simple verification using Equations (2.3) and (2.4) that respectively define c_2 and c'_2 . \square

It is the contribution of the highest weights of the form $0 + n = (n, \dots, n)$ which produces the Jacobi theta function in the unitary part of Theorem 2.1.2.(i). We will prove that the contribution of all other elements of $\widehat{\text{U}}(N)$ vanishes in the large N limit.

Proof of Theorem 2.1.2.(i) in the unitary case. Let us consider $g \geq 2$ and $T > 0$. We split the partition function $Z_N(g, T)$ into two parts

$$Z_N(g, T) = \sum_{n \in \mathbb{Z}} e^{-\frac{T}{2}n^2} + \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \neq (0, \dots, 0)}} \sum_{n \in \mathbb{Z}} e^{-\frac{T}{2}c_2(\lambda+n)} d_{\lambda+n}.$$

The first part corresponds to highest weights of the form (n, \dots, n) , which have Casimir numbers n^2 and dimension 1. The second part is the contribution of all the other highest weights. To compute it, we observe that $d_{\lambda+n} = d_\lambda$ and we use Lemma 2.3.2. We find

$$0 \leq Z_N(g, T) - \vartheta(0; iT/2\pi) \leq \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \neq (0, \dots, 0)}} \left(\sum_{n \in \mathbb{Z}} e^{-\frac{T}{2}(n+|\lambda|/N)^2} \right) e^{-\frac{T}{2}c'_2(\lambda)} d_\lambda^{2-2g}.$$

The sum between the brackets is bounded independently of N , for example, in a very elementary way, by $C = 1 + \vartheta(0; iT/2\pi)$. Hence, the right-hand side is bounded by

$$C \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \neq (0, \dots, 0)}} d_\lambda^{2-2g} = C(\zeta_{\text{su}(N)}(2g-2) - 1)$$

which, thanks to Proposition 2.2.1, converges to 0. \square

2.3.2 Genus $g = 1$

Our proof of the convergence of the partition function when $g \geq 2$ was based on our study of the dimensions of the irreducible representations of \mathfrak{su}_N , expressed in Proposition 2.2.1. A glance at (1.29) shows that when $g = 1$, these dimensions do not appear anymore in the partition function, and to treat this case we need to use completely different estimates. In this section, we will prove that $Z_N(1, T)$ still admits a finite limit for $T > 0$, but this limit will turn out to be different from the one described in Theorem 2.1.2.(i). In particular, it will involve the classical generating function of integer partitions. Recall that if we denote, for each $n \geq 0$, by $p(n)$ the number of partitions of the integer n , we have the equality of formal series in the variable t :

$$\sum_{\alpha} t^{|\alpha|} = \sum_{n \geq 0} p(n)t^n = \prod_{m=1}^{\infty} \frac{1}{1-t^m}, \quad (2.38)$$

where the first sum is over all integer partitions α . We can recognize in the right-hand side of Equation (2.38) the inverse of Euler function defined in (2.11). Before entering the technical details, let us explain the idea of the proof of Theorem 2.1.2.(ii), at least in the special unitary case. In the present situation where $g = 1$, the partition function is

$$Z'_N(1, T) = \sum_{\lambda \in \widehat{\text{SU}}(N)} e^{-c'_2(\lambda) \frac{T}{2}} = \sum_{\lambda \in \widehat{\text{SU}}(N)} q^{c'_2(\lambda)}.$$

The problem is thus to identify which highest weights of $\text{SU}(N)$ keep, in the large N limit, a bounded quadratic Casimir number, and bring a non-zero contribution to the partition function. We claim, although this statement is not very precise at this stage, that these highest weights are those depicted in Fig. 2.1 (in the special unitary case, we need to look at the right part of this figure). They are the highest weights that are flat up to a small⁵ perturbation at each end, represented by two partitions α and β of length $\leq N/2$. Let us call these highest weights *almost flat*. A similar description was proposed by Gross–Taylor in [GT93], but in the case where the perturbations remain finite, and their goal was rather to obtain a $1/N$ expansion of the partition function than to find its large N limit. The smaller the length of α and β , the flatter the highest weight: typically we will consider α and β of length $\ll \sqrt{N}$. Using the notation $\lambda(\alpha, \beta)$ introduced in Fig. 2.1, and the notation $|\alpha|$ (resp. $|\beta|$) for sum of the components of α (resp. β), the main estimate will be a refinement of the equality

$$c'_2(\lambda(\alpha, \beta)) = |\alpha| + |\beta| + O(N^{-1}) \quad (2.39)$$

⁵Small compared to N but not necessarily finite.

with an explicit expression of the error in terms of α and β . The outline of the proof is then the following

$$Z'_N(1, T) \simeq \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \text{ almost flat}}} q^{c'_2(\lambda)} \simeq \sum_{\alpha, \beta \text{ of length } \ll \sqrt{N}} q^{c'_2(\lambda_N(\alpha, \beta))} \simeq \sum_{\alpha, \beta \text{ of length } \ll \sqrt{N}} q^{|\alpha|+|\beta|}$$

and the last sum tends to the square of the generating function of integer partitions when $N \rightarrow \infty$.

The special unitary case

Proof of Theorem 2.1.2.(ii) in the special unitary case. Let us fix $\gamma \in (0, \frac{1}{2})$ and define the sets $\Lambda_{N,i}^\gamma$ for $1 \leq i \leq 4$ as in Eq. (2.34). For each $i \in \{1, 2, 3, 4\}$, we set

$$S'_{N,i} = \sum_{\lambda \in \Lambda_{N,i}^\gamma} q^{c'_2(\lambda)},$$

so that

$$Z'_N(1, T) = S'_{N,1} + S'_{N,2} + S'_{N,3} + S'_{N,4}.$$

For N large enough, any partition of an integer not greater than N^γ has less than $\frac{N}{2}$ positive parts. Thus, if α and β are any two such partitions, the highest weight $\lambda_N(\alpha, \beta)$ is well defined, and belongs to $\Lambda_{N,1}^\gamma$. As a consequence, for N large enough,

$$\Lambda_{N,1}^\gamma = \{\lambda_N(\alpha, \beta), \alpha \vdash r, \beta \vdash s : r \leq N^\gamma, s \leq N^\gamma\},$$

and

$$S'_{N,1} = \sum_{|\alpha|, |\beta| \leq N^\gamma} q^{c'_2(\lambda_N(\alpha, \beta))}.$$

From (2.35), we deduce that

$$q^{4N^{2\gamma-1}+4N^{2\gamma-2}} \sum_{|\alpha|, |\beta| \leq N^\gamma} q^{|\alpha|+|\beta|} \leq S'_{N,1} \leq q^{-4N^{2\gamma-1}} \sum_{|\alpha|, |\beta| \leq N^\gamma} q^{|\alpha|+|\beta|}.$$

Since $2\gamma - 1$ is negative, the powers of q in front of the sums on either side tend to 1 as N tends to infinity. On the other hand, the sum over α and β tends, as N tends to infinity, to the square of the generating function of partitions defined in (2.38). Hence,

$$\lim_{N \rightarrow \infty} S'_{N,1} = \lim_{N \rightarrow \infty} \sum_{|\alpha|, |\beta| \leq N^\gamma} q^{|\alpha|+|\beta|} = \left(\sum_{\alpha} q^{|\alpha|} \right)^2 = \prod_{m=1}^{\infty} (1 - q^m)^{-2}.$$

In a second step, we prove that the three other contributions to $Z'_N(1, T)$ vanish as N tends to infinity. For this, we use (2.33). Let us treat the case of $S'_{N,2}$, the case of $S'_{N,3}$ being perfectly similar, and the case of $S'_{N,4}$ even simpler. Let us remark that, as opposed to the case of $\Lambda_{N,1}^\gamma$, we only have the inclusion

$$\Lambda_{N,2}^\gamma \subset \{\lambda_N(\alpha, \beta), \alpha \vdash r, \beta \vdash s : r \leq N^\gamma, s > N^\gamma\},$$

but it will be enough to get an adequate upper bound. Indeed,

$$0 \leq S'_{N,2} \leq \sum_{|\alpha| \leq N^\gamma, |\beta| > N^\gamma} q^{c'_2(\lambda_N(\alpha, \beta))},$$

and from (2.33) we have

$$\sum_{|\alpha| \leq N^\gamma, |\beta| > N^\gamma} q^{c'_2(\lambda_N(\alpha, \beta))} \leq \sum_{|\alpha| \leq N^\gamma, |\beta| > N^\gamma} q^{\frac{1}{2}(|\alpha| + |\beta|)} \leq \sum_{\alpha} q^{\frac{1}{2}|\alpha|} \sum_{k > N^\gamma} p(k) q^{\frac{k}{2}}.$$

The first sum of right-hand side is finite, and the second sum, as a remainder of a convergent series, tends to 0 as N tends to infinity. This concludes the proof. \square

The unitary case

The proof of Theorem 2.1.2.(ii) in the unitary case will rely on the same tools as the special unitary case, that is, the use of almost flat highest weights, combined with the bijection $\Phi : (\lambda, n) \mapsto \lambda + n$ introduced in (2.29). In particular, Lemma 2.3.2 will be of great help in order to control the convergence of $Z_N(1, T)$ using the convergence of $Z'_N(1, T)$.

Proof of Theorem 2.1.2.(ii) in the unitary case. Let $\lambda \in \widehat{\text{SU}}(N)$. Thanks to Lemma 2.3.2 and Proposition 2.2.5, it appears that, for all $n \in \mathbb{Z}$,

$$c_2(\lambda + n) = c'_2(\lambda) + \left(n + \frac{|\lambda|}{N} \right)^2 = c'_2(\lambda) + \left(n + \frac{|\alpha_\lambda| - |\beta_\lambda|}{N} + (\beta_\lambda)_1 \right)^2,$$

so that we can write, up to a change of index $n \leftarrow n - (\beta_\lambda)_1$,

$$Z_N(1, T) = \sum_{\lambda \in \widehat{\text{SU}}(N)} \left(\sum_{n \in \mathbb{Z}} q^{\left(n + \frac{|\alpha_\lambda| - |\beta_\lambda|}{N} \right)^2} \right) q^{c'_2(\lambda)}. \quad (2.40)$$

The main difference with the case of $\text{SU}(N)$ is the sum over n between the brackets, and we will need to control it in order to get the convergence.

Let $\gamma \in (0, \frac{1}{2})$, and the subsets $(\Lambda_{N,i}^\gamma)_{1 \leq i \leq 4}$ of $\widehat{\text{SU}}(N)$ as in the special unitary case. We define, for $1 \leq i \leq 4$,

$$S_{N,i} = \sum_{\lambda \in \Lambda_{N,i}^\gamma} \left(\sum_{n \in \mathbb{Z}} q^{\left(n + \frac{|\alpha_\lambda| - |\beta_\lambda|}{N} \right)^2} \right) q^{c'_2(\lambda)},$$

and we obtain the following decomposition:

$$Z_N(1, T) = S_{N,1} + S_{N,2} + S_{N,3} + S_{N,4}.$$

Let λ be an element of $\Lambda_{N,1}^\gamma$. From the fact that $||\alpha_\lambda| - |\beta_\lambda|| \leq |\alpha_\lambda| + |\beta_\lambda| \leq 2N^\gamma$ we get

$$n^2 - 4nN^{\gamma-1} \leq \left(n + \frac{|\alpha_\lambda| - |\beta_\lambda|}{N} \right)^2 \leq n^2 + 4nN^{\gamma-1} + 4N^{2\gamma-2}. \quad (2.41)$$

For the same reason as in the special unitary case, for N large enough we have

$$S_{N,1} = \sum_{|\alpha|, |\beta| \leq N^\gamma} \left(\sum_{n \in \mathbb{Z}} q^{(n + \frac{|\alpha| - |\beta|}{N})^2} \right) q^{c'_2(\lambda_N(\alpha, \beta))};$$

Then, equations (2.35) and (2.41) yield

$$q^{4N^{2\gamma-1} + 8N^{2\gamma-2}} \left(\sum_{n \in \mathbb{Z}} q^{n^2 + 4nN^{\gamma-1}} \right) \sum_{|\alpha|, |\beta| \leq N^\gamma} q^{|\alpha| + |\beta|} \leq S_{N,1} \quad (2.42)$$

and

$$S_{N,1} \leq q^{-4N^{2\gamma-1}} \left(\sum_{n \in \mathbb{Z}} q^{n^2 - 4nN^{\gamma-1}} \right) \sum_{|\alpha|, |\beta| \leq N^\gamma} q^{|\alpha| + |\beta|}. \quad (2.43)$$

The sums $\sum_{n \in \mathbb{Z}} q^{n^2 \pm 4nN^{\gamma-1}}$ in both cases tend to $\sum_{n \in \mathbb{Z}} q^{n^2}$ by dominated convergence, because $\gamma - 1 < 0$. The remaining terms in both inequalities (2.42) and (2.43) behave in the same way as in the proof of Theorem 2.1.2.(ii) in the special unitary case. This proves that

$$\lim_{N \rightarrow \infty} S_{N,1} = \sum_{n \in \mathbb{Z}} q^{n^2} \prod_{m=1}^{\infty} \frac{1}{(1 - q^m)^2}.$$

Now let us treat the cases of $S_{N,2}$, $S_{N,3}$ and $S_{N,4}$. The arguments are the same for the three of them, so we only choose to detail the case of $S_{N,2}$. We have, using equation (2.33),

$$0 \leq S_{N,2} \leq \sum_{|\alpha| \leq N^\gamma, |\beta| > N^\gamma} \left(\sum_{n \in \mathbb{Z}} q^{(n + \frac{|\alpha| - |\beta|}{N})^2} \right) q^{\frac{1}{2}(|\alpha| + |\beta|)},$$

and the sum between brackets can be bounded independently from N , $|\alpha|$ and $|\beta|$ by $C = 1 + \vartheta(0; iT/2\pi)$, thus

$$\begin{aligned} 0 \leq S_{N,2} &\leq C \sum_{\alpha} q^{\frac{1}{2}|\alpha|} \sum_{|\beta| > N^\gamma} q^{\frac{1}{2}|\beta|} \\ &= C \sum_{\alpha} q^{\frac{1}{2}|\alpha|} \sum_{k > N^\gamma} p(k) q^{\frac{k}{2}} \rightarrow 0, \text{ as } N \rightarrow \infty. \end{aligned}$$

This concludes the proof as in the special unitary case. \square

2.4 LARGE N LIMIT OF PARTITION FUNCTION ON NON-ORIENTABLE SURFACES

We now turn to the study of non-orientable surfaces. Let us recall that, according to Theorem 1.4.1, any such surface can be constructed as the connected sum of projective planes. In order to estimate the large N asymptotics of its associated partition function, we need to compute the Frobenius–Schur indicator associated to any highest weight of $U(N)$ or $SU(N)$.

Let (ρ, V) a complex finite-dimensional representation of a compact group G with character χ . Recall that ρ is said to be:

- (i) *Real* if it exists a symmetric G -invariant nondegenerate bilinear form;
- (ii) *Quaternionic* if it exists a skew-symmetric G -invariant nondegenerate bilinear form;
- (iii) *Complex* if there is no G -invariant nondegenerate bilinear form.

The value of ι_χ is actually linked to this classification, as stated by the following Proposition, which can be found in [BtD95, Prop. 6.8].

Proposition 2.4.1. *Let (ρ, V) be a complex finite-dimensional representation of a compact group G , with character χ . Its Frobenius–Schur indicator is given by:*

$$\iota_\chi = \begin{cases} 1 & \text{if } \rho \text{ is real;} \\ 0 & \text{if } \rho \text{ is complex;} \\ -1 & \text{if } \rho \text{ is quaternionic.} \end{cases} \quad (2.44)$$

The next result allows us to decide when an irreducible representation of $U(N)$ or $SU(N)$ is real, complex or quaternionic, based on its highest weight.

Proposition 2.4.2. *Let $\lambda = (\lambda_1 \geq \dots \geq \lambda_N)$ be a highest weight of $U(N)$ (or $SU(N)$ if we fix $\lambda_N = 0$). Let $m_i = \lambda_i - \lambda_{i+1} \in \mathbb{N}$ for every $i \in \{1, \dots, N-1\}$. An irreducible representation of $U(N)$ or $SU(N)$ with highest weight λ is:*

- *Complex* if there exists i such that $m_i \neq m_{N-i}$;
- *Real* if $m_i = m_{N-i}$ for all $i \in \{1, \dots, \lfloor N/2 \rfloor\}$ and one of the following properties is satisfied:
 - $N \not\equiv 2[4]$;
 - $N \equiv 2[4]$ and m_{2k+1} is even;
- *Quaternionic* if $N \equiv 2[4]$, $m_i = m_{N-i}$ for all $i \in \{1, \dots, N/2\}$ and m_{2k+1} is odd.

Proof. The proof is given for $SU(N)$ in [FH91, Prop.26.24]. Now, if $\lambda_N \neq 0$, we define $\mu \in SU(N)$ by setting $\mu_i = \lambda_i - \lambda_N$. It can be verified that a representation of $U(N)$ with highest weight λ is real (resp. complex, quaternionic) if and only if a representation of $SU(N)$ with highest weight μ is real (resp. complex, quaternionic), and for any $i \in \{1, \dots, N\}$ the value of m_i is the same for λ and μ by definition. \square

If we apply this proposition to the construction $\lambda = \lambda_N(\alpha_\lambda, \beta_\lambda)$, it yields the following result.

Corollary 2.4.3. *Let $\lambda \in \widehat{SU}(N)$ and $n \in \mathbb{Z}$.*

- (i) *If $N = 2M + 1$ is odd, then an irreducible representation of $SU(N)$ (resp. $U(N)$) with highest weight λ (resp. $\lambda + n$) is complex if and only if $\alpha_\lambda \neq \beta_\lambda$.*
- (ii) *Assume that $N = 2M$ is even. Let $\alpha = \alpha_\lambda \in \widehat{SU}(M)$ and $\beta = \beta_\lambda \in \widehat{SU}(M+1)$, and set $\tilde{\beta} = (\beta_1 - \beta_M, \dots, \beta_{M-1} - \beta_M, 0) \in \widehat{SU}(M)$. Then an irreducible representation of $SU(N)$ (resp. $U(N)$) with highest weight λ (resp. $\lambda + n$) is complex if and only if $\alpha \neq \tilde{\beta}$.*
- (iii) *For all integer partitions α and β , all $n \in \mathbb{Z}$, and for N large enough, the highest weight $\lambda_N(\alpha, \beta, n)$ as defined by (2.27) is not quaternionic.*

The main point of this corollary is that highest weights that are not symmetric are complex and therefore do not contribute to the non-orientable partition function because their Frobenius–Schur indicator vanishes. We can also notice that quaternionic representations with almost flat highest weight do not appear in the large N scale, and that the partition function becomes a sum of non-negative terms.

2.4.1 Genus $g \geq 3$

The special unitary case

The proof of Theorem 2.1.3.(i) will be based on the same reasoning as for orientable surfaces of genus $g \geq 2$, that is, using Proposition 2.2.1 to show that the contribution of all other highest weights than $(0, \dots, 0)$ vanish in the large N limit. However, the case of non-orientable surfaces with $g = 3$ will need a finer control, as we will see later. In particular, for even values of N and $g = 3$ the following inequality is needed.

Proposition 2.4.4. *Let $N = 2M$ be an integer. For $\lambda \in \widehat{\text{SU}}(N)$, if we set $\alpha = \alpha_\lambda \in \widehat{\text{SU}}(M)$ and $\beta = \beta_\lambda \in \widehat{\text{SU}}(M+1)$, as well as $\tilde{\beta} = (\beta_1 - \beta_M, \dots, \beta_{M-1} - \beta_M, 0) \in \widehat{\text{SU}}(N)$ as in Corollary 2.4.3.(ii), then*

$$d_\lambda \geq \left(1 + \frac{\beta_M}{M}\right)^M d_\alpha d_{\tilde{\beta}}.$$

Proof. Using Equation (2.2) and the fact that

$$\lambda = \lambda_N(\alpha, \beta) = (\alpha_1 + \beta_1, \dots, \alpha_{M-1} + \beta_1, \beta_1, \beta_M - \beta_1, \dots, \beta_2 - \beta_1, 0),$$

it is clear that $d_\lambda \geq d_\alpha d_\beta$. Moreover,

$$\begin{aligned} d_\beta &= \prod_{1 \leq i < j \leq M+1} \left(1 + \frac{\beta_i - \beta_j}{j - i}\right) \\ &= \prod_{i=1}^M \left(\frac{\beta_i}{M+1-i}\right) d_{\tilde{\beta}} \\ &\geq \left(\frac{\beta_M}{M}\right)^M d_{\tilde{\beta}}. \end{aligned}$$

Combining both inequalities gives the expected result. \square

Proof of Theorem 2.1.3.(i). The highest weight $(0, \dots, 0)$ is associated to the trivial representation, which is real by Proposition 2.4.1 and has dimension 1 and Casimir number 0. We can then rewrite

$$Z'_N(g, T) = 1 + \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \neq (0, \dots, 0)}} q^{c'_2(\lambda)} d_\lambda^{2-g} (\iota_\lambda)^g,$$

and the remaining sum can be bounded as follows:

$$\left| \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \neq (0, \dots, 0)}} q^{c'_2(\lambda)} d_\lambda^{2-g} \iota_\lambda \right| \leq \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \neq (0, \dots, 0)}} q^{c'_2(\lambda)} d_\lambda^{2-g}.$$

If $g \geq 4$, then the right-hand side has been proved to converge to 0 as $N \rightarrow \infty$ in Section 2.3.1, hence the result follows.

Now, if $g = 3$, we need to refine the analysis in order to get the convergence. From Corollary 2.4.3, it appears that $\lambda \in \text{SU}(N)$ contributes to the partition function if and only if it is symmetric. The case $N = 2M + 1$ is easier to prove, so we start with it. As $\iota_\lambda = 0$ if λ is associated with a complex representation, we have

$$Z'_N(3, T) = 1 + \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \neq (0, \dots, 0) \\ \lambda \text{ is symmetric}}} q^{c'_2(\lambda)} d_\lambda^{-1} (\iota_\lambda)^3,$$

which means that

$$\begin{aligned} |Z'_N(3, T) - 1| &= \left| \sum_{\substack{\alpha \in \widehat{\text{SU}}(M+1) \\ \alpha \neq (0, \dots, 0)}} q^{c'_2(\lambda_N(\alpha, \alpha))} d_{\lambda_N(\alpha, \alpha)}^{-1} (\iota_{\lambda_N(\alpha, \alpha)})^3 \right| \\ &\leq \sum_{\substack{\alpha \in \widehat{\text{SU}}(M+1) \\ \alpha \neq (0, \dots, 0)}} q^{c'_2(\lambda_N(\alpha, \alpha))} d_\alpha^{-2} \\ &\leq \zeta_{\text{su}(M)}(2), \end{aligned}$$

where in the first inequality we used the fact that $d_{\lambda_N(\alpha, \alpha)} \geq d_\alpha^2$. Then, letting M tend to infinity and using Proposition 2.2.1, we have indeed

$$\lim_{M \rightarrow \infty} Z'_{2M+1}(3, T) = 1.$$

Now consider $N = 2M$. Let $\tilde{\beta} = (\beta_1 - \beta_M, \dots, \beta_{M-1} - \beta_M, 0)$. Corollary 2.4.3 states that $\lambda = \lambda_N(\alpha, \beta)$ contributes to the partition function if and only if $\alpha = \tilde{\beta}$. It implies:

$$|Z'_N(3, T) - 1| = \left| \sum_{\substack{(\alpha, \beta) \in \widehat{\text{SU}}(M) \times \widehat{\text{SU}}(M+1) \\ \alpha = \tilde{\beta}}} q^{c'_2(\lambda_N(\alpha, \beta))} d_{\lambda_N(\alpha, \beta)}^{-1} (\iota_{\lambda_N(\alpha, \beta)})^3 \right|.$$

We can then apply Proposition 2.4.4 to get

$$\begin{aligned} |Z'_N(3, T) - 1| &\leq \sum_{\substack{\alpha \in \widehat{\text{SU}}(M+1) \\ \alpha \neq (0, \dots, 0)}} \sum_{n \in \mathbb{N}} \left(1 + \frac{n}{M}\right)^{-M} d_\alpha^{-2} \\ &= \sum_{n \in \mathbb{N}} \left(1 + \frac{n}{M}\right)^{-M} \sum_{\substack{\alpha \in \widehat{\text{SU}}(M+1) \\ \alpha \neq (0, \dots, 0)}} d_\alpha^{-2}. \end{aligned}$$

The first sum is bounded because $\left(1 + \frac{n}{M}\right)^{-M} \leq e^{-n}$ for any n, M , and the second one converges, following the same argument as in the case $N = 2M + 1$. We finally get

$$\lim_{M \rightarrow \infty} Z'_{2M}(3, T) = 1.$$

Finally, we have shown that

$$\lim_{M \rightarrow \infty} Z'_{2M}{}^-(3, T) = \lim_{M \rightarrow \infty} Z'_{2M+1}{}^-(3, T) = 1,$$

which concludes the proof. \square

The unitary case

As for the special unitary case, the proof of the unitary case for non-orientable surfaces of genus $g \geq 3$ is similar to the one of orientable surfaces of genus $g \geq 2$. Indeed, the point is to show that only constant highest weights contribute to the large N limit.

Proof of Theorem 2.1.3.(i) in the unitary case. Let us consider $g \geq 3$ and $T > 0$. We split the partition function $Z_N^-(g, T)$ into two parts

$$Z_N^-(g, T) = \sum_{n \in \mathbb{Z}} e^{-\frac{T}{2}n^2} + \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \neq (0, \dots, 0)}} \sum_{n \in \mathbb{Z}} e^{-\frac{T}{2}c_2(\lambda+n)} d_{\lambda+n}^{2-g} t_{\lambda+n}^g.$$

Let us assume that $g \geq 4$. Following the arguments used in the orientable case with $g = 2$, we find

$$0 \leq |Z_N(g, T) - \vartheta(0; iT/2\pi)| \leq \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \neq (0, \dots, 0)}} \left(\sum_{n \in \mathbb{Z}} e^{-\frac{T}{2}(n+|\lambda|/N)^2} \right) e^{-\frac{T}{2}c_2(\lambda)} d_{\lambda}^{2-g}.$$

The sum between the brackets is bounded by $C = 1 + \vartheta(0; iT/2\pi)$ and the other sum is bounded in absolute value by $\zeta_{\text{su}(N)}(g-2) - 1$, which converges to 0. Hence, the whole right-hand side converges to 0.

