Thesis

On Chern-Simons-matter matrix models

PUTROV, Pavel

Abstract

The partition function and expectation values of certain BPS observables of N>=2 Chern-Simons-matter theories on the three-sphere are known to reduce to a finite dimensional matrix integral. We analyze these matrix models by different methods. We find different properties and exact results and provide various explicit and non-trivial checks of the AdS_4/CFT_3 conjecture. In particular we find the N^{3/2} behavior in the strong coupling limit predicted from string theory. In the case of ABJM theory we find a relation to topological strings on local P^1xP^1. We also find a connection between the strong coupling limit and the tropical limit for the spectral curve. We show that for theories with at least N=3 supersymmetry the partition function can be interpreted as a partition function of a certain free Fermi gas which allows us to find an explicit and exact answer for the perturbative part of the partition function.

Reference

PUTROV, Pavel. *On Chern-Simons-matter matrix models*. Thèse de doctorat : Univ. Genève, 2012, no. Sc. 4435

Université de Genève
Section de Mathématiques

Faculté des sciences
Professeur M. Mariño

On Chern–Simons–matter Matrix Models

Thèse

présentée à la Faculté des Sciences de l'Université de Genève
pour obtenir le grade de Docteur ès Sciences, mention interdisciplinaire

par
Pavel Putrov
de
Russie

Thèse № 4435

Genève
2012

La faculté des sciences, sur le préavis de Messieurs M. Mariño Beiras, professeur ordinaire et directeur de thèse (Section de mathématiques), A. Alekseev, professeur ordinaire (Section de mathématiques) et A. Klemm, professeur (Physikalisches Institut, Universität Bonn, Deutschland), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 31 mai 2012

Thèse № 4435

Résumé en français

Une percée importante dans l'étude de la correspondance $\mathrm{AdS}_4/\mathrm{CFT}_4$ a été faite par Aharony, Bergman, Jafferis et Maldacena [1] qui a trouvé une famille des théories de jauge superconformes à 3 dimensions marquées par $k \in \mathbb{Z}$ et qui ont été conjecturées être les théories de volume d'universe d'une pile de N branes M2 sondant $\mathbb{C}^4/\mathbb{Z}_k$. Une théorie de cette famille est donc duale à la théorie M dans $\mathrm{AdS}_4 \times \mathbb{S}^7/\mathbb{Z}_k$ ou, pour k grand, à la théorie des cordes de type IIA dans $\mathrm{AdS}_4 \times \mathbb{CP}^3$. Cette théorie est communément appelé la théorie ABJM. Plus tard elle a été généralisée à d'autres exemples de ce qu'on appelle les théories Chern–Simons–matter (CSM), qui ont été conjecturé pour décrire N branes M2 sondant une variété de Calabi-Yau Y_8. Pour N grandes la géométrie à cause de back-reaction devient $\mathrm{AdS}_4 \times X_7$ de sorte que Y_8 est un cône sur X_7 qui est alors une variété tri-Sasaki-Einstein. Ainsi, on peut déclarer la correspondance

$$\text{théorie CSM} \longleftrightarrow \text{M-théorie dans } \mathrm{AdS}_4 \times X_7 \tag{1}$$

Les données nécessaires pour définir une théorie CSM sont un groupe de jauge G et sa représentation R pour les champs de matière. Il faut également fixer les niveaux de CS, c.-à-d. un entier pour chaque composant simple de G. Kapustin, Willett et Yaakov dans [2] à l'aide du procédé de localisation ont réduit l'intégrale de chemin pour une théorie Chern–Simons–matter à une intégrale de matrice de la forme suivante:

$$Z \propto \int da\, e^{\frac{\mathrm{i}}{4\pi}\mathrm{Tr}\, a^2}\, \frac{\det_{\mathrm{Ad}} 2\sinh\frac{a}{2}}{\det_R 2\cosh\frac{a}{2}} \tag{2}$$

où l'intégration est effectuée sur la sous-algèbre de Cartan de l'algèbre de Lie $\mathfrak{g}$ du groupe de jauge G, et Tr est un certain produit invariant interne de $\mathfrak{g}$ qui dépend du choix des niveaux de CS, et

$$\det_R f(a) \equiv \prod_{\rho} f(\rho(a)) \tag{3}$$

où le produit est effectuée sur les poids de la représentation R (dans le cas de la représentation adjointe ($R = \mathrm{Ad}$) nous excluons les poids nuls). Le but principal de cette thèse est une analyse détaillée des modèles de matrices de type (3) et l'application des résultats obtenus à la vérification explicite non triviale de la correspondance $\mathrm{AdS}_4/\mathrm{CFT}_3$. Le premier chapitre de la thèse donne une revue des connaissances de base qui sont utilisé dans la partie principale de la thèse.

Dans le deuxième chapitre nous effectuons une analyse détaillée des modèles de matrice CSM dans la limite de 't Hooft. Dans le cas du modèle de matrice ABJM cette limite est définie comme suit:

$$N \to \infty, \qquad \lambda \equiv \frac{N}{k} \text{ est fixé.} \tag{4}$$

Alors les observables ont des séries asymptotiques en $g_s \equiv 2\pi i/k$ qui sont appelées les développements de genre, ou topologiques. Par exemple, pour l'énergie libre on a

$$F \equiv \log Z = \sum_{g \geq 0} g_s^{2g-2} F_g(\lambda). \tag{5}$$

Il y a une approche universelle pour étudier ces développements dans les modèles de matrice basée sur la notion d'une courbe spectrale. Dans cette approche la solution d'un modèle de matrice est entièrement déterminée par une courbe complexe $\mathcal{C}$, appelée courbe spectrale, une 1-forme méromorphe ω sur cette courbe et un choix d'une base dans $H_1(\mathcal{C}, \mathbb{Z})$. D'abord nous établissons la relation entre le modèle de matrice ABJM et le modèle de matrice pour la théorie de Chern–Simons sur l'espace lenticulaire $L(2,1) \cong \mathbb{RP}^3$:

$$Z_{\mathrm{ABJM}}(N_1, N_2, g_s) = Z_{L(2,1)}(N_1, -N_2, g_s) \tag{6}$$

où l'on considère la généralisation [3] de la théorie originale de ABJM au cas avec le groupe de jauge $U(N_1) \times U(N_2)$. En utilisant cette relation et la solution connue du modèle de matrice de l'espace lenticulaire nous construisons explicitement la solution du modèle de matrice ABJM en termes de la courbe spectrale. En utilisant en outre le fait que la théorie de CS sur $L(2,1)$ est un grand-N double des cordes topologiques sur $\mathbb{P}^1 \times \mathbb{P}^1$ local, nous concluons que l'espace des modules de la théorie ABJM est un sous-espace réel certain de l'espace des modules des structures de Kähler complexifiées de $\mathbb{P}^1 \times \mathbb{P}^1$.

Dans le cas de la théorie ABJM nous avons trouvé les expressions explicites exactes pour l'énergie libre et les boucle de Wilson BPS pour quelques premiers genres. Le développement de couplage faible ($\lambda \to 0$) peut être vérifiée pour être en accord avec les calculs dans la théorie des perturbations de la

théorie quantique des champs. Et nous vérifions que l'expansion de couplage fort ($\lambda \to \infty$) est en accord avec la prédiction du côté de la théorie des cordes par la correspondance AdS/CFT. En particulier, nous trouvons qu'au premier ordre

$$F_0 \approx \frac{4\pi^3\sqrt{2}}{3}\lambda^{3/2}, \qquad \lambda \to \infty. \tag{7}$$

qui fournit une dérivation de la mystérieuse loi de $N^{3/2}$ pour l'énergie libre prédite à partir de la théorie des cordes [4]. En outre, nous montrons que le coefficient coïncide avec celui qui peut être obtenu à partir du calcul de la supergravité. Nous trouvons aussi que les limites de couplage fort des boucles de Wilson 1/2 et 1/6 BPS sont en accord avec le comportement prédit par la correspondance AdS/CFT. Nous avons également étudié en détail les corrections sous-dominantes et trouvé qu'elles sont aussi en accord avec le côté de la théorie des cordes de la dualité. En particulier, les corrections non-perturbatives peuvent être interprétées comme les contributions des instantons de feuille d'univers.

Dans les sections 2.8 - 2.11, nous développons une méthode géométrique permettant le calcul direct de la limite de couplage fort sans obtenir d'abord les expressions exactes. La caractéristique principale de cette méthode est que l'ordre dominant des observables du modèle de matrice au couplage fort peut être calculé en termes de la version tropicale de la courbe spectrale. En utilisant cette méthode, nous vérifions la relation générale

$$F \approx -\sqrt{\frac{2\pi^6}{27\mathrm{vol}(X_7)}}\,N^{3/2}, \qquad N \to \infty \tag{8}$$

prédite par la correspondance (1) dans le cas de la théorie ABJM avec des saveurs.

Dans la section 2.12, nous utilisons les équations d'anomalie holomorphe pour trouver les expressions explicites de $F_g(\lambda)$ pour grand genre g dans le cas de la théorie ABJM pure. Dans la section 2.13, nous analysons le comportement de $F_g(\lambda)$ à grande genre et on trouve

$$F_g(\lambda) \sim \Gamma(2g-1)\,(A(\lambda))^{-2g}, \qquad g \gg 1. \tag{9}$$

d'où on déduit la forme des instantons de grand N, les contributions non-perturbatives à l'énergie libre qui ne peuvent pas être vues directement dans la série asymptotique (5):

$$\sim \mathrm{e}^{-\frac{A(\lambda)}{g_s}}. \tag{10}$$

Nous montrons que $A(\lambda)$, l'action de l'instanton, est déterminée par une période de ω sur la courbe spectrale. Dans la dernière section, nous montrons que, dans la limite de couplage fort l'action coïncide avec l'action de D2-brane enveloppant $\mathbb{RP}^3$ dans $\mathbb{CP}^3$ dans le double de type IIA fournissant une autre vérification non triviale de la correspondance AdS/CFT. Le deuxième chapitre est basé sur les papiers originaux [5, 6, 7, 8].

Dans le troisième chapitre, nous développons une approche complètement différente pour étudier les modèles de matrice CSM. Nous constatons que la fonction de partition du modèle de matrice d'une théorie $\mathcal{N} \geq 3$ peut être interprétée comme la fonction de partition d'un gaz parfait de Fermi avec une certaine hamiltonien à une particule à une dimension. En utilisant cette approche, nous étudions la fonction de partition dans la limite

$$N \to \infty, \qquad \text{les niveaux de CS sont fixés.} \tag{11}$$

Cette limite peut être dénommée la limite de théorie M, tandis que (4) peut être considéré comme la limite de type IIA. Du point de vue de gaz de Fermi, la limite (11) est la limite thermodynamique ordinaire qui est déterminé par le comportement asymptotique des valeurs propres de l'hamiltonien. Cela permet une dérivation élémentaire de la loi $N^{3/2}$ (8) dans le cadre standard de la mécanique statistique. Il s'avère que à l'aide de l'analyse semi-classique du spectre il est possible de trouver une expression fermée de la partie perturbative complète de la fonction de partition:

$$Z(N) \propto \mathrm{Ai}\left[C^{-1/3}(N-B)\right] \cdot \left(1 + \mathcal{O}(\mathrm{e}^{-cN})\right), \qquad N \to \infty \tag{12}$$

où les constantes C et B varient avec la théorie CSM et dépendent des niveaux de CS. Nous étudions également la structure des contributions non perturbatives en N dans la description de gaz de Fermi et les interprétons comme les instantons de membrane sur le côté de la théorie des cordes. Le troisième chapitre est basé sur le papier original [9].

Contents

Chapter 0

Introduction and summary

0.1 Motivation

Understanding the nonperturbative dynamics of gauge theories is one of the most challenging problems of theoretical physics. As usual in physics, there are essentially two ways: either one can try to study a real-life theory by some approximate methods (e.g. considering QCD on a lattice), or one can study a different theory (e.g. supersymmetric gauge theory) which has better properties, more controllable from mathematical point of view, and can be solved exactly to some extent but its non-trivial part of the dynamics is believed to be the same as in the real-life theory. Such simplified theories can also be considered as toy models where one can develop some new techniques which could be useful later for solving real-life theories.

A lot of progress has been made over the last 20 years in studying supersymmetric gauge theories. Ones of the most important breakthroughs were the solution of $\mathcal{N}=2$ theories by Seiberg and Witten [10] (and the subsequent direct calculation by Nekrasov [11]) and the discovery of the AdS/CFT correspondence by Maldacena [12]. The AdS/CFT correspondence is a conjecture that states that observables in d-dimensional supersymmetric gauge theories are related to certain observables in string theory with background geometry involving AdS_{d+1} space so that the space–time of the gauge theory is associated to the boundary of AdS_{d+1}. From the string theory point of view the $U(N)$ gauge theory can be understood as a world-volume theory of N coincident d-dimensional branes. When N is large the geometry of the ambient string theory target space backreacts and deforms into AdS_{d+1}. The direct verification of the AdS/CFT correspondence is hard since the string theory side is well understood only in the supergravity limit which corresponds to the strong coupling limit of gauge theory where one cannot use the standard perturbation theory. For example the AdS/CFT correspondence implies that the strong coupling limit of a Wilson loop observable in gauge theory is related to a string whose world-sheet boundary lies in the boundary of AdS_{d+1} and coincides with the loop. In particular this implies automatically the famous area-law behavior for the Wilson loop which is equivalent to the statement that there is a linear confining potential between probe quarks.

An important result concerning Wilson loop in supersymmetric gauge theories was made by Pestun [13] who showed that the vacuum expectation value of a certain Wilson loops in $\mathcal{N}=4$ and $\mathcal{N}=2$ 4-dimensional super Yang–Mills (SYM) theories can be reduced to a matrix integral, and thus can be calculated exactly, and the strong coupling limit is in agreement with the AdS/CFT prediction.

The AdS/CFT correspondence was mostly studied in its original formulation which states that $N=4$ SYM theory is dual to type IIB string theory in $\mathrm{AdS}_5\times\mathbb{S}^5$. For a long time it was a problem to construct a proper 3-dimensional gauge theory that would have enough amount of supersymmetry and participate in $\mathrm{AdS}_4/\mathrm{CFT}_3$ duality. A breakthrough was made by Aharony, Bergman, Jafferis and Maldacena [1] who found a family of theories labeled by $k\in\mathbb{Z}$ and which were conjectured to be the world-volume theories for a stack of N M2 branes probing $\mathbb{C}^4/\mathbb{Z}_k$ and thus dual to M-theory in $\mathrm{AdS}_4\times\mathbb{S}^7/\mathbb{Z}_k$ or, for large k, to type IIA string theory in $\mathrm{AdS}_4\times\mathbb{CP}^3$. A theory of this family is commonly referred to as ABJM theory. For the case $N=2$ it is equivalent to the previously constructed BLG theory [14, 15]. It was later generalized to other examples of so-called Chern–Simons–matter (CSM) theories (described in detail in the next chapter) which were conjectured to describe N M2-branes probing a Calabi-Yau 4-fold Y_8. For large N the geometry backreacts into $\mathrm{AdS}_4\times X_7$ so that Y_8 is a cone over X_7 which is then a tri-Sasaki-Einstein manifold. Thus one can state the correspondence

$$\text{CSM theory}\longleftrightarrow\text{M-theory in }\mathrm{AdS}_4\times X_7\qquad(0.1.1)$$

Apart from a manifold X_7 one also needs to fix a 3-form torsion flux, an element of $H_3(X_7,\mathbb{Z})$, to define completely the background of M-theory. The data needed to define a CSM theory is a gauge group G and its representation R for matter fields. One also needs to fix the CS levels, an integer for each simple component of G.

An important progress was made by Kapustin, Willett and Yaakov in [2]. Similarly to localization for 4 dimension theories considered by Pestun [13], one can perform localization in the path integral for Chern–Simons–matter and reduce it to a matrix integral. For a general supersymmetric $\mathcal{N} \geq 3$ Chern–Simons–theory the matrix integral reads as follows in condensed notations:

$$Z \propto \int \mathrm{d}a \mathrm{e}^{\frac{\mathrm{i}}{4\pi}\mathrm{Tr}a^2} \frac{\det_{\mathrm{Ad}} 2\sinh\frac{a}{2}}{\det_R 2\cosh\frac{a}{2}} \tag{0.1.2}$$

where the integration is performed over the Cartan subalgebra of the Lie algebra $\mathfrak{g}$ of the gauge group G, Tr defines some invariant inner product on $\mathfrak{g}$ which depends on the choice of CS levels, and

$$\det_R f(a) \equiv \prod_\rho f(\rho(a)) \tag{0.1.3}$$

where the product runs over the weights of the representation R (in the case of the adjoint representation ($R = \mathrm{Ad}$) we exclude zero weights). Note that localization to a matrix model can be performed for a general $\mathcal{N} = 2$ CSM theory, however such theory does not always have an M-theory dual of type $\mathrm{AdS}_4 \times X_7$.

The main purpose of this thesis is detailed analysis of matrix models of type (0.1.3) and application of the obtained results to explicit non-trivial verification of the $\mathrm{AdS}_4/\mathrm{CFT}_3$ correspondence.

0.2 Organization of the thesis and summary of the results

In the first chapter we give a review of background knowledge that is used in the main part of the thesis. Namely, we consider Chern–Simons–matter theories, their string theory duals and localization. This chapter follows the lectures [16][1].

In the second chapter we perform detailed analysis of CSM matrix models in the so-called 't Hooft limit. In the case of ABJM matrix model it is defined as follows:

$$N \to \infty, \qquad \lambda \equiv \frac{N}{k} \text{ is fixed.} \tag{0.2.1}$$

Then the observables have asymptotic series in $g_s \equiv 2\pi i/k$ which are called genus, or topological, expansions. For example, for the free energy one has

$$F \equiv \log Z = \sum_{g\geq 0} g_s^{2g-2} F_g(\lambda). \tag{0.2.2}$$

There is a universal approach to study such expansions in matrix models based on the notion of a spectral curve. In this approach solution to a matrix model is determined completely by a complex curve $\mathcal{C}$, called a spectral curve, a meromorphic 1-form ω on this curve (which for a special choice of variables reads as $\omega = y\mathrm{d}x$ while the curve is determined by $P(y,x) = 0$) and also a choice of a basis in $H_1(\mathcal{C},\mathbb{Z})$. First we establish the relation between the ABJM matrix model and the matrix model for ordinary Chern–Simons theory on the lens space $L(2,1) \cong \mathbb{RP}^3$:

$$Z_{\mathrm{ABJM}}(N_1,N_2,g_s) = Z_{L(2,1)}(N_1,-N_2,g_s) \tag{0.2.3}$$

where we consider the generalization [3] of the original ABJM theory to the case with the gauge group $U(N_1) \times U(N_2)$. Using this relation and the known solution of the lens space matrix model we build explicitly the solution of the ABJM matrix model in terms of the spectral curve. Using further the fact that Chern–Simons theory on $L(2,1)$ is large-N dual to the topological strings on local $\mathbb{P}^1 \times \mathbb{P}^1$ we conclude that the real 2-dimensional moduli space of the ABJM theory is a certain real subspace of the complex 2-dimensional moduli space of complexified Kähler structures on $\mathbb{P}^1 \times \mathbb{P}^1$. In particular, the weak coupling limit of ABJM theory corresponds to the orbifold point of local $\mathbb{P}^1 \times \mathbb{P}^1$ while the strong coupling limit corresponds to the large radius limit.

[1]The author of this thesis is grateful to his supervisor, Professor Mariño, for allowing to use the notes of his lectures for the introductory chapter.

In the case of ABJM theory we found exact explicit expressions for the free energy and Wilson loop for the first few genera. They can be considered as exact interpolation functions of λ between the weak ($\lambda \to 0$) and strong ($\lambda \to \infty$) expansions. The weak expansion can be checked to be in agreement with the perturbation theory calculations in the quantum field theory. And we verify that the strong coupling expansion is in agreement with the prediction by the AdS/CFT correspondence from the string theory side. In particular we find that in the leading order

$$F_0 \approx \frac{4\pi^3\sqrt{2}}{3}\lambda^{3/2}, \qquad \lambda \to \infty. \tag{0.2.4}$$

which provides a derivation of the mysterious $N^{3/2}$ scaling of the free energy predicted from the string theory [4]. Furthermore we show that the coefficient coincides with the one that can be obtained from the supergravity calculation. For 1/2 and 1/6 BPS Wilson loops we find

$$\langle W_\square \rangle \sim \mathrm{e}^{\pi\sqrt{2\lambda}}, \qquad \lambda \to \infty. \tag{0.2.5}$$

which is in agreement with the area law behavior predicted by AdS/CFT. We also study in detail the subleading corrections and find that they are also in agreement with the string theory side of the duality. In particular, non-perturbative corrections can be interpreted as contributions from world-sheet instantons.

In sections 2.8–2.11 we consider a generalization of the ABJM theory with additional matter multiplets in the (anti)fundamental representation [17]. Although it is again possible to find the spectral curve of the corresponding matrix model exactly, its explicit expression is no longer algebraic and contains special functions. Therefore finding the exact explicit expressions for the free energy and the Wilson loops becomes extremely hard. To avoid this problem we develop a geometrical method allowing direct calculation of the strong coupling limit without obtaining first the exact expressions. The key feature of this method is that the strong coupling limit is related to the tropical, or ultradiscretization, limit of the spectral curve. In this limit the original complex curve in $\mathbb{C}^* \times \mathbb{C}^*$ can be described by a piece-wise linear graph in $\mathbb{R}^2$. The leading order of matrix model observables at the strong coupling can be calculated in terms of this graph. Using this method we verify the general relation

$$F \approx -\sqrt{\frac{2\pi^6}{27\mathrm{vol}(X_7)}}\,N^{3/2}, \qquad N \to \infty \tag{0.2.6}$$

implied by the correspondence (0.1.1) in the case of flavored ABJM theory. For the theory with N_f flavors the dual space is the so-called Eschenburg space that can be constructed as the hyper-Kähler quotient of $\mathbb{H}^3 \cong \mathbb{C}^{12}$ by the action of $U(1)$ with the weights (N_f, N_f, k). Its volume was computed in [18]:

$$\mathrm{vol}(X_7) = \frac{k + N_f/2}{(k+N_f)^2}\,\mathrm{vol}(\mathbb{S}^7). \tag{0.2.7}$$

We also calculate the strong coupling limit of the 1/2 and 1/6 BPS Wilson loops:

$$\langle W_\square \rangle \sim \exp\left(\frac{2\pi N^{1/2}}{\sqrt{N_f + 2k}}\right) \tag{0.2.8}$$

and argue that it agrees with the prediction from the string theory side.

In section 2.12 we use holomorphic anomaly equations to find explicit expressions of $F_g(\lambda)$ for large genus g in the case of pure ABJM theory. In section 2.13 we analyze the large genus behavior of $F_g(\lambda)$ and we find

$$F_g(\lambda) \sim \Gamma(2g-1)\,(A(\lambda))^{-2g}, \qquad g \gg 1. \tag{0.2.9}$$

from which we deduce the form of large N instantons, non-perturbative contributions to the free energy that cannot be seen directly in the asymptotic series (0.2.2):

$$\sim \mathrm{e}^{-\frac{A(\lambda)}{g_s}}. \tag{0.2.10}$$

We show that $A(\lambda)$, the instanton action, is determined by a period on the spectral curve:

$$A(\lambda) = \frac{1}{2}\int_\gamma \omega \tag{0.2.11}$$

where the choice of the contour γ varies with the region of the moduli space. In the last section we show that in the strong coupling limit

$$-\mathrm{i}A(\lambda) \approx 2\pi^2\sqrt{2\lambda}, \qquad \lambda \to \infty \tag{0.2.12}$$

the instanton action coincides with the action of D2-brane wrapping $\mathbb{RP}^3$ inside $\mathbb{CP}^3$ in the type IIA dual thus providing another nontrivial check of the AdS/CFT correspondence. Conceptually, this approach to study nonperturbative effects is similar to what had been pursued in the case of noncritical strings [19, 20, 21, 22]. The second chapter is based on the original papers [5, 6, 7, 8].

In the third chapter we develop a completely different approach to study CSM matrix models. We find that the partition function of the matrix model of a $\mathcal{N} \geq 3$ theory can be interpreted as the partition function of an ideal Fermi gas with a certain 1-particle Hamiltonian in one dimension. This defines the correspondence

$$\text{CSM theory} \longleftrightarrow \text{1-dim Hamiltonian } H \tag{0.2.13}$$

Using this approach we study the partition function in the limit

$$N \to \infty, \qquad \text{CS levels fixed.} \tag{0.2.14}$$

In this limit one can verify AdS/CFT directly in the M-theory picture and not in the type IIA picture as one should do for the limit (0.2.1) as the latter assumes the expansion w.r.t. the small string coupling g_s. The limit (0.2.14) can be referred to as the M-theory limit while (0.2.1) can be considered as the type IIA limit. From the Fermi gas point of view the limit (0.2.14) is the ordinary thermodynamical limit which is determined by the asymptotics of the eigenvalues of the Hamiltonian. This allows an elementary derivation of the $N^{3/2}$ scaling (0.2.6) in the standard framework of statistical mechanics. It turns out that using the semi-classical analysis of the spectrum it is possible to find a closed expression of the full perturbative part of the partition function:

$$Z(N) \propto \mathrm{Ai}\left[C^{-1/3}(N-B)\right] \cdot \left(1 + \mathcal{O}(\mathrm{e}^{-cN})\right), \qquad N \to \infty \tag{0.2.15}$$

where the constants C and B vary with CSM theory and depend on CS levels. This generalizes the result of [23] obtained for the case of ABJM theory. We also study the structure of contributions non-perturbative in N in the Fermi gas picture and interpret them as membrane instantons on the string theory side. The third chapter is based on the original paper [9].

0.3 Current state of the field and open problems

In this section we point out advantages and disadvantages of three known approaches (two of which are considered in detail in this thesis) to study CSM matrix models and indicate some open problems.

1) In the first approach one considers the 't Hooft limit and uses the standard tools of matrix models. This is the approach described and used in the second chapter of this thesis. The main advantage of this approach is that, once the spectral curve is found, it is in principle possible to find the free energy and the Wilson loops exactly in terms of integrals over cycles on the spectral curve. This approach works very well in the case of ABJM theory but it is difficult to find explicit expression for the spectral curve for other CSM theories. However, when one is interested in the strong coupling limit, as we show in the second chapter on the example of the flavored ABJM theory, it is sufficient to find only the tropical version of the spectral curve.

It would be interesting to have a general scheme of studying CSM theories in terms of tropical geometry. There also arises an interesting mathematical problem of constructing a "tropical" version of topological recursion of Eynard and Orantin [24].

Since our description of matrix model instantons in section 2.13 is made in the language of special geometry, it should determine the large order behavior of the genus g amplitudes in general topological string models, not necessarily encoded in matrix integrals. It would be interesting to study simple models with a spectral curve description, like topological string theory on local $\mathbb{P}^2$, in order to test the method and learn about possible non-perturbative structures in topological strings. The general picture we have developed might shed further light in related contexts, like the models studied in [25].

It is puzzling that the connection between ABJM theory and topological strings discussed in the second chapter seems to be accidental and lack a generalization for other CSM theories.

2) The second approach is the Fermi gas approach described in the third chapter of the thesis. In this approach one considers CSM theory in the M-theory limit and it allows the systematic study of $1/N$

corrections. However in the case of $\mathcal{N}=2$ CSM theories the corresponding Fermi gas is no longer free and it becomes very hard to go beyond the leading order of the strong coupling limit. It is also unclear whether it is possible to generalize this approach to the theories defined by D, E quivers considered in [26] and quivers with the gauge groups of different ranks in the nodes.

It would be interesting to understand the Fermi gas picture at a deeper level and find out if it has some direct interpretation in the string theory picture. Since D-branes behave as fermions (see for example [27]), and the gauge theories we have considered have D-brane realizations, one might be able to derive this picture directly from the D-branes underlying the gauge theory.

The exact perturbative partition function (0.2.15) found with the Fermi gas approach should provide a lot of information about M-theory dual. It is an interesting problem to check that the subleading $1/N$ terms are in agreement with M-theory.

Fermi gas formalism seems appropriate to study the strong-coupling regime of matrix models and topological strings. For example, as shown in the second chapter, the ABJM matrix model corresponds to a submanifold of the moduli space of topological strings on local $\mathbb{P}^1 \times \mathbb{P}^1$. However, by looking at for example (3.4.80), it is clear that the grand canonical potential of ABJM theory seems to be directly related to the topological string free energy in the large radius frame. Therefore, our calculation of $J(\mu)$ for the ABJM theory can be interpreted as a concrete strong coupling expansion of this topological string free energy, including non-perturbative effects. The worldsheet instantons of the topological string at large radius would appear then as quantum-mechanical instantons of the Fermi gas. Notice also that the grand canonical partition function, which is the focus of this paper, involves the sum over fluxes first considered in the context of topological strings in [27], and studied from the matrix model point of view in [28, 29, 30]. Our formalism gives a concrete approach to calculate this object at strong coupling, but one should clarify the relation between the picture proposed here and the non-perturbative approach of [28, 29] involving theta functions on the spectral curve. It would be very interesting to develop further all these relationships to topological string theory.

Although the equivalence between the matrix models and the Fermi gas partition function is an exact statement at finite N and k, in this paper we have worked in the thermodynamic limit (large N) and in the semiclassical limit (expansion in powers of k around $k=0$). Fortunately, as we have seen, the expansion in $1/N$ satisfies some sort of "non-renormalization" property and we can determine it at finite k by a next-to-leading computation in the WKB expansion. An obvious challenge is to solve the Fermi gas problem at finite k, say $k=1$, to make full contact with the M-theory expansion. This would amount to a resummation of the non-perturbative effects computed in this paper order by order in k. One possible route to achieve this is to find the exact eigenvalue spectrum of the quantum Hamiltonian, or equivalently, of the density matrix $\hat{\rho}$. In the case of ABJM theory, this means solving the integral equation (3.2.19), which can be written as

$$\int \mathrm{d}x' \rho(x, x') \phi_n(x') = \lambda_n \phi_n(x), \tag{0.3.1}$$

where the kernel can be written in the form

$$\rho(x, x') = \frac{\mathrm{e}^{-\frac{1}{2}U(x)-\frac{1}{2}U('x)}}{4\pi k \cosh\left(\frac{x-x'}{2k}\right)}. \tag{0.3.2}$$

This type of kernel appears in other contexts, like the $O(2)$ matrix model [31] (albeit with a different $U(x)$), and it is connected to both the Hirota hierarchy [31, 32] and to the Thermodynamic Bethe Ansatz [33, 34]. At least for $k=1,2$ (where supersymmetry is enhanced), we anticipate a nice solution to the eigenvalue problem in terms of an integrable system. In particular, the relation to differentiable hierarchies of the Hirota type suggests that the propery 0.2.15 might be proved by performing a suitable double-scaling limit in the hierarchy.

It would be also very interesting to develop further the relationship between worldsheet instantons and quantum mechanical instantons of the Fermi gas sketched in section 3.4.4. It is clear that the solution of the ABJM theory found in the second chapter, in the 't Hooft expansion, is extremely powerful in order to capture these corrections, but it would be important for the development of the Fermi gas approach to have a better understanding of this issue.

Finally, there might be a connection between the Fermi gas of this paper and two other pictures for string/M-theory based on fermions: the droplet picture proposed in [35, 36] to analyze 1/2 BPS operators in $\mathcal{N}=4$ SYM theory, and the Fermi liquid picture of non-critical M-theory proposed in [37].

3) There is also another approach which is not used in this thesis and was first developed in [38] and was later successfully applied for a very large class of $\mathcal{N} \geq 2$ CSM theories [26, 38, 39, 40, 41, 42].

This approach considers CSM matrix model directly in the M-theory limit. The leading term of the free energy is determined in terms of variational principle w.r.t. the limiting distribution of eigenvalues:

$$F(N) = -N^{3/2} \min_{\rho, y_a} \mathcal{F}[\rho, y_a] \tag{0.3.3}$$

The disadvantage of this approach is that it is very hard to study $1/N$ corrections and go beyond the leading term in the strong coupling limit.

0.4 Acknowledgments

This thesis would not have been possible without the guidance and the help of my supervisor Professor Mariño, to whom I am sincerely grateful. I would also like to thank people with whom I collaborated while being a PhD student in University of Geneva: Nadav Drukker, Sara Pasquetti, Ricardo Couso Santamaria. Besides my thank extends to all other people with whom I had interesting and fruitful discussions.

Chapter 1

Chern–Simons–matter theories and their gravity duals

The purpose of this chapter is to provide some background knowledge about Chern–Simons–matter theories underlying the calculations performed in the rest of the thesis. This chapter is heavily based on the lectures [16]. The organization of the chapter is as follows. Section 1.1 introduces Euclidean supersymmetric Chern–Simons–matter theories on the three-sphere and their classical properties. In section 1.2 we analyze perturbative Chern–Simons theory in some detail and in some generality. In section 1.3 we calculate the free energy of Chern–Simons–matter theories on $\mathbb{S}^3$ at one-loop. Next, in section 1.4, we look at the free energy at strong coupling by using the AdS dual, and we explain the basics of holographic renormalization of the gravitational action. In section 1.5 we review the localization computation of [2] (incorporating some simplifications in [43]), which leads to a matrix model formulation of the free energy of ABJM theory.

1.1 Supersymmetric Chern–Simons–matter theories

In this section we will introduce the basic building blocks of supersymmetric Chern–Simons–matter theories. We will work in Euclidean space, and we will put the theories on the three-sphere, since we are eventually interested in computing the free energy of the gauge theory in this curved space. In this section we will closely follow the presentation of [43].

1.1.1 Conventions

Our conventions for Euclidean spinors follow essentially [44]. In Euclidean space, the fermions ψ and $\bar\psi$ are independent and they transform in the same representation of the Lorentz group. Their index structure is

$$\psi^\alpha, \quad \bar\psi^\alpha. \tag{1.1.1}$$

We will take γ_μ to be the Pauli matrices, which are hermitian, and

$$\gamma_{\mu\nu} = \frac{1}{2}[\gamma_\mu, \gamma_\nu] = \mathrm{i}\epsilon_{\mu\nu\rho}\gamma^\rho. \tag{1.1.2}$$

We introduce the usual symplectic product through the antisymmetric matrix

$$C_{\alpha\beta} = \begin{pmatrix} 0 & C \\ -C & 0 \end{pmatrix}. \tag{1.1.3}$$

In [44] we have $C = -1$ and the matrix is denoted by $\epsilon_{\alpha\beta}$. The product is

$$\bar\epsilon\lambda = \bar\epsilon^\alpha C_{\alpha\beta}\lambda^\beta. \tag{1.1.4}$$

Notice that

$$\bar\epsilon\gamma^\mu\lambda = \bar\epsilon^\beta C_{\beta\gamma}\left(\gamma^\mu\right)^\gamma{}_\alpha\,\lambda^\alpha. \tag{1.1.5}$$

It is easy to check that

$$\bar\epsilon\lambda = \lambda\bar\epsilon, \quad \bar\epsilon\gamma^\mu\lambda = -\lambda\gamma^\mu\bar\epsilon, \tag{1.1.6}$$

and in particular

$$(\gamma^\mu \bar{\epsilon})\, \lambda = -\bar{\epsilon}\gamma^\mu \lambda. \tag{1.1.7}$$

We also have the following Fierz identities

$$\bar{\epsilon}\,(\epsilon\psi) + \epsilon\,(\bar{\epsilon}\psi) + (\bar{\epsilon}\epsilon)\,\psi = 0 \tag{1.1.8}$$

and

$$\epsilon\,(\bar{\epsilon}\psi) + 2\,(\bar{\epsilon}\epsilon)\,\psi + (\bar{\epsilon}\gamma_\mu\psi)\,\gamma^\mu\epsilon = 0. \tag{1.1.9}$$

1.1.2 Vector multiplet and supersymmetric Chern–Simons theory

We first start with theories based on vector multiplets. The three dimensional Euclidean $\mathcal{N}=2$ vector superfield V has the following content

$$V: \qquad A_\mu,\ \sigma,\ \lambda,\ \bar{\lambda},\ D, \tag{1.1.10}$$

where A_μ is a gauge field, σ is an auxiliary scalar field, λ, $\bar{\lambda}$ are two-component complex Dirac spinors, and D is an auxiliary scalar. This is just the dimensional reduction of the $\mathcal{N}=1$ vector multiplet in 4 dimensions, and σ is the reduction of the fourth component of A_μ. All fields are valued in the Lie algebra $\mathfrak{g}$ of the gauge group G. For $G=U(N)$ our convention is that $\mathfrak{g}$ are Hermitian matrices. It follows that the gauge covariant derivative is given by

$$\partial_\mu + \mathrm{i}[A_\mu, .] \tag{1.1.11}$$

while the gauge field strength is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + \mathrm{i}[A_\mu, A_\nu]. \tag{1.1.12}$$

The transformations of the fields are generated by two independent complex spinors ϵ, $\bar{\epsilon}$. They are given by,

$$\begin{aligned}
\delta A_\mu &= \frac{\mathrm{i}}{2}(\bar{\epsilon}\gamma_\mu\lambda - \bar{\lambda}\gamma_\mu\epsilon),\\
\delta\sigma &= \frac{1}{2}(\bar{\epsilon}\lambda - \bar{\lambda}\epsilon),\\
\delta\lambda &= -\frac{1}{2}\gamma^{\mu\nu}\epsilon F_{\mu\nu} - D\epsilon + \mathrm{i}\gamma^\mu\epsilon D_\mu\sigma + \frac{2\mathrm{i}}{3}\sigma\gamma^\mu D_\mu\epsilon,\\
\delta\bar{\lambda} &= -\frac{1}{2}\gamma^{\mu\nu}\bar{\epsilon} F_{\mu\nu} + D\bar{\epsilon} - \mathrm{i}\gamma^\mu\bar{\epsilon} D_\mu\sigma - \frac{2\mathrm{i}}{3}\sigma\gamma^\mu D_\mu\bar{\epsilon},\\
\delta D &= -\frac{\mathrm{i}}{2}\bar{\epsilon}\gamma^\mu D_\mu\lambda - \frac{\mathrm{i}}{2}D_\mu\bar{\lambda}\gamma^\mu\epsilon + \frac{\mathrm{i}}{2}[\bar{\epsilon}\lambda, \sigma] + \frac{\mathrm{i}}{2}[\bar{\lambda}\epsilon, \sigma] - \frac{\mathrm{i}}{6}(D_\mu\bar{\epsilon}\gamma^\mu\lambda + \bar{\lambda}\gamma^\mu D_\mu\epsilon),
\end{aligned} \tag{1.1.13}$$

and we split naturally

$$\delta = \delta_\epsilon + \delta_{\bar{\epsilon}}. \tag{1.1.14}$$

Here we follow the conventions of [43], but we change the sign of the gauge connection: $A_\mu \to -A_\mu$. The derivative D_μ is covariant with respect to both the gauge field and the spin connection. On all the fields, except D, the commutator $[\delta_\epsilon, \delta_{\bar{\epsilon}}]$ becomes a sum of translation, gauge transformation, Lorentz rotation, dilation and R-rotation:

$$\begin{aligned}
[\delta_\epsilon, \delta_{\bar{\epsilon}}]A_\mu =& \mathrm{i}v^\nu\partial_\nu A_\mu + \mathrm{i}\partial_\mu v^\nu A_\nu - D_\mu\Lambda,\\
[\delta_\epsilon, \delta_{\bar{\epsilon}}]\sigma =& \mathrm{i}v^\mu\partial_\mu\sigma + \mathrm{i}[\Lambda, \sigma] + \rho\sigma,\\
[\delta_\epsilon, \delta_{\bar{\epsilon}}]\lambda =& \mathrm{i}v^\mu\partial_\mu\lambda + \frac{\mathrm{i}}{4}\Theta_{\mu\nu}\gamma^{\mu\nu}\lambda + \mathrm{i}[\Lambda, \lambda] + \frac{3}{2}\rho\lambda + \alpha\lambda,\\
[\delta_\epsilon, \delta_{\bar{\epsilon}}]\bar{\lambda} =& \mathrm{i}v^\mu\partial_\mu\bar{\lambda} + \frac{\mathrm{i}}{4}\Theta_{\mu\nu}\gamma^{\mu\nu}\bar{\lambda} + \mathrm{i}[\Lambda, \bar{\lambda}] + \frac{3}{2}\rho\bar{\lambda} - \alpha\bar{\lambda},\\
[\delta_\epsilon, \delta_{\bar{\epsilon}}]D =& \mathrm{i}v^\mu\partial_\mu D + \mathrm{i}[\Lambda, D] + 2\rho D + \frac{1}{3}\sigma(\bar{\epsilon}\gamma^\mu\gamma^\nu D_\mu D_\nu\epsilon - \epsilon\gamma^\mu\gamma^\nu D_\mu D_\nu\bar{\epsilon}),
\end{aligned} \tag{1.1.15}$$

where

$$\begin{aligned} v^\mu &= \bar\epsilon\gamma^\mu\epsilon, \\ \Theta^{\mu\nu} &= D^{[\mu}v^{\nu]} + v^\lambda\omega_\lambda^{\mu\nu}, \\ \Lambda &= v^\mu \mathrm{i} A_\mu + \sigma\bar\epsilon\epsilon, \\ \rho &= \tfrac{\mathrm{i}}{3}(\bar\epsilon\gamma^\mu D_\mu\epsilon + D_\mu\bar\epsilon\gamma^\mu\epsilon), \\ \alpha &= \tfrac{\mathrm{i}}{3}(D_\mu\bar\epsilon\gamma^\mu\epsilon - \bar\epsilon\gamma^\mu D_\mu\epsilon). \end{aligned} \tag{1.1.16}$$

Here, $\omega_\lambda^{\mu\nu}$ is the spin connection. As a check, let us calculate the commutator acting on σ. We have,

$$\begin{aligned} [\delta_\epsilon, \delta_{\bar\epsilon}]\sigma &= \delta_\epsilon\left(\frac{1}{2}\bar\epsilon\lambda\right) - \delta_{\bar\epsilon}\left(-\frac{1}{2}\bar\lambda\epsilon\right) \\ &= \frac{1}{2}\bar\epsilon\left(-\frac{1}{2}\gamma^{\mu\nu}\epsilon F_{\mu\nu} - D\epsilon + \mathrm{i}\gamma^\mu\epsilon D_\mu\sigma\right) + \frac{\mathrm{i}}{3}\bar\epsilon\gamma^\mu\left(D_\mu\epsilon\right)\sigma \\ &+ \frac{1}{2}\left(-\frac{1}{2}\gamma^{\mu\nu}\bar\epsilon F_{\mu\nu} + D\epsilon - \mathrm{i}\gamma^\mu\bar\epsilon D_\mu\sigma\right)\epsilon - \frac{\mathrm{i}}{3}\gamma^\mu\left(D_\mu\bar\epsilon\right)\epsilon\sigma \\ &= \mathrm{i}\bar\epsilon\gamma^\mu\epsilon D_\mu\sigma + \rho\sigma, \end{aligned} \tag{1.1.17}$$

where we have used (1.1.7).

In order for the supersymmetry algebra to close, the last term in the right hand side of $[\delta_\epsilon, \delta_{\bar\epsilon}]D$ must vanish. This is the case if the Killing spinors satisfy

$$\gamma^\mu\gamma^\nu D_\mu D_\nu\epsilon = h\epsilon, \qquad \gamma^\mu\gamma^\nu D_\mu D_\nu\bar\epsilon = h\bar\epsilon \tag{1.1.18}$$

for some scalar function h. A sufficient condition for this is to have

$$D_\mu\epsilon = \frac{\mathrm{i}}{2r}\gamma_\mu\epsilon, \qquad D_\mu\bar\epsilon = \frac{\mathrm{i}}{2r}\gamma_\mu\bar\epsilon \tag{1.1.19}$$

and

$$h = -\frac{9}{4r^2} \tag{1.1.20}$$

where r is the radius of the three-sphere. This condition is satisfied by one of the Killing spinors on the three-sphere (the one which is constant in the left-invariant frame). Notice that, with this choice, ρ in (1.1.16) vanishes.

The (Euclidean) SUSY Chern–Simons (CS) action, in flat space, is given by

$$\begin{aligned} S_{\mathrm{SCS}} &= -\int \mathrm{d}^3x\,\mathrm{Tr}\left(A\wedge\mathrm{d}A + \frac{2\mathrm{i}}{3}A^3 - \bar\lambda\lambda + 2D\sigma\right) \\ &= -\int \mathrm{d}^3x\,\mathrm{Tr}\left(\epsilon^{\mu\nu\rho}\left(A_\mu\partial_\nu A_\rho + \frac{2\mathrm{i}}{3}A_\mu A_\nu A_\rho\right) - \bar\lambda\lambda + 2D\sigma\right). \end{aligned} \tag{1.1.21}$$

Here Tr denotes the trace in the fundamental representation. The part of the action involving the gauge connection A is the standard, bosonic CS action in three dimensions. This action was first considered from the point of view of QFT in [45], where the total action for a non-abelian gauge field was the sum of the standard Yang–Mills action and the CS action. In [46], the CS action was considered by itself and shown to lead to a topological gauge theory.

We can check that the supersymmetric CS action is invariant under the supersymmetry generated by δ_ϵ (the proof for $\delta_{\bar\epsilon}$ is similar).The supersymmetric variation of the integrand of (1.1.21) is

$$\begin{aligned} &\left(2\delta A_\mu\partial_\nu A_\rho + 2\mathrm{i}\delta A_\mu A_\nu A_\rho\right)\epsilon^{\mu\nu\rho} - \bar\lambda\delta\lambda + 2(\delta D)\sigma + 2D\delta\sigma \\ &= -\mathrm{i}\bar\lambda\gamma_\mu\epsilon\partial_\nu A_\rho\epsilon^{\mu\nu\rho} + \bar\lambda\gamma_\mu\epsilon A_\nu A_\rho\epsilon^{\mu\nu\rho} \\ &- \bar\lambda\left(-\frac{1}{2}\gamma^{\mu\nu}F_{\mu\nu} - D + \mathrm{i}\gamma^\mu D_\mu\sigma\right)\epsilon - \frac{2\mathrm{i}}{3}\bar\lambda\gamma^\mu D_\mu\epsilon\sigma \\ &- \mathrm{i}\left(D_\mu\bar\lambda\right)\gamma^\mu\sigma\epsilon + \mathrm{i}[\bar\lambda\epsilon, \sigma]\sigma - \frac{\mathrm{i}}{3}\bar\lambda\gamma^\mu D_\mu\epsilon\sigma - \bar\lambda\epsilon D. \end{aligned} \tag{1.1.22}$$

The terms involving D cancel on the nose. Let us consider the terms involving the gauge field. After using (1.1.2) we find

$$\frac{1}{2}\bar\lambda\gamma^{\mu\nu}F_{\mu\nu}\epsilon = \mathrm{i}\bar\lambda\gamma_\rho\epsilon\epsilon^{\mu\nu\rho}\partial_\mu A_\nu - \bar\lambda\gamma_\rho\epsilon\epsilon^{\mu\nu\rho}A_\mu A_\nu \tag{1.1.23}$$

which cancel the first two terms in (1.1.22). Let us now look at the remaining terms. The covariant derivative of $\bar\lambda$ is

$$D_\mu\bar\lambda = \partial_\mu\bar\lambda + \frac{\mathrm{i}}{2r}\gamma_\mu\bar\lambda + \mathrm{i}[A_\mu,\bar\lambda]. \tag{1.1.24}$$

If we integrate by parts the term involving the derivative of λ we find in total

$$\begin{aligned}&\mathrm{i}\bar\lambda\gamma^\mu\epsilon\partial_\mu\sigma + \mathrm{i}\bar\lambda\gamma^\mu\partial_\mu\epsilon\sigma + \frac{1}{2r}\left(\gamma^\mu\bar\lambda\right)\gamma_\mu\epsilon + [A_\mu,\bar\lambda]\gamma^\mu\epsilon\sigma\\ &= \mathrm{i}\bar\lambda\gamma^\mu\epsilon\partial_\mu\sigma + \mathrm{i}\bar\lambda\gamma^\mu D_\mu\epsilon\sigma + [A_\mu,\bar\lambda]\gamma^\mu\epsilon\sigma,\end{aligned} \tag{1.1.25}$$

where we used that

$$\left(\gamma^\mu\bar\lambda\right)\gamma_\mu\epsilon = -\bar\lambda\gamma^\mu\gamma_\mu\epsilon. \tag{1.1.26}$$

The derivative of σ cancels against the corresponding term in the covariant derivative of σ. Putting all together, we find

$$\mathrm{i}\bar\lambda\gamma^\mu\left(D_\mu\epsilon\right)\sigma - \mathrm{i}\bar\lambda\gamma^\mu\left(D_\mu\epsilon\right)\sigma + [A_\mu,\bar\lambda]\gamma^\mu\epsilon\sigma + \bar\lambda\gamma^\mu\epsilon[A_\mu,\sigma] + \mathrm{i}[\bar\lambda\epsilon,\sigma]\sigma. \tag{1.1.27}$$

The last three terms cancel due to the cyclic property of the trace. This proves the invariance of the supersymmetric CS theory.

In the path integral, the supersymmetric CS action enters in the form

$$\exp\left(\frac{\mathrm{i}k}{4\pi}S_{\mathrm{SCS}}\right) \tag{1.1.28}$$

where k plays the role of the inverse coupling constant and it is referred to as the level of the CS theory. In a consistent quantum theory, k must be an integer [45]. This is due to the fact that the Chern–Simons action for the connection A is not invariant under large gauge transformations, but changes by an integer times $8\pi^2$. The quantization of k guarantees that (1.1.28) remains invariant.

Of course, there is another Lagrangian for vector multiplets, namely the Yang–Mills Lagrangian,

$$\mathcal{L}_{\mathrm{YM}} = \mathrm{Tr}\left[\frac14 F_{\mu\nu}F^{\mu\nu} + \frac12 D_\mu\sigma D^\mu\sigma + \frac12\left(D+\frac{\sigma}{r}\right)^2 + \frac{\mathrm{i}}{2}\bar\lambda\gamma^\mu D_\mu\lambda + \frac{\mathrm{i}}{2}\bar\lambda[\sigma,\lambda] - \frac{1}{4r}\bar\lambda\lambda\right]. \tag{1.1.29}$$

In the flat space limit $r\to\infty$, this becomes the standard (Euclidean) super Yang–Mills theory in three dimensions. The Lagrangian (1.1.29) is not only invariant under the SUSY transformations (1.1.13), but it can be written as a superderivative,

$$\bar\epsilon\epsilon\,\mathcal{L}_{\mathrm{YM}} = \delta_{\bar\epsilon}\delta_\epsilon\mathrm{Tr}\Big(\frac12\bar\lambda\lambda - 2D\sigma\Big). \tag{1.1.30}$$

This will be important later on.