If $g = 3$ we need a special analysis similar to the one in the special unitary case. Using Corollary 2.4.3, any $\lambda \in \text{U}(N)$ contributes to $Z_N^-(3, T)$ if and only if it is symmetric. Let us first assume that $N = 2M + 1$. Then, we can write

$$Z_N^-(3, T) = \vartheta(0; iT/2\pi) + \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \neq (0, \dots, 0) \\ \lambda \text{ is symmetric}}} \sum_{n \in \mathbb{Z}} q^{c_2(\lambda+n)} d_{\lambda}^{-1} (t_{\lambda})^3,$$

therefore

$$\begin{aligned} 0 \leq |Z_N^-(3, T) - \vartheta(0; iT/2\pi)| &\leq \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \neq (0, \dots, 0) \\ \lambda \text{ is symmetric}}} \left(\sum_{n \in \mathbb{Z}} q^{(n+|\lambda|/N)^2} \right) q^{c_2(\lambda)} d_{\lambda}^{-1} \\ &\leq \sum_{\substack{\alpha \in \widehat{\text{SU}}(M+1) \\ \alpha \neq (0, \dots, 0)}} \left(\sum_{n \in \mathbb{Z}} q^{(n+|\lambda_N(\alpha, \alpha)|/N)^2} \right) q^{c_2(\lambda_N(\alpha, \alpha))} d_{\alpha}^{-2}. \end{aligned}$$

The right-hand side converges to 0 for the same reason as in the case $g \geq 4$. Now, let us assume that $N = 2M$. Let $\tilde{\beta}$ be defined as before. We have

$$0 \leq |Z_N^-(3, T) - \vartheta(0; iT/2\pi)| \leq \sum_{\substack{(\alpha, \beta) \in \widehat{\text{SU}}(M) \times \widehat{\text{SU}}(M+1) \\ \alpha = \tilde{\beta}}} \left(\sum_{n \in \mathbb{Z}} q^{(n + |\lambda_N(\alpha, \beta)|/N)^2} \right) \frac{q^{c_2(\lambda_N(\alpha, \beta))}}{d_{\lambda_N(\alpha, \beta)}}.$$

Proposition 2.4.4, plus similar arguments as before, yield

$$|Z_N^-(3, T) - \vartheta(0; iT/2\pi)| \leq (1 + \vartheta(0; iT/2\pi)) \sum_{n \in \mathbb{N}} \left(1 + \frac{n}{M}\right)^{-M} \sum_{\substack{\alpha \in \widehat{\text{SU}}(M+1) \\ \alpha \neq (0, \dots, 0)}} d_{\alpha}^{-2},$$

and the right-hand side converges to zero. We proved the convergence for odd and even values of N to the same quantity, which concludes the case $g = 3$. \square

2.4.2 Genus $g = 2$

The Klein bottle is the non-orientable equivalent to the torus, as we will see, in the sense that the dimensions of the irreducible representations do not appear in the formula of the partition function. Hence, the proof of Theorem 2.1.3.(ii) is using almost flat highest weights as well.

The special unitary case

Proof of Theorem 2.1.3.(ii) in the special unitary case. From Corollary 2.4.3 and Proposition 2.4.1 we deduce that $\lambda \in \widehat{\text{SU}}(2M)$ (resp. $\widehat{\text{SU}}(2M + 1)$) has a nonzero Frobenius–Schur indicator if and only if $\alpha = \beta$ (resp. $\alpha = \tilde{\beta}$), where $\alpha = \alpha_\lambda$ and $\beta = \beta_\lambda$ are defined as in Section 2.2.2, and $\tilde{\beta} = (\beta_1 - \beta_M, \dots, \beta_{M-1} - \beta_M, 0) \in \widehat{\text{SU}}(M)$.

Let $\gamma \in (0, \frac{1}{2})$, and the subsets $(\Lambda_{N,i}^\gamma)_{1 \leq i \leq 4}$ of $\widehat{\text{SU}}(N)$ defined as in (2.34). We define, for $1 \leq i \leq 4$,

$$S'_{N,i} = \sum_{\lambda \in \Lambda_{N,i}^\gamma} \iota_\lambda^2 q^{c'_2(\lambda)} = \sum_{\substack{\lambda \in \Lambda_{N,i}^\gamma \\ \lambda \text{ is symmetric}}} q^{c'_2(\lambda)},$$

and we obtain the following decomposition:

$$Z'_N(1, T) = S'_{N,1} + S'_{N,2} + S'_{N,3} + S'_{N,4}.$$

- If $N = 2M + 1$, then the symmetry condition is equivalent to the fact that $\alpha = \beta$ and we can simplify equation (2.31) into

$$c'_2(\lambda_N(\alpha, \alpha)) = 2|\alpha| + \frac{4K(\alpha)}{N},$$

for any α of length r and $N \geq 2r$. Let us recall the estimation

$$|2K(\alpha)| \leq |\alpha|(|\alpha| - 1),$$

which leads, for $\lambda = \lambda(\alpha, \alpha) \in \Lambda_{N,1}^\gamma$, to

$$|c'_2(\lambda) - 2|\alpha|| \leq 4N^{2\gamma-1}. \quad (2.45)$$

Recall that we found in the proof of Theorem 2.1.2.(ii) that, for N large enough,

$$\Lambda_{N,1}^\gamma = \{\lambda_N(\alpha, \beta), \alpha \vdash r, \beta \vdash s : r \leq N^\gamma, s \leq N^\gamma\}.$$

We then get from this equality and from (2.45) the estimate

$$q^{4N^{2\gamma-1}} \sum_{|\alpha| \leq N^\gamma} q^{2|\alpha|} \leq S'_{N,1} \leq q^{-4N^{2\gamma-1}} \sum_{|\alpha| \leq N^\gamma} q^{2|\alpha|}, \quad (2.46)$$

and both bounds converge to the expected quantity $\prod_{m=1}^{\infty} \frac{1}{1-q^{2m}}$.

- If $N = 2M$, then the symmetry condition is equivalent to the fact that $\alpha = \tilde{\beta}$ and under this condition we have

$$\left\{ \begin{array}{l} |\alpha| \leq N^\gamma \\ |\beta| \leq N^\gamma \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} |\alpha| \leq N^\gamma \\ \beta_M \leq N^\gamma - |\alpha| \end{array} \right\}.$$

Furthermore, equation (2.35) becomes

$$2|\alpha| + M\beta_M - 4N^{2\gamma-1} \leq c'_2(\lambda) \leq 2|\alpha| + M\beta_M + 4N^{2\gamma-1} + 4N^{2\gamma-2}.$$

We obtain that

$$\begin{aligned} q^{4N^{2\gamma-1}+4N^{2\gamma-2}} \sum_{|\alpha| \leq N^\gamma} \left(\sum_{n \leq N^\gamma - |\alpha|} q^{Mn} \right) q^{2|\alpha|} &\leq S'_{N,1}, \\ S'_{N,1} &\leq q^{-4N^{2\gamma-1}} \sum_{|\alpha| \leq N^\gamma} \left(\sum_{n \leq N^\gamma - |\alpha|} q^{Mn} \right) q^{2|\alpha|}. \end{aligned} \quad (2.47)$$

The sums between brackets in both inequalities are bounded between 1 and $\sum_{n \in \mathbb{Z}} q^{Mn}$. The latter converges to 1 as M tends to infinity, by dominated convergence (it is clearly bounded by the geometric series $\sum_{n \in \mathbb{N}} q^n$). It finally appears that both bounds of $S'_{N,1}$ in (2.47) converge to $\prod_{m=1}^{\infty} \frac{1}{1-q^{2m}}$.

By similar arguments as the ones used in the case of the torus, we can prove that $S'_{N,2}$, $S'_{N,3}$ and $S'_{N,4}$ all converge to 0 as the remainders of convergent series. This concludes the proof. \square

The unitary case

Proof of Theorem 2.1.3.(ii) in the unitary case. Let λ be an element of $\widehat{\text{SU}}(N)$. Recall that we found in the proof of Theorem 2.1.2.(ii) in the unitary case that, for all $n \in \mathbb{Z}$,

$$c_2(\lambda + n) = c'_2(\lambda) + \left(n + \frac{|\lambda|}{N} \right)^2 = c'_2(\lambda) + \left(n + \frac{|\alpha_\lambda| - |\beta_\lambda|}{N} + (\beta_\lambda)_1 \right)^2,$$

so that we can write, modulo a change of index $n \leftarrow n - (\beta_\lambda)_1$,

$$Z_N^-(2, T) = \sum_{\lambda \in \widehat{\text{SU}}(N)} \left(\sum_{n \in \mathbb{Z}} q^{\left(n + \frac{|\alpha_\lambda| - |\beta_\lambda|}{N}\right)^2} \iota_{\lambda+n}^2 \right) q^{c'_2(\lambda)}. \quad (2.48)$$

Let $\gamma \in (0, \frac{1}{2})$, and the subsets $(\Lambda_{N,i}^\gamma)_{1 \leq i \leq 4}$ of $\widehat{\text{SU}}(N)$ as in the special unitary case. We define, for $1 \leq i \leq 4$,

$$S_{N,i} = \sum_{\lambda \in \Lambda_{N,i}^\gamma} \left(\sum_{n \in \mathbb{Z}} q^{\left(n + \frac{|\alpha_\lambda| - |\beta_\lambda|}{N}\right)^2} \iota_{\lambda+n}^2 \right) q^{c'_2(\lambda)},$$

and we obtain the following decomposition:

$$Z_N(1, T) = S_{N,1} + S_{N,2} + S_{N,3} + S_{N,4}.$$

Let λ be an element of $\Lambda_{N,1}^\gamma$. From corollary 2.4.3 we deduce that $\iota_{\lambda+n} \neq 0$ if and only if $\alpha \neq \beta$ if N is odd or $\alpha \neq \tilde{\beta}$ if N is even, where $\alpha = \alpha_\lambda$ and $\beta = \beta_\lambda$ satisfy $\lambda = \lambda_N(\alpha, \beta)$ and $\tilde{\beta} = (\beta_1 - \beta_M, \dots, \beta_{M-1} - \beta_M, 0) \in \widehat{\text{SU}}(M)$. In the following, we will assume that this condition is satisfied by $\lambda + n$ so that it contributes to the partition function.

- If $N = 2M + 1$, then $n + \frac{|\alpha| - |\beta|}{N} = n$ and

$$S_{N,1} = \left(\sum_{n \in \mathbb{Z}} q^{n^2} \right) \sum_{\lambda = \lambda(\alpha, \alpha) \in \Lambda_{N,i}^\gamma} q^{c'_2(\lambda)}. \quad (2.49)$$

We get then back to the $\text{SU}(N)$ case which was previously proved.

- If $N = 2M$, let $\tilde{\beta} = (\beta_1, \dots, \beta_{M-1}, 0)$ as in the $g \geq 3$ case, then $n + \frac{|\alpha| - |\beta|}{N} = n - \frac{M\beta_M}{N}$ and $\beta_M = 0$ if N is large enough because λ is almost flat. Hence,

$$S_{N,1} = \left(\sum_{n \in \mathbb{Z}} q^{n^2} \right) \sum_{\substack{\lambda = \lambda(\alpha, \beta) \in \Lambda_{N,i}^\gamma \\ \alpha = \tilde{\beta}}} q^{c'_2(\lambda)}. \quad (2.50)$$

Once again, we get back to the $\text{SU}(N)$ case which was previously proved.

With similar arguments as in the previous proofs, we can prove that $S_{N,2}$, $S_{N,3}$ and $S_{N,4}$ all converge to 0 as they are remainders of convergent series. Finally, using the convergence results from the $\text{SU}(N)$ case, we see that the limit of $S_{2M,1}$ and $S_{2M+1,1}$ is the same, which is the one stated in Theorem 2.1.3.(ii), and it is therefore the limit of $Z_N^-(2, T)$. \square

THE MASTER FIELD ON COMPACT ORIENTABLE SURFACES

In this chapter, we first describe the master field and recall its construction for the plane and the sphere, before presenting results about higher genus surfaces. In particular, we compute in Thm. 3.2.5 (resp. 3.2.6) the limit of Wilson loop expectation (resp. variance) for contractible simple loops, and we explain how it can be extended to a broader class of loops, including those with self-intersections. We also discuss in Thm. 3.2.16 and Prop. 3.2.17 the case of loops which have nonzero homology.

3.1 THE MASTER FIELD, FROM WILSON TO MAKEENKO–MIGDAL

It seems that the first definition of the master field was given in [Wit80], at least informally, based on the work of 't Hooft [tH74] that is considered by many to be the starting point of the study of *large N limit*. However, it has become an intense subject of research in the early 90s, with the seminal papers [Sin95, GG95]. It was initially considered as the limiting object of Yang–Mills measure on the plane, and treated as such in the first place, but physicists began quickly to generalize it to the sphere, because some limiting aspects of Yang–Mills on the sphere were already investigated by Douglas and Kazakov [DK93] or also Boulatov [Bou94].

From a mathematical perspective, it became a subject of interest in the beginning of 21st century [Seno8b, AS12, Lévi17, Dah16, CDG17] for the plane, then for the sphere [FMS11, LM15, DN17, Hal18]. The articles [Lévi17] and [DN17] can be used as references for the construction of the master field, respectively on the plane and on the sphere. In our definition, it can be defined as the limit of the Yang–Mills holonomy field defined in Thm. 1.4.3. Recall that for a surface Σ and $m \in \Sigma$ a base point, we denote $\mathcal{L}_m(\Sigma)$ (resp. $\mathcal{P}(\Sigma)$) the monoid of loops on Σ starting from and ending to m . (resp. the set of paths in Σ)¹. Our definition of the master field is then the following, which is a reminder of Def. 1.4.5.

Definition 3.1.1. Let Σ be a surface and $m \in \Sigma$ a base point. A *master field* on Σ is the datum of a $*$ -probability space (\mathcal{A}, τ) and a noncommutative stochastic process $(h_\ell)_{\ell \in \mathcal{L}_m(\Sigma)}$ on \mathcal{A} that satisfies the following properties:

¹Cf. p. 70–71.

(i) For all $(\ell_1, \ell_2) \in \mathcal{L}_m(\Sigma)^2$,

$$h_{\ell^{-1}} = h_\ell^* = h_\ell^{-1} \text{ and } h_{\ell_1 \ell_2} = h_{\ell_2} h_{\ell_1}; \quad (3.1)$$

(ii) If $(\ell_n)_{n \in \mathbb{N}}$ is a sequence of loops that converges uniformly to a loop ℓ and if the sequence of their lengths converges to the length of ℓ , then (h_{ℓ_n}) converges to h_ℓ in $L^2(\mathcal{A}, \tau)$.

(iii) Let $(H_c)_{c \in \mathcal{P}(\Sigma)}$ be the Yang–Mills holonomy field on Σ with structure group $U(N)$ or $SU(N)$. For any $\ell \in \mathcal{L}_m(M)$ we have

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_\ell)] = \tau(h_\ell). \quad (3.2)$$

The quantity $\text{tr}(H_\ell)$ given in Equation (3.2) is called *Wilson loop functional*, and its expectation with respect to Yang–Mills measure is called *Wilson loop expectation*. It was first defined by Wilson [Wil74], as a gauge-invariant observable derived from the holonomy of a principal connection around a loop; one can find in [Sen94] how this definition of Wilson loop is related to the one we use. The reason why these functionals are studied in the setting of large N Yang–Mills is the following theorem, proved by Sengupta [Sen94] for a smaller class of groups and then extended by Lévy [Lévo4].

Theorem 3.1.1. *Let G be a finite product of groups among $U(N)$, $SU(N)$, $O(N)$, $SO(N)$, $\text{Sp}(N)$, and \mathbb{G} be an oriented graph. Then the algebra generated by the Wilson loops is dense in the space of continuous functions on the configuration space \mathcal{C}_G^G .*

We will not detail the proof of this result, but highlight the fact that it relies on the following property, that all groups given in Theorem 3.1.1 share: for all n -tuples (x_1, \dots, x_n) and (y_1, \dots, y_n) of elements of G , we have the equivalence between

$$\exists g \in G, y_1 = gx_1g^{-1}, \dots, y_n = gx_ng^{-1}$$

and

$$\forall w \text{ word in } n \text{ letters and their inverses, } w(y_1, \dots, y_n) \text{ is conjugated to } w(x_1, \dots, x_n).$$

Wilson loop expectations are well established in the case of two-dimensional Yang–Mills theory, as we will see later, but in higher dimensions they remain under investigation even for lattice gauge theory, as shown by the recent results from Chatterjee [Char8], Cao [Cao20] and Forsström *et al.* [FLV20].

The main use of Wilson loops in this chapter will be the following: as stated by Definition 3.1.1, the moments of the master field applied to a loop ℓ are the limits of the expectations of Wilson loop functionals applied to the same loop.

In this section we will explain, in a somehow unified procedure, how to prove the existence of the master field in the cases of the plane and the sphere, which are the only known cases yet. Indeed, despite the fact that the Yang–Mills holonomy process has two distinct structures for each of these surfaces, the way of constructing the master field is similar. The first step is to prove the convergence of Wilson loop expectations and variances for simple loops, *i.e.* loops without crossings, and then to compute these expectations and variances for a broader class of loops using the so-called Makeenko–Migdal equations. We will develop these two steps separately, as they are rather independent in the sense that they do not rely on the same tools.

3.1.1 Wilson loop expectations for simple loops

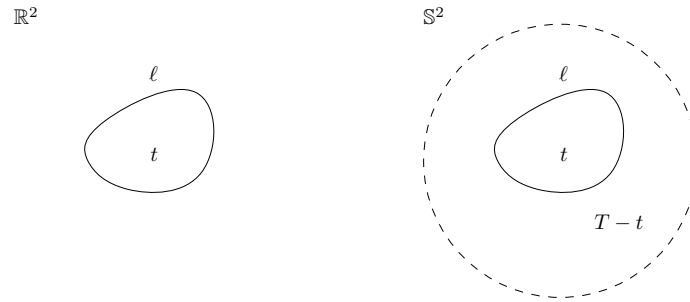


Figure 3.1: A simple loop enclosing a domain of area t , embedded in the plane (on the left) or in the sphere \mathbb{S}^2 (on the right).

The first building brick of the master field is the Wilson loop functional for a loop with no self-intersection, which we will call a *simple loop*. If it is embedded in the plane, it is the boundary of a domain with area t , and if it is embedded in the sphere it separates it into two domains of respective areas t and $T - t$, as illustrated in Fig. 3.1. If we complete this into an admissible graph, it enables us to apply Driver–Sengupta formula (1.87) in order to compute the Wilson loop expectation $W_\ell = \mathbb{E}[\text{tr}(H_\ell)]$. Recall that we denote by $(p_t)_{t \geq 0}$ the heat kernel on the structure group G .

Proposition 3.1.2. *Let Σ be either the plane \mathbb{R}^2 or the sphere \mathbb{S}^2 with area T , ℓ be a simple loop on Σ , oriented counterclockwise. The Wilson loop expectation W_ℓ is equal to*

$$W_\ell = \int_G \text{tr}(x) p_t(x) dx, \quad \text{if } \Sigma \text{ is the plane,} \quad (3.3)$$

$$W_\ell = \frac{1}{Z_T} \int_G \text{tr}(x) p_t(x) p_{T-t}(x^{-1}) dx, \quad \text{if } \Sigma \text{ is the sphere.} \quad (3.4)$$

In (3.4), the density $p_t(x) p_{T-t}(x^{-1}) / Z_T$ is nothing but the density at time t of a Brownian bridge $(B_t)_{t \in [0, T]}$ on G such that $B_0 = B_T = e$. It follows from the convolution property of the heat semigroup that $Z_T = \int_G p_t(x) p_{T-t}(x^{-1}) dx = p_T(e)$.

There are various ways of computing Wilson loop expectations, which are closely related to the computation of moments of the Brownian motion (or Brownian bridge, depending whether the underlying surface is the plane or the sphere) on $U(N)$. We will present one of them, that we will be using later: (noncommutative) harmonic analysis on $U(N)$. Let us mention some of the other existing tools, that we decided not to develop in this chapter:

- Matrix stochastic calculus, which is treated in [Gui04];
- Schur–Weyl duality, which is used in [Lévo8, CDK18] to compute the limit of moments of unitary Brownian motion;
- Free stochastic calculus, which is used in [BS01] to compute the limit of matrix integrals similar to the Wilson loops expectations (but with respect to Hermitian Brownian motion instead of the unitary one);

- Determinantal processes, which are used in [LW16] to compute the asymptotic distribution of a unitary Brownian bridge;
- Large deviations, which are depicted in [Guio4] and used in [LM15] to compute the same limit as in [LW16].

The use of noncommutative harmonic analysis in order to compute integrals such as (3.3) or (3.4) is based to the heat kernel decomposition. In our preferred setting, which is $G = U(N)$, this decomposition is derived in Section 1.3.5: let us recall it.

The irreducible representations of $U(N)$ are labelled by nonincreasing N -tuples of integers $\lambda = (\lambda_1 \geq \dots \geq \lambda_N)$ called *highest weights*, and we denote respectively by d_λ and $c_2(\lambda)$ their *dimension* and *Casimir number*, given by

$$d_\lambda = \prod_{1 \leq i < j \leq N} \frac{\lambda_i - \lambda_j + j - i}{j - i} \quad (3.5)$$

and

$$c_2(\lambda) = \frac{1}{N} \left(\sum_{i=1}^N \lambda_i^2 + \sum_{1 \leq i < j \leq N} (\lambda_i - \lambda_j) \right). \quad (3.6)$$

The set of irreducible representations is denoted by $\widehat{U}(N)$ and is in bijection with the set of highest weights. The *character* of a representation of highest weight λ is given by the *Schur function* s_λ , which is defined in Prop. 1.3.12; however, we will not need its explicit formula for our computations. Theorem 1.3.24 states that the heat kernel on $U(N)$ admits the decomposition

$$p_T(U) = \sum_{\lambda \in \widehat{U}(N)} e^{-c_2(\lambda) \frac{T}{2}} d_\lambda s_\lambda(U), \quad \forall T > 0, \quad \forall U \in U(N). \quad (3.7)$$

We can make a similar statement for the group $SU(N)$: its irreducible representations are labelled by nonincreasing N -tuples of integers $\mu = (\mu_1 \geq \dots \geq \mu_N = 0)$ also called *highest weights*, and their dimension and Casimir number are respectively given by

$$d_\mu = \prod_{1 \leq i < j \leq N} \frac{\mu_i - \mu_j + j - i}{j - i} = \prod_{1 \leq i < j \leq N} \left(1 + \frac{\mu_i - \mu_j}{j - i} \right) \quad (3.8)$$

and

$$c'_2(\mu) = \frac{1}{N} \left(\sum_{i=1}^N \mu_i^2 - \frac{1}{N} \left(\sum_{i=1}^N \mu_i \right)^2 + \sum_{1 \leq i < j \leq N} (\mu_i - \mu_j) \right). \quad (3.9)$$

The Equation (3.7) still holds for $SU(N)$ when one replaces accordingly the highest weights and their related quantities d_λ and $c_2(\lambda)$.

Let us get back to the Wilson loop expectations for the plane and the sphere: for $G = U(N)$, Eq. (3.3) becomes

$$W_\ell = \sum_{\lambda \in \widehat{U}(N)} e^{-c_2(\lambda) \frac{t}{2}} d_\lambda \int_{U(N)} \text{tr}(x) s_\lambda(x) dx, \quad (3.10)$$

and (3.4) becomes

$$W_\ell = \frac{1}{Z_T} \sum_{\lambda, \mu \in \widehat{U}(N)} e^{-c_2(\lambda)\frac{t}{2} - c_2(\mu)\frac{T-t}{2}} d_\lambda d_\mu \int_{U(N)} \text{tr}(x) s_\lambda(x) s_\mu(s^{-1}) dx. \quad (3.11)$$

We can notice that in both equations the integrals do not depend anymore on t or T . In order to compute them, we must be able to compute the product $\text{tr}(\cdot) s_\lambda(\cdot)$. It appears that this can be done thanks to the Murnaghan–Nakayama rule, which is given below. Let us first introduce two notations that will help us simplify the formulæ: if λ and μ are two highest weights, both of $U(N)$ or both of $SU(N)$, we will write $\mu \nearrow \lambda$ (or $\lambda \searrow \mu$) if λ can be obtained from μ by adding 1 to one of its parts², and $\mu \sim \lambda$ if λ can be obtained from μ by adding 1 to one of its parts and -1 to one of its parts³, *i.e.* if there exists a highest weight ν such that $\lambda \nearrow \nu$ and $\mu \nearrow \nu$. These branching rules are illustrated in Fig. 3.2.

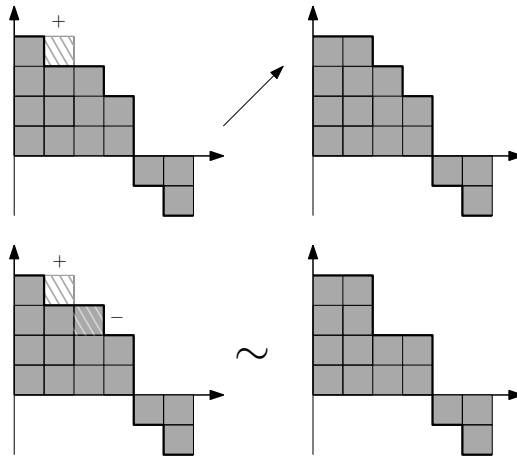


Figure 3.2: On the first row: we have $\lambda \nearrow \mu$, with the diagram of λ on the left and the one of μ on the right. On the second row: we have $\lambda \sim \mu$, with the diagram of λ on the left and the one of μ on the right.

We will also use the following notation: $\lambda \prec \mu$ or $\mu \succ \lambda$ when $\lambda \subset \mu$ and $\mu - \lambda$ is a border strip, *i.e.* it is connected and contains no 2×2 block of squares. The *height* of a border strip is defined as

$$\text{ht}(\mu - \lambda) = \#\{i : \mu_i - \lambda_i \neq 0\} - 1,$$

and its length $|\mu - \lambda|$ is simply the number of boxes it contains, and is equal to

$$|\mu - \lambda| = |\mu| - |\lambda|.$$

We note $\lambda \prec_n \mu$ (resp. $\mu \succ_n \lambda$) if $\lambda \prec \mu$ (resp. $\mu \succ \lambda$) and $|\mu - \lambda| = n$, for $n \in \mathbb{N}$.