1.1.3 Supersymmetric matter multiplets

We will now add supersymmetric matter, i.e. a chiral multiplet Φ in a representation R of the gauge group. Its components are

$$\Phi: \qquad \phi,\ \bar\phi,\ \psi,\ \bar\psi,\ F,\ \bar F. \tag{1.1.31}$$

The supersymmetry transformations are

$$\begin{aligned}\delta\phi &= \bar\epsilon\psi,\\ \delta\bar\phi &= \epsilon\bar\psi,\\ \delta\psi &= \mathrm{i}\gamma^\mu\epsilon D_\mu\phi + \mathrm{i}\epsilon\sigma\phi + \frac{2\Delta\mathrm{i}}{3}\gamma^\mu D_\mu\epsilon\phi + \bar\epsilon F,\\ \delta\bar\psi &= \mathrm{i}\gamma^\mu\bar\epsilon D_\mu\bar\phi + \mathrm{i}\bar\phi\sigma\bar\epsilon + \frac{2\Delta\mathrm{i}}{3}\bar\phi\gamma^\mu D_\mu\bar\epsilon + \bar F\epsilon,\\ \delta F &= \epsilon(\mathrm{i}\gamma^\mu D_\mu\psi - \mathrm{i}\sigma\psi - \mathrm{i}\lambda\phi) + \frac{\mathrm{i}}{3}(2\Delta-1)D_\mu\epsilon\gamma^\mu\psi,\\ \delta\bar F &= \bar\epsilon(\mathrm{i}\gamma^\mu D_\mu\bar\psi - \mathrm{i}\bar\psi\sigma + \mathrm{i}\bar\phi\bar\lambda) + \frac{\mathrm{i}}{3}(2\Delta-1)D_\mu\bar\epsilon\gamma^\mu\bar\psi,\end{aligned} \tag{1.1.32}$$

where Δ is the possible anomalous dimension of ϕ. For theories with $\mathcal{N} \geq 3$ supersymmetry, the field has the canonical dimension

$$\Delta = \frac{1}{2}, \tag{1.1.33}$$

but in general this is not the case.

The commutators of these transformations are given by

$$\begin{aligned}
[\delta_\epsilon, \delta_{\bar\epsilon}]\phi &= \mathrm{i}v^\mu\partial_\mu\phi + \mathrm{i}\Lambda\phi + \Delta\rho\phi - \Delta\alpha\phi, \\
[\delta_\epsilon, \delta_{\bar\epsilon}]\bar\phi &= \mathrm{i}v^\mu\partial_\mu\bar\phi - \mathrm{i}\bar\phi\Lambda + \Delta\rho\bar\phi + \Delta\alpha\bar\phi, \\
[\delta_\epsilon, \delta_{\bar\epsilon}]\psi &= \mathrm{i}v^\mu\partial_\mu\psi + \tfrac{1}{4}\Theta_{\mu\nu}\gamma^{\mu\nu}\psi + \mathrm{i}\Lambda\psi + \left(\Delta + \frac{1}{2}\right)\rho\psi + (1-\Delta)\alpha\psi, \\
[\delta_\epsilon, \delta_{\bar\epsilon}]\bar\psi &= \mathrm{i}v^\mu\partial_\mu\bar\psi + \tfrac{1}{4}\Theta_{\mu\nu}\gamma^{\mu\nu}\bar\psi - \mathrm{i}\bar\psi\Lambda + \left(\Delta + \frac{1}{2}\right)\rho\bar\psi + (\Delta-1)\alpha\bar\psi, \\
[\delta_\epsilon, \delta_{\bar\epsilon}]F &= \mathrm{i}v^\mu\partial_\mu F + \mathrm{i}\Lambda F + (\Delta+1)\rho F + (2-\Delta)\alpha F, \\
[\delta_\epsilon, \delta_{\bar\epsilon}]\bar F &= \mathrm{i}^v\mu\partial_\mu\bar F - \mathrm{i}\bar F\Lambda + (\Delta+1)\rho\bar F + (\Delta-2)\alpha\bar F.
\end{aligned} \tag{1.1.34}$$

The lowest components of the superfields are assigned the dimension Δ and R-charge $\mp\Delta$. The supersymmetry algebra closes off-shell when the Killing spinors $\epsilon, \bar\epsilon$ satisfy (1.1.18) and h is given by (1.1.20). As a check, we compute

$$\begin{aligned}
[\delta_\epsilon, \delta_{\bar\epsilon}]\phi &= \delta_\epsilon\left(\bar\epsilon\psi\right) \\
&= \bar\epsilon\left(\mathrm{i}\gamma^\mu\epsilon D_\mu\phi + \mathrm{i}\epsilon\sigma\phi + \frac{2\mathrm{i}\Delta}{3}\gamma^\mu\left(D_\mu\epsilon\right)\phi\right) = \mathrm{i}v^\mu D_\mu\phi + \mathrm{i}\sigma\bar\epsilon\epsilon + \frac{2\mathrm{i}\Delta}{3}\left(\bar\epsilon\gamma^\mu D_\mu\epsilon\right),
\end{aligned} \tag{1.1.35}$$

which is the wished-for result.

Let us now consider supersymmetric Lagrangians for the matter hypermultiplet. If the fields have their canonical dimensions, the Lagrangian

$$\mathcal{L} = D_\mu\bar\phi D^\mu\phi - \mathrm{i}\bar\psi\gamma^\mu D_\mu\psi + \frac{3}{4r^2}\bar\phi\phi + \mathrm{i}\bar\psi\sigma\psi + \mathrm{i}\bar\psi\lambda\phi - \mathrm{i}\bar\phi\bar\lambda\psi + \mathrm{i}\bar\phi D\phi + \bar\phi\sigma^2\phi + \bar F F \tag{1.1.36}$$

is invariant under supersymmetry if the Killing spinors $\epsilon, \bar\epsilon$ satisfy (1.1.18), with h given in (1.1.20). The quadratic part of the Lagrangian for ϕ gives indeed the standard conformal coupling for a scalar field. We recall that the action for a massless scalar field in a curved space of n dimensions contains a coupling to the curvature R given by

$$S = \int \mathrm{d}^n x\sqrt{g}\left(g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + \xi R\phi^2\right), \tag{1.1.37}$$

where ξ is a constant. The wave equation is then

$$\left(g^{\mu\nu}D_\mu D_\nu + \xi R\right)\phi = 0. \tag{1.1.38}$$

This equation is conformally invariant when

$$\xi = \frac{1}{4}\frac{n-2}{n-1}. \tag{1.1.39}$$

If the spacetime is an n-sphere of radius r, the curvature is

$$R = \frac{n(n-1)}{r^2}, \tag{1.1.40}$$

and the conformal coupling of the scalar leads to an effective mass term of the form

$$\frac{n(n-2)}{4r^2}\phi^2 \tag{1.1.41}$$

which in $n = 3$ dimensions gives the quadratic term for ϕ in (1.1.36).

If the fields have non-canonical dimensions, the Lagrangian

$$\begin{aligned}
\mathcal{L}_{\mathrm{mat}} =& D_\mu\bar\phi D^\mu\phi + \bar\phi\sigma^2\phi + \frac{\mathrm{i}(2\Delta-1)}{r}\bar\phi\sigma\phi + \frac{\Delta(2-\Delta)}{r^2}\bar\phi\phi + \mathrm{i}\bar\phi D\phi + \bar F F \\
&- \mathrm{i}\bar\psi\gamma^\mu D_\mu\psi + \mathrm{i}\bar\psi\sigma\psi - \frac{2\Delta-1}{2r}\bar\psi\psi + \mathrm{i}\bar\psi\lambda\phi - \mathrm{i}\bar\phi\bar\lambda\psi
\end{aligned} \tag{1.1.42}$$

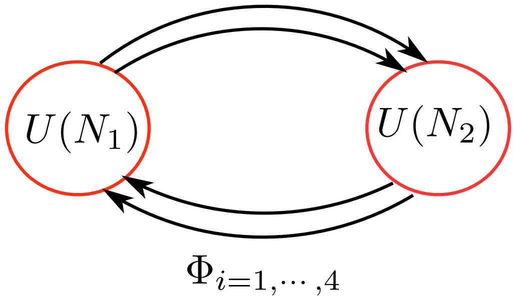

Figure 1.1: The quiver for ABJ(M) theory. The two nodes represent the $U(N_{1,2})$ Chern–Simons theories (with opposite levels) and the arrows between the nodes represent the four matter multiplets in the bifundamental representation.

is supersymmetric, provided the parameters $\epsilon, \bar{\epsilon}$ satisfy the Killing spinor conditions (1.1.19). The Lagrangian (1.1.42) is not only invariant under the supersymmetries $\delta_{\epsilon,\bar{\epsilon}}$, but it can be written as a total superderivative,

$$\bar{\epsilon}\epsilon\,\mathcal{L}_{\rm mat} = \delta_{\bar{\epsilon}}\delta_{\epsilon}\left(\bar{\psi}\psi - 2{\rm i}\bar{\phi}\sigma\phi + \frac{2(\Delta-1)}{r}\bar{\phi}\phi\right). \tag{1.1.43}$$

1.1.4 ABJM theory

The theory proposed in [1, 3] to describe N M2 branes is a particular example of a supersymmetric Chern–Simons theory. It consist of two copies of Chern–Simons theory with gauge groups $U(N_1)$, $U(N_2)$, and opposite levels k, $-k$. In addition, we have four matter supermultiplets Φ_i, $i = 1, \cdots, 4$, in the bifundamental representation of the gauge group $U(N_1) \times U(N_2)$. This theory can be represented as a quiver, with two nodes representing the Chern–Simons theories, and four edges between the nodes representing the matter supermultiplets, see Fig. 1.1. In addition, there is a superpotential involving the matter fields, which after integrating out the auxiliary fields in the Chern–Simons–matter system, reads (on $\mathbb{R}^3$)

$$W = \frac{4\pi}{k}{\rm Tr}\left(\Phi_1\Phi_2^\dagger\Phi_3\Phi_4^\dagger - \Phi_1\Phi_4^\dagger\Phi_3\Phi_2^\dagger\right), \tag{1.1.44}$$

where we have used the standard superspace notation for $\mathcal{N} = 1$ supermultiplets [44].

1.2 A brief review of Chern–Simons theory

Since one crucial ingredient in the theories we are considering is Chern–Simons theory, we review here some results concerning the perturbative structure of this theory on general three-manifolds. These results were first obtained in the seminal paper by Witten [46] and then extended and refined in various directions in [47, 48, 49, 50, 51, 52]. Chern–Simons perturbation theory on general three-manifolds is a important subject in itself, hence we will try to give a general presentation which might be useful in other contexts. This will require a rather formal development, and the reader interested in the result for the one-loop contribution might want to skip some of the derivations in the next two subsections.

1.2.1 Perturbative approach

In this section, we will denote the bosonic Chern–Simons action by

$$S = -\frac{1}{4\pi}\int_M {\rm Tr}\Big(A\wedge {\rm d}A + \frac{2{\rm i}}{3}A\wedge A\wedge A\Big) \tag{1.2.1}$$

where we use the conventions appropriate for Hermitian connections, and we included the factor $1/4\pi$ in the action for notational convenience. The group of gauge transformations $\mathcal{G}$ acts on the gauge connections as follows,

$$A \to A^U = UAU^{-1} - {\rm i}U\,{\rm d}U^{-1}, \qquad U \in \mathcal{G}. \tag{1.2.2}$$

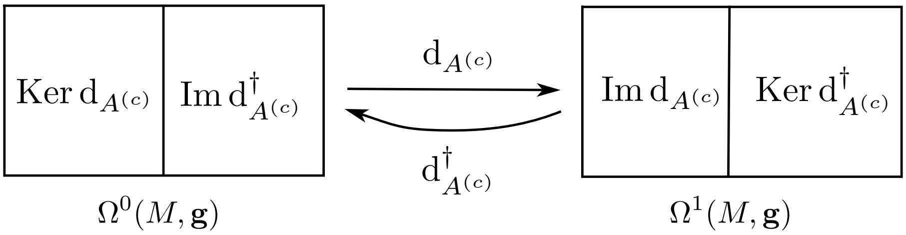

Figure 1.2: The standard elliptic decomposition of $\Omega^0(M, \mathbf{g})$ and $\Omega^1(M, \mathbf{g})$.

We will assume that the theory is defined on a compact three-manifold M. The partition function is defined as

$$Z(M) = \frac{1}{\mathrm{vol}(\mathcal{G})} \int [\mathcal{D}A] \mathrm{e}^{\mathrm{i}kS} \tag{1.2.3}$$

where we recall that $k \in \mathbb{Z}$.

There are many different approaches to the calculation of (1.2.3), but the obvious strategy is to use perturbation theory. Notice that, since the theory is defined on a compact manifold, there are no IR divergences and we just have to deal with UV divergences, as in standard QFT. Once these are treated appropriately, the partition function (1.2.3) is a well-defined observable. In perturbation theory we evaluate (1.2.3) by expanding around saddle–points. These are flat connections, which are in one-to-one correspondence with group homomorphisms

$$\pi_1(M) \to G \tag{1.2.4}$$

modulo conjugation. For example, if $M = \mathbb{S}^3/\mathbb{Z}_p$ is the lens space $L(p, 1)$, one has $\pi_1(L(p, 1)) = \mathbb{Z}_p$, and flat connections are labelled by homomorphisms $\mathbb{Z}_p \to G$. Let us assume that these are a discrete set of points (this happens, for example, if M is a rational homology sphere, since in that case $\pi_1(M)$ is a finite group). We will label the flat connections with an index c, and a flat connection will be denoted by $A^{(c)}$. Each flat connection leads to a covariant derivative

$$\mathrm{d}_{A^{(c)}} = \mathrm{d} + \mathrm{i}[A^{(c)}, \cdot], \tag{1.2.5}$$

and flatness implies that

$$\mathrm{d}^2_{A^{(c)}} = \mathrm{i}\, F_{A^{(c)}} = 0. \tag{1.2.6}$$

Therefore, the covariant derivative leads to a complex

$$0 \to \Omega^0(M, \mathbf{g}) \xrightarrow{\mathrm{d}_{A^{(c)}}} \Omega^1(M, \mathbf{g}) \xrightarrow{\mathrm{d}_{A^{(c)}}} \Omega^2(M, \mathbf{g}) \xrightarrow{\mathrm{d}_{A^{(c)}}} \Omega^3(M, \mathbf{g}). \tag{1.2.7}$$

The first two terms in this complex have a natural interpretation in the context of gauge theories: $\Omega^0(M, \mathbf{g})$ is the Lie algebra of the group of gauge transformations, and we can write a gauge transformation as

$$U = \mathrm{e}^{\mathrm{i}\phi}, \qquad \phi \in \Omega^0(M, \mathbf{g}). \tag{1.2.8}$$

The elements of $\Omega^0(M, \mathbf{g})$ generate infinitesimal gauge transformations,

$$\delta A = -\mathrm{d}_A \phi. \tag{1.2.9}$$

The second term, $\Omega^1(M, \mathbf{g})$, can be identified with the tangent space to the space of gauge connections. The first map in the complex (1.2.7) is interpreted as (minus) an infinitesimal gauge transformation in the background of $A^{(c)}$.

We recall that the space of $\mathbf{g}$-valued forms on M has a natural inner product given by

$$\langle a, b \rangle = \int_M \mathrm{Tr}\, (a \wedge *b), \tag{1.2.10}$$

where $*$ is the Hodge operator. With respect to this product we can define an adjoint operator on $\mathbf{g}$-valued p-forms in the same way that is done for the usual de Rham operator,

$$\mathrm{d}^{\dagger}_{A^{(c)}} = (-1)^{3(1+p)+1} * \mathrm{d}_{A^{(c)}} *\,. \tag{1.2.11}$$

We then have the orthogonal decompositions (see Fig. 1.2)

$$\begin{aligned}\Omega^0(M,\mathbf{g}) =&\mathrm{Ker}\,\mathrm{d}_{A^{(c)}} \oplus \mathrm{Im}\,\mathrm{d}^\dagger_{A^{(c)}},\\ \Omega^1(M,\mathbf{g}) =&\mathrm{Ker}\,\mathrm{d}^\dagger_{A^{(c)}} \oplus \mathrm{Im}\,\mathrm{d}_{A^{(c)}}.\end{aligned} \tag{1.2.12}$$

These decompositions are easily proved. For the first one, for example, we just note that

$$a \in \mathrm{Ker}\,\mathrm{d}_{A^{(c)}} \Rightarrow \langle \mathrm{d}_{A^{(c)}} a, \phi\rangle = \langle a, \mathrm{d}^\dagger_{A^{(c)}}\phi\rangle = 0, \qquad \forall\phi \tag{1.2.13}$$

therefore

$$(\mathrm{Ker}\,\mathrm{d}_{A^{(c)}})^\perp = \mathrm{Im}\,\mathrm{d}^\dagger_{A^{(c)}}. \tag{1.2.14}$$

One also has the analogue of the Laplace–Beltrami operator acting on p-forms

$$\Delta^p_{A^{(c)}} = \mathrm{d}^\dagger_{A^{(c)}}\mathrm{d}_{A^{(c)}} + \mathrm{d}_{A^{(c)}}\mathrm{d}^\dagger_{A^{(c)}}. \tag{1.2.15}$$

In the following we will assume that

$$H^1(M,\mathbf{g}) = 0. \tag{1.2.16}$$

This means that the connection $A^{(c)}$ is *isolated*. However, we will consider the possibility that $A^{(c)}$ has a non-trivial isotropy group $\mathcal{H}_c$. We recall that the isotropy group of a connection $A^{(c)}$ is the subgroup of gauge transformations which leave $A^{(c)}$ invariant,

$$\mathcal{H}_c = \{\phi\in\mathcal{G}|\phi(A^{(c)}) = A^{(c)}\}. \tag{1.2.17}$$

The Lie algebra of this group is given by zero-forms annihilated by the covariant derivative (1.2.5),

$$\mathrm{Lie}(\mathcal{H}_c) = H^0(M,\mathbf{g}) = \mathrm{Ker}\,\mathrm{d}_{A^{(c)}}, \tag{1.2.18}$$

which is in general non-trivial. A connection is *irreducible* if its isotropy group is equal to the center of the group. In particular, if $A^{(c)}$ is irreducible one has

$$H^0(M,\mathbf{g}) = 0. \tag{1.2.19}$$

It can be shown that the isotropy group $\mathcal{H}_c$ consists of constant gauge transformations that leave $A^{(c)}$ invariant,

$$\phi A^{(c)}\phi^{-1} = A^{(c)}. \tag{1.2.20}$$

They are in one-to-one correspondence with a subgroup of G which we will denote by H_c.

In the semiclassical approximation, $Z(M)$ is written as a sum of terms associated to saddle-points:

$$Z(M) = \sum_c Z^{(c)}(M), \tag{1.2.21}$$

where c labels the different flat connections $A^{(c)}$ on M. Each of the $Z^{(c)}(M)$ will be a perturbative series in $1/k$ of the form

$$Z^{(c)}(M) = Z^{(c)}_{1-\mathrm{loop}}(M)\exp\left\{\sum_{\ell=1}^\infty S^{(c)}_\ell k^{-\ell}\right\}, \tag{1.2.22}$$

where $S^{(c)}_\ell$ is the $(\ell+1)$-loop contribution around the flat connection $A^{(c)}$. In order to derive this expansion, we split the connection into a “background", which is the flat connection $A^{(c)}$, plus a “fluctuation" B:

$$A = A^{(c)} + B. \tag{1.2.23}$$

Expanding around this, we find

$$S(A) = S(A^{(c)}) + S(B), \tag{1.2.24}$$

where

$$S(B) = -\frac{1}{4\pi}\int_M \mathrm{Tr}\,(B\wedge \mathrm{d}_{A^{(c)}}B + \frac{2\mathrm{i}}{3}B^3). \tag{1.2.25}$$

The first term in (1.2.24) is the classical Chern–Simons invariant of the connection $A^{(c)}$. Since Chern–Simons theory is a gauge theory, in order to proceed we have to fix the gauge. We will follow the detailed analysis of [51]. Our gauge choice will be the standard, covariant gauge,

$$g_{A^{(c)}}(B) = \mathrm{d}^\dagger_{A^{(c)}}B = 0 \tag{1.2.26}$$

where $g_{A^{(c)}}$ is the gauge fixing function. We recall that in the standard Fadeev–Popov (FP) gauge fixing one first defines

$$\Delta^{-1}_{A^{(c)}}(B) = \int \mathcal{D}U\,\delta\left(g_{A^{(c)}}\left(B^U\right)\right), \tag{1.2.27}$$

and then inserts into the path integral

$$1 = \left[\int \mathcal{D}U\,\delta\left(g_{A^{(c)}}\left(B^U\right)\right)\right]\Delta_{A^{(c)}}(B). \tag{1.2.28}$$

The key new ingredient here is the presence of a non-trivial isotropy group $\mathcal{H}_c$ for the flat connection $A^{(c)}$. When there is a non-trivial isotropy group, the gauge-fixing condition does not fix the gauge completely, since

$$g_{A^{(c)}}(B^\phi) = \phi g_{A^{(c)}}(B)\phi^{-1}, \qquad \phi \in \mathcal{H}_c, \tag{1.2.29}$$

i.e. the basic assumption that $g(A) = 0$ only cuts the gauge orbit once is not true, and there is a residual symmetry given by the isotropy group. Another way to see this is that the standard FP determinant vanishes due to zero modes. In fact, the standard calculation of (1.2.27) (which is valid if the isotropy group of $A^{(c)}$ is trivial) gives

$$\Delta^{-1}_{A^{(c)}}(B) = \left|\det\frac{\delta g_{A^{(c)}}(B^U)}{\delta U}\right|^{-1} = \left|\det \mathrm{d}^\dagger_{A^{(c)}}\mathrm{d}_A\right|^{-1}. \tag{1.2.30}$$

But when $\mathcal{H}_c \neq 0$, the operator $\mathrm{d}_{A^{(c)}}$ has zero modes due to the nonvanishing of (1.2.18), and the FP procedure is ill-defined. The correct way to proceed in the calculation of (1.2.27) is to split the integration over the gauge group into two pieces. The first piece is the integration over the isotropy group. Due to (1.2.29), the integrand does not depend on it, and we obtain a factor of $\mathrm{Vol}(\mathcal{H}_c)$. The second piece gives an integration over the remaining part of the gauge transformations, which has as its Lie algebra

$$(\mathrm{Ker}\,\mathrm{d}_{A^{(c)}})^\perp. \tag{1.2.31}$$

The integration over this piece leads to the standard FP determinant (1.2.30) but with the zero modes removed. We then find,

$$\Delta^{-1}_{A^{(c)}}(B) = \mathrm{Vol}(\mathcal{H}_c)\left|\det \mathrm{d}^\dagger_{A^{(c)}}\mathrm{d}_A\right|^{-1}_{(\mathrm{Ker}\,\mathrm{d}_{A^{(c)}})^\perp} \tag{1.2.32}$$

This phenomenon was first observed by Rozansky in [48], and developed in this language in [51]. As usual, the determinant appearing here can be written as a path integral over ghost fields, with action

$$S_{\mathrm{ghosts}}(C,\overline{C},B) = \langle \overline{C}, \mathrm{d}^\dagger_{A^{(c)}}\mathrm{d}_A C\rangle, \tag{1.2.33}$$

where $C, \overline{C}$ are Grassmannian fields taking values in

$$(\mathrm{Ker}\,\mathrm{d}_{A^{(c)}})^\perp. \tag{1.2.34}$$

The action for the ghosts can be divided into a kinetic term plus an interaction term between the ghost fields and the fluctuation B:

$$S_{\mathrm{ghosts}}(C,\overline{C},B) = \langle \overline{C}, \Delta^0_{A^{(c)}} C\rangle + \mathrm{i}\langle \overline{C}, \mathrm{d}^\dagger_{A^{(c)}}[B,C]\rangle. \tag{1.2.35}$$

The modified FP gauge–fixing leads then to the path integral

$$\begin{aligned} Z^{(c)}(M) &= \frac{\mathrm{e}^{\mathrm{i}kS(A^{(c)})}}{\mathrm{vol}(\mathcal{G})}\int_{\Omega^1(M,\mathbf{g})}\mathcal{D}B\,\mathrm{e}^{\mathrm{i}kS(B)}\Delta_{A^{(c)}}(B)\delta\left(\mathrm{d}^\dagger_{A^{(c)}}B\right) \\ &= \frac{\mathrm{e}^{\mathrm{i}kS(A^{(c)})}}{\mathrm{Vol}(\mathcal{H}_c)}\int_{\Omega^1(M,\mathbf{g})}\mathcal{D}B\,\delta\left(\mathrm{d}^\dagger_{A^{(c)}}B\right)\int_{(\mathrm{Ker}\,\mathrm{d}_{A^{(c)}})^\perp}\mathcal{D}C\mathcal{D}\overline{C}\,\mathrm{e}^{\mathrm{i}kS(B)-S_{\mathrm{ghosts}}(C,\overline{C},B)}. \end{aligned} \tag{1.2.36}$$

Finally, we analyze the delta constraint on B. Due to the decomposition of $\Omega^1(M,\mathbf{g})$ in (1.2.12), we can write

$$B = \mathrm{d}_{A^{(c)}}\phi + B', \tag{1.2.37}$$

where

$$\phi \in (\mathrm{Ker}\, \mathrm{d}_{A^{(c)}})^{\perp}, \quad B' \in \mathrm{Ker}\, \mathrm{d}^{\dagger}_{A^{(c)}}. \tag{1.2.38}$$

The presence of the operator $\mathrm{d}_{A^{(c)}}$ in the change of variables (1.2.37) leads to a non-trivial Jacobian. Indeed, we have

$$\|B\|^2 = \langle \phi, \Delta^0_{A^{(c)}} \phi \rangle + \|B'\|^2, \tag{1.2.39}$$

and the measure in the functional integral becomes

$$\mathcal{D}B = \left(\det{}' \Delta^0_{A^{(c)}}\right)^{\frac{1}{2}} \mathcal{D}\phi\, \mathcal{D}B', \tag{1.2.40}$$

where the $'$ indicates, as usual, that we are removing zero modes. Notice that the operator in (1.2.40) is positive–definite, so the square root of its determinant is well–defined. We also have that

$$\delta\left(\mathrm{d}^{\dagger}_{A^{(c)}} B\right) = \delta\left(\Delta^0_{A^{(c)}} \phi\right) = \left(\det{}' \Delta^0_{A^{(c)}}\right)^{-1} \delta(\phi), \tag{1.2.41}$$

which is a straightforward generalization of the standard formula

$$\delta(ax) = \frac{1}{|a|}\delta(x). \tag{1.2.42}$$

We conclude that the delta function, together with the Jacobian in (1.2.40), lead to the the following factor in the path integral:

$$\left(\det{}' \Delta^0_{A^{(c)}}\right)^{-\frac{1}{2}}. \tag{1.2.43}$$

In addition, the delta function sets $\phi = 0$. It only remains the integration over B', which we relabel $B' \to B$. The final result for the gauge-fixed path integral is then

$$Z^{(c)}(M) = \frac{\mathrm{e}^{\mathrm{i}kS(A^{(c)})}}{\mathrm{Vol}(\mathcal{H}_c)} \left(\det{}' \Delta^0_{A^{(c)}}\right)^{-\frac{1}{2}} \int_{\mathrm{Ker}\, \mathrm{d}^{\dagger}_{A^{(c)}}} \mathcal{D}B \int_{(\mathrm{Ker}\, \mathrm{d}_{A^{(c)}})^{\perp}} \mathcal{D}C\mathcal{D}\overline{C}\, \mathrm{e}^{\mathrm{i}kS(B) - S_{\mathrm{ghosts}}(C,\overline{C},B)}. \tag{1.2.44}$$

This is starting point to perform gauge-fixed perturbation theory in Chern–Simons theory.

1.2.2 The one–loop contribution

We now consider the one-loop contribution of a saddle-point to the path integral. This has been studied in many papers [47, 48, 49, 50]. We will follow the detailed presentation in [52].

Before proceeding, we should specify what is the regularization method that we will use to define the functional determinants appearing in our calculation. A natural and useful regularization for quantum field theories in curved space is zeta-function regularization [53]. We recall that the zeta function of a self-adjoint operator T with eigenvalues $\lambda_n > 0$ is defined as

$$\zeta_T(s) = \sum_n \lambda_n^{-s}. \tag{1.2.45}$$

Under appropriate conditions, this defines a meromorphic function on the complex s-plane which is regular at $s = 0$. Since

$$-\zeta'_T(0) = \sum_n \log \lambda_n \tag{1.2.46}$$

we can define the determinant of T as

$$\det(T) = \mathrm{e}^{-\zeta'_T(0)}. \tag{1.2.47}$$

This is the regularization we will adopt in the following. It has the added advantage that, when used in Chern–Simons theory, it leads to natural topological invariants like the Ray–Singer torsion, as we will see.

The main ingredients in the one-loop contribution to the path integral (1.2.44) are the determinants of the operators appearing in the kinetic terms for B, C and $\overline{C}$. Putting together the determinant (1.2.43) and the determinant coming from the ghost fields, we obtain

$$\left(\det{}' \Delta^0_{A^{(c)}}\right)^{1/2} \tag{1.2.48}$$

since the ghosts are restricted to (1.2.34) and their determinant is also primed. The operator appearing in the kinetic term for B is $\mathrm{i}Q/2$, where

$$Q = -\frac{k}{2\pi} * \mathrm{d}_{A^{(c)}} \tag{1.2.49}$$

is a self-adjoint operator acting on $\Omega^1(M, \mathbf{g})$ which has to be restricted to

$$\mathrm{Ker}\, \mathrm{d}^{\dagger}_{A^{(c)}} = (\mathrm{Im}\, \mathrm{d}_{A^{(c)}})^{\perp} \tag{1.2.50}$$

due to the gauge fixing. Notice that, if (1.2.16) holds, one has

$$H^1(M, \mathbf{g}) = 0 \Rightarrow \mathrm{Ker}\, \mathrm{d}_{A^{(c)}} = \mathrm{Im}\, \mathrm{d}_{A^{(c)}}, \tag{1.2.51}$$

and due to the restriction to (1.2.50), Q has no zero modes. However, the operator Q is *not* positive definite, and one has to be careful in order to define its determinant. We will now do this following the discussion in [46, 57]. A natural definition takes as its starting point the trivial Gaussian integral

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \exp\left(\mathrm{i}\frac{\lambda x^2}{2}\right) = \sqrt{\frac{2\pi}{|\lambda|}} \exp\left(\frac{\mathrm{i}\pi}{4} \mathrm{sign}\, \lambda\right). \tag{1.2.52}$$

If we want to have a natural generalization of this, the integration over B should be

$$\exp\left(\frac{\mathrm{i}\pi}{4}\mathrm{sign}(Q)\right) \left|\det\left(\frac{Q}{2\pi}\right)\right|^{-1/2}. \tag{1.2.53}$$

In order to compute the determinant in absolute value, we can consider the square of the operator $- * \mathrm{d}_{A^{(c)}}$, which is given by

$$* \mathrm{d}_{A^{(c)}} * \mathrm{d}_{A^{(c)}} = \mathrm{d}^{\dagger}_{A^{(c)}} \mathrm{d}_{A^{(c)}}, \tag{1.2.54}$$

and it is positive definite when restricted to (1.2.50). It is the Laplacian on one-forms, restricted to (1.2.50). We then define

$$\left|\det \frac{Q}{2\pi}\right|^2 = \det{}' \left[\left(\frac{k}{4\pi^2}\right)^2 \mathrm{d}^{\dagger}_{A^{(c)}} \mathrm{d}_{A^{(c)}}\right] \tag{1.2.55}$$

where we have subtracted the zero–modes (i.e., we consider the restriction to (1.2.50)). In order to take into account the sign in (1.2.53), we need the η invariant of the operator $- * \mathrm{d}_{A^{(c)}}$. We recall that the η invariant is defined as

$$\eta_T(s) = \sum_j \frac{1}{\left(\lambda_j^+\right)^s} - \sum_j \frac{1}{\left(-\lambda_j^-\right)^s} \tag{1.2.56}$$

where $\lambda_j^{\pm}$ are the strictly positive (negative, respectively) eigenvalues of T. The regularized difference of eigenvalues is then $\eta_T(0)$. In our case, this gives

$$\eta(A^{(c)}) \equiv \eta_{-*\mathrm{d}_{A^{(c)}}}(0). \tag{1.2.57}$$

Finally, we have to take into account that for each eigenvalue of the operator (1.2.54) we have a factor of

$$\left(\frac{k}{4\pi^2}\right)^{-1/2} \tag{1.2.58}$$

appearing in the final answer. The regularized number of eigenvalues of the operator is simply

$$\zeta(A^{(c)}) \equiv \zeta_{\mathrm{d}^{\dagger}_{A^{(c)}} \mathrm{d}_{A^{(c)}}}(0), \tag{1.2.59}$$

restricted again to (1.2.50). Putting all together we obtain,

$$\left(\det \frac{\mathrm{i}Q}{2\pi}\right)^{-\frac{1}{2}} = \left(\frac{k}{4\pi^2}\right)^{-\zeta(A^{(c)})/2} \exp\Bigl(\frac{\mathrm{i}\pi}{4}\eta(A^{(c)})\Bigr) \left(\det{}' \mathrm{d}^{\dagger}_{A^{(c)}} \mathrm{d}_{A^{(c)}}\right)^{-\frac{1}{4}}. \tag{1.2.60}$$

Here we assumed that $k > 0$. For a negative level $-k < 0$ the answer is still given by (1.2.60), but the phase involving the eta invariant has the opposite sign. We can now combine this result with (1.2.48).

The quotient of the determinants of the Laplacians gives the square root of the so-called *Ray–Singer torsion* of the flat connection $A^{(c)}$ [54],

$$\frac{(\det' \Delta^0_{A^{(c)}})^{\frac{1}{2}}}{\left(\det' \mathrm{d}^\dagger_{A^{(c)}} \mathrm{d}_{A^{(c)}}\right)^{\frac{1}{4}}} = \sqrt{\tau'_R(M, A^{(c)})}. \tag{1.2.61}$$

This was first observed by Schwarz in the Abelian theory [55]. When the connection $A^{(c)}$ is isolated and irreducible, this quotient is a topological invariant of M, but in general it is not. However, for a reducible and isolated flat connection, the dependence on the metric is just given by an overall factor, equal to the volume of the manifold M:

$$\tau'_R(M, A^{(c)}) = (\mathrm{vol}(M))^{\dim(\mathcal{H}_c)}\, \tau_R(M, A^{(c)}). \tag{1.2.62}$$

where $\tau_R(M, A^{(c)})$ is now metric-indepedendent. For an explanation of this fact, see for example Appendix B in [56]. However, this volume, metric-dependent factor cancels in the final answer for the one-loop path integral [52]. The isotropy group $\mathcal{H}_c$ is the space of constant zero forms, taking values in a subgroup $H_c \subset G$ of the gauge group. Each generator of its Lie algebra has a norm given by its norm as an element of $\mathbf{g}$, times

$$\left(\int_M *1\right)^{1/2} = (\mathrm{vol}(M))^{1/2}\,. \tag{1.2.63}$$

Therefore,

$$\mathrm{vol}(\mathcal{H}_c) = (\mathrm{vol}(M))^{\dim(\mathcal{H}_c)/2}\, \mathrm{vol}(H_c), \tag{1.2.64}$$

and

$$\frac{\sqrt{\tau'_R(M, A^{(c)})}}{\mathrm{vol}(\mathcal{H}_c)} = \frac{\sqrt{\tau_R(M, A^{(c)})}}{\mathrm{vol}(H_c)} \tag{1.2.65}$$

which does not depend on the metric of M. Finally, in order to write down the answer, we take into account that, for isolated flat connections, $\zeta(A^{(c)})$ can be evaluated as [58]

$$\zeta(A^{(c)}) = \dim H^0(M, \mathbf{g}). \tag{1.2.66}$$

Putting everything together, we find for the one–loop contribution to the path integral

$$Z^{(c)}_{1-\mathrm{loop}}(M) = \frac{1}{\mathrm{vol}(H_c)} \left(\frac{k}{4\pi^2}\right)^{-\frac{1}{2}\dim H^0(M,\mathbf{g})} \mathrm{e}^{\mathrm{i}kS(A^{(c)}) + \frac{\mathrm{i}\pi}{4}\eta(A^{(c)})} \sqrt{\tau_R(M, A^{(c)})}. \tag{1.2.67}$$

As noticed above, this expression is valid for $k > 0$. For a negative level $-k < 0$, the phase involving the gravitational η invariant has the opposite sign. It was pointed out in [46] that this phase can be written in a more suggestive form by using the Atiyah–Patodi–Singer theorem, which says that

$$\eta(A^{(c)}) = \eta(0) + \frac{4y}{\pi} S(A^{(c)}). \tag{1.2.68}$$

Here y is the dual Coxeter number of G (for $U(N)$, $y = N$), and $\eta(0)$ is the eta invariant of the trivial connection. Let us denote by

$$d_G = \dim(G), \tag{1.2.69}$$

the dimension of the gauge group. The operator involved in the calculation of $\eta(0)$ is just d_G copies of the “gravitational" operator $- * \mathrm{d}$, which is only coupled to the metric. We can then write

$$\eta(0) = d_G\, \eta_{\mathrm{grav}}, \tag{1.2.70}$$

where η_{grav} is the “gravitational" eta invariant of $- * \mathrm{d}$. We then find,

$$Z^{(c)}_{1-\mathrm{loop}}(M) = \frac{1}{\mathrm{vol}(H_c)} \left(\frac{k}{4\pi^2}\right)^{-\frac{1}{2}\dim H^0(M,\mathbf{g})} \mathrm{e}^{\mathrm{i}(k+y)S(A^{(c)}) + \frac{\mathrm{i}\pi}{4} d_G \eta_{\mathrm{grav}}} \sqrt{\tau_R(M, A^{(c)})}. \tag{1.2.71}$$

This formula exhibits a one–loop renormalization of k

$$k \to k + y \tag{1.2.72}$$

which is simply a shift by the dual Coxeter number [46]. However, different regularizations of the theory seem to lead to different shifts [59].

When $A^{(c)} = 0$ is the trivial flat connection, one has that $H_c = G$, where G is the gauge group, and the cohomology twisted by $A^{(c)}$ reduces to the ordinary cohomology. The Ray–Singer torsion is the torsion $\tau_R(M)$ of the ordinary de Rham differential, to the power d_G. We then obtain, for the trivial connection,

$$Z_{1-\text{loop}}(M) = \frac{1}{\text{vol}(G)} \left(\frac{k}{4\pi^2}\right)^{-d_G/2} e^{\frac{i\pi}{4} d_G \eta_{\text{grav}}} \left(\tau_R(M)\right)^{d_G/2}. \tag{1.2.73}$$

As explained in detail in [46], the phase in (1.2.71) and (1.2.73) involving the η invariant is metric-dependent. In constructing a topological field theory out of Chern–Simons gauge theory, as in [46], one wants to preserve topological invariance, and an appropriate counterterm has to be added to the action. The counterterm involves the gravitational Chern–Simons action $S(\omega)$, where ω is the spin connection. However, this action is ambiguous, and it depends on a trivialization of the tangent bundle to M. Such a choice of trivialization is called a *framing* of the three-manifold. The difference between two trivializations can be encoded in an integer s, and when one changes the trivialization the gravitational Chern–Simons action changes as

$$S(\omega) \to S(\omega) + 2\pi s, \tag{1.2.74}$$

similarly to the gauge Chern–Simons action. According to the Atiyah–Patodi–Singer theorem, the combination

$$\frac{1}{4}\eta_{\text{grav}} + \frac{1}{12}\frac{S(\omega)}{2\pi} \tag{1.2.75}$$

is a topological invariant. It depends on the choice of framing of M, but not on its metric. Therefore, if we modify (1.2.73) by including in the phase an appropriate multiple of the gravitational Chern–Simons action,

$$\exp\left(\frac{i\pi}{4} d_G \eta_{\text{grav}}\right) \to \exp\left[i\pi d_G\left(\frac{\eta_{\text{grav}}}{4} + \frac{1}{12}\frac{S(\omega)}{2\pi}\right)\right], \tag{1.2.76}$$

the resulting one-loop partition function is a topological invariant of the framed three-manifold M. If we change the framing of M by s units, the above factor induces a change in the partition function which at one-loop is of the form

$$Z(M) \to \exp\left(2\pi i s \cdot \frac{d_G}{24}\right) Z(M). \tag{1.2.77}$$

One of the most beautiful results of [46] is that Chern–Simons theory is exactly solvable on any three-manifold M, and its partition function can be computed exactly as a function of k, for any gauge group G, by using current algebra in two dimensions. In particular, one can compute the exact change of the partition function under a change of framing, and one finds

$$Z(M) \to \exp\left(2\pi i s \cdot \frac{c}{24}\right) Z(M), \tag{1.2.78}$$

where

$$c = \frac{k d_G}{k + y}. \tag{1.2.79}$$

A detailed explanation of the exact solution of CS theory would take us too far, and we refer the reader to the original paper [46] or to the presentation in [60] for more details. We will however list later on the relevant results when $M = \mathbb{S}^3$.

1.3 The free energy at weak coupling

The partition function of a CFT on $\mathbb{S}^3$ should encode information about the number of degrees of freedom of the theory, in the sense that at weak coupling it should scale as the number $\mathcal{N}$ of elementary constituents. This follows simply from the factorization property of the partition function in the absence of interactions:

$$Z(\mathbb{S}^3, \mathcal{N}) \approx \left(Z(\mathbb{S}^3, 1)\right)^{\mathcal{N}}. \tag{1.3.1}$$

For example, a gauge theory with gauge group $U(N)$ has at weak coupling N^2 degrees of freedom, and we should expect the free energy on the three-sphere to scale in this regime as

$$F(\mathbb{S}^3) \sim \mathcal{O}(N^2). \tag{1.3.2}$$

Of course, at strong coupling this is not necessarily the case.

In this section we will compute the partition function on $\mathbb{S}^3$ of supersymmetric Chern–Simons–matter theories in the weak coupling approximation, i.e. at one loop. First, we will do the computation in Chern–Simons theory, and then we will consider the much simpler case of supersymmetric matter multiplets.

1.3.1 Chern–Simons theory on $\mathbb{S}^3$

In the previous section we presented the general procedure to calculate the one-loop contribution of Chern–Simons theory on any three-manifold, around an isolated flat connection. This procedure can be made very concrete when the manifold is $\mathbb{S}^3$. In this case there is only one flat connection, the trivial one $A^{(c)} = 0$, and we can use (1.2.73). Therefore, we just have to compute the Ray–Singer torsion $\tau(\mathbb{S}^3)$ for the standard de Rham differential, i.e. the quotient of determinants appearing in (1.2.61) with $A^{(c)} = 0$ (a similar calculation was made in Appendix A of [56]).

We endow $\mathbb{S}^3$ with its standard metric (the one induced by its standard embedding in $\mathbb{R}^4$ with Euclidean metric), and we choose the radius $r = 1$ (it is easy to verify explicitly that the final result is independent of r). The determinant of the scalar Laplacian on the sphere can be computed very explicitly, since its eigenvalues are known to be (see the Appendix)

$$\lambda_n = n(n+2), \qquad n = 0, 1, \cdots \tag{1.3.3}$$

where n is related to j in (A.45) by $n = 2j$. The degeneracy of this eigenvalue is

$$d_n = (n+1)^2. \tag{1.3.4}$$

Removing the zero eigenvalue just means that we remove $n = 0$ from the spectrum. To calculate the determinant we must calculate the zeta function,

$$\zeta_{\Delta^0}(s) = \sum_{n=1}^{\infty} \frac{d_n}{\lambda_n^s} = \sum_{n=1}^{\infty} \frac{(n+1)^2}{(n(n+2))^s} = \sum_{m=2}^{\infty} \frac{m^2}{(m^2-1)^s}. \tag{1.3.5}$$

This zeta function can not be written in closed form, but its derivative at $s = 0$ is easy to calculate. The calculation can be done in many ways, and general results for the determinant of Laplacians on $\mathbb{S}^m$ can be found in for example [61, 62]. We will follow a simple procedure inspired by [63]. We split

$$\frac{m^2}{(m^2-1)^s} = \frac{1}{m^{2(s-1)}} + \frac{s}{m^{2s}} + R(m,s), \tag{1.3.6}$$

where

$$R(m,s) = \frac{m^2}{(m^2-1)^s} - \frac{1}{m^{2(s-1)}} - \frac{s}{m^{2s}} \tag{1.3.7}$$

which decreases as m^{-2s-2} for large m, and therefore leads to a convergent series for all $s \geq -1/2$ which is moreover uniformly convergent. Therefore, it is possible to exchange sums with derivatives. The derivative of $R(m,s)$ at $s = 0$ can be calculated as

$$\left.\frac{\mathrm{d}R(m,s)}{\mathrm{d}s}\right|_{s=0} = -1 - m^2 \log\left(1 - \frac{1}{m^2}\right). \tag{1.3.8}$$

The sum of this series can be explicitly calculated by using the Hurwitz zeta function, and one finds

$$-\sum_{m=2}^{\infty}\left[1 + m^2 \log\left(1 - \frac{1}{m^2}\right)\right] = \frac{3}{2} - \log(\pi). \tag{1.3.9}$$

We then obtain

$$\zeta_{\Delta^0}(s) = \zeta(2s-2) - 1 + s\left(\zeta(2s) - 1\right) + \sum_{m=2}^{\infty} R(m,s), \tag{1.3.10}$$

where $\zeta(s)$ is Riemann's zeta function, and

$$-\zeta'_{\Delta^0}(0) = \log(\pi) - 2\zeta'(-2). \tag{1.3.11}$$

The final result can be expressed in terms of $\zeta(3)$, since

$$\zeta'(-2) = -\frac{\zeta(3)}{4\pi^2}. \tag{1.3.12}$$

We conclude that the determinant of the scalar Laplacian on $\mathbb{S}^3$ is given by

$$\log \det{}' \Delta^0 = \log(\pi) + \frac{\zeta(3)}{2\pi^2}. \tag{1.3.13}$$

We now compute the determinant in the denominator of (1.2.61). We must consider the space of one-forms on $\mathbb{S}^3$, and restrict to the ones that are not in the image of d. These forms are precisely the vector spherical harmonics, whose properties are reviewed in the Appendix A. The eigenvalues of the operator $\mathrm{d}^\dagger \mathrm{d}$ are given in (A.52), and they read

$$\lambda_n = (n+1)^2, \qquad n = 1, 2, \cdots, \tag{1.3.14}$$

with degeneracies

$$d_n = 2n(n+2). \tag{1.3.15}$$

The zeta function associated to this Laplacian (restricted to the vector spherical harmonics) is

$$\zeta_{\Delta^1}(s) = \sum_{n=1}^{\infty} \frac{2n(n+2)}{(n+1)^{2s}} = 2\sum_{m=1}^{\infty} \frac{m^2-1}{m^{2s}} = 2\zeta(2s-2) - 2\zeta(2s), \tag{1.3.16}$$

and

$$\log \det{}' \Delta^1 = -4\zeta'(-2) - 2\log(2\pi) = -2\log(2\pi) + \frac{\zeta(3)}{\pi^2}. \tag{1.3.17}$$

Here we have used that

$$\zeta'(0) = -\frac{1}{2}\log(2\pi). \tag{1.3.18}$$

We conclude that

$$\log \tau'_R(\mathbb{S}^3) = \log \det{}' \Delta^{(0)} - \frac{1}{2}\log \det{}' \Delta^{(1)} = \log(2\pi^2). \tag{1.3.19}$$

This is in agreement with the calculation of the analytic torsion for general spheres in for example [64]. In view of (1.2.62), and since

$$\mathrm{vol}(\mathbb{S}^3) = 2\pi^2, \tag{1.3.20}$$

we find

$$\tau_R(\mathbb{S}^3) = 1. \tag{1.3.21}$$

One can also calculate the invariant (1.2.59) directly, since this is d_G times

$$\zeta_{\Delta^1}(0) = -2\zeta(0) = 1, \tag{1.3.22}$$

and it agrees with (1.2.66). Since the eigenvalues of $*\mathrm{d}$ on the vector spherical harmonics come in pairs with the same absolute value but opposite signs (see (A.51)), $\eta_{\mathrm{grav}} = 0$. We conclude that, for $k > 0$,

$$Z_{1-\mathrm{loop}}(\mathbb{S}^3) = \frac{1}{\mathrm{vol}(G)}\left(\frac{k}{4\pi^2}\right)^{-\frac{d_G}{2}}. \tag{1.3.23}$$

In particular, for $G = U(N)$ we have

$$Z_{1-\mathrm{loop}}(\mathbb{S}^3) = \frac{1}{\mathrm{vol}(U(N))}\left(\frac{k}{4\pi^2}\right)^{-\frac{N^2}{2}} \tag{1.3.24}$$

The volume of $U(N)$ is given by

$$\mathrm{vol}(U(N)) = \frac{(2\pi)^{\frac{1}{2}N(N+1)}}{G_2(N+1)}, \tag{1.3.25}$$

where $G_2(z)$ is the Barnes function, defined by

$$G_2(z+1) = \Gamma(z)G_2(z), \quad G_2(1) = 1. \tag{1.3.26}$$

Notice that

$$G_2(N+1) = (N-1)!(N-2)! \cdots 2!1!. \tag{1.3.27}$$

As we mentioned before, the partition function of Chern–Simons theory on $\mathbb{S}^3$ can be computed exactly for any gauge group. In the case of $U(N)$, the answer is, for $k > 0$ [46]

$$Z(\mathbb{S}^3) = \frac{1}{(k+N)^{N/2}} \prod_{j=1}^{N-1} \left(2 \sin \frac{\pi j}{k+N}\right)^{N-j}. \tag{1.3.28}$$

Here an explicit choice of framing has been made, but one can compute the partition function for any choice of framing by simply using (1.2.78). The expansion for large k should reproduce the perturbative result, and in particular the leading term should agree with the result (1.3.24). Indeed, we have for k large

$$Z(\mathbb{S}^3) \approx k^{-N/2} \prod_{j=1}^{N-1} \left(\frac{2\pi j}{k}\right)^{N-j} = (2\pi)^{\frac{1}{2}N(N-1)}\, k^{-N^2/2} G_2(N+1), \tag{1.3.29}$$

which is exactly (1.3.24).

1.3.2 Matter fields

The supersymmetric multiplet contains a conformally coupled complex scalar and a fermion, both in a representation R of the gauge group. The partition function at one loop is just given by the quotient of functional determinants

$$Z^{\text{matter}}_{1-\text{loop}} = \left(\frac{\det(-\mathrm{i}\not{D})}{\det \Delta_c}\right)^{d_R} \tag{1.3.30}$$

where d_R is the dimension of the representation, and

$$\Delta_c = \Delta^0 + \frac{3}{4} \tag{1.3.31}$$

is the conformal Laplacian (we have set again $r = 1$). We now compute these determinants.