We can now state the Murnaghan–Nakayama rule, whose proof can be found for instance in [Mac15, I.7, example 5].

²In the language of diagrams, it means that we add a box to a positive part or remove one to a negative part.
³it can be the same one!

Lemma 3.1.3 (Murnaghan–Nakayama rule). *Let $\lambda \in \widehat{U}(N)$ be a highest weight, r a positive integer. Then we have*

$$\mathrm{Tr}(x^r) s_\lambda(x) = \sum_{\substack{\mu \in \widehat{U}(N) \\ \mu \succ_r \lambda}} (-1)^{\mathrm{ht}(\mu-\lambda)} s_\mu(x), \quad \forall x \in \mathrm{SU}(N). \quad (3.12)$$

In particular, for $r = 1$, we have⁴

$$\mathrm{Tr}(x) s_\lambda(x) = \sum_{\substack{\mu \in \widehat{U}(N) \\ \mu \searrow \lambda}} s_\mu(x), \quad \forall x \in \mathrm{U}(N). \quad (3.13)$$

Using (3.13), we have in the plane

$$W_\ell = \frac{1}{N} \sum_{\lambda \in \widehat{U}(N)} e^{-c_2(\lambda) \frac{t}{2}} d_\lambda \sum_{\substack{\mu \in \widehat{U}(N) \\ \mu \searrow \lambda}} \int_{\mathrm{U}(N)} s_\mu(x) dx. \quad (3.14)$$

As the integral $\int_{\mathrm{U}(N)} s_\mu(x) dx$ is equal to 1 if $\mu = (0, \dots, 0)$ and 0 otherwise, it appears that

$$W_\ell = e^{-c_2((0, \dots, 0, -1)) \frac{t}{2}} d_{(0, \dots, 0, -1)}. \quad (3.15)$$

A direct computation of the Casimir number and the dimension of the representation yields

$$W_\ell = e^{-\frac{t}{2}}, \quad (3.16)$$

which trivially converges to $e^{-\frac{t}{2}}$ when $N \rightarrow \infty$.

In the case of the sphere, with similar arguments we obtain

$$W_\ell = \frac{1}{N Z_T} \sum_{\substack{\lambda, \mu \in \widehat{U}(N) \\ \lambda \nearrow \mu}} e^{-c_2(\lambda) \frac{t}{2} - c_2(\mu) \frac{T-t}{2}} d_\lambda d_\mu. \quad (3.17)$$

It is much more complicated to compute the limit of such a quantity, because it is a sum over a set of indices whose size depends on N , and thus it cannot be treated as a simple series. Dahlqvist and Norris [DN17] found a way to pass through this obstacle, using the empirical distribution of the highest weights

$$\mu_\lambda = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i}, \quad \forall \lambda \in \widehat{U}(N).$$

Indeed, they applied a large deviation principle found by Guionnet and Maïda in [GM05], and they used some concentration results as well as contour integrals making rigorous the arguments already present in [Bou94], to show that

$$\lim_{N \rightarrow \infty} W_\ell = \frac{2}{n\pi} \int_0^\infty \cosh\left((2t - T) \frac{x}{2}\right) \sin(\pi \rho_T(x)) dx. \quad (3.18)$$

⁴This $n = 1$ case is also a particular case of another (simpler) rule, which is Pieri's rule, given in Prop.1.3.14.

In the equation above, ρ_T denotes the density, with respect to Lebesgue measure, of the minimizer of the functional \mathcal{I}_T on the set $\mathcal{M}_1(\mathbb{R})$ of probability measures on \mathbb{R} having a density with respect to Lebesgue measure such that this density takes values in $[0, 1]$, as

$$\mathcal{I}_T(\mu) = \begin{cases} \int_{\mathbb{R}^2} (x^2 + y^2)^{\frac{T}{2}} - 2 \log |x - y| d\mu(x) d\mu(y) & \text{if } \mu([a, b]) \leq b - a \forall [a, b] \subset \mathbb{R}, \\ +\infty & \text{otherwise.} \end{cases}$$

It appears that this minimizer actually is the semicircle distribution with variance $\frac{1}{T}$ when $T \leq \pi^2$, and a much more complicated distribution otherwise. The fact that this distribution changes at the critical value π^2 is called the *Douglas–Kazakov transition phase*, named after the physicists who conjectured it in [DK93]. This conjecture was proved independently by Liechty and Wang [LW16] and Lévy and Maïda [LM15].

The Wilson loop expectation W_ℓ and its limit admit a generalization, in the sense that there is a closed formula for $\mathbb{E}[\text{tr}(H_\ell^n)]$ for any $n \in \mathbb{N}$ and an explicit expression of its limit, for both the plane and the sphere; they are given respectively by the moments of the unitary Brownian motion and the unitary Brownian bridge, and their limits are respectively computed in [Bia97] and [DN17], based on formula (3.12). However, as we will see in the next section, these formulæ are superfluous to know the existence and the expression of the master field.

3.1.2 Makeenko–Migdal equations

The *Makeenko–Migdal* equations are named after the authors of [MM79] who derived them using an informal integration by parts. Nevertheless, they could also be named after Kazakov and Kostov who gave them in [KK80, §4] the form we will use. Given a loop in a surface with a self-intersection point⁵, these equations give an expression of the partial derivatives of its Wilson loop average with respect to the areas of the faces that are adjacent to the self-intersection point. Let us be more explicit about this expression. Let Σ is either the plane or a compact surface and $\mathbb{G} = (\mathbb{E}, \mathbb{V}, \mathbb{F})$ a graph embedded in Σ . If

- ℓ is a loop on Σ obtained as a path in \mathbb{G} with a simple crossing on a vertex $v \in \mathbb{V}$,
- ℓ_1 and ℓ_2 are subloops of ℓ with basepoint v such that $\ell = \ell_1 \ell_2$,
- $(e_1, e_2, e_3, e_4) \in \mathbb{E}^4$ are edges that start or end with v , and such that $\ell_1 = e_1 w_1 e_2$ and $\ell_2 = e_4 w_2 e_3$ with w_1 and w_2 some words in edges of $\mathbb{E} \setminus \{e_1, e_2, e_3, e_4\}$,
- t_1, t_2, t_3 and t_4 are the areas of the faces bordered by ℓ that share the common vertex v , labelled counterclockwise and starting from the face between e_1 and e_4 ,

then the Makeenko–Migdal equation reads

$$\left(\frac{\partial}{\partial t_1} - \frac{\partial}{\partial t_2} + \frac{\partial}{\partial t_3} - \frac{\partial}{\partial t_4} \right) W_\ell = \mathbb{E}[\text{tr}(H_{\ell_1}) \text{tr}(H_{\ell_2})]. \quad (3.19)$$

⁵From now on, we will only consider such points to be simple crossings, *i.e.* vertices that only have 2 incoming edges and 2 outgoing edges.

The splitting of ℓ into ℓ_1 and ℓ_2 is illustrated in Fig. 3.3. The dashed parts of the loops are here to highlight the fact that the rest of the loops can be arbitrary, as complex and with as many self-intersection points as needed, without changing the equation. In that sense, the Makeenko–Migdal equation can be seen as local. In fact, it is this point of view that made Driver, Gabriel, Hall and Kemp [DGHK17] able to prove Eq. (3.19) for compact surfaces after Lévy [Lév17] and Dahlqvist [Dah16] proved them for the plane. Three of them also found in [DHK13] several new and simpler proofs of the Makeenko–Migdal equations in the plane. All these proofs are mainly based on the heat equation and an integration by parts on the structure group G .

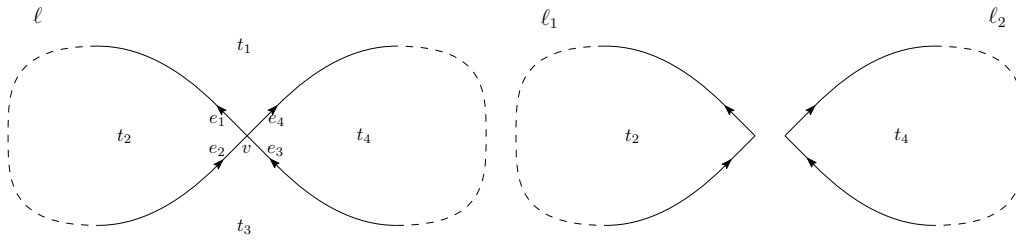


Figure 3.3: The loop ℓ (on the left) can be split into two subloops ℓ_1 and ℓ_2 (on the right).

Now, let us explain how to use Makeenko–Migdal equations to extend the master field from simple loops to more general loops. This analysis was done in [Lév17] for the plane and in [DN17, Hal18] for the sphere, with little variations. The common steps of this analysis are the following:

- Given a general loop L on a graph \mathbb{G} , we use Makeenko–Migdal equations to simultaneously vary areas of the faces delimited by L such that all areas but 2 tend to 0.
- The new loop produced with the previous step becomes a loop that winds n times around a simple loop ℓ , with n being an integer depending on L .

Instead of paraphrasing the original articles, let us explain how it works on a simple example that can also be found in [DK94] and [Hal18]: the eight-shaped loop. We consider a loop as in Fig. 3.4, on the left, embedded on the sphere with area T , and that produces a graph with 3 faces with respective areas a , b and c such that $a + b + c = T$. If one wants to understand the case of the plane, it suffices to let $c \rightarrow \infty$ ⁶.

If we apply the Makeenko–Migdal equation to the self-intersection point of the loop, we may shrink the faces with areas a and b such that the smaller one becomes a point: if we assume $a < b$ it leads to the new loop on the right of Fig. 3.4. More precisely, if we denote $\Phi_L(a, b, c)$ the large N limit of the Wilson loop expectation of the loop on the left, the Makeenko–Migdal equation yields

$$\frac{d}{dt} \Phi_L(a - t, b - t, c + 2t) = \lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_{\ell_1(t)}) \text{tr}(H_{\ell_2(t)})], \quad (3.20)$$

where $\ell_1(t)$ and $\ell_2(t)$ are subloops of the initial loops obtained by an identical splitting as the one in Fig. 3.3. We are able to compute the limit of the right-hand side of (3.20), although it might not be

⁶This choice is actually arbitrary: because of the topology of the sphere, the ‘external’ face of the loop could be any other face, but we chose this one because it justifies the name ‘eight-shaped’.

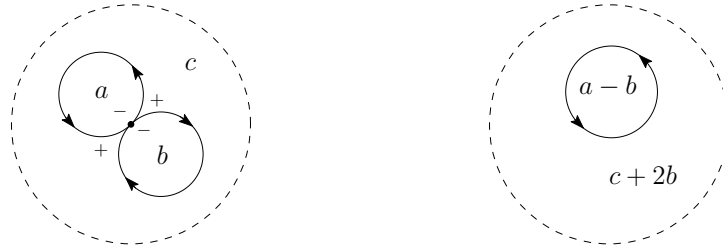


Figure 3.4: The loop on the left can be continuously deformed into the loop on the right, provided that $a > b$, and the Makeenko–Migdal equations link their Wilson loop expectations.

obvious at this point: it is possible to show⁷ that

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_{\ell_1})] \mathbb{E}[\text{tr}(H_{\ell_2})] = \lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_{\ell_1})] \lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_{\ell_2})].$$

Once we have this result, if we note $\Phi_\ell = \lim_{N \rightarrow \infty} W_\ell$ for any simple loop, we obtain

$$\frac{d}{dt} \Phi_L(a - t, b - t, c + 2t) = \Phi_{\ell_1(t)} \Phi_{\ell_2(t)}. \tag{3.21}$$

If we integrate this equation we can finally link the value $\Phi_L(a, b, c)$ (which we want to know) to $\Phi_L(a - b, 0, c + 2b)$ which we can compute with the help of (3.18).

The general results for arbitrary loops is based on similar arguments, which also make use of the *winding number* of a loop with respect to a face of the graph. This number is defined by how many times the loop winds around the considered face, and can be canonically defined, up to an additive constant, for any compact surface. This was proved in [Lévi17] for the plane, [DN17] for the sphere, and [Hal18] for any compact surface. Let us finish this section with the general theorem that follows from the previous discussion.

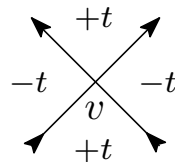


Figure 3.5: A zoom on the checkerboard pattern around the crossing at v .

Theorem 3.1.4. *Let Σ be the \mathbb{R}^2 or a sphere, and L be a loop on in Σ with a finite number of simple self-interesection points, eventually completed into a graph.*

(i) *The limit of its Wilson loop expectation*

$$\Phi_L = \lim_{N \rightarrow \infty} W_L \tag{3.22}$$

exists and is a continuous function of the areas of the faces bordered by L .

⁷We will not give the proof of this statement, but it comes from the fact that $\text{Var}(\text{tr}(H_\ell))$ tends to 0 for any simple loop ℓ as $N \rightarrow \infty$, and it is proved in [DN17] using similar arguments as for the Wilson loop expectations.

(ii) The limit of its Wilson loop variance is zero:

$$\lim_{N \rightarrow \infty} \text{Var}(\text{tr}(H_L)) = 0. \quad (3.23)$$

(iii) The limit Φ_L satisfy a large- N Makeenko–Migdal equation as follows: if v is a self-intersection point of L and if we vary the areas of the faces in a checkerboard pattern as in Fig. 3.5, resulting in a family of curves $L(t)$, then

$$\frac{d}{dt} \Phi_{L(t)} = \Phi_{L_1(t)} \Phi_{L_2(t)}, \quad (3.24)$$

where L_1 and L_2 are the subloops of L obtained by the usual splitting of Makeenko–Migdal.

Note that it defines a function Φ over $\mathcal{P}(\Sigma)$ as in Def. 1.4.6, which we can also call *master field*. Thanks to Eq. (3.23), we obtain that

$$\text{tr}(H_\ell) \xrightarrow{\mathbb{P}} \Phi_L$$

when N tends to infinity; in other terms, Wilson loops converge in probability to the master field (or to its trace, depending on the definition we take), not only in expectation. Note that Lévy [Lév17] proved that this convergence even holds *almost surely*. With Theorem 3.1.4 and Equations (3.16) and (3.17), we have all the needed tools to construct the master field on the plane and on the sphere, and to compute its value on loops with a finite number of self-intersections.

3.2 COMPACT SURFACES OF HIGHER GENUS

After the master field was constructed in the plane and in the sphere, one could expect it to be constructed on other compact surfaces than the sphere, possibly depending on their genus. This construction still remains unsolved, although Hall [Hal15] did a first step towards it, and the purpose of this section is to make another step.

We start by giving an expression of Wilson loop expectation and variance for a contractible simple loop in surfaces with genus 1 and higher and with structure group $U(N)$ or $SU(N)$, then we compute their limit using the theory of *almost flat highest weights* developed in Chapter 2. We also prove that the fluctuations are of order $\frac{1}{N^{1-\varepsilon}}$ for $\varepsilon > 0$ arbitrarily small, and we give a simpler proof of the convergence of Wilson loop expectation and variance for $g \geq 2$ with structure group $U(N)$.

Then, we discuss the case of loops with nontrivial homology, which is a case that does not appear in the plane or the sphere, as well as the joint distribution of the holonomy of the generators of the group of reduced loops on a graph embedded in the surface.

Finally, we discuss what is still missing in order to obtain the master field.

3.2.1 Contractible simple loops

Let us start with simple loops that are contractible, that is, homotopy equivalent to a point. We will show in this section that their associated Wilson loop expectation and variance behave exactly as in the plane when N tends to infinity. In other words, the global structure of the surface has no influence on such loops – provided the surface is not the sphere.

Formulæ for the expectation and variance

Before we give the formulæ of the Wilson loop expectation and variance, we start by giving a more precise idea of the loops we consider.

Definition 3.2.1. Let $\Sigma_{g,T}$ be an orientable compact connected surface of genus $g \geq 1$ and of area T . A loop ℓ in $\Sigma_{g,T}$ is *simple* if there exists an injective continuous map

$$\phi : S^1 \rightarrow \Sigma_{g,T}$$

such that $\phi(S^1) = \ell$.

Given a simple loop ℓ on $\Sigma_{g,T}$, there are two possibilities: either $\Sigma_{g,T} \setminus \ell$ is connected, and ℓ is said to be *nonseparating*, or $\Sigma_{g,T} \setminus \ell$ contains at least two connected components and ℓ is said to be *separating*. It can be proved (cf. [Sti93, §6.3]) that any nonseparating loop is canonically homeomorphic to an edge of the fundamental domain of $\Sigma_{g,T}$ ⁸ and that any separating loop splits $\Sigma_{g,T}$ into two components, which are homeomorphic to compact connected orientable surfaces with boundary, and these surfaces have respectively genus g_1 and g_2 such that $g_1 + g_2 = g$. Furthermore, ℓ is contractible if and only if $g_1 = 0$ or $g_2 = 0$.

If ℓ is a contractible simple loop on a surface $\Sigma_{g,T}$, then it is the boundary of a topological disk⁹ D of area $t \in (0, T)$; t will be called the *interior area* of ℓ . If we remark that $\Sigma_{g,T} \setminus D$ is homeomorphic to a surface Σ' with boundary, then we can choose a set of generators $\{a_1, b_1, \dots, a_g, b_g\}$ of $\pi_1(\Sigma')$, and we can pull them back by homeomorphism into generators of $\pi_1(\Sigma_{g,T})$. By taking the base point v_1 of these generators and the base point v_2 of ℓ , and considering a simple curve e from v_1 to v_2 , we have that $\{a_1, b_1, \dots, a_g, b_g, e, \ell\}$ is the set of edges of an admissible graph \mathbb{G} with two faces of respective areas t and $T - t$. Such a graph is illustrated in Fig. 3.6 for a surface $\Sigma_{2,T}$ of genus 2.

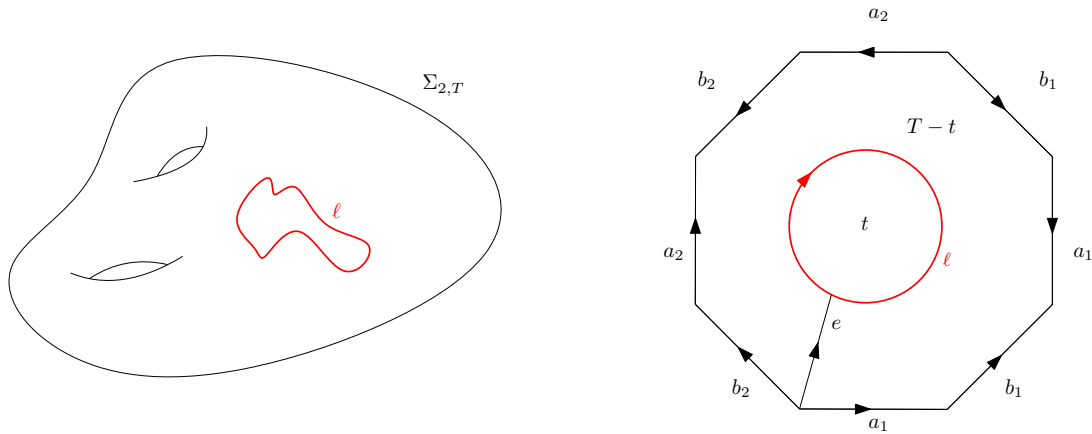


Figure 3.6: An contractible simple loop ℓ (on the left) and the oriented admissible graph associated to it (on the right).

If ℓ is an admissible simple loop of interior area t , then we can compute its Wilson loops expectation $W_\ell = \mathbb{E}[\text{tr}(H_\ell)]$, using either the Driver–Sengupta formula (1.87) or directly Eq. (1.90) from Proposition 1.4.5¹⁰:

⁸It means in particular that such a loop can be completed into a set of generators of $\pi_1(\Sigma_{g,T})$.
⁹A *topological disk* is a two-dimensional topological manifold with boundary homeomorphic to a closed disk.
¹⁰This formula depends on the orientation of the graph, but the orientation of ℓ is the only one that actually matters.

$$W_\ell = \frac{1}{Z_N(g, T)} \int_{G^{2g+1}} \text{tr}(x) p_t(x^{-1}) p_{T-t}(x[y_1, z_1] \cdots [y_g, z_g]) dx \prod_{i=1}^g dy_i dz_i, \quad (3.25)$$

where $(p_t)_{t \in \mathbb{R}_+}$ is the heat kernel on G .

We can also define the *Wilson loop variance* as the variance of Wilson loop functional – as tautological as it seems. It leads to the following formula, because the Wilson loop is a complex random variable:

$$\text{Var}[\text{tr}(H_\ell)] = \mathbb{E}[|\text{tr}(H_\ell)|^2] - |\mathbb{E}[\text{tr}(H_\ell)]|^2.$$

It appears that

$$\text{Var}[\text{tr}(H_\ell)] = \mathbb{E}[\text{tr}(H_\ell) \overline{\text{tr}(H_\ell)}] - |\mathbb{E}[\text{tr}(H_\ell)]|^2 = \mathbb{E}[\text{tr}(H_\ell) \text{tr}(H_\ell^*)] - |\mathbb{E}[\text{tr}(H_\ell)]|^2, \quad (3.26)$$

so that the variance can be explicitly computed as long as we know the expectations $\mathbb{E}[\text{tr}(H_\ell)]$ and $\mathbb{E}[\text{tr}(H_\ell) \text{tr}(H_\ell^*)]$.

In order to push further the computation of Wilson loop expectations when the gauge group is $U(N)$ or $SU(N)$, we will use the harmonic analysis tools that we described in Section 3.1.1, based on irreducible representations. Our goal will be to prove the following Proposition.

Proposition 3.2.1 (Wilson loop expectation and variance). *Let $\Sigma_{g,T}$ be an orientable compact connected surface of genus $g \geq 1$ and of area T , ℓ be a contractible loop of interior area t , and $G = SU(N)$ or $U(N)$ be the structure group. If we set $q = e^{-T/2}$, then we have the following formulae:*

(i) *If $G = SU(N)$, then*

$$\mathbb{E}[\text{tr}(H_\ell)] = \frac{1}{N Z'_N(g, T)} \sum_{\substack{\lambda, \mu \in \widehat{SU}(N) \\ \mu \nearrow \lambda}} \frac{q^{c'_2(\mu)}}{(d_\mu)^{2g-2}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c'_2(\mu) - c'_2(\lambda))}, \quad (3.27)$$

$$\mathbb{E}[\text{tr}(H_\ell) \text{tr}(H_\ell^*)] = \frac{1}{N^2 Z'_N(g, T)} \sum_{\substack{\lambda, \mu \in \widehat{SU}(N) \\ \lambda \sim \mu}} \frac{q^{c'_2(\mu)}}{(d_\mu)^{2g-2}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c'_2(\mu) - c'_2(\lambda))}. \quad (3.28)$$

(ii) *If $G = U(N)$, then*

$$\mathbb{E}[\text{tr}(H_\ell)] = \frac{1}{N Z_N(g, T)} \sum_{\substack{\lambda, \mu \in \widehat{U}(N) \\ \mu \nearrow \lambda}} \frac{q^{c_2(\mu)}}{(d_\mu)^{2g-2}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c_2(\mu) - c_2(\lambda))}, \quad (3.29)$$

$$\mathbb{E}[\text{tr}(H_\ell) \text{tr}(H_\ell^*)] = \frac{1}{N^2 Z_N(g, T)} \sum_{\substack{\lambda, \mu \in \widehat{U}(N) \\ \lambda \sim \mu}} \frac{q^{c_2(\mu)}}{(d_\mu)^{2g-2}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c_2(\mu) - c_2(\lambda))}. \quad (3.30)$$

Before we prove Prop. 3.2.1, let us introduce the following lemma, which enables to integrate Schur functions involving commutators.

Lemma 3.2.2. *Let G be a compact group and dg its normalized Haar measure. If (ρ, V) is an irreducible representation of G , we have*

$$\int_{G^2} \chi_\rho(x[y, z]) dy dz = \frac{\chi_\rho(x)}{d_\rho^2}, \quad \forall x \in G. \quad (3.31)$$

In order to prove Lemma 3.2.2, we need two intermediary propositions, which we will not prove because they are quite standard.

Proposition 3.2.3 ([Faro8], Prop.5.2). *Let G be a compact group, and dg its normalized Haar measure. For any irreducible representation (ρ, V) of G , we have*

$$\int_G \chi_\rho(xgyg^{-1}) dg = \frac{1}{d_\rho} \chi_\rho(x) \chi_\rho(y), \quad \forall (x, y) \in G^2.$$

Proposition 3.2.4. *Let G be a compact group and $(\rho, V), (\pi, W)$ two irreducible representations of G . Then*

$$\chi_\rho * \chi_\pi = \begin{cases} \frac{\chi_\rho}{d_\rho} & \text{if } \rho \sim \pi, \\ 0 & \text{otherwise.} \end{cases}$$

Proof of Lemma 3.2.2. First, according to Proposition 3.2.3, we have for any $(x, y) \in G^2$:

$$\int_G \chi_\rho(xyzzy^{-1}z^{-1}) dz = \frac{1}{d_\rho} \chi_\rho(xy) \chi_\rho(y^{-1}).$$

If we integrate out $y \in G$ it appears that

$$\int_{G^2} \chi_\rho(x[y, z]) dy dz = \frac{(\chi_\rho * \chi_\rho)(x)}{d_\rho},$$

which yields (3.31) using Proposition 3.2.4. □

We now have all the tools to prove Prop. 3.2.1.

Proof of Prop. 3.2.1. We will prove it in the case $G = U(N)$, the case $G = SU(N)$ being the same. Let us start from Eq. (3.25). We can decompose the heat kernels following (3.7):

$$W_\ell = \frac{1}{Z_N(g, T)} \sum_{\lambda, \mu \in \widehat{U}(N)} d_\lambda d_\mu e^{-\frac{c_2(\lambda)t}{2} - \frac{c_2(\mu)(T-t)}{2}} \int_{U(N)^{2g+1}} \text{tr}(x) s_\lambda(x^{-1}) s_\mu(x[y_1, z_1] \cdots [y_g, z_g]) dx \prod_{i=1}^g dy_i dz_i.$$

We can then apply Lemma 3.2.2 g times, which transforms the commutators into dimensions:

$$W_\ell = \frac{1}{Z_N(g, T)} \sum_{\lambda, \mu \in \widehat{U}(N)} d_\lambda (d_\mu)^{1-2g} e^{-\frac{c_2(\lambda)t}{2} - \frac{c_2(\mu)(T-t)}{2}} \int_{U(N)} \text{tr}(x) s_\lambda(x^{-1}) s_\mu(x) dx.$$

Then, using Pieri's rule and the fact that $\text{tr} = \frac{1}{N} \text{Tr}$ gives

$$W_\ell = \frac{1}{Z_N(g, T)} \sum_{\lambda, \mu \in \widehat{U}(N)} \frac{d_\lambda(d_\mu)^{1-2g}}{N} e^{-\frac{c_2(\lambda)t}{2} - \frac{c_2(\mu)(T-t)}{2}} \sum_{\substack{\nu \in \widehat{U}(N) \\ \nu \searrow \lambda \mu}} \int_{U(N)} s_\lambda(x^{-1}) s_\nu(x) dx.$$

If we set $q = e^{-T/2}$ and use the orthogonality relations of characters, it yields Eq. (3.29).