The eigenvalues of the conformal Laplacian are simply

$$n(n+2) + \frac{3}{4}, \qquad n = 0, 1, \cdots, \tag{1.3.32}$$

with the same multiplicity as the standard Laplacian, namely $(n+1)^2$. We then have

$$\zeta_{\Delta_c}(s) = \sum_{n=0} \frac{(n+1)^2}{\left(n(n+2)+\frac{3}{4}\right)^s} = \sum_{m=1}^{\infty} \frac{m^2}{\left(m^2 - \frac{1}{4}\right)^s}. \tag{1.3.33}$$

As in the case of the standard Laplacian, we split

$$\frac{m^2}{\left(m^2-\frac{1}{4}\right)^s} = \frac{1}{m^{2(s-1)}} + \frac{s}{4m^{2s}} + R_c(m,s), \tag{1.3.34}$$

where

$$R_c(m,s) = \frac{m^2}{\left(m^2-\frac{1}{4}\right)^s} - \frac{1}{m^{2(s-1)}} - \frac{s}{4m^{2s}}. \tag{1.3.35}$$

The derivative of $R_c(m,s)$ at $s = 0$ is

$$\left.\frac{\mathrm{d}R_c(m,s)}{\mathrm{d}s}\right|_{s=0} = -\frac{1}{4} - m^2 \log\left(1 - \frac{1}{4m^2}\right). \tag{1.3.36}$$

The sum of this series can be explicitly calculated as

$$-\sum_{m=1}^{\infty} \left[\frac{1}{4} + m^2 \log\left(1 - \frac{1}{4m^2}\right)\right] = \frac{1}{8} - \frac{1}{4}\log(2) + \frac{7\zeta(3)}{8\pi^2}. \tag{1.3.37}$$

We then find

$$\zeta_{\Delta_c}(s) = \zeta(2s-2) + \frac{s}{4}\zeta(2s) + \sum_{m=1}^{\infty} R_c(m,s), \tag{1.3.38}$$

and we conclude that the determinant of the conformal Laplacian on $\mathbb{S}^3$ is given by

$$\log \det \Delta_c = -\zeta'_{\Delta_c}(0) = \frac{1}{4}\log(2) - \frac{3\zeta(3)}{8\pi^2}. \tag{1.3.39}$$

This is in agreement with the result quoted in the Erratum to [65][1].

Let us now consider the determinant (in absolute value) for the spinor field. We have, using (A.66),

$$\zeta_{|\not D|}(s) = 2\sum_{n=1}^{\infty} \frac{n(n+1)}{\left(n+\frac{1}{2}\right)^s}. \tag{1.3.40}$$

After a small manipulation we can write it as

$$\begin{aligned}\zeta_{|\not D|}(s) =& 2\cdot 2^{s-2}\left\{\sum_{m\geq 1}\frac{1}{(2m+1)^{s-2}} - \sum_{m\geq 1}\frac{1}{(2m+1)^s}\right\}\\ =& 2\left(2^{s-2}-1\right)\zeta(s-2) - \frac{1}{2}\left(2^s-1\right)\zeta(s),\end{aligned} \tag{1.3.41}$$

where we have used that

$$\sum_{m\geq 0}\frac{1}{(2m+1)^s} = \sum_{n\geq 1}\frac{1}{n^s} - \sum_{k\geq 1}\frac{1}{(2k)^s} = \left(1-2^{-s}\right)\zeta(s). \tag{1.3.42}$$

The regularized number of negative eigenvalues of this operator is given by $\zeta_{|\not D|}(0)/2$ and it vanishes, so the determinant of the Dirac operator equals its absolute value. We deduce

$$\log \det\left(-\mathrm{i}\not D\right) = -\zeta'_{|\not D|}(0) = -\frac{3}{8\pi^2}\zeta(3) - \frac{1}{4}\log 2. \tag{1.3.43}$$

Combining the conformal scalar determinant with the spinor determinant we obtain,

$$\log \det\left(-\mathrm{i}\not D\right) - \log \det \Delta_c = -\frac{1}{2}\log 2. \tag{1.3.44}$$

This can be seen directly at the level of eigenvalues. The quotient of determinants is

$$\prod_{m=1}^{\infty}\frac{\left(m+\frac{1}{2}\right)^{m(m+1)}\left(m-\frac{1}{2}\right)^{m(m-1)}}{\left(m^2-\frac{1}{4}\right)^{m^2}} = \prod_{m=1}^{\infty}\frac{\left(m+\frac{1}{2}\right)^m}{\left(m-\frac{1}{2}\right)^m} \tag{1.3.45}$$

and its regularization leads directly to the result above (see Appendix C). We conclude that

$$Z^{\text{matter}}_{1-\text{loop}} = 2^{-d_R/2}. \tag{1.3.46}$$

1.3.3 ABJM theory at weak coupling

We can now calculate the free energy on $\mathbb{S}^3$ of ABJM theory. We will restrict ourselves to the "ABJM slice" where the two gauge groups have the same rank, i.e. the theory originally considered in [1]. We have two copies of CS theory with gauge group $U(N)$ and opposite levels k, $-k$, together with four chiral multiplets in the bifundamental representation of $U(N)\times U(N)$. Keeping the first term (one-loop) in perturbation theory we find, at one-loop,

$$F_{\text{ABJM}}(\mathbb{S}^3) \approx -N^2\log\left(\frac{k}{4\pi^2}\right) - 2\log\left(\text{vol}(U(N))\right) - 2N^2\log(2) \tag{1.3.47}$$

where the first two terms come from the CS theories, and the last term comes from the supersymmetric matter. Here we assume $k>0$. Notice that the theory with opposite level $-k$ gives the same contribution

[1] Beware: the arXiv version of this paper gives a wrong result for this determinant.

as the theory with level k. In order to obtain the planar limit of this quantity, we have to expand the volume of $U(N)$ at large N. Using the asymptotic expansion of the Barnes function

$$\begin{aligned}\log G_2(N+1) &= \frac{N^2}{2}\log N - \frac{1}{12}\log N - \frac{3}{4}N^2 + \frac{1}{2}N\log 2\pi + \zeta'(-1) \\ &+ \sum_{g=2}^{\infty}\frac{B_{2g}}{2g(2g-2)}N^{2-2g},\end{aligned} \tag{1.3.48}$$

where B_{2g} are the Bernoulli numbers, we finally obtain the weakly coupled, planar result

$$F_{\rm ABJM}(\mathbb{S}^3) \approx N^2\left\{\log(2\pi\lambda) - \frac{3}{2} - 2\log(2)\right\}. \tag{1.3.49}$$

1.4 Strong coupling and AdS duals

Some SCFTs in three dimensions have AdS duals given by M-theory/string theory on backgrounds of the form

$$X = {\rm AdS}_4 \times X_{6,7}, \tag{1.4.1}$$

where $X_{6,7}$ is a six-dimensional or seven-dimensional compactification manifold, depending on whether we consider a superstring or an M-theory dual, respectively. One of the consequences of the AdS/CFT duality is that the partition function of the Euclidean gauge theory on $\mathbb{S}^3$ should be equal to the partition function of the Euclidean version of M-theory/string theory on the dual AdS background [66], i.e.

$$Z_{\rm CFT}(\mathbb{S}^3) = Z(X). \tag{1.4.2}$$

In the large N limit, we can compute the r.h.s. in classical (i.e. genus zero) string theory, and at strong coupling it is sufficient to consider the supergravity approximation. This means that the partition function of the strongly coupled gauge theory on $\mathbb{S}^3$ in the planar limit should be given by

$$Z_{\rm CFT}(\mathbb{S}^3) \approx {\rm e}^{-I({\rm AdS}_4)} \tag{1.4.3}$$

where I is the classical gravity action evaluated on the $\rm AdS_4$ metric. This gives a prediction for the strongly coupled behavior of the gauge theory. However, the gravitational action on $\rm AdS_4$ is typically divergent, and it has to be regularized in order to obtain finite results. We will now review the method of holographic renormalization and its application to the calculation of the free energy of ABJM theory on $\mathbb{S}^3$.

1.4.1 Holographic renormalization

The gravitational action in an Euclidean space with boundary has two contributions. The first contribution is the bulk term, given by the Einstein–Hilbert action

$$I_{\rm bulk} = -\frac{1}{16\pi G_N}\int_M {\rm d}^{n+1}x\,\sqrt{G}\,(R-2\Lambda) \tag{1.4.4}$$

where G_N is Newton's constant in $n+1$ dimensions and G is the $(n+1)$-dimensional metric. The second contribution is the surface term [67]

$$I_{\rm surf} = -\frac{1}{8\pi G_N}\int_{\partial M} K|\gamma|^{1/2}{\rm d}^n x, \tag{1.4.5}$$

where ∂M is the boundary of spacetime, γ is the metric induced by G on the boundary, and K is the extrinsic curvature of the boundary. K satisfies the useful relation (see for example [68])

$$\sqrt{\gamma}K = \mathcal{L}_n\sqrt{\gamma} \tag{1.4.6}$$

where n is the normal unit vector to ∂M, and $\mathcal{L}_n$ is the Lie derivative along this vector. Both actions, when computed on an AdS background, diverge due to the non-compactness of the space. For example, after using Einstein's equations, the bulk action of an AdS space of radius L can be written as

$$I_{\rm bulk} = \frac{n}{8\pi G_N L^2}\int {\rm d}^{n+1}x\sqrt{G} \tag{1.4.7}$$

which is proportional to the volume of space-time, and it is divergent.

In order to use the AdS/CFT correspondence, we have to regularize the gravitational action in an appropriate way. The procedure which has emerged in studies of the AdS/CFT correspondence is to introduce a set of *universal counterterms*, depending only on the induced metric on the boundary, which lead to finite values of the gravitational action, energy-momentum tensor, etc. This procedure gives values for the gravitational quantities in agreement with the corresponding quantities computed in the CFT side, and we will adopt it here. It is sometimes called "holographic renormalization" and it has been developed in for example [69, 70, 71, 72]. We now present the basics of holographic renormalization in AdS. Useful reviews, focused on AdS_5, can be found in for example [73, 74].

An asymptotically AdS metric in $n+1$ dimensions with radius L and cosmological constant

$$\Lambda = -\frac{n(n-1)}{2L^2} \tag{1.4.8}$$

can be written near its boundary at $u=0$ as

$$\mathrm{d}s^2 = L^2\left[\frac{\mathrm{d}u^2}{u^2} + \frac{1}{u^2}g_{ij}(u^2,x)\mathrm{d}x^i\mathrm{d}x^j\right]. \tag{1.4.9}$$

The metric $g_{ij}(u^2,x)$ can be expanded in a power series in u near $u=0$,

$$g_{ij}(u^2,x) = g_{ij}^{(0)}(x) + u^2 g_{ij}^{(2)}(x) + u^4\left[g^{(4)}(x) + \log(u^2)h^{(4)}(x)\right] + \cdots \tag{1.4.10}$$

where the first term, $g_{ij}^{(0)}$, is the metric of the CFT on the boundary. The coefficients $g_{ij}^{(2n)}$ appearing here can be solved recursively in terms of $g_{ij}^{(0)}$ by plugging (1.4.10) in Einstein's equations. One finds, for example [72][2]

$$g_{ij}^{(2)} = -\frac{1}{n-2}\left(R_{ij} - \frac{1}{2(n-1)}Rg_{ij}^{(0)}\right), \tag{1.4.11}$$

where R_{ij} and R are the Ricci tensor and curvature of $g_{ij}^{(0)}$. The resulting metric is then used to compute the gravitational action with a cut-off at $u=\epsilon$ which regulates the divergences,

$$I_\epsilon = -\frac{1}{16\pi G_N}\int_{M_\epsilon}\mathrm{d}^{n+1}x\,\sqrt{G}\left(R + \frac{n(n-1)}{L^2}\right) - \frac{1}{8\pi G_N}\int_{\partial M_\epsilon}K|\gamma|^{1/2}\mathrm{d}^n x. \tag{1.4.12}$$

Here, M_ϵ is the manifold with $u \geq \epsilon$ and a boundary ∂M_ϵ at $u=\epsilon$. To calculate the boundary term, we consider the normal vector to the hypersurfaces of constant u,

$$n^u = -\frac{u}{L}. \tag{1.4.13}$$

The minus sign is due to the fact that the boundary is at $u=0$, so that the normal vector points towards the origin. The induced metric is

$$\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j = \frac{L^2}{u^2}g_{ij}(u^2,x)\mathrm{d}x^i\mathrm{d}x^j, \tag{1.4.14}$$

with element of volume

$$\sqrt{\gamma} = \left(\frac{L}{u}\right)^n\sqrt{g}. \tag{1.4.15}$$

The intrinsic curvature of the hypersurface at constant u is then

$$\sqrt{\gamma}K = \mathcal{L}_n\sqrt{\gamma} = -\frac{u}{L}\partial_u\left[\left(\frac{L}{u}\right)^n\sqrt{g}\right] = \frac{nL^{n-1}}{u^n}\left(1 - \frac{1}{n}u\partial_u\right)\sqrt{g}. \tag{1.4.16}$$

We then find,

$$I_\epsilon = \frac{nL^{n-1}}{8\pi G_N}\int\mathrm{d}^n x\int_\epsilon\frac{\mathrm{d}u}{u^{n+1}}\sqrt{g} - \frac{nL^{n-1}}{8\pi G_N\epsilon^n}\int\mathrm{d}^n x\left(1 - \frac{1}{n}u\partial_u\right)\sqrt{g}\Big|_{u=\epsilon}. \tag{1.4.17}$$

[2]The sign in the curvature is opposite to the conventions in [72], which give a positive curvature to AdS.

The singularity structure of this regulated action is [69, 72]

$$I_\epsilon = \frac{L^{n-1}}{16\pi G_N}\int \mathrm{d}^n x\sqrt{g^{(0)}}\left(\epsilon^{-n}a_{(0)} + \epsilon^{-n+2}a_{(2)} + \cdots - 2\log(\epsilon)a_{(n)}\right) + \mathcal{O}(\epsilon^0). \tag{1.4.18}$$

The logarithmic divergence appears only when n is even. In order to regularize power-type divergences in $n = 3$ and $n = 4$, it suffices to calculate the first two coefficients, $a_{(0)}$ and $a_{(2)}$. Let us now calculate these coefficients (the next two are computed in [72]). We first expand,

$$\det g = \det g^{(0)}\left(1 + u^2\mathrm{Tr}\left(g^{(0)-1}g^{(2)}\right) + \cdots\right), \tag{1.4.19}$$

so that

$$\begin{aligned}\sqrt{g(u^2,x)} &= \sqrt{g^{(0)}}\left(1 + \frac{u^2}{2}\mathrm{Tr}\left(g^{(0)-1}g^{(2)}\right) + \cdots\right),\\ \left(1 - \frac{1}{n}u\partial_u\right)\sqrt{g(u^2,x)} &= \sqrt{g^{(0)}}\left(1 + \frac{n-2}{2n}u^2\mathrm{Tr}\left(g^{(0)-1}g^{(2)}\right) + \cdots\right).\end{aligned} \tag{1.4.20}$$

The regulated Einstein–Hilbert action gives

$$\frac{nL^{n-1}}{8\pi G_N}\int \mathrm{d}^n x\,\sqrt{g^{(0)}}\left[\frac{1}{n\epsilon^n} + \frac{1}{2(n-2)\epsilon^{n-2}}\mathrm{Tr}\left(g^{(0)-1}g^{(2)}\right) + \cdots\right], \tag{1.4.21}$$

while the regulated Gibbons–Hawking term gives

$$-\frac{nL^{n-1}}{8\pi G_N}\int \mathrm{d}^n x\,\sqrt{g^{(0)}}\left[\frac{1}{\epsilon^n} + \frac{n-2}{2n\epsilon^{n-2}}\mathrm{Tr}\left(g^{(0)-1}g^{(2)}\right) + \cdots\right]. \tag{1.4.22}$$

In total, we find

$$I_\epsilon = \frac{L^{n-1}}{16\pi G_N}\int \mathrm{d}^n x\,\sqrt{g^{(0)}}\left[\frac{2(1-n)}{\epsilon^n} - \frac{n^2-5n+4}{(n-2)\epsilon^{n-2}}\mathrm{Tr}\left(g^{(0)-1}g^{(2)}\right) + \cdots\right], \tag{1.4.23}$$

and we deduce,

$$\begin{aligned} a_{(0)} &= 2(1-n),\\ a_{(2)} &= -\frac{(n-4)(n-1)}{n-2}\mathrm{Tr}\left(g^{(0)-1}g^{(2)}\right), \quad n \neq 2.\end{aligned} \tag{1.4.24}$$

The counterterm action is obtained by using a gravitational analogue of the minimal subtraction scheme, and it is given by minus the divergent part of I_ϵ,

$$\begin{aligned} I_{\mathrm{ct}} = \frac{L^{n-1}}{16\pi G_N}\int \mathrm{d}^n x\,\sqrt{g^{(0)}}\Bigg[&\frac{2(n-1)}{\epsilon^n} + \frac{(n-4)(n-1)}{(n-2)\epsilon^{n-2}}\mathrm{Tr}\left(g^{(0)-1}g^{(2)}\right) + \cdots\\ &+ 2\log(\epsilon)a_{(n)}\Bigg].\end{aligned} \tag{1.4.25}$$

As pointed out in [70], we should re-write this in terms of the induced metric in the boundary (1.4.14), evaluated at $u = \epsilon$. From (1.4.20) we deduce

$$\sqrt{g^{(0)}} = \left(\frac{\epsilon}{L}\right)^n\left(1 - \frac{\epsilon^2}{2}\mathrm{Tr}\left(g^{(0)-1}g^{(2)}\right) + \mathcal{O}(\epsilon^4)\right)\sqrt{\gamma}. \tag{1.4.26}$$

On the other hand, from (1.4.11) we obtain

$$\begin{aligned}\mathrm{Tr}\left(g^{(0)-1}g^{(2)}\right) &= -\frac{1}{n-2}\left(g^{(0)}_{ij}R^{ij} - \frac{1}{2(n-1)}Rg^{(0)}_{ij}g^{(0)ij}\right) = -\frac{1}{2(n-1)}R\\ &= -\frac{L^2}{2(n-1)\epsilon^2}R[\gamma] + \cdots,\end{aligned} \tag{1.4.27}$$

where in the first line the Ricci tensor and curvature are computed for $g^{(0)}_{ij}$, while in the second line the curvature is computed for the induced metric γ. Plugging these results into the counterterm action we find

$$
\begin{aligned}
I_{\rm ct} = & \frac{1}{16\pi G_N L}\int \mathrm{d}^n x\,\sqrt{\gamma}\left(1+\frac{L^2}{4(n-1)}R[\gamma]+\cdots\right)\\
& \times\left[2(n-1)-\frac{n-4}{2(n-2)}L^2R[\gamma]+2\log(\epsilon)a_{(n)}[\gamma]+\cdots\right]\\
= & \frac{1}{8\pi G_N}\int \mathrm{d}^n x\,\sqrt{\gamma}\left(2\log(\epsilon)a_{(n)}[\gamma]+\frac{n-1}{L}+\frac{L}{2(n-2)}R[\gamma]+\cdots\right),
\end{aligned}
\tag{1.4.28}
$$

which is the result written down in [71, 72] (for Euclidean signature). This is the counterterm action which is relevant for AdS in four and five dimensions, and the dots denote higher order counterterms (in the Riemann tensor of the induced metric) which are needed for higher dimensional spaces [70, 71, 72]. The total, regularized gravitational action is then

$$
I = I_{\rm bulk}+I_{\rm surf}+I_{\rm ct} \tag{1.4.29}
$$

and it yields a finite result by construction. The removal of these IR divergences in the gravitational theory is dual to the removal of UV divergences in the CFT theory. It can be verified in examples that the gravity answers obtained by holographic renormalization match the answers obtained in CFT on a curved background after using zeta-function regularization [70]. In the next subsection we work out a beautiful example of this matching closely related to the techniques developed here, namely the Casimir energy for $\mathcal{N}=4$ super Yang–Mills on $\mathbb{R}\times\mathbb{S}^3$, which was first derived in [70].

1.4.2 Free energy in AdS$_4$

We are interested in studying CFTs on $\mathbb{S}^n$. Therefore, in the AdS dual we need the Euclidean version of the AdS metric with that boundary, which can be written as [71]

$$
\mathrm{d}s^2=\frac{\mathrm{d}r^2}{1+r^2/L^2}+r^2\mathrm{d}\Omega_n^2 \tag{1.4.30}
$$

so that the boundary is at $r\to\infty$. This metric can be also written as [66, 71]

$$
\mathrm{d}s^2=L^2\left(\mathrm{d}\rho^2+\sinh^2(\rho)\mathrm{d}\Omega_n^2\right). \tag{1.4.31}
$$

Let us compute the regularized gravitational action (1.4.29), with a cutoff at the boundary ∂M located at constant r. The element of volume of the metric G given in (1.4.30) is

$$
\sqrt{G}=\frac{r^n}{\sqrt{1+r^2/L^2}}\sqrt{g_{\mathbb{S}^n}}, \tag{1.4.32}
$$

where $g_{\mathbb{S}^n}$ is the metric on an n-sphere of unit radius. The bulk action is just (1.4.7), i.e.

$$
I_{\rm bulk}=\frac{n\mathrm{vol}(\mathbb{S}^n)}{8\pi G_N L}\int_0^r \mathrm{d}\rho\frac{\rho^n}{\sqrt{L^2+\rho^2}}. \tag{1.4.33}
$$

To calculate the surface action, we notice that the unit normal vector to ∂M is

$$
n=\sqrt{1+r^2/L^2}\frac{\partial}{\partial r} \tag{1.4.34}
$$

while the induced metric is

$$
\gamma=r^2 g_{\mathbb{S}^n} \tag{1.4.35}
$$

which has the element of volume

$$
\sqrt{\gamma}=r^n\sqrt{g_{\mathbb{S}^n}} \tag{1.4.36}
$$

and scalar curvature

$$
R[\gamma]=\frac{R_{\mathbb{S}^n}}{r^2}=\frac{n(n-1)}{r^2}. \tag{1.4.37}
$$

We then obtain

$$
\mathcal{L}_n\sqrt{\gamma}=nr^{n-1}\sqrt{1+r^2/L^2}\sqrt{g_{\mathbb{S}^n}}, \tag{1.4.38}
$$

and the surface term is

$$I_{\rm surf} = -\frac{nr^{n-1}}{8\pi G_N}\sqrt{1+r^2/L^2}\,{\rm vol}(\mathbb{S}^n). \tag{1.4.39}$$

Finally, the first two counterterms are given by

$$\frac{{\rm vol}(\mathbb{S}^n)}{8\pi G_N}\left[\frac{n-1}{L}r^n + \frac{Lr^{n-2}n(n-1)}{2(n-2)}\right] = \frac{{\rm vol}(\mathbb{S}^n)}{8\pi G_N}\frac{r^n(n-1)}{L}\left[1+\frac{n}{2(n-2)}\frac{L^2}{r^2}\right]. \tag{1.4.40}$$

Putting everything together we obtain,

$$\begin{aligned} I = \frac{{\rm vol}(\mathbb{S}^n)}{8\pi G_N L}\Bigg[& nL^n\int_0^{r/L}{\rm d}u\frac{u^n}{\sqrt{1+u^2}} - nr^{n-1}\sqrt{r^2+L^2} \\ & + r^n(n-1)\left(1+\frac{n}{2(n-2)}\frac{L^2}{r^2}\right)\Bigg]. \end{aligned} \tag{1.4.41}$$

We should now take the limit of this expression when $r\to\infty$. For $n=3$ we find

$$I = \frac{{\rm vol}(\mathbb{S}^3)}{8\pi G_N L}\left(2L^3 + \mathcal{O}(r^{-1})\right), \tag{1.4.42}$$

therefore we obtain a *finite* action given by

$$I = \frac{\pi L^2}{2G_N}. \tag{1.4.43}$$

This will give us the strong coupling prediction for the free energy on $\mathbb{S}^3$ of supersymmetric Chern–Simons–matter theories with an AdS dual.

1.4.3 ABJM theory and its AdS dual

In order to compute the free energy of ABJM theory at strong coupling we have to be more precise about the gauge/gravity dictionary. We will now write down this dictionary for supersymmetric Chern–Simons–matter theories, which was first established for ABJM theory in [1]. We will not attempt here to review the derivation of the duality. A pedagogical introduction can be found in [75].

The AdS duals to the theories we will consider are given by M-theory on

$${\rm AdS}_4\times X_7, \tag{1.4.44}$$

where X_7 is a seven-dimensional manifold. In the case of ABJM theory,

$$X_7 = \mathbb{S}^7/\mathbb{Z}_k. \tag{1.4.45}$$

The eleven-dimensional metric and four-form flux are given by the Freund–Rubin background (see [76] for a review)

$$\begin{aligned} {\rm d}s^2_{11} &= L^2_{X_7}\left(\frac{1}{4}{\rm d}s^2_{{\rm AdS}_4} + {\rm d}s^2_{X_7}\right), \\ F &= \frac{3}{8}L^3_{X_7}\omega_{{\rm AdS}_4}, \end{aligned} \tag{1.4.46}$$

where $\omega_{{\rm AdS}_4}$ is the volume form with unit radius. The radius L_{X_7} is determined by the flux quantization condition

$$(2\pi\ell_p)^6 Q = \int_{X_7}\star_{11}F = 6L^6_{X_7}{\rm vol}(X_7). \tag{1.4.47}$$

In this equation, ℓ_p is the eleven-dimensional Planck length. The charge Q is given, at large radius, by the number of M2 branes N, but it receives corrections [77, 78]. In ABJM theory we have

$$Q = N - \frac{1}{24}\left(k-\frac{1}{k}\right). \tag{1.4.48}$$

This extra term comes from the coupling

$$\int C_3\wedge I_8 \tag{1.4.49}$$

in M-theory, which contributes to the charge of M2 branes. Here, I_8 is proportional to the Euler density in eight dimensions, and it satisfies

$$\int_{M_8} I_8 = -\frac{\chi}{24} \tag{1.4.50}$$

where M_8 is a compact eight-manifold. In ABJM theory, the relevant eight-manifold is $\mathbb{C}^4/\mathbb{Z}_k$, with regularized Euler characteristic

$$\chi\left(\mathbb{C}^4/\mathbb{Z}_k\right) = k - \frac{1}{k}. \tag{1.4.51}$$

This leads to the shift in (1.4.48).

One final ingredient that we will need is Newton's constant in four dimensions. It can be obtained from the Einstein–Hilbert action in eleven dimensions, which leads to its four-dimensional counterpart by standard Kaluza–Klein reduction,

$$\frac{1}{16\pi G_{11}} \int \mathrm{d}^{11}x\, \sqrt{g_{11}} R_{11} \to \frac{1}{4} \cdot \frac{L_{X_7}^7}{16\pi G_{11}} \mathrm{vol}\,(X_7) \int \mathrm{d}^4x\, \sqrt{g_4} R_4 = \frac{1}{16\pi G_N} \int \mathrm{d}^4x\, \sqrt{g_4} R_4, \tag{1.4.52}$$

where G_{11}, G_N denote the eleven-dimensional and the four-dimensional Newton's constant, respectively, and the volume of X_7 is calculated for unit radius. In the resulting Einstein–Hilbert action in four dimensions, the metric and scalar curvature refer to an AdS_4 space of radius L_{X_7}, and not $L_{X_7}/2$ as in (1.4.46). This is the source for the extra factor of $1/4$ in the second term. Recalling that

$$16\pi G_{11} = \frac{1}{2\pi} \left(2\pi \ell_p\right)^9, \tag{1.4.53}$$

we obtain

$$\frac{1}{G_N} = \frac{2\sqrt{6}\pi^2 Q^{3/2}}{9\sqrt{\mathrm{vol}(X_7)}} \frac{1}{L_{X_7}^2}. \tag{1.4.54}$$

It follows that the regularized gravitational action (1.4.43) is given by

$$I = \frac{\pi L_{X_7}^2}{2G_N} = Q^{3/2} \sqrt{\frac{2\pi^6}{27\mathrm{vol}(X_7)}}. \tag{1.4.55}$$

In particular, for ABJM theory we have

$$I = \frac{\pi\sqrt{2}}{3} k^{1/2} Q^{3/2}. \tag{1.4.56}$$

In the supergravity and planar approximation we can just set $Q = N$, and we find indeed that the planar free energy is given by

$$-\frac{1}{N^2} F(\mathbb{S}^3) \approx \sqrt{\frac{2\pi^6}{27\mathrm{vol}(X_7)}} \frac{1}{N^{1/2}}. \tag{1.4.57}$$

In the next chapter we will reproduce this strong coupling result from the gauge theory side.

1.5 Localization

Localization is an ubiquitous technique in supersymmetric QFT which makes possible to reduce an infinite-dimensional path integral to a finite dimensional integral. It features prominently in Witten's topological quantum field theories of the cohomological type, where one can argue that the semiclassical approximation is exact, see [79, 83] for reviews and a list of references.

The basic idea of localization is the following. Let δ be a Grassmann-odd symmetry of a theory with action $S(\phi)$, where ϕ denotes the set of fields in the theory. We assume that the measure of the path integral is invariant under δ as well (i.e. δ is not anomalous), and that

$$\delta^2 = \mathcal{L}_B \tag{1.5.1}$$

where $\mathcal{L}_B$ is a Grassmann-even symmetry. In a Lorentz-invariant, gauge invariant theory, $\mathcal{L}_B$ could be a combination of a Lorentz and a gauge transformation. Consider now the perturbed partition function

$$Z(t) = \int \mathcal{D}\phi\, \mathrm{e}^{-S - t\delta V}, \tag{1.5.2}$$

where V is a Grassmann-odd operator which is invariant under $\mathcal{L}_B$. It is easy to see that $Z(t)$ is independent of t, since

$$\frac{\mathrm{d}Z}{\mathrm{d}t} = -\int \mathcal{D}\phi\, \delta V\, \mathrm{e}^{-S-t\delta V} = -\int \mathcal{D}\phi\, \delta \left(V \mathrm{e}^{-S-t\delta V}\right) = 0. \tag{1.5.3}$$

Here we have used the fact that $\delta^2 V = \mathcal{L}_B V = 0$. In the final step we have used the fact that δ is a symmetry of the path integral, in order to interpret the integrand as a total derivative. In some cases, the integral of the total derivative does not vanish due to boundary terms (a closely related example appears in section 11.3 of [80]), but if the integral decays sufficiently fast in field space one expects the perturbed partition function $Z(t)$ to be independent of t. This means that it can be computed at $t = 0$ (where one recovers the original partition function) but also for other values of t, like $t \to \infty$. In this regime, simplifications typically occur. For example, if δV has a positive definite bosonic part $(\delta V)_B$, the limit $t \to \infty$ localizes the integral to a submanifold of field space where

$$(\delta V)_B = 0. \tag{1.5.4}$$

It turns out that, in many interesting examples, this submanifold is finite-dimensional. This leads to a “collpase" of the path integral to a finite-dimensional integral. It is easy to see that this method also makes it possible to calculate the correlation functions of δ-invariant operators.

In order to see how the method of localization works, let us briefly review a beautiful and simple example, namely the field theoretical version of the Poincaré–Hopf theorem.

1.5.1 A simple example of localization

The Poincaré–Hopf theorem has been worked out from the point of view of supersymmetric localization in many references, like for example [79, 81, 83]. Let X be a Riemannian manifold of even dimension n, with metric $g_{\mu\nu}$, vierbein e^a_μ, and let V_μ be a vector field on X. We will consider the following “supercoordinates" on the tangent bundle TX

$$\left(x^\mu, \psi^\mu\right), \qquad \left(\bar\psi_\mu, B_\mu\right), \tag{1.5.5}$$

where the first doublet represents supercoordinates on the base X, and the second doublet represents supercoordinates on the fiber. ψ^μ and $\bar\psi_\mu$ are Grasmann variables. The above supercoordinates are related by the Grasmannian symmetry

$$\begin{aligned} \delta x^\mu &= \psi^\mu, \qquad & \delta\bar\psi_\mu &= B_\mu, \\ \delta\psi^\mu &= 0, \qquad & \delta B_\mu &= 0, \end{aligned} \tag{1.5.6}$$

which squares to zero, $\delta^2 = 0$. With these fields we construct the “action"

$$S(t) = \delta\Psi, \qquad \Psi = \frac{1}{2}\bar\psi_\mu \left(B^\mu + 2\mathrm{i}tV^\mu + \Gamma^\sigma_{\tau\nu}\bar\psi_\sigma\psi^\nu g^{\mu\tau}\right), \tag{1.5.7}$$

and we define the partition function of the theory as

$$Z_X(t) = \frac{1}{(2\pi)^n}\int_X \mathrm{d}x\, \mathrm{d}\psi\, \mathrm{d}\bar\psi\, \mathrm{d}B\, \mathrm{e}^{-S(t)}. \tag{1.5.8}$$

Using that

$$\frac{\partial g^{\mu\sigma}}{\partial x^\tau} = -\Gamma^\mu_{\tau\sigma} g^{\tau\sigma} - \Gamma^\sigma_{\tau\nu} g^{\tau\mu}, \tag{1.5.9}$$

one finds that, in the resulting theory, B_μ is a Gaussian field with mean value

$$B^\mu = -\mathrm{i}tV^\mu - g^{\mu\tau}\Gamma^\sigma_{\tau\nu}\bar\psi_\sigma\psi^\nu. \tag{1.5.10}$$

If we integrate it out, we obtain an overall factor

$$\frac{(2\pi)^{n/2}}{\sqrt{g}}, \tag{1.5.11}$$

and the action becomes

$$\frac{t^2}{2} g_{\mu\nu}V^\mu V^\nu - \frac{1}{4}R^{\rho\sigma}{}_{\mu\nu}\bar\psi_\rho\bar\psi_\sigma\psi^\mu\psi^\nu - \mathrm{i}t\nabla_\mu V^\nu\bar\psi_\nu\psi^\mu. \tag{1.5.12}$$

We can define orthonormal coordinates on the fiber by using the inverse vierbein,

$$\chi_a = E_a^\mu \bar{\psi}_\mu, \tag{1.5.13}$$

so that the partition function reads

$$Z_X(t) = \frac{1}{(2\pi)^{n/2}} \int_X \mathrm{d}x\, \mathrm{d}\psi\, \mathrm{d}\chi\, \mathrm{e}^{-\frac{t^2}{2} g_{\mu\nu} V^\mu V^\nu + \frac{1}{4} R^{ab}{}_{\mu\nu} \chi_a \chi_b \psi^\mu \psi^\nu + \mathrm{i}t \nabla_\mu V^\nu e^a_\nu \chi_a \psi^\mu}. \tag{1.5.14}$$

It is clear that this partition function should be independent of t, since the action can be written as

$$S(t) = S(0) + t\delta V, \qquad V = \mathrm{i}\bar{\psi}_\mu V^\mu. \tag{1.5.15}$$

We can then evaluate it in different regimes: $t \to 0$ or $t \to \infty$. The calculation when $t = 0$ is very easy, since we just have

$$Z_X(0) = \frac{1}{(2\pi)^{n/2}} \int_X \mathrm{d}x\, \mathrm{d}\psi\, \mathrm{d}\chi\, \mathrm{e}^{\frac{1}{4} R^{ab}{}_{\mu\nu} \chi_a \chi_b \psi^\mu \psi^\nu} = \frac{1}{(2\pi)^{n/2}} \int_X \mathrm{d}x\, \mathrm{Pf}(R), \tag{1.5.16}$$

where we have integrated over the Grassmann variables χ_a to obtain the Pfaffian of the matrix R^{ab}. The resulting top form in the integrand,

$$e(X) = \frac{1}{(2\pi)^{n/2}} \mathrm{Pf}(R) \tag{1.5.17}$$

is nothing but the Chern–Weil representative of the Euler class, therefore the evaluation at $t = 0$ produces the Euler characteristic of X,

$$Z_X(0) = \chi(X). \tag{1.5.18}$$

Let us now calculate the partition function in the limit $t \to \infty$. We will now assume that V^μ has isolated, simple zeroes p_k where $V^\mu(p_k) = 0$. These are the saddle–points of the "path integral," so we can write $Z_X(t)$ as a sum over saddle–points p_k, and for each saddle–point we have to perform a perturbative expansion. Let ξ^μ be coordinates around the point p_k. We have the expansion,

$$V^\mu(x) = \sum_{n \geq 1} \frac{1}{n!} \partial_{\mu_1} \cdots \partial_{\mu_n} V^\mu(p_k) \xi^{\mu_1} \cdots \xi^{\mu_n}. \tag{1.5.19}$$

After rescaling the variables as

$$\xi \to t^{-1}\xi, \qquad \psi \to t^{-1/2}\psi, \qquad \chi \to t^{-1/2}\chi, \tag{1.5.20}$$

the theory becomes Gaussian in the limit $t \to \infty$, since higher order terms in the fluctuating fields ξ, ψ, χ contain at least a power $t^{-1/2}$. Interactions are suppressed, and the partition function is one-loop exact:

$$\lim_{t\to\infty} Z_X(t) = \sum_{p_k} \frac{1}{(2\pi)^{n/2}} \int_X \mathrm{d}\xi\, \mathrm{d}\psi\, \mathrm{d}\chi\, \mathrm{e}^{-\frac{1}{2} g_{\mu\nu} H^{(k)\mu}_\alpha H^{(k)\nu}_\beta \xi^\alpha \xi^\beta + \mathrm{i} H^{(k)\nu}_\mu e^a_\nu \chi_a \psi^\mu} \tag{1.5.21}$$

where we denoted,

$$H^{(k)\mu}_\sigma = \partial_\sigma V^\mu \big|_{p_k}. \tag{1.5.22}$$

Each term in this sum can now be computed as a product of a bosonic Gaussian integral, times a Grassmann integral, and we obtain

$$\lim_{t\to\infty} Z_X(t) = \sum_{p_k} \frac{1}{\sqrt{g} |\det\, H^{(k)}|} \det(e^a_\mu) \det H^{(k)} = \sum_{p_k} \frac{\det\, H^{(k)}}{|\det\, H^{(k)}|}. \tag{1.5.23}$$

The equality between (1.5.18) and (1.5.23) is the famous Poincaré–Hopf theorem.

Conceptually, the localization analysis in [2] that we will review now is not very different from this example, although technically it is more complicated. The key common ingredient in the analysis of the $t \to \infty$ limit is that the localization locus becomes very simple, and all Feynman diagrams involving at least two loops are suppressed by a factor $t^{-1/2}$, so that the one-loop approximation is exact.

1.5.2 Localization in Chern–Simons–matter theories: gauge sector

We are now ready to use the ideas of localization in supersymmetric Chern–Simons–matter theories on $\mathbb{S}^3$, following [2]. The Grassmann-odd symmetry is simply $\mathcal{Q}$, defined by $\delta_\epsilon = \epsilon\mathcal{Q}$, where ϵ is the conformal Killing spinor satisfying (1.1.19). This symmetry satisfies $\mathcal{Q}^2 = 0$, and then it is a suitable symmetry for localization. To localize in the gauge sector, we add to the CS-matter theory the term

$$-\,tS_{\mathrm{YM}}, \tag{1.5.24}$$

which thanks to (1.1.30) and (1.1.15) is of the form $\mathcal{Q}V$, and its bosonic part is positive definite. By the localization argument, the partition function of the theory (as well as the correlators of $\mathcal{Q}$-invariant operators) does not depend on t, and we can take $t \to \infty$. This forces the fields to take the values that make the bosonic part of (1.1.29) to vanish. Since this is a sum of positive definite terms, they have to vanish separately. We then have the localizing locus,

$$F_{\mu\nu} = 0, \qquad D_\mu\sigma = 0, \qquad D + \frac{\sigma}{r} = 0. \tag{1.5.25}$$

The first equation says that the gauge connection A_μ must be flat, but since we are on $\mathbb{S}^3$ the only flat connection is $A_\mu = 0$. Plugging this into the second equation, we obtain

$$\partial_\mu\sigma = 0 \Rightarrow \sigma = \sigma_0, \tag{1.5.26}$$

a constant. Finally, the third equation says that

$$D = -\frac{\sigma_0}{r}. \tag{1.5.27}$$

The localizing locus is indeed finite-dimensional: it is just the submanifold where σ and D are constant Hermitian matrices, and $A_\mu = 0$.

Let us now calculate the path integral over the vector multiplet in the limit $t \to \infty$. We have to perform a gauge fixing, and we will choose the standard covariant gauge (1.2.26) as in the case of Chern–Simons theory. The path integral to be calculated is

$$\frac{1}{\mathrm{Vol}(G)}(\det{}'\,\Delta^0)^{-\frac{1}{2}}\int_{\mathrm{Ker}\,\mathrm{d}^\dagger}\mathcal{D}A\int_{(\mathrm{Ker}\,\mathrm{d})^\perp}\mathcal{D}C\mathcal{D}\overline{C}\,\mathrm{e}^{\frac{\mathrm{i}k}{4\pi}S_{\mathrm{SCS}}-tS_{\mathrm{YM}}(A)+S_{\mathrm{ghosts}}(C,\overline{C},A)}, \tag{1.5.28}$$

where $C, \overline{C}$ are ghosts fields. As in the example of the Poincaré–Hopf theorem, we expand the fields around the localizing locus, and we set

$$\begin{aligned}\sigma &= \sigma_0 + \frac{1}{\sqrt{t}}\sigma', \\ D &= -\frac{\sigma_0}{r} + \frac{1}{\sqrt{t}}D', \\ A,\ \lambda,\ C &\to \frac{1}{\sqrt{t}}A,\ \frac{1}{\sqrt{t}}\lambda,\ \frac{1}{\sqrt{t}}C,\end{aligned} \tag{1.5.29}$$

where the factors of t are chosen to remove the overall factor of t in the Yang–Mills action. In the Yang–Mills Lagrangian, only the terms which are quadratic in the fluctuations survive in this limit, namely,

$$\begin{aligned}\frac{1}{2}\int\sqrt{g}\,\mathrm{d}^3x\,\mathrm{Tr}\Big(&-A^\mu\Delta A_\mu - [A_\mu,\sigma_0]^2 + \partial_\mu\sigma'\partial^\mu\sigma' + (D'+\sigma')^2 \\ &+ \mathrm{i}\bar{\lambda}\gamma^\mu\nabla_\mu\lambda + \mathrm{i}\bar{\lambda}[\sigma_0,\lambda] - \frac{1}{2}\bar{\lambda}\lambda + \partial_\mu\overline{C}\partial^\mu C\Big),\end{aligned} \tag{1.5.30}$$

where we set $r = 1$. We are then left with a Gaussian theory, but with non-trivial quadratic operators for the fluctuations. In the same way, when we expand (1.1.21) around the fixed-point limit (1.5.29), we obtain

$$\frac{\mathrm{i}k}{4\pi}S_{\mathrm{SCS}} = \frac{\mathrm{i}k}{\pi}\mathrm{Tr}(\sigma_0^2)\mathrm{vol}(\mathbb{S}^3) + \mathcal{O}(t^{-1/2}), \tag{1.5.31}$$

so only the first term survives in the $t \to \infty$ limit.

Let us now calculate the path integral when $t \to \infty$. Like in the example of the Poincaré–Hopf theorem, we just have to compute the one–loop determinants. In this calculation we will only take into account the factors which depend explicitly on σ_0. The remaining, numerical factors (which might depend on N, but not on the coupling constant k) can be incorporated afterwards by comparing to the weak coupling results. The integral over the fluctuation D' can be done immediately. It just eliminates the term $(D'+\sigma')^2$. The integral over σ' and over the ghost field $C, \overline{C}$ gives

$$(\det{}' \Delta^0)^{\frac{1}{2}} \tag{1.5.32}$$

which cancels the overall factor in (1.5.28).

Before proceeding, we just note that due to gauge invariance we can diagonalize σ_0 so that it takes values in the Cartan subalgebra. This introduces the usual Vandermonde factor in the integral over σ_0, namely

$$\prod_{\alpha>0} (\alpha(\sigma_0))^2, \tag{1.5.33}$$

where α denote the roots of the Lie algebra $\mathbf{g}$, and $\alpha > 0$ are the positive roots. Using the Cartan decomposition of $\mathbf{g}$, we can write A_μ as

$$A_\mu = \sum_\alpha A_\mu^\alpha X_\alpha + h_\mu \tag{1.5.34}$$

In this equation, X_α are representatives of the root spaces of G, normalized as

$$\mathrm{Tr}(X_\alpha X_\beta) = \delta_{\alpha+\beta}, \tag{1.5.35}$$

where $\delta_{\alpha+\beta}$ is one if $\alpha+\beta=0$, and zero otherwise. In (1.5.34), h_μ is the component of A_μ along the Cartan subalgebra. Notice that this part of A_μ will only contribute a σ_0-independent factor to the one-loop determinant, so we will ignore it. We have

$$[\sigma_0, A_\mu] = \sum_\alpha \alpha(\sigma_0) A_\mu^\alpha X_\alpha \tag{1.5.36}$$

and a similar equation for λ. Plugging this into the action, we can now write it in terms of ordinary (as opposed to matrix valued) vectors and spinors

$$\frac{1}{2}\int \sqrt{g}\mathrm{d}^3x \sum_\alpha \left(g^{\mu\nu} A_\mu^{-\alpha}\left(-\Delta + \alpha(\sigma_0)^2\right) A_\nu^\alpha + \bar{\lambda}^{-\alpha}\left(\mathrm{i}\gamma^\mu\nabla_\mu + \mathrm{i}\alpha(\sigma_0) - \frac{1}{2}\right)\lambda^\alpha\right). \tag{1.5.37}$$

We now have to calculate the determinants of the above operators. The integration over the fluctuations of the gauge field is restricted, as in the Chern–Simons case, to the vector spherical harmonics. Using the results (1.3.14), (1.3.15), we find that the bosonic part of the determinant is:

$$\det(\text{bosons}) = \prod_\alpha \prod_{n=1}^\infty \left((n+1)^2 + \alpha(\sigma_0)^2\right)^{2n(n+2)}. \tag{1.5.38}$$

For the gaugino, we can use (A.66) to write the fermion determinant as:

$$\det(\text{fermions}) = \prod_\alpha \prod_{n=1}^\infty \Big((n+\mathrm{i}\alpha(\sigma_0))(-n-1+\mathrm{i}\alpha(\sigma_0))\Big)^{n(n+1)}, \tag{1.5.39}$$

and the quotient gives

$$\begin{aligned} Z_{\text{1-loop}}^{\text{gauge}}[\sigma_0] &= \prod_\alpha \prod_{n=1}^\infty \frac{(n+\mathrm{i}\alpha(\sigma_0))^{n(n+1)}(-n-1+\mathrm{i}\alpha(\sigma_0))^{n(n+1)}}{((n+1)^2+\alpha(\sigma_0)^2)^{n(n+2)}} \\ &= \prod_\alpha \prod_{n=1}^\infty \frac{(n+\mathrm{i}\alpha(\sigma_0))^{n(n+1)}(-n-1+\mathrm{i}\alpha(\sigma_0))^{n(n+1)}}{(n+\mathrm{i}\alpha(\sigma_0))^{(n-1)(n+1)}(n+1-i\alpha(\sigma_0))^{n(n+2)}}, \end{aligned} \tag{1.5.40}$$

up to a σ_0-independent sign. We see there is partial cancellation between the numerator and the denominator, and this becomes:

$$\begin{aligned} Z_{\text{1-loop}}^{\text{gauge}}[\sigma_0] &= \prod_\alpha \prod_{n=1}^\infty \frac{(n+\mathrm{i}\alpha(\sigma_0))^{n+1}}{(n-\mathrm{i}\alpha(\sigma_0))^{n-1}} = \prod_{\alpha>0} \prod_{n=1}^\infty \frac{(n^2+\alpha(\sigma_0)^2)^{n+1}}{(n^2+\alpha(\sigma_0)^2)^{n-1}} \\ &= \prod_{\alpha>0} \prod_{n=1}^\infty (n^2+\alpha(\sigma_0)^2)^2, \end{aligned} \tag{1.5.41}$$

where we used the fact that the roots split into positive roots $\alpha > 0$ and negative roots $-\alpha$, $\alpha > 0$. We finally obtain

$$Z^{\text{gauge}}_{\text{1-loop}}[\sigma_0] = \left(\prod_{n=1}^{\infty} n^4\right) \prod_{\alpha>0} \prod_{n=1}^{\infty} \left(1 + \frac{\alpha(\sigma_0)^2}{n^2}\right)^2 . \tag{1.5.42}$$

We can regularize this infinite product with the zeta function. This will lead to a finite, numerical result for the infinite product

$$\prod_{n=1}^{\infty} n^4. \tag{1.5.43}$$

On the other hand, we can use the well-known formula

$$\frac{\sinh(\pi z)}{\pi z} = \prod_{n=1} \left(1 + \frac{z^2}{n^2}\right) \tag{1.5.44}$$

to write

$$Z^{\text{gauge}}_{\text{1-loop}}[\sigma_0] \propto \prod_{\alpha>0} \left(\frac{\sinh(\pi\alpha(\sigma_0))}{\pi\alpha(\sigma_0)}\right)^2 , \tag{1.5.45}$$

where the proportionality factor is independent of σ_0. We conclude that the localization of the vector multiplets leads to a total contribution to the partition function

$$\int \mathrm{d}\mu \prod_{\alpha>0} \left(2\sinh\left(\alpha\left(\frac{\mu}{2}\right)\right)\right)^2 \mathrm{e}^{-\frac{1}{2g_s}\mathrm{Tr}(\mu^2)} \tag{1.5.46}$$

where we defined the convenient coupling,

$$g_s = \frac{2\pi \mathrm{i}}{k} \tag{1.5.47}$$

and we wrote

$$\sigma_0 = \frac{\mu}{2\pi}, \tag{1.5.48}$$

where μ takes values in the Cartan subalgebra.