In the same manner as in Eq. (3.25), we can compute $\mathbb{E}[\text{tr}(H_\ell) \text{tr}(H_\ell^*)]$ as

$$\begin{aligned} \mathbb{E}[\text{tr}(H_\ell) \text{tr}(H_\ell^*)] &= \\ &= \frac{1}{Z_N(g, T)} \int_{U(N)^{2g+1}} \text{tr}(x) \text{tr}(x^{-1}) p_t(x^{-1}) p_{T-t}(x[y_1, z_1] \cdots [y_g, z_g]) dx \prod_{i=1}^g dy_i dz_i, \end{aligned}$$

which can be rewritten, using the heat kernel decomposition, Lemma 3.2.2 and Pieri's rule:

$$\begin{aligned} \mathbb{E}[\text{tr}(H_\ell) \text{tr}(H_\ell^*)] &= \frac{1}{N^2 Z_N(g, T)} \sum_{\lambda, \mu \in \widehat{U}(N)} d_\lambda(d_\mu)^{1-2g} e^{-\frac{c_2(\lambda)t}{2} - \frac{c_2(\mu)(T-t)}{2}} \\ &\quad \times \sum_{\substack{\nu, \tau \in \widehat{U}(N) \\ \nu \searrow \lambda \mu, \tau \searrow \lambda}} \int_{U(N)} s_\tau(x^{-1}) s_\nu(x) dx. \end{aligned}$$

Setting $q = e^{-T/2}$ as before and using the orthogonality of Schur functions, we obtain Eq. (3.30) as expected. \square

It is now time to state the main results of this section, which give the limits of the Wilson loop expectation and variance for a simple loop in a closed topological disk.

Theorem 3.2.5. *Let $\Sigma_{g,T}$ be an orientable compact connected surface of genus $g \geq 1$ and of area T , ℓ be a contractible simple loop of interior area t , and $G = \text{SU}(N)$ or $\text{U}(N)$ be the structure group. The associated Wilson loop expectation converges, as $N \rightarrow \infty$, and its limit is*

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_\ell)] = e^{-\frac{t}{2}}. \quad (3.32)$$

Note that the limit does actually not depend on the genus of the surface, as long as it is greater or equal to 1. The value of the limit is the same as in the plane. The result about the variance is the following.

Theorem 3.2.6. *Let $\Sigma_{g,T}$ be an orientable compact connected surface of genus $g \geq 1$ and of area T , ℓ be a contractible simple loop of interior area t , and $G = \text{SU}(N)$ or $\text{U}(N)$ be the structure group. The associated Wilson loop variance satisfies the following limit:*

$$\lim_{N \rightarrow \infty} \text{Var}[\text{tr}(H_\ell)] = 0. \quad (3.33)$$

Tools for asymptotic study

In order to compute the limits of Equations (3.27) to (3.30), we will need the theory of *almost flat highest weights* developed in [Lem19] and introduced in Chapter 2. Let us recall a few results about it, sometimes without proof.

From two integer partitions $\alpha = (\alpha_1 \geq \dots \geq \alpha_r > 0)$ and $\beta = (\beta_1 \geq \dots \geq \beta_s > 0)$ of respective lengths r and s , and an integer $n \in \mathbb{Z}$, we can form, for all $N \geq r + s + 1$, the highest weight

$$\lambda_N(\alpha, \beta, n) = (\alpha_1 + n, \dots, \alpha_r + n, \underbrace{n, \dots, n}_{N-r-s}, n - \beta_s, \dots, n - \beta_1) \in \widehat{U}(N). \quad (3.34)$$

We extend this definition in the obvious way to the cases where one or both of the partitions α and β are the empty partition.

We can also form the highest weight

$$\lambda_N(\alpha, \beta) = \lambda_N(\alpha, \beta, \beta_1) \in \widehat{SU}(N),$$

with the convention that $\lambda_N(\alpha, \emptyset) = \lambda_N(\alpha, \emptyset, 0) = (\alpha_1, \dots, \alpha_r, 0)$.

We have seen in Section 2.2.2 that these constructions can be reversed, in the sense that given a highest weight $\lambda \in \widehat{U}(N)$, then we can define unambiguously α, β and n such that $\lambda = \lambda_N(\alpha, \beta, n)$ [†]. In this case, we will denote them by $\alpha_\lambda, \beta_\lambda$ and n_λ to emphasize the fact that they are determined by λ . As we will see, this decomposition will be of great help to control the Casimir numbers of highest weights, as well as their dimensions. Let us start with estimations about Casimir numbers.

Lemma 3.2.7. *Let $\lambda \in \widehat{SU}(N)$. Set $k = |\alpha_\lambda| + |\beta_\lambda|$. Then the following inequalities hold:*

$$k - \frac{k^2}{N} \leq c'_2(\lambda) \leq k + \frac{k^2}{N} + \frac{k^2}{N^2}, \quad (3.35)$$

$$\frac{k}{2} \leq c'_2(\lambda). \quad (3.36)$$

This result was already proved in Chapter 2: it is actually Lemma 2.2.6. We can complete it with the following proposition.

Proposition 3.2.8. *Let $(\lambda, \mu) \in \widehat{SU}(N)^2$ be two highest weights and set $\alpha = \alpha_\mu$ and $\beta = \beta_\mu$. If $\lambda \searrow \mu$ or if $\lambda \sim \mu$, then we have for N large enough*

$$-\frac{T}{2}c'_2(\mu) + \frac{t}{2}(c'_2(\mu) - c'_2(\lambda)) \leq -\frac{T}{8}(|\alpha| + |\beta|) + t. \quad (3.37)$$

[†]More precisely, it comes from the mappings given in (2.28) and (2.29).

Proof. Let us start with The case when $\lambda \searrow \mu$. Let i_0 be the index such that $\lambda_{i_0} = \mu_{i_0} + 1$. From the definitions of $c'_2(\mu)$ and $c'_2(\lambda)$ (see (3.9)) and the fact that $i_0 \leq N$, we have the estimation

$$c'_2(\lambda) - c'_2(\mu) = 1 + \frac{2(\mu_{i_0} + 1 - i_0)}{N} - \frac{2|\mu| + 1}{N^2} \geq -2 - \frac{2|\mu|}{N^2}.$$

From (3.36) and the fact that $|\mu| = |\alpha| - |\beta| + N\beta_1 \leq |\alpha| + N|\beta|$, we then get

$$\begin{aligned} -\frac{T}{2}c'_2(\mu) + \frac{t}{2}(c'_2(\mu) - c'_2(\lambda)) &\leq -\frac{T}{4}(|\alpha| + |\beta|) + \frac{t}{2N^2}(|\alpha| + N|\beta|) + t \\ &\leq (|\alpha| + |\beta|) \left(\frac{t}{2N} - \frac{T}{4} \right) + t, \end{aligned}$$

and the inequality (3.37) is satisfied for N such that $\frac{t}{2N} < \frac{T}{8}$.

Now let us prove the case when $\lambda \sim \mu$. If $\lambda = \mu$ then the result directly follows from Lemma 3.2.7. Otherwise, there are $i_0 \neq j_0$ such that $\lambda_{i_0} = \mu_{i_0} + 1$, $\lambda_{j_0} = \mu_{j_0} - 1$ and $\lambda_i = \mu_i$, $\forall i \notin \{i_0, j_0\}$. Using the definition of Casimir number, we have

$$c'_2(\mu) - c'_2(\lambda) = \frac{2(\mu_{j_0} - \mu_{i_0} + i_0 - j_0 - 1)}{N}.$$

As $\mu_1 \geq \dots \geq \mu_N$, we have $\mu_1 - 1 > \dots > \mu_N - N$ and in particular

$$0 < \mu_i - \mu_j + j - i \leq \mu_1 + N - 1, \forall i < j.$$

If $i_0 < j_0$ then we have $c'_2(\mu) - c'_2(\lambda) \leq 0$ and the bound given by the case $\lambda = \mu$ still holds. Otherwise, we have

$$c'_2(\mu) - c'_2(\lambda) \leq \frac{2(\mu_1 + N - 2)}{N} \leq 2 + \frac{2\mu_1}{N}.$$

Since $\mu_1 = \alpha_1 + \beta_1 \leq |\alpha| + |\beta|$ and using Lemma 3.2.7 we obtain

$$-\frac{T}{2}c'_2(\mu) + \frac{t}{2}(c'_2(\mu) - c'_2(\lambda)) \leq (|\alpha| + |\beta|) \left(\frac{t}{N} - \frac{T}{4} \right) + t.$$

The inequality follows, when N satisfies $\frac{t}{N} < \frac{T}{8}$. \square

We can also mention a similar result, that will be used to prove Thm. 3.2.5 in the unitary case, and which is a direct consequence of Prop. 2.2.5.

Lemma 3.2.9. *Let $\alpha = (\alpha_1 \geq \dots \geq \alpha_r)$ and $\beta = (\beta_1 \geq \dots \geq \beta_s)$ be integer partitions, $(n, N) \in \mathbb{Z}^2$ be two integers, such that $N \geq r + s + 1$.*

(i) *If α' is a partition such that $\alpha' \searrow \alpha$ and i_0 is the index such that $\alpha'_{i_0} = \alpha_{i_0} + 1$, then*

$$c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha', \beta, n)) = -1 - \frac{2}{N}(\alpha_{i_0} + n + 1 - i_0). \quad (3.38)$$

(ii) *If β' is a partition such that $\beta' \nearrow \beta$ and i_0 is the index such that $\beta'_{i_0} = \beta_{i_0} + 1$, then*

$$c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha, \beta', n)) = 1 + \frac{2}{N}(\beta_{i_0} - n - i_0). \quad (3.39)$$

From now on, let us fix a real $\gamma \in (0, \frac{1}{3})$, that we can consider as a control parameter¹². We split the set of highest weights of $\mathrm{SU}(N)$ in four disjoint subsets:

$$\begin{aligned}\Lambda_{N,1}^\gamma &= \{\lambda \in \widehat{\mathrm{SU}}(N) : |\alpha_\lambda| \leq N^\gamma, |\beta_\lambda| \leq N^\gamma\}, \\ \Lambda_{N,2}^\gamma &= \{\lambda \in \widehat{\mathrm{SU}}(N) : |\alpha_\lambda| > N^\gamma, |\beta_\lambda| \leq N^\gamma\}, \\ \Lambda_{N,3}^\gamma &= \{\lambda \in \widehat{\mathrm{SU}}(N) : |\alpha_\lambda| \leq N^\gamma, |\beta_\lambda| > N^\gamma\}, \\ \Lambda_{N,4}^\gamma &= \{\lambda \in \widehat{\mathrm{SU}}(N) : |\alpha_\lambda| > N^\gamma, |\beta_\lambda| > N^\gamma\}.\end{aligned}\tag{3.40}$$

We can do the same for highest weights of $\mathrm{U}(N)$, but the subsets will be denoted by $\Omega_{N,i}^\gamma$ instead of $\Lambda_{N,i}^\gamma$. In this framework, (3.35) can be refined as the following for any highest weight $\lambda \in \Lambda_{N,1}^\gamma$, called *almost flat*:

$$|\alpha_\lambda| + |\beta_\lambda| - 4N^{2\gamma-1} \leq c'_2(\lambda) \leq |\alpha_\lambda| + |\beta_\lambda| + 4N^{2\gamma-1} + 4N^{2\gamma-2}.\tag{3.41}$$

We can rewrite this as

$$|c'_2(\lambda) - (|\alpha_\lambda| + |\beta_\lambda|)| \leq 8N^{2\gamma-1}.\tag{3.42}$$

Another crucial point is the following: for N large enough, any partition of an integer not greater than N^γ has less than $\frac{N}{2}$ positive parts. Thus, if α and β are any two such partitions, the highest weight $\lambda_N(\alpha, \beta)$ is well defined, and belongs to $\Lambda_{N,1}^\gamma$. As a consequence, for N large enough,

$$\Lambda_{N,1}^\gamma = \{\lambda_N(\alpha, \beta), \alpha \vdash r, \beta \vdash s : r \leq N^\gamma, s \leq N^\gamma\}.\tag{3.43}$$

Using the bijection Φ given in (2.29), we get as well for almost flat highest weights of $\mathrm{U}(N)$ with N large enough,

$$\Omega_{N,1}^\gamma = \{\lambda_N(\alpha, \beta, n), \alpha \vdash r, \beta \vdash s : r \leq N^\gamma, s \leq N^\gamma, n \in \mathbb{Z}\}.\tag{3.44}$$

The dimension of almost flat highest weights can also be related to the dimension of irreducible representations of the symmetric group.

Proposition 3.2.10. *Let $\alpha = (\alpha_1 \geq \dots \geq \alpha_r)$ and $\beta = (\beta_1 \geq \dots \geq \beta_s)$ be two integer partitions, $N \geq r + s + 1$ an integer and $\gamma \in (0, \frac{1}{3})$ a real number. Let us assume that $|\alpha| \leq N^\gamma$ and $|\beta| \leq N^\gamma$. The partitions α and β induce two highest weights of $\mathrm{SU}(N)$, $\tilde{\alpha} = \lambda_N(\alpha, \emptyset)$ and $\tilde{\beta} = \lambda_N(\beta, \emptyset)$. We have the following facts.*

- (i) *If we set d^α as the dimension of the irreducible representation of $\mathfrak{S}_{|\alpha|}$ associated with α , then, assuming that N is large enough,*

$$\frac{d^\alpha N^{|\alpha|}}{|\alpha|!} (1 - 2N^{2\gamma-1}) \leq d_{\tilde{\alpha}} \leq \frac{d^\alpha N^{|\alpha|}}{|\alpha|!} (1 + 2N^{2\gamma-1}),\tag{3.45}$$

and the same result holds for β .

- (ii) *For any $n \in \mathbb{Z}$, the dimension of $\lambda_N(\alpha, \beta, n)$ satisfies the following estimation, assuming that N is large enough:*

$$\frac{d^\alpha d^\beta N^{|\alpha|+|\beta|}}{|\alpha|!|\beta|!} (1 - 24N^{3\gamma-1}) \leq d_{\lambda_N(\alpha, \beta, n)} \leq \frac{d^\alpha d^\beta N^{|\alpha|+|\beta|}}{|\alpha|!|\beta|!} (1 + 24N^{3\gamma-1})\tag{3.46}$$

¹²Intuitively, as we will discuss it later, we want it to be as small as possible, while remaining positive.

Note that this proposition generalizes a result by Gross and Taylor: in [GT93], they derived similar asymptotic expansions but in the case where $|\alpha|$ and $|\beta|$ were finite and not depending on N . However, we really need the stronger assumption $|\alpha|, |\beta| \leq N^\gamma$ as we will see later.

Proof of Proposition 3.2.10. (i) let us first recall that (cf. [GT93, VO04])

$$d_{\tilde{\alpha}} = \frac{d^\alpha N^{|\alpha|}}{|\alpha|!} \prod_{\substack{1 \leq i \leq r \\ 1 \leq j \leq \alpha_i}} \left(1 + \frac{j-i}{N}\right). \quad (3.47)$$

But for any $1 \leq i \leq r$ and $1 \leq j \leq \alpha_i$, we have $1-r \leq j-i \leq \alpha_i-1$, and under the assumption $|\alpha| \leq N^\gamma$ it implies that $|j-i| \leq N^\gamma$. Thus,

$$(1 - N^{\gamma-1})^{|\alpha|} \leq \prod_{\substack{1 \leq i \leq r \\ 1 \leq j \leq \alpha_i}} \left(1 + \frac{j-i}{N}\right) \leq (1 + N^{\gamma-1})^{|\alpha|}.$$

From the convexity inequality of the exponential function, we have

$$(1 + N^{\gamma-1})^{|\alpha|} \leq e^{|\alpha|N^{\gamma-1}}.$$

We can use the following reverse inequalities, that hold for any $x \in (0, \frac{1}{2})$:

$$e^x \leq 1 + 2x, \quad \log(1-x) \geq -2x.$$

It implies that, for N such that $N^{\gamma-1} < \frac{1}{2}$ (which is true for N large enough),

$$1 - 2N^{2\gamma-1} \leq e^{-2|\alpha|N^{\gamma-1}} \leq (1 - N^{\gamma-1})^{|\alpha|} \leq \prod_{\substack{1 \leq i \leq r \\ 1 \leq j \leq \alpha_i}} \left(1 + \frac{j-i}{N}\right) \leq e^{|\alpha|N^{\gamma-1}} \leq 1 + 2N^{2\gamma-1}.$$

This estimation, applied to (3.47), gives the expected result.

(ii) Let us first remark that, from the Weyl dimension formula, we have for any $n \in \mathbb{Z}$

$$d_{\lambda_N(\alpha, \beta, n)} = d_{\tilde{\alpha}} d_{\tilde{\beta}} Q(\alpha, \beta), \quad (3.48)$$

with

$$Q(\alpha, \beta) = \prod_{\substack{1 \leq i \leq r \\ 1 \leq j \leq s}} \frac{(N+1-i-j)(\alpha_i + \beta_j + N+1-i-j)}{(\alpha_i + N+1-i-j)(\beta_j + N+1-i-j)}.$$

As $r \leq |\alpha|$ and $s \leq |\beta|$, the assumption $|\alpha|, |\beta| \leq N^\gamma$ implies that we also have $r, s \leq N^\gamma$. For any $1 \leq i \leq r$ and $1 \leq j \leq s$, we have therefore

$$-2N^\gamma \leq 3 - 2N^\gamma \leq \alpha_i + \beta_j - i - j + 1 \leq 2N^\gamma - 1 \leq 2N^\gamma.$$

It implies that

$$\left| \frac{1 + \alpha_i + \beta_j - i - j}{N} \right| \leq 2N^{\gamma-1},$$

and we have the same bound for $\left| \frac{1+\alpha_i-i-j}{N} \right|$, $\left| \frac{1+\beta_j-i-j}{N} \right|$ and $\left| \frac{1-i-j}{N} \right|$, so that

$$Q(\alpha, \beta) = \prod_{\substack{1 \leq i \leq r \\ 1 \leq j \leq s}} \frac{(1 + A_N(i, j))(1 + B_N(i, j))}{(1 + C_N(i, j))(1 + D_N(i, j))},$$

with $|A_N(i, j)|, |B_N(i, j)|, |C_N(i, j)|, |D_N(i, j)| \leq 2N^{\gamma-1}$.

For any (i, j) we have

$$\frac{1}{C_N(i, j)} = 1 - \frac{1 + C_N(i, j)}{1 + C_N(i, j)} = 1 + C'_N(i, j),$$

with $|C'_N(i, j)| \leq 2|C_N(i, j)|$, and the same result holds for $D_N(i, j)$. It implies that

$$Q(\alpha, \beta) = \prod_{\substack{1 \leq i \leq r \\ 1 \leq j \leq s}} (1 + A_N(i, j))(1 + B_N(i, j))(1 + C'_N(i, j))(1 + D'_N(i, j)),$$

with $|A_N(i, j)|, |B_N(i, j)|, |C'_N(i, j)|, |D'_N(i, j)| \leq 4N^{\gamma-1}$. Hence, using the same inequalities as in (i) we get the estimation

$$e^{-8N^{3\gamma-1}} \leq (1 - 4N^{\gamma-1})^{N^{2\gamma}} \leq Q(\alpha, \beta) \leq (1 + 4N^{\gamma-1})^{N^{2\gamma}} \leq e^{4N^{3\gamma-1}},$$

which implies

$$1 - 8N^{3\gamma-1} \leq Q(\alpha, \beta) \leq 1 + 8N^{3\gamma-1}.$$

We can apply this, as well as the point (i), to get for $N^{2\gamma-1} < \frac{1}{4}$ (which is in particular true for N large enough)

$$d_{\lambda_N(\alpha, \beta, n)} \leq \frac{d^\alpha d^\beta N^{|\alpha|+|\beta|}}{|\alpha|!|\beta|!} (1 + 2N^{2\gamma-1})^2 (1 + 8N^{3\gamma-1}),$$

which can be simplified considering that for $(x, y, z) \in (0, \frac{1}{4})^3$ such that $x + y + z < \frac{1}{4}$,

$$(1 + x)(1 + y)(1 + z) \leq e^{x+y+z} \leq 1 + 2(x + y + z).$$

Indeed, for N such that $N^{3\gamma-1} + 2N^{2\gamma-1} < \frac{1}{4}$ we obtain

$$d_{\lambda_N(\alpha, \beta, n)} \leq \frac{d^\alpha d^\beta N^{|\alpha|+|\beta|}}{|\alpha|!|\beta|!} (1 + 8N^{2\gamma-1} + 16N^{3\gamma-1}) \leq \frac{d^\alpha d^\beta N^{|\alpha|+|\beta|}}{|\alpha|!|\beta|!} (1 + 24N^{3\gamma-1}),$$

and

$$d_{\lambda_N(\alpha, \beta, n)} \geq \frac{d^\alpha d^\beta N^{|\alpha|+|\beta|}}{|\alpha|!|\beta|!} (1 - 24N^{3\gamma-1}),$$

which proves the result. \square

Branching rules

Before proving Theorem 3.2.5, we still have to discuss a bit about branching rules. Indeed, in (3.27) (resp. (3.29)), a sum over $\lambda \searrow \mu$ appears, with λ and μ being highest weights of $SU(N)$ (resp. $U(N)$). We will show how this branching is transformed in the decomposition $\lambda = \lambda_N(\alpha, \beta, n)$ that we introduced in the previous section.

Proposition 3.2.11. *Let $\alpha = (\alpha_1 \geq \dots \geq \alpha_r)$, $\beta = (\beta_1 \geq \dots \geq \beta_s)$, $\alpha' = (\alpha'_1 \geq \dots \geq \alpha'_{r'})$ and $\beta' = (\beta'_1 \geq \dots \geq \beta'_{s'})$ be integer partitions, and $(n, n', N) \in \mathbb{Z}^3$ three integers such that $N \geq \max(r + s, r' + s')$. Then the following assertions are equivalent:*

- (i) $\lambda_N(\alpha', \beta', n') \searrow \lambda_N(\alpha, \beta, n)$,
- (ii) $(\alpha' \searrow \alpha, \beta' = \beta \text{ and } n' = n)$ or $(\beta' \nearrow \beta, \alpha' = \alpha \text{ and } n' = n)$.

Proof. In order to see the equivalence, recall the construction of $\lambda_N(\alpha, \beta, n)$ given in (3.34):

$$\lambda_N(\alpha, \beta, n) = (\alpha_1 + n, \dots, \alpha_r + n, \underbrace{n, \dots, n}_{N-r-s}, n - \beta_s, \dots, n - \beta_1) = (\lambda_1, \dots, \lambda_N).$$

The only way of having $\lambda_N(\alpha', \beta', n') \searrow \lambda_N(\alpha, \beta, n)$ is to increment a coefficient λ_i such that $\lambda_i > \lambda_{i+1}$. It clearly excludes the coefficients $\lambda_{r+1}, \dots, \lambda_{r+s}$. Two only ways remain: either we increment one of the coefficients $\lambda_1, \dots, \lambda_r$, or we increment one of the coefficients $\lambda_{r+s+1}, \dots, \lambda_N$. The first case corresponds to $\alpha' \searrow \alpha$ and the second one to $\beta' \nearrow \beta$ (while leaving the other parameters unchanged), according to the description of the coefficients in terms of α, β and n . The equivalence follows immediately. \square

The main consequence of this proposition, combined with (3.43) and (3.44), is that for N large enough,

$$\{(\lambda, \mu) \in \widehat{SU}(N) \times \Lambda_{N,1}^\gamma : \lambda \searrow \mu\}$$

splits into two disjoint sets

$$\{(\lambda_N(\alpha', \beta), \lambda_N(\alpha, \beta)) : |\alpha| \leq N^\gamma, |\beta| \leq N^\gamma, \alpha' \searrow \alpha\}$$

and

$$\{(\lambda_N(\alpha, \beta'), \lambda_N(\alpha, \beta)) : |\alpha| \leq N^\gamma, |\beta| \leq N^\gamma, \beta' \nearrow \beta\}.$$

From (3.44) we also have that, for N large enough,

$$\{(\lambda, \mu) \in \widehat{U}(N) \times \Omega_{N,1}^\gamma : \lambda \searrow \mu\}$$

splits into

$$\{(\lambda_N(\alpha', \beta, n), \lambda_N(\alpha, \beta, n)) : |\alpha| \leq N^\gamma, |\beta| \leq N^\gamma, \alpha' \searrow \alpha, n \in \mathbb{Z}\}$$

and

$$\{(\lambda_N(\alpha, \beta', n), \lambda_N(\alpha, \beta, n)) : |\alpha| \leq N^\gamma, |\beta| \leq N^\gamma, \beta' \nearrow \beta, n \in \mathbb{Z}\}.$$

The main advantage of these decompositions is that we make fully use of Lemma 3.2.9 that uses branching over partitions rather than highest weights. However, using Prop. 3.2.10 will somehow convert dimensions of representations of $U(N)$ or $SU(N)$ into dimensions of representations of \mathfrak{S}_n with some integer n . We will therefore need the following branching rules.

Proposition 3.2.12. *Let $\lambda \vdash n$ for any positive integer n . We have*

$$\sum_{\substack{\mu \vdash (n+1) \\ \mu \searrow \lambda}} \frac{d^\mu}{(n+1)d^\lambda} = 1, \quad (3.49)$$

and

$$\sum_{\substack{\mu \vdash (n-1) \\ \mu \nearrow \lambda}} \frac{d^\mu}{d^\lambda} = 1. \quad (3.50)$$

Proof. Let us recall the so-called branching rules on \mathfrak{S}_n , cf. [Sago1] for example:

$$\chi^\lambda \uparrow_{\mathfrak{S}_{n+1}} = \sum_{\substack{\mu \vdash (n+1) \\ \mu \searrow \lambda}} \chi^\mu,$$

and

$$\chi^\lambda \downarrow_{\mathfrak{S}_{n-1}} = \sum_{\substack{\mu \vdash (n-1) \\ \mu \nearrow \lambda}} \chi^\mu.$$

As the character of a restricted representation is equal to the restriction of the character, the second branching rule directly implies (3.50). For the character of an induced representation we have the following result [FH91, Eq.(3.18)]: if G is a finite group and H a subgroup of G , then for any character χ of a representation of H we have

$$\chi \uparrow^G (g) = \frac{1}{|H|} \sum_{\substack{x \in G \\ xgx^{-1} \in H}} \chi(xgx^{-1}), \quad \forall g \in G.$$

If we apply this formula to $G = \mathfrak{S}_{n+1}$, $H = \mathfrak{S}_n$, $\chi = \chi^\lambda$ and $g = 1$ we get (3.49) as expected. \square

Asymptotics of the expectation

We can now turn to the proof of Theorem 3.2.5. We will split it into one dedicated to $SU(N)$ and one dedicated to $U(N)$, as the proofs are slightly different.

Proof of Theorem 3.2.5 in the special unitary case. Let $\gamma \in (0, \frac{1}{3})$ be a fixed real number. For any $i \in \{1, 2, 3, 4\}$ we define $\mathbb{E}_i^\gamma[\text{tr}(H_\ell)]$ as follows:

$$\mathbb{E}_i^\gamma[\text{tr}(H_\ell)] = \frac{1}{NZ'_N(g, T)} \sum_{\mu \in \Lambda_{N,i}^\gamma} \frac{q^{c'_2(\mu)}}{(d_\mu)^{2g-2}} \sum_{\substack{\lambda \in \widehat{SU}(N) \\ \lambda \searrow \mu}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c'_2(\mu) - c'_2(\lambda))},$$

with the sets $\Lambda_{N,i}^\gamma$ being as in (3.40).

From Equation (3.27) and the definition of each $\Lambda_{N,i}^\gamma$ we have $\mathbb{E}[\text{tr}(H_\ell)] = \sum_{i=1}^4 \mathbb{E}_i^\gamma[\text{tr}(H_\ell)]$. We will show first that

$$\lim_{N \rightarrow \infty} \mathbb{E}_1^\gamma[\text{tr}(H_\ell)] = e^{-t/2},$$

and then that for $2 \leq i \leq 4$,

$$\lim_{N \rightarrow \infty} \mathbb{E}_i^\gamma[\text{tr}(H_\ell)] = 0,$$

which will imply Equation (3.32).

From Equation (3.43) we have, for N large enough,

$$\mathbb{E}_1^\gamma[\text{tr}(H_\ell)] = \frac{1}{NZ'_N(g, T)} \sum_{|\alpha|, |\beta| \leq N^\gamma} \frac{q^{c'_2(\lambda_N(\alpha, \beta))}}{(d_{\lambda_N(\alpha, \beta)})^{2g-2}} \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \searrow \lambda_N(\alpha, \beta)}} \frac{d_\lambda}{d_{\lambda_N(\alpha, \beta)}} e^{\frac{t}{2}(c'_2(\lambda_N(\alpha, \beta)) - c'_2(\lambda))}.$$

Furthermore, we can notice that adding a box to $\lambda_N(\alpha, \beta)$ is equivalent to adding a box to the partition α to get $\alpha' \searrow \alpha$ or removing one from the partition β to get $\beta' \nearrow \beta$ such that α' or β' is another partition. It means that

$$\begin{aligned} \mathbb{E}_1^\gamma[\text{tr}(H_\ell)] &= \frac{1}{NZ'_N(g, T)} \sum_{|\alpha|, |\beta| \leq N^\gamma} \frac{q^{c'_2(\lambda_N(\alpha, \beta))}}{(d_{\lambda_N(\alpha, \beta)})^{2g-2}} \\ &\quad \times \left(\sum_{\alpha' \searrow \alpha} \frac{d_{\lambda_N(\alpha', \beta)}}{d_{\lambda_N(\alpha, \beta)}} e^{\frac{t}{2}(c'_2(\lambda_N(\alpha, \beta)) - c'_2(\lambda_N(\alpha', \beta)))} \right. \\ &\quad \left. + \sum_{\beta' \nearrow \beta} \frac{d_{\lambda_N(\alpha, \beta')}}{d_{\lambda_N(\alpha, \beta)}} e^{\frac{t}{2}(c'_2(\lambda_N(\alpha, \beta)) - c'_2(\lambda_N(\alpha, \beta')))} \right). \end{aligned} \quad (3.51)$$

We will first control the differences of Casimir numbers, then the ratios of dimensions, and show that only the sum over $\alpha' \searrow \alpha$ contributes to the large N limit. From (3.42) we obtain that for any α and β such that $|\alpha|, |\beta| \leq N^\gamma$,

$$-1 - 16N^{2\gamma-1} \leq c'_2(\lambda_N(\alpha, \beta)) - c'_2(\lambda_N(\alpha', \beta)) \leq -1 + 16N^{2\gamma-1}, \quad \forall \alpha' \searrow \alpha$$

and

$$1 - 16N^{2\gamma-1} \leq c'_2(\lambda_N(\alpha, \beta)) - c'_2(\lambda_N(\alpha, \beta')) \leq 1 + 16N^{2\gamma-1}, \quad \forall \beta' \nearrow \beta.$$

We obtain the following estimation for $\mathbb{E}_1^\gamma[\text{tr}(H_\ell)]$:

$$\mathbb{E}_1^\gamma[\text{tr}(H_\ell)] = \frac{e^{\varepsilon(N, \gamma)}}{Z'_N(g, T)} \sum_{|\alpha|, |\beta| \leq N^\gamma} \frac{q^{|\alpha|+|\beta|}}{(d_{\lambda_N(\alpha, \beta)})^{2g-2}} \left(\frac{e^{-\frac{t}{2}}}{N} \sum_{\alpha' \searrow \alpha} \frac{d_{\lambda_N(\alpha', \beta)}}{d_{\lambda_N(\alpha, \beta)}} + \frac{e^{\frac{t}{2}}}{N} \sum_{\beta' \nearrow \beta} \frac{d_{\lambda_N(\alpha, \beta')}}{d_{\lambda_N(\alpha, \beta)}} \right), \quad (3.52)$$

with

$$|\varepsilon(N, \gamma)| \leq (4T + 8t)N^{2\gamma-1}.$$

Using Prop. 3.2.10 and the fact that for any $x \in (0, \frac{1}{4})$

$$\frac{1}{1-2x} \leq 1+4x \quad \text{and} \quad \frac{1}{1+2x} \geq 1-4x,$$

we have for any $\alpha' \searrow \alpha$ and N large enough

$$\frac{d_{\lambda_N(\alpha',\beta)}}{d_{\lambda_N(\alpha,\beta)}} \leq \frac{Nd^{\alpha'}}{(|\alpha|+1)d^\alpha} (1+24N^{3\gamma-1})(1+48N^{3\gamma-1}) \leq \frac{Nd^{\alpha'}}{(|\alpha|+1)d^\alpha} (1+144N^{3\gamma-1})$$

and

$$\frac{d_{\lambda_N(\alpha',\beta)}}{d_{\lambda_N(\alpha,\beta)}} \geq \frac{Nd^{\alpha'}}{(|\alpha|+1)d^\alpha} (1-144N^{3\gamma-1}).$$

Combined with Prop. 3.2.12 these equations yield

$$1-144N^{3\gamma-1} \leq \frac{1}{N} \sum_{\alpha' \searrow \alpha} \frac{d_{\lambda_N(\alpha',\beta)}}{d_{\lambda_N(\alpha,\beta)}} \leq 1+144N^{3\gamma-1}. \quad (3.53)$$

With similar arguments we also have, because $|\beta| \leq N^\gamma$:

$$0 \leq \frac{1}{N} \sum_{\beta' \nearrow \beta} \frac{d_{\lambda_N(\alpha,\beta')}}{d_{\lambda_N(\alpha,\beta)}} \leq \frac{|\beta|}{N^2} (1+144N^{3\gamma-1}) \leq 37N^{\gamma-2}. \quad (3.54)$$

Combining this with (3.52) we find

$$\mathbb{E}_1^\gamma[\text{tr}(H_\ell)] = e^{-\frac{t}{2}} \eta(N, \gamma) \frac{e^{\varepsilon(N, \gamma)}}{Z'_N(g, T)} \sum_{|\alpha|, |\beta| \leq N^\gamma} \frac{q^{|\alpha|+|\beta|}}{(d_{\lambda_N(\alpha, \beta)})^{2g-2}}, \quad (3.55)$$

with

$$1-144N^{2\gamma-1} \leq \eta(N, \gamma) \leq 1+144N^{3\gamma-1} + 37e^t N^{\gamma-2}.$$

If we let $N \rightarrow \infty$, the remaining sum has the same limit as $Z'_N(g, T)$ as we have seen in the proof of Thm. 2.1.2 (see page 98 for $g \geq 2$ and page 101 for $g = 1$). Moreover, $\varepsilon(N, \gamma)$ tends to 0 and $\eta(N, \gamma)$ to 1. From all of this we can deduce that $\lim_{N \rightarrow \infty} \mathbb{E}_1^\gamma[\text{tr}(H_\ell)] = e^{-\frac{t}{2}}$.

Now we have to show that the other \mathbb{E}_i^γ all tend to 0 when $N \rightarrow \infty$. We have

$$\mathbb{E}_i^\gamma[\text{tr}(H_\ell)] = \frac{1}{NZ'_N(g, T)} \sum_{\mu \in \Lambda_{N,i}^\gamma} \frac{q^{c_2(\mu)}}{(d_\mu)^{2g-2}} \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \searrow \mu}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c_2(\mu) - c_2(\lambda))}.$$

Using Proposition 3.2.8, if we set $\alpha = \alpha_\mu$ and $\beta = \beta_\mu$, we have the following inequality for N large enough:

$$\mathbb{E}_i^\gamma[\text{tr}(H_\ell)] \leq \frac{1}{NZ'_N(g, T)} \sum_{\mu \in \Lambda_{N,i}^\gamma} \frac{e^{-\frac{T}{8}(|\alpha|+|\beta|)+t}}{(d_\mu)^{2g-2}} \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \searrow \mu}} \frac{d_\lambda}{d_\mu}.$$

Furthermore, using (3.13) we have

$$\sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \searrow \mu}} \frac{d_\lambda}{d_\mu} = \text{Tr}(I_N) = N,$$

therefore

$$\mathbb{E}_i^\gamma[\mathrm{tr}(H_\ell)] \leq \frac{1}{Z'_N(g, T)} \sum_{\mu \in \Lambda_{N,i}^\gamma} \frac{e^{-\frac{T}{8}(|\alpha|+|\beta|)+t}}{(d_\mu)^{2g-2}} \leq \frac{1}{Z'_N(g, T)} \sum_{\mu \in \Lambda_{N,i}^\gamma} e^{-\frac{T}{8}(|\alpha|+|\beta|)+t},$$

where in the second inequality we used the fact that for any $\mu \in \Lambda_{N,i}^\gamma$, $d_\mu \geq 1$.

From now on, we will set $i = 2$, but the arguments will be similar for $i = 3$ and $i = 4$. For N large enough, we have

$$\begin{aligned} \mathbb{E}_2^\gamma[\mathrm{tr}(H_\ell)] &\leq \frac{e^t}{Z'_N(g, T)} \sum_{|\alpha| > N\gamma, |\beta| \leq N\gamma} e^{-\frac{T}{8}(|\alpha|+|\beta|)+t} \\ &= \frac{e^t}{Z'_N(g, T)} \sum_{|\alpha| > N\gamma} e^{-\frac{T}{8}|\alpha|} \sum_{|\beta| \leq N\gamma} e^{-\frac{T}{8}|\beta|}. \end{aligned}$$

The fraction $\frac{e^t}{Z'_N(g, T)}$ is bounded because $(Z'_N(g, T))_{N \geq 1}$ is a convergent sequence, according to Thm. 2.1.2; the first sum converges to 0 as the remainder of the convergent series defining the generating function of partitions. The second sum is bounded as the partial sum of a similar generating function. We obtain that $\mathbb{E}_2^\gamma[\mathrm{tr}(H_\ell)] \rightarrow 0$ as $N \rightarrow \infty$. We have the same convergence for $i = 3$ and $i = 4$ and the result follows. \square

Proof of Theorem 3.2.5 in the unitary case. Let $\gamma \in (0, \frac{1}{3})$ be a fixed real number. As in the special unitary case, we define for $1 \leq i \leq 4$ the quantity $\mathbb{E}_i^\gamma[\mathrm{tr}(H_\ell)]$ as

$$\mathbb{E}_i^\gamma[\mathrm{tr}(H_\ell)] = \frac{1}{NZ_N(g, T)} \sum_{\mu \in \Omega_{N,i}^\gamma} \frac{q^{c_2(\mu)}}{(d_\mu)^{2g-2}} \sum_{\substack{\lambda \in \widehat{U}(N) \\ \lambda \searrow \mu}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c_2(\mu) - c_2(\lambda))}. \quad (3.56)$$

As we have seen right after Prop. 3.2.II, we have for N large enough

$$\begin{aligned} \mathbb{E}_1^\gamma[\mathrm{tr}(H_\ell)] &= \frac{1}{Z_N(g, T)} \sum_{|\alpha|, |\beta| \leq N\gamma} \sum_{n \in \mathbb{Z}} \frac{q^{c_2(\lambda_N(\alpha, \beta, n))}}{(d_{\lambda_N(\alpha, \beta, n)})^{2g-2}} \\ &\quad \times \left(\frac{1}{N} \sum_{\alpha' \searrow \alpha} \frac{d_{\lambda_N(\alpha', \beta, n)}}{d_{\lambda_N(\alpha, \beta, n)}} e^{\frac{t}{2}(c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha', \beta, n)))} \right. \\ &\quad \left. + \frac{1}{N} \sum_{\beta' \nearrow \beta} \frac{d_{\lambda_N(\alpha, \beta', n)}}{d_{\lambda_N(\alpha, \beta, n)}} e^{\frac{t}{2}(c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha, \beta', n)))} \right). \end{aligned}$$

Let us introduce two intermediary quantities, depending on α , β and n :

$$A_N(\alpha, \beta, n) = \frac{1}{N} \sum_{\alpha' \searrow \alpha} \frac{d_{\lambda_N(\alpha', \beta, n)}}{d_{\lambda_N(\alpha, \beta, n)}} \sum_{n \in \mathbb{Z}} e^{\frac{t}{2}(c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha', \beta, n)))},$$

$$B_N(\alpha, \beta, n) = \frac{1}{N} \sum_{\beta' \nearrow \beta} \frac{d_{\lambda_N(\alpha, \beta', n)}}{d_{\lambda_N(\alpha, \beta, n)}} \sum_{n \in \mathbb{Z}} e^{\frac{t}{2}(c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha, \beta', n)))}.$$

We will show that $A_N(\alpha, \beta, n)$ produces the limit we are trying to get, and that $B_N(\alpha, \beta, n)$ do not contribute to this limit. Let us first consider $A_N(\alpha, \beta, n)$ and use Lemma 3.2.9: from (3.38) and the fact that for any $1 \leq i \leq N$

$$-N^{\gamma-1} \leq \frac{\alpha_i + \beta_1 + 1 - i}{N} \leq N^{\gamma-1} + \frac{1}{N} \leq 2N^{\gamma-1},$$

we deduce

$$c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha', \beta, n)) = -1 - \frac{2n}{N} + \varepsilon_1(N, \gamma)$$

with $|\varepsilon_1(N, \gamma)| \leq 2N^{\gamma-1}$. Combining this estimation with Equation (3.53) yields

$$A_N(\alpha, \beta, n) = e^{-\frac{t}{2} + \varepsilon_1(N, \gamma)} (1 + \eta_1(N, \gamma)) e^{-\frac{tn}{N}},$$

with $|\eta_1(N, \gamma)| \leq 144N^{2\gamma-1}$. Analogously, we have from (3.39) and (3.54)

$$B_N(\alpha, \beta, n) = e^{\frac{t}{2} + \varepsilon_2(N, \gamma)} \eta_2(N, \gamma) e^{-\frac{tn}{N}},$$

with $|\varepsilon_2(N, \gamma)| \leq 0$ and $|\eta_2(N, \gamma)| \leq 36N^{\gamma-2}$. In particular, we have

$$\mathbb{E}_1^\gamma[\text{tr}(H_\ell)] = \frac{C_N}{Z_N(g, T)} \left(e^{-\frac{t}{2} + \varepsilon_1(N, \gamma)} (1 + \eta_1(N, \gamma)) + e^{\frac{t}{2} + \varepsilon_2(N, \gamma)} \eta_2(N, \gamma) \right), \quad (3.57)$$

with

$$C_N = \sum_{n \in \mathbb{Z}} e^{-\frac{tn}{N}} \sum_{|\alpha|, |\beta| \leq N^\gamma} \frac{q^{c_2(\lambda_N(\alpha, \beta, n))}}{(d_{\lambda_N(\alpha, \beta, n)})^{2g-2}}.$$

We would like to show that $\lim_{N \rightarrow \infty} C_N = \lim_{N \rightarrow \infty} Z_N(g, T)$. According to Lemma 2.3.2 we have

$$C_N = \sum_{|\alpha|, |\beta| \leq N^\gamma} \left(\sum_{n \in \mathbb{Z}} q^{(n + \frac{|\alpha| - |\beta|}{N})^2 + \frac{2tn}{N}} \right) \frac{q^{c'_2(\lambda_N(\alpha, \beta))}}{(d_{\lambda_N(\alpha, \beta)})^{2g-2}},$$

and from Eq. (2.41) we get for any α and β such that $|\alpha|, |\beta| \leq N^\gamma$

$$n^2 - n \left(4N^{\gamma-1} - \frac{2t}{TN} \right) \leq \left(n + \frac{|\alpha| - |\beta|}{N} \right)^2 + \frac{2tn}{TN} \leq n^2 + n \left(4N^{\gamma-1} + \frac{2t}{TN} \right) + 4N^{2\gamma-2}.$$

It leads to the following estimates:

$$\left(\sum_{n \in \mathbb{Z}} q^{n^2 + n(4N^{\gamma-1} + \frac{2t}{TN}) + 4N^{2\gamma-2}} \right) \sum_{|\alpha|, |\beta| \leq N^\gamma} \frac{q^{c'_2(\lambda_N(\alpha, \beta))}}{(d_{\lambda_N(\alpha, \beta)})^{2g-2}} \leq C_N$$

and

$$C_N \leq \left(\sum_{n \in \mathbb{Z}} q^{n^2 - n(4N^{\gamma-1} - \frac{2t}{TN})} \right) \sum_{|\alpha|, |\beta| \leq N^\gamma} \frac{q^{c'_2(\lambda_N(\alpha, \beta))}}{(d_{\lambda_N(\alpha, \beta)})^{2g-2}}.$$

The quantity $\sum_{|\alpha|, |\beta| \leq N^\gamma} \frac{q^{c'_2(\lambda_N(\alpha, \beta))}}{(d_{\lambda_N(\alpha, \beta)})^{2g-2}}$ has the same limit as $Z'_N(g, T)$ as we have seen in the proof of Thm. 2.1.2 (see page 98 for $g \geq 2$ and page 101 for $g = 1$). Moreover we have, by dominated convergence,

$$\lim_{N \rightarrow \infty} \sum_{n \in \mathbb{Z}} q^{n^2 - n(4N^{\gamma-1} - \frac{2t}{TN})} = \lim_{N \rightarrow \infty} \sum_{n \in \mathbb{Z}} q^{n^2 + n(4N^{\gamma-1} + \frac{2t}{TN}) + 4N^{2\gamma-2}} = \sum_{n \in \mathbb{Z}} q^{n^2},$$

therefore

$$\lim_{N \rightarrow \infty} C_N = \sum_{n \in \mathbb{Z}} q^{n^2} \lim_{N \rightarrow \infty} Z'_N(g, T) = Z_N(g, T).$$

Plugging this limit into (3.57) and using the estimates of $\varepsilon_1(N, \gamma)$, $\varepsilon_2(N, \gamma)$, $\eta_1(N, \gamma)$ and $\eta_2(N, \gamma)$, we finally get

$$\lim_{N \rightarrow \infty} \mathbb{E}_1^\gamma[\text{tr}(H_\ell)] = e^{-\frac{t}{2}}.$$

Now we have to show that the other \mathbb{E}_i^γ all tend to 0 when $N \rightarrow \infty$. We will need the following estimations, which are direct consequences of (3.38): if $\alpha, \beta, \alpha', \beta'$ are partitions such that $\alpha' \searrow \alpha$ and $\beta' \nearrow \beta$, then for any $n \in \mathbb{Z}$ we have

$$c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha', \beta, n)) \leq 1 - \frac{2n}{N}$$

and

$$c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha, \beta', n)) \leq 1 - \frac{2n}{N}.$$

In particular, combined with Prop. 3.2.II, these estimations imply that for any $(\lambda, \mu) \in \widehat{U}(N)^2$ such that $\lambda \searrow \mu$,

$$c_2(\mu) - c_2(\lambda) \leq 1 - \frac{2n}{N}, \quad (3.58)$$

with $n = n_\mu$. Recall that from (3.13) we have

$$\sum_{\substack{\lambda \in \widehat{U}(N) \\ \lambda \searrow \mu}} \frac{d_\lambda}{d_\mu} = N.$$

If we combine these results with (3.56), we get the following estimation:

$$0 \leq \mathbb{E}_i^\gamma[\text{tr}(H_\ell)] \leq \frac{e^{\frac{t}{2}}}{Z_N(g, T)} \sum_{\mu \in \Omega_{N,i}^\gamma} \frac{q^{c_2(\mu) + \frac{2tn_\mu}{TN}}}{(d_\mu)^{2g-2}}. \quad (3.59)$$

Now let us specialize our computation to a given i . We will do it for $i = 2$, the other cases being similar. We have

$$0 \leq \mathbb{E}_2^\gamma[\text{tr}(H_\ell)] \leq \frac{C_{N,2}}{Z_N(g, T)},$$

with

$$C_{N,2} = e^{\frac{t}{2}} \sum_{|\alpha| > N^\gamma, |\beta| \leq N^\gamma} \sum_{n \in \mathbb{Z}} \frac{q^{c'_2(\lambda_N(\alpha, \beta)) + (n + \frac{|\alpha| - |\beta|}{N})^2 + \frac{2tn}{TN}}}{(d_{\lambda_N(\alpha, \beta)})^{2g-2}}.$$

Recall that for any α and β we have $d_{\lambda_N(\alpha,\beta)} \geq 1$. Besides, from (3.36) we have $c'_2(\lambda_N(\alpha, \beta)) \geq \frac{1}{2}|\alpha| + |\beta|$, and we also have the following estimation:

$$\begin{aligned} -\frac{T}{2} \left(\left(n + \frac{|\alpha| - |\beta|}{N} \right)^2 + \frac{2tn}{TN} \right) &= -\frac{T}{2} \left(n + \frac{|\alpha| - |\beta| + \frac{t}{T}}{N} \right)^2 - \frac{(|\alpha| - |\beta|)t}{2TN^2} + \frac{t^2}{2TN^2} \\ &\leq -\frac{T}{2} \left(n + \frac{|\alpha| - |\beta| + \frac{t}{T}}{N} \right)^2 + \frac{t^2}{2TN^2} + \frac{(|\alpha| + |\beta|)t}{2TN^2}. \end{aligned}$$

It means that

$$C_{N,2} \leq e^{\frac{t}{2}} \sum_{|\alpha| > N^\gamma, |\beta| \leq N^\gamma} \left(\sum_{n \in \mathbb{Z}} q^{\left(n + \frac{|\alpha| - |\beta|}{N} + \frac{t}{TN} \right)^2} \right) q^{(|\alpha| + |\beta|) \left(\frac{1}{2} - \frac{t}{T^2 N^2} \right) - \frac{t^2}{T^2 N^2}}.$$

The sum between parentheses is bounded independently from N , $|\alpha|$ and $|\beta|$ by $C = 1 + \vartheta(0; \frac{iT}{2\pi})$, and we have for N large enough the inequality $\frac{1}{2} - \frac{t}{T^2 N^2} > \frac{1}{4}$, therefore

$$C_{N,2} \leq C e^{\frac{t}{2} + \frac{2t}{TN^2}} \sum_{|\alpha| > N^\gamma} q^{\frac{|\alpha|}{4}} \sum_{|\beta| \leq N^\gamma} q^{\frac{|\beta|}{4}},$$

and it is clear that this quantity converges to zero when $N \rightarrow \infty$, because $\sum_{|\alpha| > N^\gamma} q^{\frac{|\alpha|}{4}}$ converges to zero and the other terms are uniformly bounded in N . We obtain that $\lim_{N \rightarrow \infty} \mathbb{E}_2^\gamma[\text{tr}(H_\ell)] = 0$, and we have the same limit for $i = 3$ and $i = 4$. This concludes the proof. \square

Asymptotics of the variance

We would like to prove Thm. 3.2.6 in this section. Before that, let us remark that Equation (3.26) implies that Equation (3.33) is equivalent to

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_\ell) \text{tr}(H_\ell^*)] = \lim_{N \rightarrow \infty} |\mathbb{E}[\text{tr}(H_\ell)]|^2,$$

which can be rewritten, thanks to Thm. 3.2.5:

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_\ell) \text{tr}(H_\ell^*)] = e^{-t}. \quad (3.60)$$

We will prove Thm. 3.2.6 by proving this limit, using similar arguments as in the proof of Thm. 3.2.5.