1.5.3 Localization in Chern–Simons–matter theories: matter sector

Let us now consider the matter sector. We will follow the computation in [43], which simplifies a little bit the original computation in [2]. As shown in (1.1.43), the matter Lagrangian is in itself a total superderivative, so we can introduce a coupling t in the form

$$-\,tS_{\text{matter}}. \tag{1.5.49}$$

By the by now familiar localization argument, the partition function is independent of t, as long as $t > 0$, and we can compute it for $t = 1$ (which is the original case) or for $t \to \infty$. We can also restrict this Lagrangian to the localization locus of the gauge sector. The matter kinetic terms are then

$$\begin{aligned}\mathcal{L}_\phi =& g^{\mu\nu}\partial_\mu\bar{\phi}\partial_\nu\phi + \bar{\phi}\sigma_0^2\phi + \frac{2\mathrm{i}(\Delta-1)}{r}\bar{\phi}\sigma_0\phi + \frac{\Delta(2-\Delta)}{r^2}\bar{\phi}\phi,\\ \mathcal{L}_\psi =& -\mathrm{i}\bar{\psi}\gamma^\mu\partial_\mu\psi + \mathrm{i}\bar{\psi}\sigma_0\psi - \frac{\Delta-2}{r}\bar{\psi}\psi.\end{aligned}$$

The real part of the bosonic Lagrangian is positive definite, and it is minimized (and equal to zero) when

$$\phi = 0. \tag{1.5.50}$$

Like before, in the $t \to \infty$ limit, only quadratic terms in the matter fields contribute to the localization computation. In particular, there is no contribution from the superpotential terms involving the matter multiplets, like (1.1.44). After using (A.43) and (A.61), we find that the operators governing the quadratic fluctuations around this fixed point are given by the operators

$$\begin{aligned}\mathcal{O}_\phi &= \frac{1}{r^2}\left\{4\mathbf{L}^2 - (\Delta - \mathrm{i}r\sigma_0)(\Delta - 2 - \mathrm{i}r\sigma_0)\right\},\\ \mathcal{O}_\psi &= \frac{1}{r}\left\{4\mathbf{L}\cdot\mathbf{S} + \mathrm{i}r\sigma_0 + 2 - \Delta\right\}.\end{aligned} \tag{1.5.51}$$

Their eigenvalues are, for the bosons,

$$\lambda_\phi(n) = r^{-2}(n+2+\mathrm{i}r\sigma_0-\Delta)(n-\mathrm{i}r\sigma_0+\Delta), \qquad n=0,1,2,\cdots, \tag{1.5.52}$$

with multiplicity $(n+1)^2$, and for the fermions

$$\lambda_\psi(n) = r^{-1}(n+1+\mathrm{i}r\sigma_0-\Delta), \quad r^{-1}(-n+\mathrm{i}r\sigma_0-\Delta), \qquad n=1,2,\cdots, \tag{1.5.53}$$

with multiplicity $n(n+1)$. We finally obtain, after setting $r=1$,

$$\frac{|\mathrm{det}\Delta_\psi|}{\mathrm{det}\Delta_\phi} = \prod_{m>0} \frac{(m+1+\mathrm{i}r\sigma_0-\Delta)^{m(m+1)}(m-\mathrm{i}r\sigma_0+\Delta)^{m(m+1)}}{(m+1+\mathrm{i}r\sigma_0-\Delta)^{m^2}(m-1-\mathrm{i}r\sigma_0+\Delta)^{m^2}}, \tag{1.5.54}$$

and we conclude

$$Z^{\mathrm{matter}}_{\text{1-loop}}[\sigma_0] = \prod_{m>0}\left(\frac{m+1-\Delta+\mathrm{i}r\sigma_0}{m-1+\Delta-\mathrm{i}r\sigma_0}\right)^m. \tag{1.5.55}$$

As a check, notice that, when $\Delta=1/2$ and $\sigma_0=0$, we recover the quotient of determinants (1.3.45) of the free theory. The quantity (1.5.55) can be easily computed by using ζ-function regularization [43, 82]. Denote

$$z = 1-\Delta+\mathrm{i}r\sigma_0 \tag{1.5.56}$$

and

$$\ell(z) = \log Z^{\mathrm{matter}}_{\text{1-loop}}[\sigma_0]. \tag{1.5.57}$$

We can regularize this quantity as

$$\ell(z) = -\frac{\partial}{\partial s}\bigg|_{s=0} \sum_{m=1}^{\infty}\left(\frac{m}{(m+z)^s}-\frac{m}{(m-z)^s}\right). \tag{1.5.58}$$

On the other hand,

$$\sum_{m=1}^{\infty}\left(\frac{m}{(m+z)^s}-\frac{m}{(m-z)^s}\right) = \zeta_H(s-1,z)-z\zeta_H(s,z)-\zeta_H(s-1,-z)-z\zeta_H(s,-z), \tag{1.5.59}$$

where

$$\zeta_H(s,z) = \sum_{m=0}^{\infty}\frac{1}{(m+z)^s} \tag{1.5.60}$$

is the Hurwitz zeta function. Using standard properties of this function (see for example [63]), one finally finds the regularized result

$$\ell(z) = -z\log\left(1-\mathrm{e}^{2\pi\mathrm{i}z}\right)+\frac{\mathrm{i}}{2}\left(\pi z^2+\frac{1}{\pi}\mathrm{Li}_2(\mathrm{e}^{2\pi\mathrm{i}z})\right)-\frac{\mathrm{i}\pi}{12}. \tag{1.5.61}$$

As a check of this, notice that

$$\ell\left(\frac{1}{2}\right) = -\frac{1}{2}\log 2 \tag{1.5.62}$$

in agreement with (1.3.44).

There is an important property of $\ell(z)$, namely when $\Delta=1/2$ (canonical dimension) one has

$$\frac{1}{2}\left(\ell(z)+\ell(z^*)\right) = -\frac{1}{2}\log\left(2\cosh(\pi r\sigma_0)\right). \tag{1.5.63}$$

To prove this, we write

$$z = \frac{1}{2}+\mathrm{i}\theta, \tag{1.5.64}$$

and we compute

$$\frac{1}{2}\left(\ell(z)+\ell(z^*)\right) = -\frac{1}{2}\log\left(2\cosh(\pi\theta)\right)+\frac{1}{2}\pi\mathrm{i}\theta^2+\frac{\mathrm{i}\pi}{24}+\frac{\mathrm{i}}{4\pi}\left(\mathrm{Li}_2(-\mathrm{e}^{-2\pi\theta})+\mathrm{Li}_2(-\mathrm{e}^{2\pi\theta})\right). \tag{1.5.65}$$

After using the following property of the dilogarithm,

$$\mathrm{Li}_2(-x) + \mathrm{Li}_2(-x^{-1}) = -\frac{\pi^2}{6} - \frac{1}{2}\left(\log(x)\right)^2, \tag{1.5.66}$$

we obtain (1.5.63).

When the matter is in a self-conjugate representation of the gauge group, the set of eigenvalues of σ_0 is invariant under change of sign, therefore we can calculate the contribution of such a multiplet by using (1.5.63). We conclude that for such a matter multiplet,

$$Z^{\text{matter}}_{\text{1-loop}}[\mu] = \prod_{\Lambda} \left(2\cosh\frac{\Lambda(\mu)}{2}\right)^{-1/2}, \tag{1.5.67}$$

where we set $r = 1$ and we used the variable μ in the Cartan defined in (1.5.48). The product is over the weights Λ of the representation of the matter multiplet. For general representations and anomalous dimensions, one has to use the more complicated result above for $\ell(z)$.

1.5.4 The Chern–Simons matrix model

As a first application of the results of localization, let us consider pure supersymmetric Chern–Simons theory, defined by the action (1.1.21). If we don't add matter to the theory, the fields D, σ and $\lambda, \bar{\lambda}$ are auxiliary and they can be integrated out. In other words, supersymmetric Chern–Simons theory on $\mathbb{S}^3$ should be equivalent to pure (bosonic) Chern–Simons theory. There is however an important difference: in super-Chern–Simons theories with at least $\mathcal{N} = 2$ supersymmetry, there is no renormalization of the coupling k due to the extended supersymmetry [84]. The localization argument developed above says that the partition function of Chern–Simons theory on $\mathbb{S}^3$ with gauge group G should be proportional to the matrix model (1.5.46):

$$Z_{\text{CS}}(\mathbb{S}^3) \propto \int \mathrm{d}\mu \prod_{\alpha>0} \left(2\sinh\frac{\alpha\cdot\mu}{2}\right)^2 \mathrm{e}^{-\frac{1}{2g_s}\mu^2}, \tag{1.5.68}$$

where we regard μ as a weight and we use the standard Cartan–Killing inner product in the space of weights. For example, in the case of $G = U(N)$, if we write μ and the positive roots in terms of an orthonormal basis e_i of the weight lattice,

$$\mu = \sum_{i=1}^{N} \mu_i e_i, \qquad \alpha_{ij} = e_i - e_j, \quad i < j, \tag{1.5.69}$$

we find

$$Z_{\text{CS}}(\mathbb{S}^3) \propto \int \prod_{i=1}^{N} \mathrm{d}\mu_i \prod_{i<j} \left(2\,\sinh\frac{\mu_i - \mu_j}{2}\right)^2 \mathrm{e}^{-\frac{1}{2g_s}\sum_{i=1}^{N}\mu_i^2}. \tag{1.5.70}$$

The proportionality constant appearing in (1.5.68) should be independent of the coupling constant k, and it is only a function of N. The matrix model (1.5.46) is a "deformation" of the standard Gaussian matrix model. It has a Gaussian weight, but instead of displaying the standard Vandermonde interaction between eigenvalues (1.5.33) it has a "trigonometric" deformation involving the sinh. This interaction reduces to the standard one for small $\alpha \cdot \mu$, which corresponds in the $U(N)$ case to a small separation between eigenvalues.

The matrix model (1.5.46), with a sinh kernel, was first introduced in [85]. It was later rederived using geometric localization techniques in [86], and abelianization techniques in [87]. As we have seen following [2], it can be derived in an elegant and simple way by using supersymmetric localization. Actually, the matrix integral appearing in the r.h.s. of (1.5.68) can be calculated in a very simple way by using Weyl's denominator formula, as pointed out in for example [88]. This formula reads,

$$\sum_{w\in\mathcal{W}} \epsilon(w)\mathrm{e}^{w(\rho)\cdot u} = \prod_{\alpha>0} 2\sinh\frac{\alpha\cdot u}{2}. \tag{1.5.71}$$

In this formula, $\mathcal{W}$ is the Weyl group of G, $\epsilon(w)$ is the signature of w, and ρ is the Weyl vector, given by the sum of the fundamental weights. Using this formula, the matrix integral reduces to a sum of Gaussian integrals which can be calculated immediately, and one finds

$$\left(\det(C)\right)^{1/2} (2\pi g_s)^{r/2}\, |\mathcal{W}|\, \mathrm{e}^{g_s\rho^2} \sum_{w\in\mathcal{W}} \epsilon(w)\mathrm{e}^{g_s\rho\cdot w(\rho)}, \tag{1.5.72}$$

where C is the inverse matrix of the inner product in the space of weights (for simply connected G, this is the Cartan matrix), and r is the rank of G. Using again Weyl's denominator formula we find,

$$\sum_{w\in\mathcal{W}} \epsilon(w)\mathrm{e}^{g_s \rho\cdot w(\rho)} = \mathrm{i}^{|\Delta_+|} \prod_{\alpha>0} 2\sin\left(\frac{\pi\alpha\cdot\rho}{k}\right) \tag{1.5.73}$$

where $|\Delta_+|$ is the number of positive roots of G. The matrix integral then gives,

$$(\det(C))^{1/2}\,(2\pi)^r\,|\mathcal{W}|\,\frac{\mathrm{i}^{|\Delta_+|-r/2}}{k^{r/2}}\mathrm{e}^{\frac{\pi \mathrm{i}}{6k}d_G y}\prod_{\alpha>0} 2\sin\left(\frac{\pi\alpha\cdot\rho}{k}\right) \tag{1.5.74}$$

where we have used Freudenthal–de Vries formula

$$\rho^2 = \frac{1}{12}d_G y. \tag{1.5.75}$$

The result (1.5.74) is indeed proportional to the partition function of Chern–Simons theory on $\mathbb{S}^3$, and we can use the result to fix the normalization, N-dependent factor in the matrix integral. Let us particularize for $G = U(N)$. Then, the matrix integral is

$$\mathrm{i}^{-\frac{N^2}{2}}(2\pi)^N N!\,\mathrm{e}^{\frac{\pi\mathrm{i}}{6k}N(N^2-1)}k^{-N/2}\prod_{j=1}^N \left[2\sin\left(\frac{\pi j}{k}\right)\right]^{N-j}, \tag{1.5.76}$$

which is indeed proportional to (1.3.28) (after changing $k \to k-N$), up to an overall factor

$$\mathrm{i}^{-\frac{N^2}{2}}(2\pi)^N N!\,\mathrm{e}^{\frac{\pi\mathrm{i}}{6k}N(N^2-1)}. \tag{1.5.77}$$

The phase appearing here depends on k, and it has the right dependence on k, N to be understood as a change of framing of $\mathbb{S}^3$ in the result (1.3.28). We can now use the above result to fix the normalization in the matrix model describing supersymmetric Chern–Simons theory, and we find

$$Z_{\rm CS}\left(\mathbb{S}^3\right) = \frac{\mathrm{i}^{-\frac{N^2}{2}}}{N!}\int \prod_{i=1}^N \frac{\mathrm{d}\mu_i}{2\pi}\prod_{i<j}\left(2\,\sinh\frac{\mu_i-\mu_j}{2}\right)^2 \mathrm{e}^{-\frac{1}{2g_s}\sum_{i=1}^N \mu_i^2}. \tag{1.5.78}$$

We will refer to this model as the Chern–Simons matrix model. Later on we will study its large N limit.

Chapter 2

’t Hooft limit and genus expansion

In this chapter we investigate in detail the matrix model of ABJM theory using the standard and powerful tools of matrix models such as spectral curve and genus expansion. The organization of the chapter is as follows. First section 2.1 we give a review the standard matrix model techniques for conventional matrix models. In section 2.2 we relate ABJM matrix model to the lens space matrix model and construct the planar solution. In section 2.3 we consider the moduli space of the ABJM theory and its relation to topological strings in local $\mathbb{P}^1 \times \mathbb{P}^1$. Then in the sections 2.4-2.6 we analyze the behavior near three different special points of the moduli space. In particular in section 2.5 find the asympotics of the free energy and Wilson loops in the strong coupling limit which agrees with the predicitons from the string theory side. In section 2.7 we find non-planar corrections to the v.e.v.’s of Wilson loops and also consider generalization to giant Wilson loops. In sections 2.8-2.11 we study the ABJM theory with flavors and we develop a geometrical method to analyze the strong coupling behavoir without finding exact interpolation functions. In section 2.12 we find higher genus free energies of ABJM theory. In section 2.13, from the analysis of their lagre order behaviour, we are able to find instanton-type corrections which are otherwise invisible in the ’t Hooft expansion. In section 2.14 we argue that they are related to membrane instantons on the string theory side.

2.1 Matrix models at large N

In this section we will review some standard techniques of conventional matrix models, and in the next sections we will use them to analyze the matrix models appearing in supersymmetric Chern–Simons–matter theories. A more detailed treatment of matrix models in the large N expansion, as well as a complete list of references, can be found in [24, 89, 90].

2.1.1 Saddle-point equations and one-cut solution

Let us consider the matrix model partition function

$$Z = \frac{1}{N!}\frac{1}{(2\pi)^N}\int \prod_{i=1}^{N} \mathrm{d}\lambda_i \, \Delta^2(\lambda) \mathrm{e}^{-\frac{1}{g_s}\sum_{i=1}^N V(\lambda_i)}. \qquad (2.1.1)$$

$V(\lambda)$, called the *potential* of the matrix model, will be taken to be a polynomial

$$V(\lambda) = \frac{1}{2}\lambda^2 + \sum_{p\geq 3} \frac{g_p}{p}\lambda^p \qquad (2.1.2)$$

where the g_p are coupling constants of the model. In (2.1.1),

$$\Delta^2(\lambda) = \prod_{i<j}(\lambda_i - \lambda_j)^2 \qquad (2.1.3)$$

is the Vandermonde determinant (1.5.33) for the group $U(N)$. The integral (2.1.1) is typically obtained as a reduction to eigenvalues of integrals over the space of $N \times N$ Hermitian matrices, see [89, 90] for more details. We want to study Z in the so-called ’t Hooft limit, in which

$$g_s \to 0, \quad N \to \infty, \qquad (2.1.4)$$

but the ‘t Hooft parameter of the matrix model

$$t = g_s N \tag{2.1.5}$$

is fixed. In particular, we want to study the leading asymptotic behavior of the free energy

$$F = \log Z \tag{2.1.6}$$

in this limit. Let us write the partition function (2.1.1) as follows:

$$Z = \frac{1}{N!} \int \prod_{i=1}^{N} \frac{\mathrm{d}\lambda_i}{2\pi} \mathrm{e}^{g_s^{-2} S_{\mathrm{eff}}(\lambda)} \tag{2.1.7}$$

where the effective action is given by

$$S_{\mathrm{eff}}(\lambda) = -\frac{t}{N} \sum_{i=1}^{N} V(\lambda_i) + \frac{2t^2}{N^2} \sum_{i<j} \log|\lambda_i - \lambda_j|. \tag{2.1.8}$$

We can now regard g_s^2 as a sort of $\hbar$, in such a way that, as $g_s \to 0$ with t fixed, the integral (2.1.7) will be dominated by a saddle-point configuration that extremizes the effective action. Notice that, since a sum over N eigenvalues is roughly of order N, in the 't Hooft limit the effective action is of order $\mathcal{O}(1)$, and the free energy scales as

$$F(g_s, t) \approx g_s^{-2} F_0(t). \tag{2.1.9}$$

$F_0(t)$ is called the genus zero, or planar, free energy of the matrix model, and it is obtained by evaluating the effective action at the saddle point. This dominant contribution is just the first term in an asymptotic expansion around $g_s = 0$,

$$F = \sum_{g=0}^{\infty} F_0(t) g_s^{2g-2}. \tag{2.1.10}$$

In order to obtain the saddle-point equation, we just vary $S_{\mathrm{eff}}(\lambda)$ w.r.t. the eigenvalue λ_i. We obtain the equation

$$\frac{1}{2t} V'(\lambda_i) = \frac{1}{N} \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j}, \quad i = 1, \cdots, N. \tag{2.1.11}$$

This equation can be given a simple interpretation: we can regard the eigenvalues as coordinates of a system of N classical particles moving on the real line. (2.1.11) says that these particles are subject to an effective potential

$$V_{\mathrm{eff}}(\lambda_i) = V(\lambda_i) - \frac{2t}{N} \sum_{j \neq i} \log|\lambda_i - \lambda_j| \tag{2.1.12}$$

which involves a logarithmic Coulomb repulsion between eigenvalues. For small 't Hooft parameter, the potential term dominates over the Coulomb repulsion, and the particles tend to be at a critical point x_* of the potential: $V'(x_*) = 0$. As t grows, the logarithmic Coulomb interaction will force the eigenvalues to repel each other and to spread out away from the critical point.

To encode this information about the equlibrium distribution of the particles, it is convenient to define an *eigenvalue distribution* (for finite N) as

$$\rho(\lambda) = \frac{1}{N} \sum_{i=1}^{N} \langle \delta(\lambda - \lambda_i) \rangle, \tag{2.1.13}$$

where the λ_i solve (2.1.11) in the saddle-point approximation. In the large N limit, it is reasonable to expect that this distribution becomes a continuous distribution $\rho_0(\lambda)$. As we will see in a moment, this distribution has a compact support. The simplest case occurs when $\rho_0(\lambda)$ vanishes outside a connected interval $\mathcal{C} = [a, b]$. This is the so-called *one-cut solution.* Based on the considerations above, we expect $\mathcal{C}$ to be centered around a critical point x_* of the potential. In particular, as $t \to 0$, the interval $\mathcal{C}$ should collapse to the point x_*.

We can now write the saddle-point equation in terms of continuum quantities, by using the rule

$$\frac{1}{N} \sum_{i=1}^{N} f(\lambda_i) \to \int_{\mathcal{C}} f(\lambda) \rho_0(\lambda) \mathrm{d}\lambda. \tag{2.1.14}$$

Notice that the distribution of eigenvalues $\rho_0(\lambda)$ satisfies the normalization condition

$$\int_{\mathcal{C}} \rho_0(\lambda) d\lambda = 1. \tag{2.1.15}$$

The equation (2.1.11) then becomes

$$\frac{1}{2t} V'(\lambda) = \mathrm{P} \int_{\mathcal{C}} \frac{\rho_0(\lambda') \mathrm{d}\lambda'}{\lambda - \lambda'} \tag{2.1.16}$$

where P denotes the principal value of the integral. The above equation is an integral equation that allows one in principle to compute $\rho_0(\lambda)$, given the potential $V(\lambda)$, as a function of the 't Hooft parameter t and the coupling constants. Once $\rho_0(\lambda)$ is known, one can easily compute $F_0(t)$, since the effective action in the continuum limit is a functional of ρ_0:

$$S_{\text{eff}}(\rho_0) = -t \int_{\mathcal{C}} \mathrm{d}\lambda \, \rho_0(\lambda) V(\lambda) + t^2 \int_{\mathcal{C} \times \mathcal{C}} \mathrm{d}\lambda \, \mathrm{d}\lambda' \, \rho_0(\lambda) \rho_0(\lambda') \log |\lambda - \lambda'|. \tag{2.1.17}$$

The planar free energy is given by

$$F_0(t) = S_{\text{eff}}(\rho_0). \tag{2.1.18}$$

We can obtain (2.1.11) directly in the continuum formulation by computing the extremum of the functional

$$S(\rho_0, \Gamma) = S_{\text{eff}}(\rho_0) + \Gamma \Big(t \int_{\mathcal{C}} \mathrm{d}\lambda \, \rho_0(\lambda) - t \Big) \tag{2.1.19}$$

with respect to ρ_0. Here, Γ is a Lagrange multiplier that imposes the normalization condition of the density of eigenvalues (times t). This leads to

$$V(\lambda) = 2t \int \mathrm{d}\lambda' \, \rho_0(\lambda') \log |\lambda - \lambda'| + \Gamma, \tag{2.1.20}$$

which can be also obtained by integrating (2.1.16) with respect to λ. It is convenient to introduce the *effective potential on an eigenvalue* as

$$V_{\text{eff}}(\lambda) = V(\lambda) - 2t \int \mathrm{d}\, \lambda' \rho_0(\lambda') \log |\lambda - \lambda'|. \tag{2.1.21}$$

This is of course the continuum counterpart of (2.1.12). In terms of this quantity, the saddle–point equation (2.1.20) says that the effective potential is *constant* on the interval $\mathcal{C}$:

$$V_{\text{eff}}(\lambda) = \Gamma, \qquad \lambda \in \mathcal{C}. \tag{2.1.22}$$

The Lagrange multiplier Γ appears in this way as an integration constant that only depends on t and the coupling constants. As in any other Lagrange minimization problem, the multiplier is obtained by taking minus the derivative of the target function w.r.t. the constraint, which in this case is t. We then find the very useful equation

$$\partial_t F_0(t) = -\Gamma = -V_{\text{eff}}(b). \tag{2.1.23}$$

where b is the endpoint of the cut $\mathcal{C}$.

The density of eigenvalues is obtained as a solution to the saddle-point equation (2.1.16). This equation is a singular integral equation which has been studied in detail in other contexts of physics (see, for example, [91]). The way to solve it is to introduce an auxiliary function called the *resolvent*. The resolvent is defined, at finite N, as

$$\omega(p) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \frac{1}{p - \lambda_i} \right\rangle, \tag{2.1.24}$$

and we will denote its large N limit by $\omega_0(p)$, which is also called the genus zero resolvent. This can be written in terms of the eigenvalue density as

$$\omega_0(p) = \int d\lambda \frac{\rho_0(\lambda)}{p - \lambda}. \tag{2.1.25}$$

The genus zero resolvent (2.1.25) has three important properties. First of all, due to the normalization property of the eigenvalue distribution (2.1.15), it has the asymptotic behavior

$$\omega_0(p) \sim \frac{1}{p}, \qquad p \to \infty. \tag{2.1.26}$$

Second, as a function of p it is an analytic function on the whole complex plane except on the interval $\mathcal{C}$, where it has a discontinuity as one crosses the interval $\mathcal{C}$. This discontinuity can be computed by standard contour deformations. We have

$$\omega_0(p + \mathrm{i}\epsilon) = \int_{\mathbb{R}} \mathrm{d}\lambda \frac{\rho_0(\lambda)}{p + \mathrm{i}\epsilon - \lambda} = \int_{\mathbb{R}-\mathrm{i}\epsilon} \mathrm{d}\lambda \frac{\rho_0(\lambda)}{p - \lambda} = \mathrm{P} \int \mathrm{d}\lambda \frac{\rho_0(\lambda)}{p - \lambda} + \int_{C_\epsilon} \mathrm{d}\lambda \frac{\rho_0(\lambda)}{p - \lambda}, \tag{2.1.27}$$

where C_ϵ is a contour around $\lambda = p$ in the lower half plane, and oriented counterclockwise. The last integral can be evaluated as a residue, and we finally obtain,

$$\omega_0(p + \mathrm{i}\epsilon) = \mathrm{P} \int \mathrm{d}\lambda \frac{\rho_0(\lambda)}{p - \lambda} - \pi \mathrm{i} \rho_0(p). \tag{2.1.28}$$

Similarly

$$\omega_0(p - \mathrm{i}\epsilon) = \int_{\mathbb{R}+\mathrm{i}\epsilon} \mathrm{d}\lambda \frac{\rho_0(\lambda)}{p - \lambda} = \mathrm{P} \int \mathrm{d}\lambda \frac{\rho_0(\lambda)}{p - \lambda} + \pi \mathrm{i} \rho_0(p). \tag{2.1.29}$$

One then finds the key equation

$$\rho_0(\lambda) = -\frac{1}{2\pi \mathrm{i}} \big(\omega_0(\lambda + \mathrm{i}\epsilon) - \omega_0(\lambda - \mathrm{i}\epsilon) \big). \tag{2.1.30}$$

From these equations we deduce that, if the resolvent at genus zero is known, the planar eigenvalue distribution follows from (2.1.30), and one can compute the planar free energy. On the other hand, by using again (2.1.27) and (2.1.29) we can compute

$$\omega_0(p + \mathrm{i}\epsilon) + \omega_0(p - \mathrm{i}\epsilon) = 2\mathrm{P} \int \mathrm{d}\lambda \frac{\rho_0(\lambda)}{p - \lambda} \tag{2.1.31}$$

and we then find

$$\omega_0(p + \mathrm{i}\epsilon) + \omega_0(p - \mathrm{i}\epsilon) = \frac{1}{t} V'(p), \qquad p \in \mathcal{C}, \tag{2.1.32}$$

which determines the resolvent in terms of the potential. In this way we have reduced the original problem of computing $F_0(t)$ to the Riemann-Hilbert problem of computing $\omega_0(\lambda)$. In order to solve (2.1.32), we write it as a sum of an analytic or regular part $\omega_r(p)$, and a singular part $\omega_s(p)$,

$$\omega_0(p) = \omega_r(p) + \omega_s(p), \tag{2.1.33}$$

where

$$\omega_r(p) = \frac{1}{2t} V'(p). \tag{2.1.34}$$

It follows that the singular part satisfies

$$\omega_s(p + \mathrm{i}\epsilon) + \omega_s(p - \mathrm{i}\epsilon) = 0, \qquad p \in \mathcal{C}. \tag{2.1.35}$$

This is automatically satisfied if $\omega_s(p)$ has a square-root branch cut across $\mathcal{C}$, and we find

$$\omega_s(p) = -\frac{1}{2t} M(p) \sqrt{(p-a)(p-b)}, \tag{2.1.36}$$

where a, b are the endpoints of $\mathcal{C}$, and $M(p)$ is a polynomial, which is fully determined by the asymptotic condition (2.1.26). There is in fact a closed expression for the planar resolvent in terms of a contour integral [92] which reads

$$\omega_0(p) = \frac{1}{2t} \oint_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi \mathrm{i}} \frac{V'(z)}{p - z} \left(\frac{(p-a)(p-b)}{(z-a)(z-b)} \right)^{1/2}, \tag{2.1.37}$$

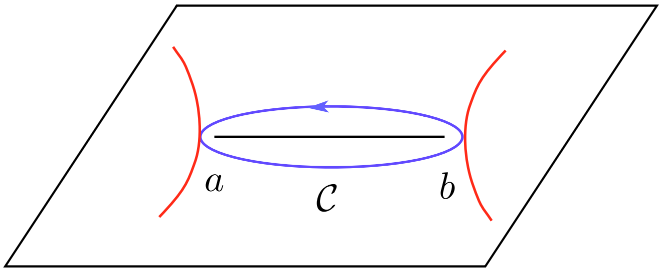

Figure 2.1: The contour $\mathcal{C}$ encircling the support of the density of eigenvalues, which can be regarded as a contour in the spectral curve $y(p)$.

where $\mathcal{C}$ denotes now a contour encircling the interval. The r.h.s. of (2.1.37) behaves like $c+d/p+\mathcal{O}(1/p^2)$. Requiring the asymptotic behavior (2.1.26) imposes $c=0$ and $d=1$, and this leads to

$$\begin{aligned} \oint_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi \mathrm{i}} \frac{V'(z)}{\sqrt{(z-a)(z-b)}} &= 0, \\ \oint_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi \mathrm{i}} \frac{zV'(z)}{\sqrt{(z-a)(z-b)}} &= 2t. \end{aligned} \tag{2.1.38}$$

These equations are enough to determine the endpoints of the cuts, a and b, as functions of the 't Hooft coupling t and the coupling constants of the model. Equivalently, after deforming the contour in (2.1.37) to infinity, we pick a pole at $z=p$, which gives the regular piece, and we find the equation

$$\omega_0(p) = \frac{1}{2t} V'(p) - \frac{1}{2t} M(p) \sqrt{(p-a)(p-b)}, \tag{2.1.39}$$

where

$$M(p) = \oint_{\infty} \frac{\mathrm{d}z}{2\pi i} \frac{V'(z)}{z-p} \frac{1}{\sqrt{(z-a)(z-b)}}. \tag{2.1.40}$$

A useful way to encode the solution to the matrix model is to define the *spectral curve* of the matrix model by

$$y(p) = V'(p) - 2t\,\omega_0(p) = M(p)\sqrt{(p-a)(p-b)}. \tag{2.1.41}$$

Notice that, up to a constant,

$$\int^{\lambda} \mathrm{d}p\, y(p) = V_{\mathrm{eff}}(\lambda). \tag{2.1.42}$$

If we regard $\omega_0(p)\mathrm{d}p$ as a differential on the spectral curve, the 't Hooft parameter can be written as a contour integral

$$t = \oint_{\mathcal{C}} \frac{\mathrm{d}p}{4\pi \mathrm{i}} 2t\omega_0(p). \tag{2.1.43}$$

This contour on the spectral curve (regarded as a complex curve) is represented in Fig. 2.1.