Proof of Theorem 3.2.6 in the special unitary case. First, using Equation (3.28), we have

$$\mathbb{E}[\text{tr}(H_\ell) \text{tr}(H_\ell^*)] = \frac{1}{N^2 Z'_N(g, T)} \sum_{\substack{\lambda, \mu \in \widehat{\text{SU}}(N) \\ \lambda \sim \mu \\ \lambda \neq \mu}} \frac{q^{c'_2(\mu)}}{(d_\mu)^{2g-2}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c'_2(\mu) - c'_2(\lambda))} + \frac{1}{N^2}.$$

We deduce from this the fact that

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_\ell) \text{tr}(H_\ell^*)] = \lim_{N \rightarrow \infty} \frac{1}{N^2 Z'_N(g, T)} \sum_{\substack{\lambda, \mu \in \widehat{\text{SU}}(N) \\ \lambda \sim \mu \\ \lambda \neq \mu}} \frac{q^{c'_2(\mu)}}{(d_\mu)^{2g-2}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c'_2(\mu) - c'_2(\lambda))}.$$

Let us define, for $\gamma \in (0, \frac{1}{3})$ and $i \in \{1, 2, 3, 4\}$:

$$\mathbb{E}_i^\gamma[\text{tr}(H_\ell)\text{tr}(H_\ell^*)] = \frac{1}{N^2 Z'_N(g, T)} \sum_{\substack{\lambda, \mu \in \Lambda_{N,i}^\gamma \\ \lambda \sim \mu \\ \lambda \neq \mu}} \frac{q^{c'_2(\mu)}}{(d_\mu)^{2g-2}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c'_2(\mu) - c'_2(\lambda))}.$$

Using similar arguments as in the proof of Theorem 3.2.5, we have for N large enough

$$\mathbb{E}_1^\gamma[\text{tr}(H_\ell)\text{tr}(H_\ell^*)] = \frac{1}{N^2 Z'_N(g, T)} \sum_{|\alpha|, |\beta| \leq N^\gamma} \frac{q^{c'_2(\lambda_N(\alpha, \beta))}}{(d_{\lambda_N(\alpha, \beta)})^{2g-2}} \sum_{\substack{\lambda \in \widehat{\text{SU}}(N) \\ \lambda \sim \lambda_N(\alpha, \beta) \\ \lambda \neq \lambda_N(\alpha, \beta)}} \frac{d_\lambda e^{\frac{t}{2}(c'_2(\lambda_N(\alpha, \beta)) - c'_2(\lambda))}}{d_{\lambda_N(\alpha, \beta)}}.$$

We can notice that adding a box and removing another box¹³ to $\lambda_N(\alpha, \beta)$ is equivalent to one of these 4 cases:

- Adding a box to α and β ;
- Adding a box and removing another one to α ;
- removing a box and adding another one to β ;
- Removing a box to α and β .

Remark that the third case is equivalent to “adding a box and removing another one to β ” because the operations “adding a box” and “removing a box” commute. Remark also that all these operations are under the implicit condition that they are mappings from the set of integer partitions to itself. Hence, if we define

$$\begin{aligned} S_{N,1} &= \frac{1}{N^2} \sum_{\substack{\alpha' \searrow \alpha \\ \beta' \searrow \beta}} \frac{d_{\lambda_N(\alpha', \beta')}}{d_{\lambda_N(\alpha, \beta)}} e^{\frac{t}{2}(|\alpha| + |\beta| - |\alpha'| - |\beta'| + \varepsilon_1(N, \gamma))}, \\ S_{N,2} &= \frac{1}{N^2} \sum_{\substack{\alpha' \nearrow \alpha \\ \beta' \nearrow \beta}} \frac{d_{\lambda_N(\alpha', \beta')}}{d_{\lambda_N(\alpha, \beta)}} e^{\frac{t}{2}(|\alpha| + |\beta| - |\alpha'| - |\beta'| + \varepsilon_2(N, \gamma))}, \\ S_{N,3} &= \frac{1}{N^2} \sum_{\substack{\alpha' \sim \alpha \\ \alpha' \neq \alpha}} \frac{d_{\lambda_N(\alpha', \beta)}}{d_{\lambda_N(\alpha, \beta)}} e^{\frac{t}{2}(|\alpha| - |\alpha'| + \varepsilon_3(N, \gamma))}, \\ S_{N,4} &= \frac{1}{N^2} \sum_{\substack{\beta' \sim \beta \\ \beta' \neq \beta}} \frac{d_{\lambda_N(\alpha, \beta')}}{d_{\lambda_N(\alpha, \beta)}} e^{\frac{t}{2}(|\beta| - |\beta'| + \varepsilon_4(N, \gamma))}, \end{aligned}$$

with $|\varepsilon_i(N, \gamma)| \leq 16N^{2\gamma-1}$ for $i \in \{1, 2, 3, 4\}$, we have

$$\mathbb{E}_1^\gamma[\text{tr}(H_\ell)\text{tr}(H_\ell^*)] = \frac{1}{Z'_N(g, T)} \sum_{|\alpha|, |\beta| \leq N^\gamma} \frac{q^{c'_2(\lambda_N(\alpha, \beta))}}{(d_{\lambda_N(\alpha, \beta)})^{2g-2}} (S_{N,1} + S_{N,2} + S_{N,3} + S_{N,4}).$$

¹³A different one, this time, because we assume that the new highest weight is different from the initial one.

We will prove that only $S_{N,1}$ contributes to the limit. Using Prop. 3.2.10, we have for N large enough

$$S_{N,1} = \sum_{\substack{\alpha' \searrow \alpha \\ \beta' \searrow \beta}} \frac{d^{\alpha'} d^{\beta'}}{|\alpha'| |\beta'| d^{\alpha} d^{\beta}} e^{\frac{t}{2}(2+\varepsilon_1(N,\gamma))} \eta_1(N, \gamma),$$

with $|\eta_1(N, \gamma) - 1| \leq 144N^{3\gamma-1}$ (following the same arguments as in the proof of Thm. 3.2.5). We can apply Prop. 3.2.12 and get

$$S_{N,1} = e^{t+\frac{t}{2}\varepsilon_1(N,\gamma)} \eta_1(N, \gamma).$$

Similarly, we have

$$S_{N,2} = \frac{|\alpha| |\beta|}{N^4} e^{t+\frac{t}{2}\varepsilon_2(N,\gamma)} \eta_2(N, \gamma),$$

with $|\eta_2(N, \gamma) - 1| \leq 144N^{3\gamma-1}$. Now, in order to compute $S_{N,3}$ and $S_{N,4}$, let us notice that

$$\sum_{\substack{\alpha' \sim \alpha \\ \alpha' \neq \alpha}} \frac{d^{\alpha'}}{d^{\alpha}} = -1 + \sum_{\alpha'' \searrow \alpha} \frac{d^{\alpha''}}{d^{\alpha}} \sum_{\alpha' \nearrow \alpha''} \frac{d^{\alpha'}}{d^{\alpha''}} = -1 + \sum_{\alpha'' \searrow \alpha} \frac{d^{\alpha''}}{d^{\alpha}} = |\alpha|.$$

We can apply this equality and use the same arguments as above to get

$$\begin{aligned} S_{N,3} &= \frac{|\alpha|}{N^2} e^{\varepsilon_3(N,\gamma)} \eta_3(N, \gamma), \\ S_{N,4} &= \frac{|\beta|}{N^2} e^{\varepsilon_4(N,\gamma)} \eta_4(N, \gamma), \end{aligned}$$

with $|\eta_3(N, \gamma) - 1| \leq 144N^{3\gamma-1}$ and $|\eta_4(N, \gamma) - 1| \leq 144N^{3\gamma-1}$. As we have $|\alpha|, |\beta| \leq N^{\gamma}$, it appears that

$$S_{N,2} \leq 145 e^{t+\frac{t}{2}\varepsilon_2(N,\gamma)} N^{2\gamma-1},$$

and $S_{N,2}$ tends to 0 when N tends to infinity. We come to the same conclusion for $S_{N,3}$ and $S_{N,4}$, and we also find that $S_{N,1}$ tends to e^t when N tends to infinity. It follows that

$$\mathbb{E}_1^{\gamma}[\mathrm{tr}(H_{\ell})\mathrm{tr}(H_{\ell}^*)] = \frac{e^{-t} + o(1)}{Z'_N(g, T)} \sum_{|\alpha|, |\beta| \leq N^{\gamma}} \frac{q^{c'_2(\lambda_N(\alpha, \beta))}}{(d_{\lambda_N(\alpha, \beta)})^{2g-2}},$$

and the right-hand side converges to e^{-t} as $N \rightarrow \infty$.

Now let us prove that $\mathbb{E}_i^{\gamma}[\mathrm{tr}(H_{\ell})\mathrm{tr}(H_{\ell}^*)]$ tends to 0 for $i \in \{2, 3, 4\}$. Recall that

$$\mathbb{E}_i^{\gamma}[\mathrm{tr}(H_{\ell})\mathrm{tr}(H_{\ell}^*)] = \frac{1}{N^2 Z'_N(g, T)} \sum_{\substack{\lambda, \mu \in \Lambda_{N,i}^{\gamma} \\ \lambda \sim \mu \\ \lambda \neq \mu}} \frac{q^{c'_2(\mu)}}{(d_{\mu})^{2g-2}} \frac{d_{\lambda}}{d_{\mu}} e^{\frac{t}{2}(c'_2(\mu) - c'_2(\lambda))}.$$

Using Lemma 3.2.8 and setting $\alpha = \alpha_\mu$ and $\beta = \beta_\mu$, we have

$$\begin{aligned} \mathbb{E}_i^\gamma[\mathrm{tr}(H_\ell)\mathrm{tr}(H_\ell^*)] &\leq \frac{e^t}{N^2 Z'_N(g, T)} \sum_{\substack{\lambda, \mu \in \Lambda_{N,i}^\gamma \\ \lambda \sim \mu \\ \lambda \neq \mu}} \frac{1}{(d_\mu)^{2g-2}} \frac{d_\lambda}{d_\mu} e^{-\frac{T}{8}(|\alpha|+|\beta|)} \\ &= \frac{e^t}{N^2 Z'_N(g, T)} \sum_{\mu \in \Lambda_{N,i}^\gamma} \frac{e^{-\frac{T}{8}(|\alpha|+|\beta|)}}{(d_\mu)^{2g-2}} \sum_{\substack{\lambda \in \Lambda_{N,i}^\gamma \\ \lambda \sim \mu \\ \lambda \neq \mu}} \frac{d_\lambda}{d_\mu}. \end{aligned}$$

Let us turn to the sum of $\frac{d_\lambda}{d_\mu}$: we can write

$$\sum_{\substack{\lambda \in \Lambda_{N,i}^\gamma \\ \lambda \sim \mu}} \frac{d_\lambda}{d_\mu} = \frac{1}{d_\mu} \sum_{\nu \searrow \mu} d_\nu \sum_{\lambda: \nu \searrow \lambda} \frac{d_\lambda}{d_\nu},$$

and it is not hard to find out that for any $\nu \in \widehat{\mathrm{SU}}(N)$

$$\sum_{\lambda: \nu \searrow \lambda} \frac{d_\lambda}{d_\nu} \leq \sum_{\lambda: \nu \searrow \lambda} 1 \leq N,$$

so that

$$\sum_{\substack{\lambda \in \Lambda_{N,i}^\gamma \\ \lambda \sim \mu}} \frac{d_\lambda}{d_\mu} = \frac{N}{d_\mu} \sum_{\nu \searrow \mu} d_\nu.$$

Lemma 3.1.3 then gives us

$$\sum_{\substack{\lambda \in \Lambda_{N,i}^\gamma \\ \lambda \sim \mu}} \frac{d_\lambda}{d_\mu} \leq \frac{N}{d_\mu} N d_\mu = N^2. \quad (3.61)$$

Finally, we have

$$\begin{aligned} \mathbb{E}_i^\gamma[\mathrm{tr}(H_\ell)\mathrm{tr}(H_\ell^*)] &\leq \frac{e^t}{Z'_N(g, T)} \sum_{\mu \in \Lambda_{N,i}^\gamma} \frac{e^{-\frac{T}{8}(|\alpha|+|\beta|)}}{(d_\mu)^{2g-2}} \left(1 - \frac{1}{N^2}\right) \\ &\leq \frac{e^t(1 - \frac{1}{N^2})}{Z'_N(g, T)} \sum_{\mu \in \Lambda_{N,i}^\gamma} e^{-\frac{T}{8}(|\alpha|+|\beta|)}. \end{aligned}$$

Now it is clear that, following the same arguments as in the proof of Theorem 3.2.5, that the quantity $\mathbb{E}_i^\gamma[\mathrm{tr}(H_\ell)\mathrm{tr}(H_\ell^*)]$ tends to 0 as N tends to infinity, for $i \in \{2, 3, 4\}$. This proves Equation (3.60), and therefore Theorem 3.2.6. \square

Proof of Thm. 3.2.6 in the unitary case. According to Equation (3.30), we have

$$\mathbb{E}[\mathrm{tr}(H_\ell)\mathrm{tr}(H_\ell^*)] = \frac{1}{N^2 Z_N(g, T)} \sum_{\substack{\lambda, \mu \in \widehat{\mathrm{U}}(N) \\ \lambda \sim \mu \\ \lambda \neq \mu}} \frac{q^{c_2(\mu)}}{(d_\mu)^{2g-2}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c_2(\mu) - c_2(\lambda))} + \frac{1}{N^2},$$

and as in the special unitary case, we see that

$$\lim_{N \rightarrow \infty} \mathbb{E}[\mathrm{tr}(H_\ell)\mathrm{tr}(H_\ell^*)] = \lim_{N \rightarrow \infty} \frac{1}{N^2 Z_N(g, T)} \sum_{\substack{\lambda, \mu \in \widehat{U}(N) \\ \lambda \sim \mu \\ \lambda \neq \mu}} \frac{q^{c_2(\mu)}}{(d_\mu)^{2g-2}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c_2(\mu) - c_2(\lambda))}.$$

We set, for $\gamma \in (0, \frac{1}{3})$ and $i \in \{1, 2, 3, 4\}$:

$$\mathbb{E}_i^\gamma[\mathrm{tr}(H_\ell)\mathrm{tr}(H_\ell^*)] = \frac{1}{N^2 Z_N(g, T)} \sum_{\substack{\lambda, \mu \in \Omega_{N,i}^\gamma \\ \lambda \sim \mu \\ \lambda \neq \mu}} \frac{q^{c_2(\mu)}}{(d_\mu)^{2g-2}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c_2(\mu) - c_2(\lambda))}.$$

Let us define, for α and β two partitions and $n \in \mathbb{Z}$ an integer,

$$\begin{aligned} S_{N,1} &= \frac{1}{N^2} \sum_{\substack{\alpha' \searrow \alpha \\ \beta' \searrow \beta}} \frac{d_{\lambda_N(\alpha', \beta')}}{d_{\lambda_N(\alpha, \beta)}} e^{\frac{t}{2}(c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha', \beta', n)))}, \\ S_{N,2} &= \frac{1}{N^2} \sum_{\substack{\alpha' \nearrow \alpha \\ \beta' \nearrow \beta}} \frac{d_{\lambda_N(\alpha', \beta')}}{d_{\lambda_N(\alpha, \beta)}} e^{\frac{t}{2}(c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha', \beta', n)))}, \\ S_{N,3} &= \frac{1}{N^2} \sum_{\substack{\alpha' \sim \alpha \\ \alpha' \neq \alpha}} \frac{d_{\lambda_N(\alpha', \beta)}}{d_{\lambda_N(\alpha, \beta)}} e^{\frac{t}{2}(c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha', \beta, n)))}, \\ S_{N,4} &= \frac{1}{N^2} \sum_{\substack{\beta' \sim \beta \\ \beta' \neq \beta}} \frac{d_{\lambda_N(\alpha, \beta')}}{d_{\lambda_N(\alpha, \beta)}} e^{\frac{t}{2}(c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha, \beta', n)))}. \end{aligned}$$

We have, for N large enough,

$$\mathbb{E}_1^\gamma[\mathrm{tr}(H_\ell)\mathrm{tr}(H_\ell^*)] = \frac{1}{N^2 Z_N(g, T)} \sum_{|\alpha|, |\beta| \leq N^\gamma} \sum_{n \in \mathbb{Z}} \frac{q^{c_2(\lambda_N(\alpha, \beta, n))}}{(d_{\lambda_N(\alpha, \beta, n)})^{2g-2}} (S_{N,1} + S_{N,2} + S_{N,3} + S_{N,4}).$$

We can compute the differences of Casimir numbers in each $S_{N,i}$ using Prop. 2.2.5, in the same way as we did in Lemma 3.2.9. For instance, if $\alpha' \searrow \alpha$ and $\beta' \searrow \beta$ and i_0 and j_0 are such that $\alpha'_{i_0} = \alpha_{i_0} + 1$ and $\beta'_{j_0} = \beta_{j_0} + 1$, then

$$c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha', \beta', n)) = -2 - \frac{2}{N}(\alpha_{i_0} + \beta_{j_0} - i_0 - j_0).$$

In particular, if $|\alpha|, |\beta| \leq N^\gamma$, we have

$$-2 - 4N^{\gamma-1} \leq c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha', \beta', n)) \leq -2 + 4N^{\gamma-1}.$$

Following the same argument, if $\alpha' \nearrow \alpha$ and $\beta' \nearrow \beta$ then

$$2 - 4N^{\gamma-1} \leq c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha', \beta', n)) \leq 2 + 4N^{\gamma-1}.$$

If $\alpha' \sim \alpha$ then

$$|c_2(\lambda_N(\alpha, \beta, n)) - c_2(\lambda_N(\alpha', \beta, n))| \leq 4N^{\gamma-1},$$

and it is the same if we consider $\beta' \sim \beta$. Now $S_{N,i}$ can be estimated the same way as in the special unitary case: we have

$$\begin{aligned} S_{N,1} &= e^{-t+\varepsilon_1(N,\gamma)} \eta_1(N, \gamma), \\ S_{N,2} &= e^{t+\varepsilon_2(N,\gamma)} \frac{|\alpha||\beta|}{N^4} \eta_2(N, \gamma), \\ S_{N,3} &= e^{\varepsilon_3(N,\gamma)} \frac{|\alpha|}{N^2} \eta_3(N, \gamma), \\ S_{N,4} &= e^{\varepsilon_4(N,\gamma)} \frac{|\beta|}{N^2} \eta_4(N, \gamma), \end{aligned}$$

with $|\varepsilon_i(N, \gamma)| \leq 8tN^{\gamma-1}$ and $|\eta_i(N, \gamma) - 1| \leq 144N^{3\gamma-1}$ for $1 \leq i \leq 4$. Then, still using similar arguments, we find that

$$\mathbb{E}_1^\gamma[\text{tr}(H_\ell)\text{tr}(H_\ell^*)] = \frac{e^{-t} + o(1)}{Z_N(g, T)} \sum_{|\alpha|, |\beta| \leq N^\gamma} \sum_{n \in \mathbb{Z}} \frac{q^{c'_2(\lambda_N(\alpha, \beta, n))}}{(d_{\lambda_N(\alpha, \beta, n)})^{2g-2}},$$

which tends to e^{-t} as $N \rightarrow \infty$.

It remains to prove that $\mathbb{E}_i^\gamma[\text{tr}(H_\ell)\text{tr}(H_\ell^*)]$ converges to 0 for $i \in \{2, 3, 4\}$. Recall that we have

$$\mathbb{E}_i^\gamma[\text{tr}(H_\ell)\text{tr}(H_\ell^*)] = \frac{1}{N^2 Z_N(g, T)} \sum_{\substack{\lambda, \mu \in \Omega_{N,i}^\gamma \\ \lambda \sim \mu \\ \lambda \neq \mu}} \frac{q^{c_2(\mu)}}{(d_\mu)^{2g-2}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c_2(\mu) - c_2(\lambda))}.$$

For any $\lambda \sim \mu$, if we set $\alpha = \alpha_\mu, \alpha' = \alpha_\lambda, \beta = \beta_\mu, \beta' = \beta_\lambda, n = n_\mu$ and $n' = n_\lambda$, we get in particular that $n = n'$ and $|\alpha| - |\beta| = |\alpha'| - |\beta'|$. In particular, if we use the fact that

$$c_2(\mu) = c'_2(\lambda_N(\alpha, \beta)) + \left(n + \frac{|\alpha| - |\beta|}{N}\right)^2 \quad \text{and} \quad c_2(\lambda) = c'_2(\lambda_N(\alpha', \beta')) + \left(n' + \frac{|\alpha'| - |\beta'|}{N}\right)^2,$$

then it follows that

$$c_2(\mu) - c_2(\lambda) = c'_2(\lambda_N(\alpha, \beta)) - c'_2(\lambda_N(\alpha', \beta')).$$

If we define

$$\begin{aligned}
S_{N,1} &= \frac{1}{N^2} \sum_{\substack{\alpha' \searrow \alpha \\ \beta' \searrow \beta}} \frac{d_{\lambda_N(\alpha', \beta')}}{d_{\lambda_N(\alpha, \beta)}} e^{-\frac{T}{2} c'_2(\lambda_N(\alpha, \beta)) + \frac{t}{2} (c'_2(\lambda_N(\alpha, \beta)) - c'_2(\lambda_N(\alpha', \beta')))}, \\
S_{N,2} &= \frac{1}{N^2} \sum_{\substack{\alpha' \nearrow \alpha \\ \beta' \nearrow \beta}} \frac{d_{\lambda_N(\alpha', \beta')}}{d_{\lambda_N(\alpha, \beta)}} e^{-\frac{T}{2} c'_2(\lambda_N(\alpha, \beta)) + \frac{t}{2} (c'_2(\lambda_N(\alpha, \beta)) - c'_2(\lambda_N(\alpha', \beta')))}, \\
S_{N,3} &= \frac{1}{N^2} \sum_{\substack{\alpha' \sim \alpha \\ \alpha' \neq \alpha}} \frac{d_{\lambda_N(\alpha', \beta)}}{d_{\lambda_N(\alpha, \beta)}} e^{-\frac{T}{2} c'_2(\lambda_N(\alpha, \beta)) + \frac{t}{2} (c'_2(\lambda_N(\alpha, \beta)) - c'_2(\lambda_N(\alpha', \beta)))}, \\
S_{N,4} &= \frac{1}{N^2} \sum_{\substack{\beta' \sim \beta \\ \beta' \neq \beta}} \frac{d_{\lambda_N(\alpha, \beta')}}{d_{\lambda_N(\alpha, \beta)}} e^{-\frac{T}{2} c'_2(\lambda_N(\alpha, \beta)) + \frac{t}{2} (c'_2(\lambda_N(\alpha, \beta)) - c'_2(\lambda_N(\alpha, \beta')))},
\end{aligned}$$

then we get

$$\mathbb{E}_2^\gamma[\mathrm{tr}(H_\ell)\mathrm{tr}(H_\ell^*)] \leq \frac{1}{Z_N(g, T)} \sum_{|\alpha| > N^\gamma, |\beta| \leq N^\gamma} \sum_{n \in \mathbb{Z}} \frac{e^{-\frac{T}{2} (n + \frac{|\alpha| - |\beta|}{N})^2}}{(d_{\lambda_N(\alpha, \beta, n)})^{2g-2}} (S_{N,1} + S_{N,2} + S_{N,3} + S_{N,4}).$$

From Prop. 3.2.8 we have then

$$S_{N,1} \leq \frac{e^{-\frac{T}{8} (|\alpha| + |\beta|)}}{N^2} \sum_{\substack{\alpha' \searrow \alpha \\ \beta' \searrow \beta}} \frac{d_{\lambda_N(\alpha', \beta')}}{d_{\lambda_N(\alpha, \beta)}},$$

and from Prop. 3.2.10 and 3.2.12 we have for N large enough

$$S_{N,1} \leq e^{-\frac{T}{8} (|\alpha| + |\beta|)} (1 + 144N^{3\gamma-1}).$$

Similarly, we have

$$\begin{aligned}
S_{N,2} &\leq e^{-\frac{T}{8} (|\alpha| + |\beta|)} (1 + 144N^{3\gamma-1}), \\
S_{N,3} &\leq e^{-\frac{T}{8} (|\alpha| + |\beta|)} \frac{|\alpha|}{N^2} (1 + 144N^{3\gamma-1}), \\
S_{N,4} &\leq e^{-\frac{T}{8} (|\alpha| + |\beta|)} \frac{|\beta|}{N^2} (1 + 144N^{3\gamma-1}).
\end{aligned}$$

As $\sum_{n \in \mathbb{Z}} e^{-\frac{T}{2} (n + \frac{|\alpha| - |\beta|}{N})^2}$ is uniformly bounded for every α, β and N by $C = 1 + \vartheta(0; iT/2\pi)$, we get

$$\mathbb{E}_2^\gamma[\mathrm{tr}(H_\ell)\mathrm{tr}(H_\ell^*)] \leq \frac{C(1 + 144N^{3\gamma-1})}{Z_N(g, T)} \sum_{|\alpha| > N^\gamma, |\beta| \leq N^\gamma} \frac{4(|\alpha| + |\beta|) e^{-\frac{T}{8} (|\alpha| + |\beta|)}}{N^2 (d_{\lambda_N(\alpha, \beta)})^{2g-2}}.$$

Let us also recall that for any α and β we have $d_{\lambda_N(\alpha, \beta)} \geq 1$, therefore the sum of the right-hand side is bounded by

$$\sum_{|\alpha| > N^\gamma, |\beta| \leq N^\gamma} 4(|\alpha| + |\beta|) e^{-\frac{T}{8} (|\alpha| + |\beta|)}.$$

This sum converges to 0 when $N \rightarrow \infty$, thus so does $\mathbb{E}_2^\gamma[\mathrm{tr}(H_\ell)\mathrm{tr}(H_\ell^*)]$. We can apply the same trick for $i \in \{3, 4\}$, the Thm. 3.2.6 follows. \square

Fluctuations of the variance

Let us discuss the convergence rate of $\text{Var}[\text{tr}(H_\ell)]$, because the proofs involved many different estimations. Among them, the most restrictive ones are Prop. 3.2.10 and (3.42). Their common feature is the use of the control parameter γ which is assumed to be either in $(0, \frac{1}{3})$ or $(0, \frac{1}{2})$ depending on the estimation; the convergence does not depend on the specific value of γ . Indeed, the whole point of this parameter is to control $|\alpha|$ and $|\beta|$ by a sequence $(x_N)_{N \geq 1}$ that satisfies the following assumptions:

$$\begin{aligned} \lim_{N \rightarrow \infty} x_N &= \infty, \\ \lim_{N \rightarrow \infty} N^{-\frac{1}{3}} x_N &= 0. \end{aligned}$$

These assumptions were already implicit in Chapter 2, in the informal reasoning made in p. 101. The first one is here to ensure that the weights that are not almost flat have a contribution which is bounded by the remainder of a convergent series, whereas the second one is here to ensure that the contribution of almost flat highest weights converges to the right limit. We took $x_N = N^\gamma$, but we could have taken $x_N = \log(N)$ as well. The outline of our proofs would probably have remained unchanged, but without giving any better information on the convergence. Indeed, we can take γ as small as needed, thus the rate of convergence can be as close to $\frac{1}{N}$ as we want: it is precisely $\frac{1}{N^{1-\varepsilon}}$ with $\varepsilon > 0$ arbitrarily small. It can be summarized in the following proposition.

Proposition 3.2.13. *Let $\Sigma_{g,T}$ be a surface and ℓ be a simple loop satisfying the same assumptions as in Thm. 3.2.6. Then for any $\varepsilon > 0$, there exists a constant $C_\varepsilon > 0$ such that, for N large enough,*

$$\text{Var}(\text{tr}(H_\ell)) \leq C_\varepsilon N^{\varepsilon-1}. \quad (3.62)$$

Simpler proofs for $g \geq 2$

Although we directly proved Thm. 3.2.5 and 3.2.6 for any $g \geq 1$ using almost flat highest weights, it appears to be an interesting to see if the proof can be simplified for $g \geq 2$. Indeed, as we saw in Chapter 2 for the limit of partition function, the case $g \geq 2$ did not require the use of almost flat highest weights, but rather ‘flat’ weights (*i.e.* constant weights). We will show here that it is still true for the Wilson loop expectation and variance, and that the proofs of Thm. 3.2.5 and 3.2.6 can be simplified for $g \geq 2$.

Proof of Thm. 3.2.5 in the unitary case with $g \geq 2$. Let us start with (3.27). The sum over $\mu \in \widehat{U}(N)$ can be split into two terms, the one associated with $\mu = (n, \dots, n)$ for $n \in \mathbb{Z}$ (we will call μ a *flat* highest weight and denote by Λ_N the set of such weights) and the sum of the remaining terms. The main point is that for any $\mu = (n, \dots, n) \in \Lambda_N$, the only $\lambda \in \widehat{U}(N)$ such that $\lambda \searrow \mu$ is $\lambda = (n+1, n, \dots, n)$, which has dimension N and Casimir number

$$c_2((n+1, n, \dots, n)) = n^2 + 1 + \frac{2n}{N}.$$

Furthermore, it is straightforward that (n, \dots, n) has dimension 1 and Casimir number n^2 . It yields

$$\begin{aligned} \mathbb{E}[\mathrm{tr}(H_\ell)] &= \frac{1}{Z_N(g, T)} \sum_{n \in \mathbb{Z}} e^{-\frac{T}{2}n^2 - \frac{t}{2}(1 + \frac{2n}{N})} \\ &\quad + \frac{1}{NZ_N(g, T)} \sum_{\mu \in \widehat{\mathrm{SU}}(N) \setminus \Lambda_N} \frac{q^{c'_2(\mu)}}{d_\mu^{2g-2}} \sum_{\substack{\lambda \in \widehat{\mathrm{U}}(N) \\ \lambda \searrow \mu}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c'_2(\mu) - c'_2(\lambda))}. \end{aligned}$$

The first sum is equal to

$$e^{-\frac{t}{2}} \sum_{n \in \mathbb{Z}} e^{-\frac{T}{2}n^2 - \frac{t}{N}n} = e^{-\frac{t}{2} + \frac{t^2}{2TN^2}} \sum_{n \in \mathbb{Z}} e^{-\frac{T}{2}(n + \frac{t}{TN})^2},$$

and the right-hand side converges to $e^{-\frac{t}{2}} \vartheta(0; \frac{iT}{2\pi})$ as $N \rightarrow \infty$ by dominated convergence. Recall that we also have $Z_N(g, T) \rightarrow \vartheta(0; \frac{iT}{2\pi})$, therefore we get

$$\lim_{N \rightarrow \infty} \frac{1}{Z_N(g, T)} \sum_{n \in \mathbb{Z}} e^{-\frac{T}{2}n^2 - \frac{t}{2}(1 + \frac{2n}{N})} = e^{-\frac{t}{2}}.$$

The rest of the proof will be dedicated to bound the remainder by a term that tends to 0 when $N \rightarrow \infty$:

$$\Delta(N) = \frac{1}{NZ_N(g, T)} \sum_{\mu \in \widehat{\mathrm{SU}}(N) \setminus \Lambda_N} \frac{q^{c'_2(\mu)}}{d_\mu^{2g-2}} \sum_{\substack{\lambda \in \widehat{\mathrm{U}}(N) \\ \lambda \searrow \mu}} \frac{d_\lambda}{d_\mu} e^{\frac{t}{2}(c'_2(\mu) - c'_2(\lambda))}.$$

From Lemma 2.2.4 and the fact that adding 1 to all parts of a highest weight does not change the dimension, we get that

$$d_\mu \geq N, \quad \forall \mu \in \widehat{\mathrm{U}}(N) \setminus \Lambda_N.$$

Furthermore, it is clear that $c'_2(\mu) \geq 0$ for any μ , from the definition of Casimir element. Thus,

$$0 \leq \Delta(N) \leq \frac{1}{NZ_N(g, T)} \sum_{\mu \in \widehat{\mathrm{U}}(N) \setminus \Lambda_N} \frac{1}{N^{2g-2}} \sum_{\substack{\lambda \in \widehat{\mathrm{U}}(N) \\ \lambda \searrow \mu}} \frac{d_\lambda}{d_\mu} e^{-\frac{T}{2}c'_2(\mu) + \frac{t}{2}(c'_2(\mu) - c'_2(\lambda))}.$$

Eq. (3.58) implies that, for N large enough,

$$e^{-\frac{T}{2}c_2(\mu) + \frac{t}{2}(c_2(\mu) - c_2(\lambda))} \leq e^{-\frac{T}{2}c_2(\mu) + \frac{t}{2} - \frac{tn}{N}},$$

with $n = n_\mu$ in the sense that there exist unique partitions α and β with less than $\frac{N}{2}$ parts such that $\mu = \lambda_N(\alpha, \beta, n)$. From (3.13) we have for any $\mu \in \widehat{\mathrm{U}}(N)$

$$\sum_{\lambda \searrow \mu} \frac{d_\lambda}{d_\mu} = N.$$

These equations yield

$$0 \leq \Delta(N) \leq \frac{e^{\frac{t}{2}}}{Z_N(g, T) N^{2g-2}} \sum_{\mu \in \widehat{\mathrm{U}}(N) \setminus \Lambda_N} e^{-\frac{T}{2}c_2(\mu) - \frac{tn}{N}},$$

and the right-hand side tends to 0 as N tends to infinity because $2g - 2 > 0$ and the sum on the right is bounded independently from N . Finally, as $Z_N(g, T) \rightarrow \vartheta(0; \frac{iT}{2\pi})$, we get that

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_\ell)] = e^{-\frac{t}{2}},$$

as expected. □

Proof of Thm. 3.2.6 in the unitary case with $g \geq 2$. We will prove (3.60) as previously, and this will imply that the Wilson loop variance tends to 0. Let us set

$$\Lambda_N = \{(n, \dots, n) \in \widehat{U}(N), n \in \mathbb{Z}\}$$

as in the previous proof. We have from Equation (3.30)

$$\begin{aligned} \mathbb{E}[\text{tr}(H_\ell)\text{tr}(H_\ell^*)] &= \frac{1}{N^2 Z_N(g, T)} \sum_{\mu \in \Lambda_N} \sum_{\substack{\lambda \in \widehat{U}(N) \\ \lambda \sim \mu}} \frac{q^{c_2(\mu)} d_\lambda}{(d_\mu)^{2g-2} d_\mu} e^{\frac{t}{2}(c_2(\mu) - c_2(\lambda))} \\ &\quad + \frac{1}{N^2 Z_N(g, T)} \sum_{\mu \in \widehat{U}(N) \setminus \Lambda_N} \sum_{\substack{\lambda \in \widehat{U}(N) \\ \lambda \sim \mu}} \frac{q^{c_2(\mu)} d_\lambda}{(d_\mu)^{2g-2} d_\mu} e^{\frac{t}{2}(c_2(\mu) - c_2(\lambda))}. \end{aligned}$$

If $\mu = (n, \dots, n) \in \Lambda_N$ is a flat highest weight, then there are only two highest weights equivalent to μ : $\lambda = (n+1, n, \dots, n, n-1)$ and μ itself. We have $c_2(\lambda) = n^2 + 2$ and $d_\lambda = N^2 - 1$, therefore

$$\frac{1}{N^2 Z_N(g, T)} \sum_{\mu \in \Lambda_N} \sum_{\substack{\lambda \in \widehat{U}(N) \\ \lambda \sim \mu}} \frac{q^{c_2(\mu)} d_\lambda}{(d_\mu)^{2g-2} d_\mu} e^{\frac{t}{2}(c_2(\mu) - c_2(\lambda))} = \frac{e^{-t}}{Z_N(g, T)} \sum_{n \in \mathbb{Z}} \left(\frac{1}{N^2} + q^{n^2} \frac{N^2 - 1}{N^2} \right).$$

It is clear that the right-hand side converges to e^{-t} when $N \rightarrow \infty$.

Now, let us prove that the following remainder converges to 0:

$$\Delta(N) = \frac{1}{N^2 Z_N(g, T)} \sum_{\mu \in \widehat{U}(N) \setminus \Lambda_N} \sum_{\substack{\lambda \in \widehat{U}(N) \\ \lambda \sim \mu}} \frac{q^{c_2(\mu)} d_\lambda}{(d_\mu)^{2g-2} d_\mu} e^{\frac{t}{2}(c_2(\mu) - c_2(\lambda))}.$$

Recall that for any $\mu \in \widehat{U}(N) \setminus \Lambda_N$, we have $d_\mu \geq N$; using similar arguments as in Prop. 3.2.8, we can prove that for any $\lambda \sim \mu$

$$c_2(\mu) - c_2(\lambda) \leq 2 + \frac{2\mu_1}{N}.$$

We can also reproduce the proof of (3.61) to get

$$\sum_{\substack{\lambda \in \widehat{U}(N) \\ \lambda \not\sim \mu}} \frac{d_\lambda}{d_\mu} \leq N^2.$$

It yields

$$0 \leq \Delta(N) \leq \frac{e^t}{N^{2g-1} Z_N(g, T)} \sum_{\mu \in \widehat{U}(N) \setminus \Lambda_N} e^{-\frac{T}{2} c_2(\mu) + \frac{t\mu_1}{N}}.$$

From the definition of $c_2(\mu)$ it is straightforward to check that

$$c_2(\mu) \geq \frac{\mu_1^2}{N} \geq \frac{\mu_1}{N},$$

hence

$$-\frac{T}{2} c_2(\mu) + \frac{t\mu_1}{N} \leq \left(\frac{t}{N} - \frac{T}{2} \right) c_2(\mu),$$

and for $N > 4t/T$ we have

$$\Delta(N) \leq \frac{e^t}{N^{2g-2} Z_N(g, T)} \sum_{\mu \in \widehat{U}(N) \setminus \Lambda_N} e^{-\frac{T}{4} c_2(\mu)} \leq \frac{1}{N^{2g-2}} \frac{Z_N(1, \frac{T}{2})}{Z_N(g, T)}.$$

The right-hand side converges to 0 when $N \rightarrow \infty$, therefore it is also the case for $\Delta(N)$. This concludes the proof. \square

3.2.2 Loops with self-intersections

From our asymptotic analysis of simple loops and Makeenko–Migdal equations we can deduce the limit of Wilson loops for a broader class of loops, still included in a topological disk: it was proved by Hall in [Hal18], using the same kind of arguments as in the plane or the sphere. Before we state this theorem, let us give a few definitions that it will rely on.

Definition 3.2.2. Let Σ be a closed compact connected surface of total area T , and L a loop on Σ with a finite number of simple self-intersections included in a topological disk $U \subset \Sigma$. Suppose that there is an admissible graph $\mathbb{G} = (\mathbb{E}, \mathbb{V}, \mathbb{F})$ such that L is traced out on \mathbb{G} , and that there exists a face $F_0 \in \mathbb{F}^{14}$ such that ∂F_0 is contained in L .

- (i) If x is a point of $U \setminus L([0, 1])$, we define the *winding number* $w(L, x)$ of L around x as the homotopy class of L in $U \setminus \{x\}$.
- (ii) For any face $F \in \mathbb{F}$ we define the *winding number* of L around F as the quantity $w(L, F) = w(L, x)$ for a given $x \in F^{15}$.
- (iii) We call *admissible subloop* of L any subloop obtained by splitting of L according to Makeenko–Migdal equations. By convention we also consider L to be an admissible subloop of itself.
- (iv) The *maximal winding number* of L is defined as

$$|w_{\max}| = \max_{L', F'} |w(L', F')|, \quad (3.63)$$

where the maximum is taken among the indices $F' \in \mathbb{F}$ and the admissible subloops L' of L .

¹⁴This face can be seen as the ‘external face’ of the loop.

¹⁵It is not hard to see that this definition does not depend on the $x \in F$ we choose.

(v) The loop L is said to be *small* if

$$A|w_{\max}| < T, \quad (3.64)$$

where $A = T - \text{area}(F_0)$ is the difference between the total area of Σ and the area of the external face F_0 .

Under the ‘smallness’ assumption described in the previous definition, we have the following result.

Theorem 3.2.14 ([Hal18], Thm.18). *Let Σ be a closed compact connected surface of total area T . If for any simple loop c contained in a topological disk of Σ the Wilson loop expectation $\mathbb{E}[\text{tr}(H_c)]$ admits a finite limit (as $N \rightarrow \infty$) which is continuous with respect to the area of the domain enclosed by c , and if for any such curve c the variance of the Wilson loop $\text{Var}(\text{tr}(H_c))$ tends to 0, then for any small loop L with a finite number of simple self-intersections included in a topological disk:*

(i) *The limit $\Phi_L := \lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_L)]$ exists and depends continuously on the areas of the faces of L .*

(ii) *The associated variance vanishes in the large N limit:*

$$\lim_{N \rightarrow \infty} \text{Var}(\text{tr}(H_L)) = 0.$$

(iii) *The limiting expectation values satisfy the following large- N Makeenko–Migdal equations. Let us vary the areas of the faces surrounding a crossing v in a checkerboard pattern as in Fig. 3.5, resulting in a family of curves $L(t)$. Then*

$$\frac{d}{dt} \Phi_{L(t)} = \Phi_{L_1(t)} \Phi_{L_2(t)}, \quad (3.65)$$

where L_1 and L_2 are derived from L in the usual way.

What we proved in Section 3.2.1 imply that, for $G = \text{U}(N)$ or $\text{SU}(N)$, the assumptions of Thm. 3.2.14 are satisfied. In particular, we are able to compute the values of the master field (assuming that it exists) for any small loop contained in a topological disk, with a finite number of self-intersections. This is a good start but obviously not satisfactory. It would be at least interesting to get rid of the smallness assumption. Let us illustrate the limits of this Theorem with Fig. 3.7. The loop considered here can be deformed into a simple loop only if the external face has a sufficient area. If its area z is less than y , where y is the area of the smallest internal face, then we are not able to get rid of the self-intersection.

There is still hope, because Hall also proved in [Hal18] the following theorem, based on stronger assumptions.

Theorem 3.2.15 ([Hal15], Thm.19). *Let Σ be a closed compact connected surface of total area T . If for any closed curve c contained in a topological disk of Σ and any $n \in \mathbb{Z}$ the Wilson loop expectation $\mathbb{E}[\text{tr}(H_c^n)]$ admits a finite limit (as $N \rightarrow \infty$) which is continuous with respect to the area of the domain enclosed by c , and if for any such curve c the variance of the Wilson loop $\text{Var}(\text{tr}(H_c^n))$ tends to 0, then for any loop L with a finite number of simple self-intersections included in a topological disk:*

(i) *The limit $\Phi_L := \lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_L)]$ exists and depends continuously on the areas of the faces of L .*

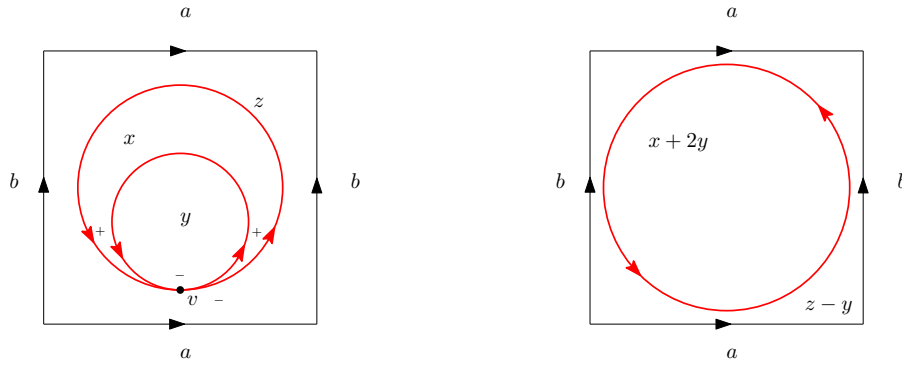


Figure 3.7: On the left: a loop in the torus with maximal winding number 2 and simple crossing at v . It can be deformed to become the simple loop on the right if $z \geq y$, otherwise the face with area y cannot be completely shrunk and the crossing at v remains.

(ii) The associated variance goes to zero:

$$\lim_{N \rightarrow \infty} \text{Var}(\text{tr}(H_L)) = 0.$$

(iii) The limiting expectations satisfy the following large- N Makeenko–Migdal equations. Let us vary the areas of the faces surrounding a crossing v in a checkerboard pattern as in Fig. 3.5, resulting in a family of curves $L(t)$. Then

$$\frac{d}{dt} \Phi_{L(t)} = \Phi_{L_1(t)} \Phi_{L_2(t)}, \tag{3.66}$$

where L_1 and L_2 are derived from L in the usual way.

It might be possible, although quite gruesome, to generalize Theorem 3.2.5 and Theorem 3.2.6 to the traces of higher moments of the holonomy process. It seems, fortunately, that using the methods from [DN17] actually give these limits, according to a private communication from Antoine Dahlqvist. This would enable to treat all (reasonable) loops contained in a topological disk, according to Thm. 3.2.15.

Even if we consider this, the construction of the master field is far from being complete yet, because for all surfaces with genus 1 and higher there are a lot of loops that cannot be contained in a topological disk: for instance, nonseparating loops have no chances to be in a topological disk, as they are not even homotopically equivalent to a point. The next sections will be devoted to describe the Wilson loops for loops that do not fit to the assumptions of Theorems 3.2.5, 3.2.6, 3.2.14 or 3.2.15.

3.2.3 Loops with nontrivial homology

In this section, we will prove that the Wilson loop expectation of a loop with nontrivial homology is a Haar unitary, and then we will study the convergence of its variance.

Theorem 3.2.16. *Let Σ be a closed compact connected surface of genus $g \geq 1$ and \mathbb{G} an admissible graph on Σ with \mathbf{f} faces, and v a vertex in \mathbb{G} . For any $\ell \in \mathcal{L}_v^{\text{red}}(\mathbb{G})$ with nontrivial homology and for any $n \neq 0$ we have*

$$\mathbb{E}[\text{tr}(H_\ell^n)] = 0. \tag{3.67}$$

Proof. Let $a_1, b_1, \dots, a_g, b_g, \ell_1, \dots, \ell_f$ the tame generators of $\mathcal{L}_v^{\text{red}}(\mathbb{G})$ (cf. Prop. 1.4.5). Without a loss of generality we can set the loop $\ell = \prod_{j=1}^k \prod_{i=1}^g (a_i^{m_{ij}} b_i^{n_{ij}} w_{ij}(\ell_1, \dots, \ell_{f-1}))$ and its homology class $(p_1, q_1, \dots, p_g, q_g)$ in $H_1(\Sigma) \simeq \mathbb{Z}^{2g}$. If the latter is nontrivial, it implies that at least one of the p_i or q_i is nonzero. Recall that $H_1(\Sigma)$ is abelian, therefore the roles of p_i and q_i are symmetric and we can assume without a loss of generality that $p_{i_0} \neq 0$ for a fixed i_0 . Given the definition we choose for ℓ , it is clear that $p_i = \sum_j m_{ij}$ and $q_i = \sum_j n_{ij}$ for all i . Let us now compute $\mathbb{E} [\text{tr}(H_\ell^n)]$. From the Driver–Sengupta formula we have

$$\begin{aligned} \mathbb{E} [\text{tr}(H_\ell^n)] &= \frac{1}{Z} \int_{G^{2g+f-1}} \text{tr} \left[\left(\prod_{j=1}^k \prod_{i=1}^g (x_i^{m_{ij}} y_i^{n_{ij}} w_{ij}(z_1, \dots, z_{f-1})) \right)^n \right] \prod_{t=1}^{f-1} p_{|F_t|}(z_t) \\ &\quad \times p_{|F_f|}([x_1, y_1] \dots [x_g, y_g](z_1 \dots z_{f-1})^{-1}) \prod_{i=1}^g dx_i dy_i \prod_{j=1}^{f-1} dz_j. \end{aligned} \quad (3.68)$$

Let $u_1, v_1, \dots, u_g, v_g \in Z(G)$ some elements of the center of $U(N)$. As $Z(U(N))$ is the set of scalar matrices we can set $u_i = \mathbf{u}_i I_N$ and $v_i = \mathbf{v}_i I_N$ with $|\mathbf{u}_i| = |\mathbf{v}_i| = 1$. Let us consider now \tilde{H}_ℓ the random variable obtained from H_ℓ by multiplying the generators of $\pi_1(\Sigma)$ by u_i and v_i respectively. In other terms, it is characterized by

$$\begin{aligned} \mathbb{E} [\text{tr}(\tilde{H}_\ell^n)] &= \frac{1}{Z} \int_{G^{2g+f-1}} \text{tr} \left[\left(\prod_{j=1}^k \prod_{i=1}^g ((u_i x_i)^{m_{ij}} (v_i y_i)^{n_{ij}} w_{ij}(z_1, \dots, z_{f-1})) \right)^n \right] \prod_{t=1}^{f-1} p_{|F_t|}(z_t) \\ &\quad \times p_{|F_f|}([u_1 x_1, v_1 y_1] \dots [u_g x_g, v_g y_g](z_1 \dots z_{f-1})^{-1}) \prod_{i=1}^g dx_i dy_i \prod_{j=1}^{f-1} dz_j. \end{aligned}$$

We will compute $\mathbb{E} [\text{tr}(\tilde{H}_\ell^n)]$ in two different ways.

(i) Using the fact that $[u_i x_i, v_i y_i] = [x_i, y_i]$, we get by linearity

$$\begin{aligned} \mathbb{E} [\text{tr}(\tilde{H}_\ell^n)] &= \frac{\prod_{i=1}^g \mathbf{u}_i^{np_i} \mathbf{v}_i^{nq_i}}{Z} \int_{G^{2g+f-1}} \text{tr} \left[\left(\prod_{j=1}^k \prod_{i=1}^g ((x_i)^{m_{ij}} (y_i)^{n_{ij}} w_{ij}(z_1, \dots, z_{f-1})) \right)^n \right] \\ &\quad \times \left(\prod_{t=1}^{f-1} p_{|F_t|}(z_t) \right) p_{|F_f|}([x_1, y_1] \dots [x_g, y_g](z_1 \dots z_{f-1})^{-1}) \prod_{i=1}^g dx_i dy_i \prod_{j=1}^{f-1} dz_j \\ &= \left(\prod_{i=1}^g \mathbf{u}_i^{np_i} \mathbf{v}_i^{nq_i} \right) \mathbb{E} [\text{tr}(H_\ell^n)]. \end{aligned}$$

(ii) With the change of variables $x'_i = u_i x_i, y'_i = v_i y_i$ in Equation (3.68) we directly get $\mathbb{E} [\text{tr}(H_\ell^n)]$.

These computations lead to

$$\left(\prod_{i=1}^g \mathbf{u}_i^{np_i} \mathbf{v}_i^{nq_i} \right) \mathbb{E} [\text{tr}(H_\ell^n)] = \mathbb{E} [\text{tr}(H_\ell^n)],$$

hence

$$\left(1 - \left(\prod_{i=1}^g \mathbf{u}_i^{np_i} \mathbf{v}_i^{nq_i}\right)\right) \mathbb{E}[\mathrm{tr}(H_\ell^n)] = 0.$$

As we assumed $n \neq 0$ and $p_{i_0} \neq 0$, by taking \mathbf{u}_{i_0} such that $(\mathbf{u}_{i_0}^{p_{i_0}})^n \neq 1$ and fixing the other \mathbf{u}_i and \mathbf{v}_i to 1, we finally get

$$(1 - \mathbf{u}_{i_0}^{np_{i_0}}) \mathbb{E}[\mathrm{tr}(H_\ell^n)] = 0,$$

which concludes the proof. \square

Now let us turn to the computation of the variance. It is more complicated because we cannot use the same trick as in the expectation: indeed, the multiplication of edge variables by scalar matrices will not produce a nontrivial factor, as the coefficient from the word will cancel out with the one from its inverse. However, it is still possible to prove a concentration result for simple nonseparating loops.

Proposition 3.2.17. *Let $\Sigma_{g,T}$ be a closed compact connected surface of genus $g \geq 1$ and total area $T > 0$, and ℓ be a nonseparating simple loop. We have*

$$\lim_{N \rightarrow \infty} \mathrm{Var}[\mathrm{tr}(H_\ell)] = 0.$$

Proof. We can complete ℓ into a set of generators $\{a_1, b_1, \dots, a_g, b_g\}$ of $\pi_1(\Sigma_{g,T})$, with $\ell = a_1$. From Driver–Sengupta formula we have

$$\mathbb{E}[\mathrm{tr}(H_\ell) \overline{\mathrm{tr}(H_\ell)}] = \frac{1}{Z_N(g, T)} \int_{\mathrm{U}(N)^{2g}} \mathrm{tr}(x_1) \overline{\mathrm{tr}(x_1)} p_T([x_1, y_1] \cdots [x_g, y_g]) dx_1 dy_1 \cdots dx_N dy_N,$$

and we can use, as always, the Fourier decomposition of the heat kernel, which yields

$$\begin{aligned} \mathbb{E}[\mathrm{tr}(H_\ell) \overline{\mathrm{tr}(H_\ell)}] &= \frac{1}{Z_N(g, T)} \sum_{\lambda \in \widehat{\mathrm{U}}(N)} e^{-\frac{c_2(\lambda)T}{2}} d_\lambda \\ &\quad \times \int_{\mathrm{U}(N)^{2g}} \mathrm{tr}(x_1) \overline{\mathrm{tr}(x_1)} s_\lambda([x_1, y_1] \cdots [x_g, y_g]) dx_1 dy_1 \cdots dx_N dy_N. \end{aligned}$$

Then, using $g - 1$ times Lemma 3.2.2 and once Prop. 3.2.3, we find

$$\mathbb{E}[\mathrm{tr}(H_\ell) \overline{\mathrm{tr}(H_\ell)}] = \frac{1}{Z_N(g, T)} \sum_{\lambda \in \widehat{\mathrm{U}}(N)} e^{-\frac{c_2(\lambda)T}{2}} d_\lambda^{2-2g} \int_{\mathrm{U}(N)} \mathrm{tr}(x_1) \overline{\mathrm{tr}(x_1)} s_\lambda(x_1) s_\lambda(x_1^*) dx_1.$$

Let us notice that the summand is exactly the one from the partition function, multiplied by the remaining integral. The latter can be reinterpreted as follows: because $\overline{\mathrm{tr}(x_1)} = \mathrm{tr}(x_1^*)$ we have

$$\int_{\mathrm{U}(N)} \mathrm{tr}(x_1) \overline{\mathrm{tr}(x_1)} s_\lambda(x_1) s_\lambda(x_1^*) dx_1 = \|\mathrm{tr}(\cdot) s_\lambda(\cdot)\|_{L^2(\mathrm{U}(N))}^2,$$

and we will prove that this norm tends to 0 as N tends to infinity. Indeed, using formula (3.13) and the character orthogonality relations, we have

$$\begin{aligned} \|\mathrm{tr}(\cdot) s_\lambda(\cdot)\|_{L^2(\mathrm{U}(N))}^2 &= \frac{1}{N^2} \sum_{\substack{\mu \searrow \lambda \\ \nu \searrow \lambda}} \int_{\mathrm{U}(N)} s_\mu(x) s_\nu(x^*) dx \\ &= \frac{1}{N^2} \#\{\mu \in \widehat{\mathrm{U}}(N) : \mu \searrow \lambda\}. \end{aligned}$$

Besides, as there are at most N ways of adding a box to the diagram of λ , we indeed find out that $\|\mathrm{tr}(\cdot) s_\lambda(\cdot)\|_{L^2(\mathrm{U}(N))}^2 \leq \frac{1}{N}$, and therefore $\mathbb{E}[\mathrm{tr}(H_\ell) \overline{\mathrm{tr}(H_\ell)}] \rightarrow 0$ as $N \rightarrow \infty$. From Theorem 3.2.16 we deduce that the variance also converges to 0. \square

We deduce from Theorem 3.2.16 and Proposition 3.2.17 that simple nonseparating loops are Haar unitaries such that the variance of their trace vanishes in the large N limit. However, that does not give us any information about the variance of more general loops with nonzero homology.

The conclusion to this section is that the master field, assuming its existence, is equal to zero when applied to the generators of the fundamental group; the moments of the limit of Yang–Mills holonomy process also converge to 0 for all other loops that are homologically nontrivial, but we don’t know yet if the limit of the process itself is 0, nor if the convergence holds in probability.

3.2.4 Joint distribution for several loops

Till now, we only considered the noncommutative distribution of a single loop. However it might be also interesting to study the joint distribution of several loops, as in the case of the plane. Indeed, Lévy [Lév17] proved – among other things – that the variables $(H_\ell)_{\ell \in \mathcal{B}}$ associated to the lasso basis of a graph on the plane are independent and asymptotically free. We will see how this result is modified in our setup, for a compact surface.

Proposition 3.2.18. *Let Σ be a closed compact connected surface of genus $g \geq 1$ and \mathbb{G} an admissible graph on Σ with \mathfrak{f} faces, and v a vertex in \mathbb{G} . Let $\mathcal{B} = \{a_1, b_1, \dots, a_g, b_g, \ell_1, \dots, \ell_{\mathfrak{f}}\}$ the set of tame generators of $\mathcal{L}_v^{\mathrm{red}}(\mathbb{G})$, then all these generators are asymptotically uncorrelated, that is,*

$$\lim_{N \rightarrow \infty} \mathrm{Cov}(\mathrm{tr}(H_{c_1}), \mathrm{tr}(H_{c_2})) = 0, \quad \forall (c_1, c_2) \in \mathcal{B}^2. \tag{3.69}$$

Proof. From Cauchy–Schwarz inequality, we have

$$|\mathrm{Cov}(\mathrm{tr}(H_{c_1}), \mathrm{tr}(H_{c_2}))| \leq \sqrt{\mathrm{Var}(\mathrm{tr}(H_{c_1}))} \sqrt{\mathrm{Var}(\mathrm{tr}(H_{c_2}))}.$$

Then, to conclude the proof, we only need to check that for any couple of (c_1, c_2) both Wilson loops have finite variance and at least one has a variance that vanishes in the large N limit; but this is a direct consequence of Theorem 3.2.6 for the lasso generators and Prop. 3.2.17 for the generators of $\pi_1(\Sigma)$, which are indeed simple nonseparating loops. \square

This Proposition implies in particular that for all $(c_1, c_2) \in \mathcal{B}^2$,

$$\lim_{N \rightarrow \infty} \mathbb{E}[\mathrm{tr}(H_{c_1}) \mathrm{tr}(H_{c_2})] = \lim_{N \rightarrow \infty} \mathbb{E}[\mathrm{tr}(H_{c_1})] \mathbb{E}[\mathrm{tr}(H_{c_2})].$$

By induction, we can prove the following result.

Corollary 3.2.19. *Let Σ be a closed compact connected surface of genus $g \geq 1$ and \mathbb{G} an admissible graph on Σ with \mathbf{f} faces, and v a vertex in \mathbb{G} . Let $\mathcal{B} = \{a_1, b_1, \dots, a_g, b_g, \ell_1, \dots, \ell_{\mathbf{f}}\}$ the set of tame generators of $\mathcal{L}_v^{\text{red}}(\mathbb{G})$, then we have for all $1 \leq k \leq \mathbf{f} + 2g$ and $(c_1, \dots, c_k) \in \mathcal{B}^k$:*

$$\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_{c_1}) \cdots \text{tr}(H_{c_k})] = \lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(H_{c_1})] \cdots \mathbb{E}[\text{tr}(H_{c_k})]. \quad (3.70)$$

This guarantees that the large- N Makeenko–Migdal equations (3.65) remain satisfied for all the tame generators of $\mathcal{L}_v^{\text{red}}(\mathbb{G})$. However, this is not enough yet to prove that these generators are asymptotically free, and we do not even know if it could be the case.

3.2.5 Remaining loops: the ‘missing link’

The purpose of this section is to offer an informal overview of what is left to prove in order to get a robust construction of the master field for general compact surfaces. We will give examples of situations where the values of the master field remain unknown, or even where it is not clear whether the Wilson loops converge.

Theorems 3.2.5, 3.2.6 and 3.2.14 enabled us to know the existence and the value of the master field for small loops contained in a topological disk, and Theorem 3.2.15 made us able to extend it to more general loops, still contained in a disk, assuming that we could compute all moments of the Yang–Mills field and not only its Wilson loop expectations. For a compact surface of genus 1 or higher, there are unfortunately loops that are homotopically trivial but cannot be contained into a topological disk. Let us illustrate it with Fig. 3.8.

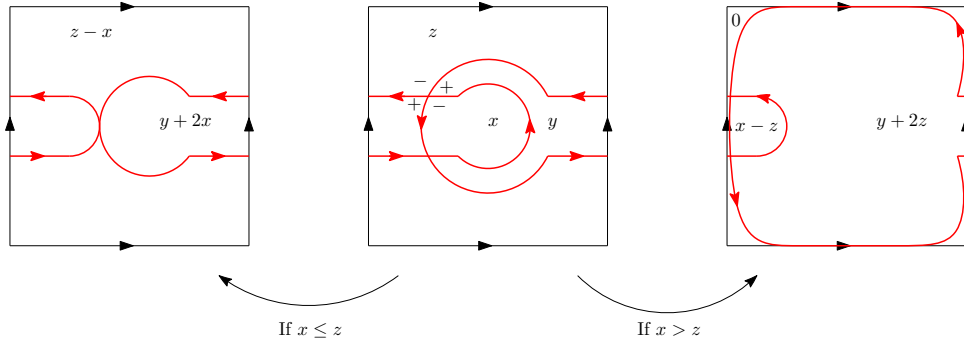


Figure 3.8: In the middle: a self-overlapping loop on the torus. It can be transformed into the one on the left if the area z is large enough to ‘absorb’ the area x , otherwise it will be stuck in the configuration in the right.

The loop considered here in a torus can be deformed into a simple loop only if the external face has a sufficient area (and in this case, it is possible to make it fit into a topological disk). If its area z is less than x , where x is the area of the smallest internal face. Otherwise, we are stuck in a configuration like the one in the bottom right of Fig. 3.8, which still has self-intersection points. It appears that even if we know all moments of the Yang–Mills field on a simple loop, it will not be enough to compute its value in the remaining loop, nor its limit. For this kind of situation, the key could be to use analytic properties of the limit of Wilson loop expectations. It might be possible to extend them to negative values of the variable areas, although it would not correspond to a physical situation anymore; another (maybe more promising) possibility would be to use the fact that the large N limit of Wilson

loop expectations and variance do not depend on the total area of the surface, or even on the area of the ‘external face’, and perhaps one could deform this face so that the smallness condition on the loop is satisfied. To this day, we are still investigating such theories.

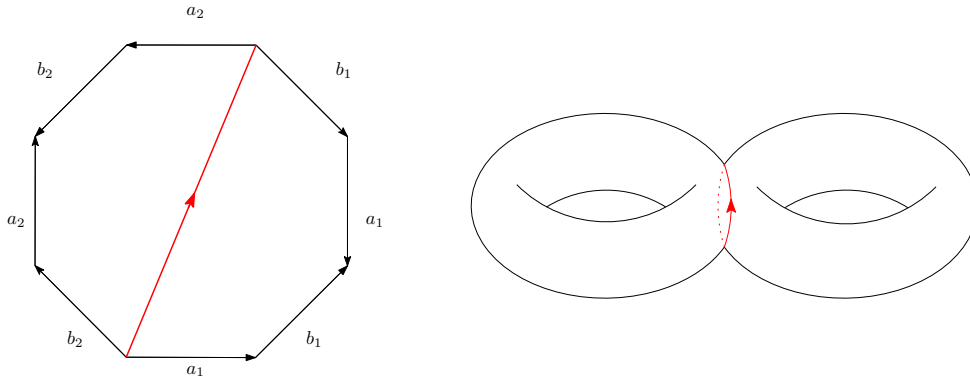


Figure 3.9: The loop in red, traced out on a surface of genus 2, is homologically trivial but not homotopically trivial.

In the case of genus 2 and higher, another kind of loop is also of interest: as the homology group and the fundamental group are only isomorphic for surfaces with genus 1, it appears that for higher genus there exist loops which are separating but not contractible (in other terms, they have zero homology but nonzero homotopy). For instance, the loop ℓ illustrated in Fig. 3.9 has zero homology but cannot be continuously deformed into a single point; it splits the surface into two faces with areas T_1 and T_2 . Such a loop is kind of critical and it is not clear how to compute the limit of its Wilson loop.

Finally, the last class of loops which must be investigated is the class of loops which have nonzero homology. Indeed, although we proved that they have the distribution of Haar unitaries, we only proved the convergence of the Wilson loop variance for a particular subset. The case of more general loops with nonzero homology remains open.

BIBLIOGRAPHIE

- [AGZ10] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.
- [AS12] Michael Anshelevich and Ambar N. Sengupta. Quantum free Yang–Mills on the plane. *J. Geom. Phys.*, 62(2):330–343, 2012.
- [Bia97] Philippe Biane. Free Brownian motion, free stochastic calculus and random matrices. In *Free probability theory (Waterloo, ON, 1995)*, volume 12 of *Fields Inst. Commun.*, pages 1–19. Amer. Math. Soc., Providence, RI, 1997.
- [BO05] Alexei Borodin and Grigori Olshanski. Harmonic analysis on the infinite-dimensional unitary group and determinantal point processes. *Ann. of Math. (2)*, 161(3):1319–1422, 2005.
- [Bou63] N. Bourbaki. *Éléments de mathématique. Fascicule XXIX. Livre VI: Intégration. Chapitre 7: Mesure de Haar. Chapitre 8: Convolution et représentations*. Actualités Scientifiques et Industrielles, No. 1306. Hermann, Paris, 1963.
- [Bou94] D. V. Boulatov. Wilson loop on a sphere. *Modern Phys. Lett. A*, 9(4):365–374, 1994.
- [BP13] Alexei Borodin and Leonid Petrov. Integrable probability: from representation theory to macdonald processes. arXiv preprint: <https://arxiv.org/abs/1310.8007v3>, 2013.
- [BS01] Philippe Biane and Roland Speicher. Free diffusions, free entropy and free Fisher information. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(5):581–606, 2001.
- [BtD95] Theodor Bröcker and Tammo tom Dieck. *Representations of compact Lie groups*, volume 98 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.
- [Cao20] Sky Cao. Wilson loop expectations in lattice gauge theories with finite gauge groups. arXiv preprint: <https://arxiv.org/abs/2001.05627>, 2020.
- [CDG17] Guillaume Cébron, Antoine Dahlqvist, and Franck Gabriel. The generalized master fields. *J. Geom. Phys.*, 119:34–53, 2017.
- [CDK18] Benoît Collins, Antoine Dahlqvist, and Todd Kemp. The spectral edge of unitary Brownian motion. *Probab. Theory Related Fields*, 170(1-2):49–93, 2018.

- [Chai8] Sourav Chatterjee. Wilson loops in Ising lattice gauge theory. arXiv preprint: <https://arxiv.org/abs/1811.09770>, 2018.
- [Con90] Alain Connes. *Géométrie non commutative*. InterEditions, Paris, 1990.
- [Dah16] Antoine Dahlqvist. Free energies and fluctuations for the unitary Brownian motion. *Comm. Math. Phys.*, 348(2):395–444, 2016.
- [Dah17] Antoine Dahlqvist. Integration formulas for Brownian motion on classical compact Lie groups. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(4):1971–1990, 2017.
- [DEF⁺99] Pierre Deligne, Pavel Etingof, Daniel S. Freed, Lisa C. Jeffrey, David Kazhdan, John W. Morgan, David R. Morrison, and Edward Witten. *Quantum fields and strings: a course for mathematicians*. American Mathematical Society, 1999.
- [DGHK17] Bruce K. Driver, Franck Gabriel, Brian C. Hall, and Todd Kemp. The Makeenko-Migdal equation for Yang-Mills theory on compact surfaces. *Comm. Math. Phys.*, 352(3):967–978, 2017.
- [DHK13] Bruce K. Driver, Brian C. Hall, and Todd Kemp. The large- N limit of the Segal-Bargmann transform on \mathbb{U}_N . *J. Funct. Anal.*, 265(11):2585–2644, 2013.
- [Dix64] Jacques Dixmier. *Les C^* -algèbres et leurs représentations*. Cahiers Scientifiques. Gauthier-Villars, 1964.
- [DK93] M. R. Douglas and V. A. Kazakov. Large N Phase Transition in continuum QCD₂. *Physics Letters B*, 319, 1993.
- [DK94] Jean-Marc Daul and Vladimir A. Kazakov. Wilson loop for large N Yang-Mills theory on a two-dimensional sphere. *Phys. Lett. B*, 335(3-4):371–376, 1994.
- [DN17] Antoine Dahlqvist and James Norris. Yang–Mills measure and the master field on the sphere. arXiv preprint: <https://arxiv.org/abs/1703.10578>, 2017.
- [Dri89] Bruce K. Driver. YM₂: continuum expectations, lattice convergence, and lassos. *Comm. Math. Phys.*, 123(4):575–616, 1989.
- [Dys62] Freeman J. Dyson. A Brownian-motion model for the eigenvalues of a random matrix. *J. Mathematical Phys.*, 3:1191–1198, 1962.
- [Far80] Jacques Faraut. *Analyse sur les paires de Gelfand et les espaces hyperboliques*. Cours du CIMPA. 1980.
- [Far08] Jacques Faraut. *Analysis on Lie groups*, volume 110 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2008.
- [FH91] William Fulton and Joe Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991.
- [FLV20] Malin P. Forsström, Jonatan Lenells, and Fredrik Viklund. Wilson loops in finite Abelian lattice gauge theories. arXiv preprint: <https://arxiv.org/abs/2001.07453>, 2020.

- [FMS11] Peter J. Forrester, Satya N. Majumdar, and Grégory Schehr. Non-intersecting Brownian walkers and Yang-Mills theory on the sphere. *Nuclear Phys. B*, 844(3):500–526, 2011.
- [GG95] Rajesh Gopakumar and David J. Gross. Mastering the master field. *Nuclear Phys. B*, 451(1-2):379–415, 1995.
- [GMO5] Alice Guionnet and Mylène Maïda. Character expansion method for the first order asymptotics of a matrix integral. *Probab. Theory Related Fields*, 132(4):539–578, 2005.
- [GN43] I. Gelfand and M. Neumark. On the imbedding of normed rings into the ring of operators in Hilbert space. *Rec. Math. [Mat. Sbornik] N.S.*, 12(54):197–213, 1943.
- [GO16] Vadim Gorin and Grigori Olshanski. A quantization of the harmonic analysis on the infinite-dimensional unitary group. *J. Funct. Anal.*, 270(1):375–418, 2016.
- [God03] Roger Godement. *Analyse mathématique IV*. Universitext. Springer-Verlag, 2003.
- [Gop96] Rajesh Gopakumar. The master field in generalised QCD₂. *Nuclear Phys. B*, 471(1-2):246–260, 1996.
- [GP15] Vadim Gorin and Greta Panova. Asymptotics of symmetric polynomials with applications to statistical mechanics and representation theory. *Ann. Probab.*, 43(6):3052–3132, 2015.
- [Gro88] Leonard Gross. The Maxwell equations for Yang-Mills theory. In *Mathematical quantum field theory and related topics (Montreal, PQ, 1987)*, volume 9 of *CMS Conf. Proc.*, pages 193–203. Amer. Math. Soc., Providence, RI, 1988.
- [GT93] David J. Gross and Washington Taylor, IV. Two-dimensional QCD is a string theory. *Nuclear Phys. B*, 400(1-3):181–208, 1993.
- [Gui04] Alice Guionnet. Large deviations and stochastic calculus for large random matrices. *Probab. Surv.*, 1:72–172, 2004.
- [Haa33] Alfred Haar. Der Massbegriff in der Theorie der kontinuierlichen Gruppen. *Ann. of Math. (2)*, 34(1):147–169, 1933.
- [Hal15] Brian C. Hall. *Lie groups, Lie algebras, and representations*, volume 222 of *Graduate Texts in Mathematics*. Springer, Cham, second edition, 2015.
- [Hal18] Brian C. Hall. The large- N limit for two-dimensional Yang–Mills theory. *Comm. Math. Phys.*, 363(3):789–828, 2018.
- [HL10] Ben Hambly and Terry Lyons. Uniqueness for the signature of a path of bounded variation and the reduced path group. *Ann. of Math. (2)*, 171(1):109–167, 2010.
- [Kak38] Shizuo Kakutani. On the uniqueness of Haar’s measure. *Proc. Imp. Acad. Tokyo*, 14(2):27–31, 1938.
- [KK80] V. A. Kazakov and I. K. Kostov. Nonlinear strings in two-dimensional $U(\infty)$ gauge theory. *Nuclear Phys. B*, 176(1):199–215, 1980.

- [KN96] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of differential geometry. Vol. I.* Wiley Classics Library. John Wiley & Sons, Inc., New York, 1996.
- [Laf12] J. Lafontaine. *Introduction aux variétés différentielles.* GRENOBLE SCIENCES. EDP Sciences, 2012.
- [Lee13] John M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.
- [Lem19] Thibaut Lemoine. Large N behaviour of the two-dimensional Yang–Mills partition function. 2019. arXiv preprint: <https://arxiv.org/abs/1911.07010>.
- [Lévo3] Thierry Lévy. Yang-Mills measure on compact surfaces. *Mem. Amer. Math. Soc.*, 166(790):xiv+122, 2003.
- [Lévo4] Thierry Lévy. Wilson loops in the light of spin networks. *J. Geom. Phys.*, 52(4):382–397, 2004.
- [Lévo8] Thierry Lévy. Schur-Weyl duality and the heat kernel measure on the unitary group. *Adv. Math.*, 218(2):537–575, 2008.
- [Lévo10] Thierry Lévy. Two-dimensional Markovian holonomy fields. *Astérisque*, (329):172, 2010.
- [Lévo11] Thierry Lévy. Topological quantum field theories and Markovian random fields. *Bull. Sci. Math.*, 135(6-7):629–649, 2011.
- [Lévo17] Thierry Lévy. The master field on the plane. *Astérisque*, (388):ix+201, 2017.
- [LM15] Thierry Lévy and Mylène Maïda. On the Douglas-Kazakov phase transition. Weighted potential theory under constraint for probabilists. In *Modélisation Aléatoire et Statistique—Journées MAS 2014*, volume 51 of *ESAIM Proc. Surveys*, pages 89–121. EDP Sci., Les Ulis, 2015.
- [LW16] Karl Liechty and Dong Wang. Nonintersecting Brownian motions on the unit circle. *Ann. Probab.*, 44(2):1134–1211, 2016.
- [LZo4] Sergei K. Lando and Alexander K. Zvonkin. *Graphs on surfaces and their applications*, volume 141 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004.
- [Mac15] I. G. Macdonald. *Symmetric functions and Hall polynomials.* Oxford Classic Texts in the Physical Sciences. The Clarendon Press, Oxford University Press, New York, second edition, 2015.
- [Mas91] William S. Massey. *A basic course in algebraic topology*, volume 127 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991.
- [Max65] James C. Maxwell. A dynamical theory of the electromagnetic field. 1865.
- [Mey86] Paul-André Meyer. Éléments de probabilités quantiques (exposés i à v). *Séminaire de probabilités de Strasbourg*, 20:186–312, 1986.

- [Mig75] Alexander A. Migdal. Recursion equations in gauge field theories. *Sov. Phys. JETP*, pages 413–418, 1975.
- [MM79] Yuri Makeenko and Alexander A. Migdal. Exact equation for the loop average in multi-color QCD. *Physics Letters B*, 88:135–137, 1979.
- [Mor01] Shigeyuki Morita. *Geometry of differential forms*, volume 201 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 2001.
- [NO06] Nikita A. Nekrasov and Andrei Okounkov. Seiberg-Witten theory and random partitions. In *The unity of mathematics*, volume 244 of *Progr. Math.*, pages 525–596. Birkhäuser Boston, Boston, MA, 2006.
- [NS06] A. Nica and R. Speicher. *Lectures on the Combinatorics of Free Probability*. Number vol. 13 in *Lectures on the combinatorics of free probability*. Cambridge University Press, 2006.
- [Olso3] Grigori Olshanski. The problem of harmonic analysis on the infinite-dimensional unitary group. *J. Funct. Anal.*, 205(2):464–524, 2003.
- [Pau14] Frédéric Paugam. *Towards the mathematics of quantum field theory*, volume 59 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer, Cham, 2014.
- [Proo7] Claudio Procesi. *Lie groups*. Universitext. Springer, New York, 2007. An approach through invariants and representations.
- [Rai97] E. M. Rains. Combinatorial properties of Brownian motion on the compact classical groups. *J. Theoret. Probab.*, 10(3):659–679, 1997.
- [Sag01] Bruce E. Sagan. *The symmetric group*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001.
- [Sen92] Ambar Sengupta. The Yang-Mills measure for S^2 . *J. Funct. Anal.*, 108(2):231–273, 1992.
- [Sen94] Ambar Sengupta. Gauge invariant functions of connections. *Proc. Amer. Math. Soc.*, 121(3):897–905, 1994.
- [Sen97] Ambar Sengupta. Yang-Mills on surfaces with boundary: quantum theory and symplectic limit. *Comm. Math. Phys.*, 183(3):661–705, 1997.
- [Seno8a] Ambar N. Sengupta. Gauge theory in two dimensions: topological, geometric and probabilistic aspects. In *Stochastic analysis in mathematical physics*, pages 109–129. World Sci. Publ., Hackensack, NJ, 2008.
- [Seno8b] Ambar N. Sengupta. Traces in two-dimensional QCD: the large- N limit. In *Traces in number theory, geometry and quantum fields*, Aspects Math., E38, pages 193–212. Friedr. Vieweg, Wiesbaden, 2008.
- [Ser78] Jean-Pierre Serre. *Représentations linéaires des groupes finis*. Hermann, Paris, revised edition, 1978.

- [Sero6] Jean-Pierre Serre. *Lie algebras and Lie groups*, volume 1500 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006.
- [Sin95] Isadore M. Singer. On the master field in two dimensions. In *Functional analysis on the eve of the 21st century, Vol. 1 (New Brunswick, NJ, 1993)*, volume 131 of *Progr. Math.*, pages 263–281. Birkhäuser Boston, Boston, MA, 1995.
- [Sta99] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.
- [Ste51] Norman Steenrod. *The Topology of Fibre Bundles*. Princeton Mathematical Series, vol. 14. Princeton University Press, Princeton, N. J., 1951.
- [Sti93] John Stillwell. *Classical topology and combinatorial group theory*, volume 72 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1993.
- [tH74] Gerard 't Hooft. A planar diagram theory for strong interactions. *Nuclear Physics B*, 72(3):461–473, 1974.
- [vN36] John von Neumann. The uniqueness of Haar's measure. *Math. Sbornik*, 1(43):721–734, 1936.
- [VO04] Anatoli M. Vershik and Andrei Yu. Okounkov. A new approach to representation theory of symmetric groups. II. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 307(Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 10):57–98, 281, 2004.
- [Wei40] André Weil. *L'intégration dans les groupes topologiques et ses applications*. Actual. Sci. Ind., no. 869. Hermann et Cie., Paris, 1940.
- [Wil74] Kenneth G. Wilson. Confinement of quarks. *Phys. Rev.*, 10, 1974.
- [Wit80] Edward Witten. The $1/N$ expansion in atomic and particle physics. *NATO Sci. Ser. B*, 59:403–419, 1980.
- [Wit91] Edward Witten. On quantum gauge theories in two dimensions. *Comm. Math. Phys.*, 141(1):153–209, 1991.
- [Xu97] Feng Xu. A random matrix model from two-dimensional Yang–Mills theory. *Comm. Math. Phys.*, 190(2):287–307, 1997.
- [YM54] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev. (2)*, 96:191–195, 1954.
- [Zag94] Don Zagier. Values of zeta functions and their applications. In *First European Congress of Mathematics, Vol. II (Paris, 1992)*, volume 120 of *Progr. Math.*, pages 497–512. Birkhäuser, Basel, 1994.