Example 2.1.1. *The Gaussian matrix model.* Let us now apply these results to the simplest case, the Gaussian model with $V(z) = z^2/2$. We first look for the position of the endpoints from (2.1.38). After deforming the contour to infinity and changing $z \to 1/z$, the first equation in (2.1.38) becomes

$$\oint_0 \frac{\mathrm{d}z}{2\pi \mathrm{i}} \frac{1}{z^2} \frac{1}{\sqrt{(1-az)(1-bz)}} = 0, \tag{2.1.44}$$

where the contour is now around $z=0$. Therefore $a+b=0$, in accord with the symmetry of the potential. Taking this into account, the second equation becomes:

$$\oint_0 \frac{\mathrm{d}z}{2\pi \mathrm{i}} \frac{1}{z^3} \frac{1}{\sqrt{1-a^2z^2}} = 2t, \tag{2.1.45}$$

and gives

$$a = 2\sqrt{t}. \tag{2.1.46}$$

We see that the interval $\mathcal{C} = [-a, a] = [-2\sqrt{t}, 2\sqrt{t}]$ opens as the 't Hooft parameter grows up, and as $t \to 0$ it collapses to the minimum of the potential at the origin, as expected. We immediately find from (2.1.39)

$$\omega_0(p) = \frac{1}{2t}\Big(p - \sqrt{p^2 - 4t}\Big), \tag{2.1.47}$$

and from the discontinuity equation we derive the density of eigenvalues

$$\rho_0(\lambda) = \frac{1}{2\pi t}\sqrt{4t - \lambda^2}. \tag{2.1.48}$$

The graph of this function is a semicircle of radius $2\sqrt{t}$, and the above eigenvalue distribution is the famous Wigner-Dyson semicircle law. Notice also that the equation (2.1.41) is in this case

$$y^2 = p^2 - 4t. \tag{2.1.49}$$

This is the equation for a curve of genus zero, which resolves the singularity $y^2 = p^2$. We then see that the opening of the cut as we turn on the 't Hooft parameter can be interpreted as a deformation of a geometric singularity.

2.1.2 Multi–cut solutions

So far we have considered the so-called one-cut solution to the one-matrix model. This is not, however, the most general solution, and we will now consider the so-called multi-cut solution, in the saddle-point approximation. Recall from our previous discussion that the cut appearing in the one-matrix model was centered around a critical point of the potential. If the potential has many critical points, one can have a saddle–point solution with various cuts, centered around different critical points. The most general solution has then n cuts (where n is lower or equal than the number of critical points), and the support of the eigenvalue distribution is a disjoint union of n intervals

$$\mathcal{C} = \bigcup_{i=1}^{n} \mathcal{C}_i. \tag{2.1.50}$$

The total number of eigenvalues N splits into n integers N_i,

$$N = N_1 + \cdots + N_n, \tag{2.1.51}$$

where N_i is the number of eigenvalues in the interval $\mathcal{C}_i$. We introduce the *filling fractions*

$$\epsilon_i = \frac{N_i}{N} = \int_{\mathcal{C}_i} \mathrm{d}\lambda\, \rho_0(\lambda), \qquad i = 1, \cdots, n. \tag{2.1.52}$$

Notice that

$$\sum_{i=1}^{n} \epsilon_i = 1. \tag{2.1.53}$$

A closely related set of variables are the partial 't Hooft parameters

$$t_i = t\epsilon_i = g_s N_i, \qquad i = 1, \cdots, n. \tag{2.1.54}$$

Notice that there are only $g = n - 1$ independent filling fractions, but the partial 't Hooft parameters are all independent.

The multi-cut solution is just a more general solution of the saddle-point equations that we derived above. It can be found by extremizing the functional (2.1.17) with the condition that the partial 't Hooft parameters are *fixed*,

$$S(\rho_0, \epsilon^I) = S_{\text{eff}}(\rho_0) + \sum_{i=1}^{n} \Gamma_i \Big(t \int_{\mathcal{C}_i} \mathrm{d}\lambda\, \rho_0(\lambda) - t_i \Big), \tag{2.1.55}$$

where Γ_i are Lagrange multipliers. If we take the variation w.r.t. the density $\rho_0(\lambda)$ we find the equation

$$V(\lambda) = 2t \int_{\mathcal{C}} \mathrm{d}\lambda' \rho_0(\lambda') \log|\lambda - \lambda'| + \Gamma_i, \qquad \lambda \in \mathcal{C}_i \tag{2.1.56}$$

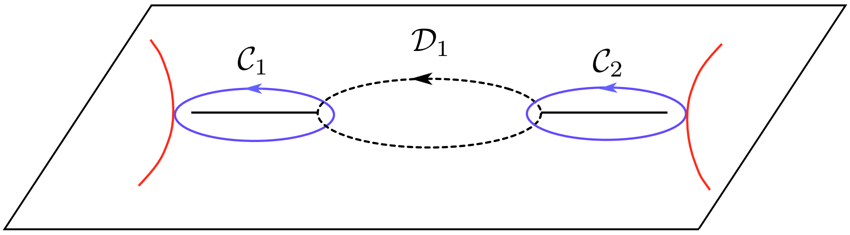

Figure 2.2: A two-cut spectral curve, showing two contours $\mathcal{C}_{1,2}$ around the cuts where $N_{1,2}$ eigenvalues sit. The "dual" cycle $\mathcal{D}_1$ goes from $\mathcal{C}_2$ to $\mathcal{C}_1$.

which can be rewritten as

$$V_{\text{eff}}(\lambda) = \Gamma_i, \qquad \lambda \in \mathcal{C}_i. \tag{2.1.57}$$

The planar resolvent still solves (2.1.32), and the way to implement the multi–cut solution is to require $\omega_0(p)$ to have $2n$ branch points. Therefore we have

$$\omega_0(p) = \frac{1}{2t}V'(p) - \frac{1}{2t}M(p)\sqrt{\prod_{k=1}^{2n}(p - x_k)}, \tag{2.1.58}$$

which can be solved in a compact way by

$$\omega_0(p) = \frac{1}{2t}\oint_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{V'(z)}{p - z} \left(\prod_{k=1}^{2n} \frac{p - x_k}{z - x_k}\right)^{1/2}. \tag{2.1.59}$$

In order to satisfy the asymptotics (2.1.26) the following conditions must hold:

$$\delta_{\ell n} = \frac{1}{2t}\oint_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{z^\ell V'(z)}{\prod_{k=1}^{2n}(z - x_k)^{\frac{1}{2}}}, \qquad \ell = 0, 1, \cdots, n. \tag{2.1.60}$$

In contrast to the one-cut case, these are only $n + 1$ conditions for the $2n$ variables x_k representing the endpoints of the cut. The remaining $n - 1$ conditions are obtained by fixing the values of the filling fractions through (2.1.52) (or, equivalently, by fixing the partial 't Hooft parameters). The multipliers in (2.1.55) are obtained, as before, by taking derivatives w.r.t. the constraints, and we find the equation

$$\frac{\partial F_0}{\partial t_i} - \frac{\partial F_0}{\partial t_{i+1}} = \Gamma_{i+1} - \Gamma_i, \tag{2.1.61}$$

which generalizes (2.1.23) to the multi–cut situation.

We can write the multi-cut solution in a very elegant way by using contour integrals. First, the partial 't Hooft parameters are given by

$$t_i = \frac{1}{4\pi\mathrm{i}}\oint_{\mathcal{C}_i} 2t\omega_0(p)\mathrm{d}p. \tag{2.1.62}$$

We now introduce dual cycles $\mathcal{D}_i$ cycles, $i = 1, \cdots, n-1$, going from the $\mathcal{C}_{i+1}$ cycle to the $\mathcal{C}_i$ cycle counterclockwise, see Fig. 2.2. In terms of these, we can write (2.1.61) as

$$\frac{\partial F_0}{\partial t_i} - \frac{\partial F_0}{\partial t_{i+1}} = -\frac{1}{2}\oint_{\mathcal{D}_i} 2t\omega_0(p)\mathrm{d}p. \tag{2.1.63}$$

2.2 The ABJM matrix model and Wilson loops

In this section we consider the relation of the ABJM matrix model to the lens space matrix model and discuss in detail the planar solution.

2.2.1 Relation to the lens space matrix model

Using the localization procedure reviewed in the previous chapter let us now consider the matrix model calculating the partition function on $\mathbb{S}^3$ of ABJM theory, or rather its generalization [3] to the gauge group $U(N_1) \times U(N_2)$. The contribution of the vector multiplets gives in the integrand

$$\prod_{1\le i<j\le N_1} \left(2\sinh\frac{\mu_i-\mu_j}{2}\right)^2 \mathrm{e}^{-\frac{1}{2g_s}\sum_{i=1}^{N_1}\mu_i^2} \prod_{1\le a<b\le N_2} \left(2\sinh\frac{\nu_a-\nu_b}{2}\right)^2 \mathrm{e}^{\frac{1}{2g_s}\sum_{a=1}^{N_2}\nu_a^2}, \tag{2.2.1}$$

where the opposite signs in the Gaussian exponents are due to the opposite signs in the levels. Since there are four hypermultiplets in the bifundamental representation, we have an extra factor due to (1.5.67),

$$\prod_{i=1}^{N_1}\prod_{a=1}^{N_2}\left(2\cosh\frac{\mu_i-\nu_a}{2}\right)^{-2}. \tag{2.2.2}$$

The normalization of the matrix model can be fixed by using the normalization for the Chern–Simons matrix model (1.5.78), and by comparing to the perturbative one-loop result. In this way we find,

$$\begin{aligned} Z_{\rm ABJM}(\mathbb{S}^3) =& \frac{\mathrm{i}^{-\frac{1}{2}(N_1^2-N_2^2)}}{N_1!N_2!}\int\prod_{i=1}^{N_1}\frac{\mathrm{d}\mu_i}{2\pi}\prod_{a=1}^{N_2}\frac{\mathrm{d}\nu_a}{2\pi}\prod_{1\le i<j\le N_1}\left(2\sinh\left(\frac{\mu_i-\mu_j}{2}\right)\right)^2 \\ &\times\prod_{1\le a<b\le N_2}\left(2\sinh\left(\frac{\nu_i-\nu_j}{2}\right)\right)^2\prod_{i,a}\left(2\cosh\left(\frac{\mu_i-\nu_a}{2}\right)\right)^{-2}\mathrm{e}^{-\frac{1}{2g_s}\left(\sum_i\mu_i^2-\sum_a\nu_a^2\right)}. \end{aligned} \tag{2.2.3}$$

This model is closely related to a matrix model that computes the partition function of Chern–Simons theory on lens spaces $L(p,1)$, in particular to the model with $p=2$. These models were introduced in [85], and the case $p=2$ was extensively studied in [88]. The matrix integral for $p=2$ is given by,

$$\begin{aligned} Z_{\rm CS}(L(2,1)) =& \frac{\mathrm{i}^{-\frac{1}{2}(N_1^2+N_2^2)}}{N_1!N_2!}\int\prod_{i=1}^{N_1}\frac{\mathrm{d}\mu_i}{2\pi}\prod_{a=1}^{N_2}\frac{\mathrm{d}\nu_a}{2\pi}\prod_{1\le i<j\le N_1}\left(2\sinh\left(\frac{\mu_i-\mu_j}{2}\right)\right)^2 \\ &\times\prod_{1\le a<b\le N_2}\left(2\sinh\left(\frac{\nu_i-\nu_j}{2}\right)\right)^2\prod_{i,a}\left(2\cosh\left(\frac{\mu_i-\nu_a}{2}\right)\right)^{2}\mathrm{e}^{-\frac{1}{2g_s}\left(\sum_i\mu_i^2+\sum_a\nu_a^2\right)}. \end{aligned} \tag{2.2.4}$$

We will refer to this matrix model as the lens space matrix model. We will now use our knowledge of the solution of the lens space matrix model to solve the ABJM model. It is clear that the matrix models (2.2.3) and(2.2.4) are very similar, but there are some obvious differences: in (2.2.3) the interaction between the μ and the ν eigenvalues is in the denominator, and the Gaussian action for the νs has the opposite sign. These ingredients are precisely the ones needed to make (2.2.3) a *supergroup* extension of (2.2.4). We will now quickly review some results on supermatrix models, following [93, 94, 95]. A Hermitian supermatrix has the form

$$\Phi = \begin{pmatrix} A & \Psi \\ \Psi^\dagger & C \end{pmatrix} \tag{2.2.5}$$

where A (C) are $N_1\times N_1$ ($N_2\times N_2$) Hermitian, Grassmann even matrices, and Ψ is a complex, Grassmann odd matrix. The supermatrix model is defined by the partition function

$$Z_{\rm s}(N_1|N_2) = \int \mathcal{D}\Phi\,\mathrm{e}^{-\frac{1}{g_s}\mathrm{Str}V(\Phi)} \tag{2.2.6}$$

where we consider a polynomial potential $V(\Phi)$, and Str is the supertrace

$$\mathrm{Str}\,\Phi = \mathrm{Tr}\,A - \mathrm{Tr}\,C. \tag{2.2.7}$$

There are two types of supermatrix models with supergroup symmetry $U(N_1|N_2)$: the ordinary supermatrix model, and the physical supermatrix model [94]. The ordinary supermatrix model is obtained by requiring A, C to be real Hermitian matrices, while the physical model is obtained by requiring that, after diagonalizing Φ by a superunitary transformation, the resulting eigenvalues are real. Here we will be interested in the physical supermatrix model. Its partition function reads, in terms of eigenvalues [94, 95]

$$Z_{\rm s}(N_1|N_2) = \int \prod_{i=1}^{N_1} {\rm d}\mu_i \prod_{j=1}^{N_2} {\rm d}\nu_j \frac{\prod_{i<j} (\mu_i - \mu_j)^2 (\nu_i - \nu_j)^2}{\prod_{i,j} (\mu_i - \nu_j)^2} {\rm e}^{-\frac{1}{g_s}\left(\sum_i V(\mu_i) - \sum_j V(\nu_j)\right)}. \tag{2.2.8}$$

When the two groups of eigenvalues μ_i, ν_j are expanded around two different critical points, the partition function (2.2.8) is well-defined as an asymptotic expansion in g_s. It is easy to show that (2.2.8) is related to the partition function of the corresponding bosonic, two-cut matrix model

$$Z_{\rm b}(N_1, N_2) = \int \prod_{i=1}^{N_1} {\rm d}\mu_i \prod_{j=1}^{N_2} {\rm d}\nu_j \prod_{i<j} (\mu_i - \mu_j)^2 (\nu_i - \nu_j)^2 \prod_{i,j} (\mu_i - \nu_j)^2 \, {\rm e}^{-\frac{1}{g_s}\left(\sum_i V(\mu_i) + \sum_j V(\nu_j)\right)} \tag{2.2.9}$$

after changing $N_2 \to -N_2$:

$$Z_{\rm s}(N_1|N_2) = Z_{\rm b}(N_1, -N_2). \tag{2.2.10}$$

Such a flip of sign is trivially performed if one knows the exact solution of the model in the $1/N$ expansion. The relation (2.2.10) can be proved diagramatically by introducing Faddeev–Popov ghosts as in [95, 96].

We now see that the relationship between the ABJM matrix model and the lens space matrix model is identical to the one we have between supergroup matrix models and multi-cut bosonic matrix models, with the only difference that the interaction between the eigenvalues has been promoted to the sinh interaction typical of Chern–Simons matrix models. Indeed, the lens space matrix model is a two-cut matrix model where the μ, ν eigenvalues are expanded around two different saddle points, $z = 0$ and $z = \pi{\rm i}$. The ABJM matrix model is just its supergroup version. We then conclude that

$$Z_{\rm ABJM}(N_1, N_2, g_s) = Z_{L(2,1)}(N_1, -N_2, g_s). \tag{2.2.11}$$

The appearance of a hidden supergroup structure in the matrix model of [2] is not surprising, since $\mathcal{N} = 4$ Chern–Simons–matter theories are classified by supergroups [97]. In fact, the ABJM theory can be constructed as an $\mathcal{N} = 4$ theory with supergroup $U(N_1|N_2)$ and containing both hypermultiplets and twisted hypermultiplets [98]. This hidden supergroup structure in the ABJM theory is explicitly used in the construction of half-BPS Wilson loops in [99].

2.2.2 The planar solution of the lens space matrix model

Let us now discuss the large N solution of the lens space matrix model, following [88, 100, 101]. At large N, the two sets of eigenvalues, μ_i, ν_j, condense around two cuts. The cut of the μ_i eigenvalues is centered around $z = 0$, while that of the ν_i eigenvalues is centered around $z = \pi{\rm i}$. We will write the cuts as

$$\mathcal{C}_1 = (-A, A), \qquad \mathcal{C}_2 = (\pi{\rm i} - B, \pi{\rm i} + B), \tag{2.2.12}$$

in terms of the endpoints A, B. It is also useful to use the exponentiated variable

$$Z = {\rm e}^z, \tag{2.2.13}$$

In the Z plane the cuts (2.2.12) get mapped to

$$(1/a, a), \qquad (-1/b, -b), \qquad\qquad a = {\rm e}^A, \qquad b = {\rm e}^B, \tag{2.2.14}$$

which are centered around $Z = 1$, $Z = -1$, respectively, see Fig. 2.3. We will use the same notation $\mathcal{C}_{1,2}$ for the cuts in the Z plane. The large N solution is encoded in the total resolvent of the matrix model, $\omega(z)$. It is defined as [100]

$$\omega(z) = g_s \left\langle {\rm Tr} \left(\frac{Z + U}{Z - U} \right) \right\rangle = g_s \left\langle \sum_{i=1}^{N_1} \coth\left(\frac{z - \mu_i}{2} \right) \right\rangle + g_s \left\langle \sum_{j=1}^{N_2} \tanh\left(\frac{z - \nu_j}{2} \right) \right\rangle \tag{2.2.15}$$

where

$$U = \begin{pmatrix} {\rm e}^{\mu_i} & 0 \\ 0 & -{\rm e}^{\nu_j} \end{pmatrix}. \tag{2.2.16}$$

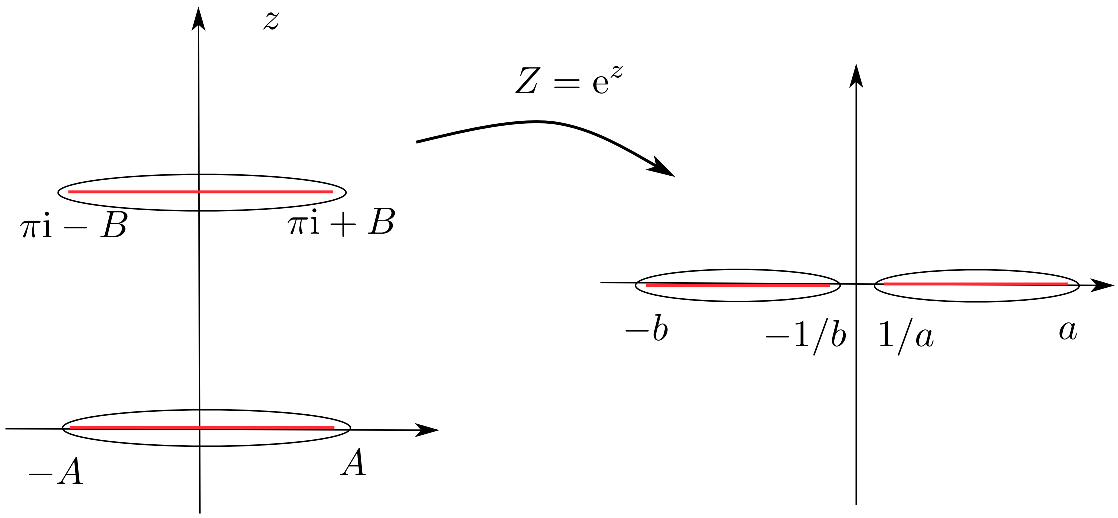

Figure 2.3: The cuts for the CS lens space matrix model in the z plane and in the $Z = \mathrm{e}^z$ plane.

We will denote by $\omega_0(z)$ the planar limit of the resolvent, which was found in explicit form in [100]. It reads,

$$\omega_0(z) = 2\log\left(\frac{\mathrm{e}^{-t/2}}{2}\left[\sqrt{(Z+b)(Z+1/b)} - \sqrt{(Z-a)(Z-1/a)}\right]\right), \tag{2.2.17}$$

where

$$t = t_1 + t_2 \tag{2.2.18}$$

is the total 't Hooft parameter. It is useful to introduce the variables

$$\zeta = \frac{1}{2}\left(a + \frac{1}{a} - b - \frac{1}{b}\right), \qquad \beta = \frac{1}{4}\left(a + \frac{1}{a} + b + \frac{1}{b}\right). \tag{2.2.19}$$

β is related to the total 't Hooft parameter through

$$\beta = \mathrm{e}^t. \tag{2.2.20}$$

All the relevant planar quantities can be expressed in terms of period integrals of the one-form $\omega_0(z)\mathrm{d}z$. The 't Hooft parameters are given by

$$t_i = \frac{1}{4\pi \mathrm{i}} \oint_{\mathcal{C}_i} \omega_0(z)\mathrm{d}z, \qquad i = 1, 2. \tag{2.2.21}$$

The planar free energy F_0 satisfies the equation

$$\mathcal{I} \equiv \frac{\partial F_0}{\partial t_1} - \frac{\partial F_0}{\partial t_2} - \frac{\pi \mathrm{i} t}{2} = -\frac{1}{2}\oint_{\mathcal{D}} \omega_0(z)\mathrm{d}z, \tag{2.2.22}$$

where the $\mathcal{D}$ cycle encloses, in the Z plane, the interval between $-1/b$ and $1/a$ (see Fig. 2.3).[1]

The derivatives of these periods can be calculated in closed form by adapting a trick from [102]. One finds,

$$\frac{\partial t_{1,2}}{\partial \zeta} = -\frac{1}{4\pi \mathrm{i}} \oint_{\mathcal{C}_{1,2}} \frac{\mathrm{d}Z}{\sqrt{(Z^2 - \zeta Z + 1)^2 - 4\beta^2 Z^2}} = \pm \frac{\sqrt{ab}}{\pi(1+ab)} K(k), \tag{2.2.23}$$

and similarly

$$\frac{\partial t_1}{\partial \beta} = -2\frac{\sqrt{ab}}{\pi(1+ab)}\left(K(k) - \frac{2ab}{1+ab}\Pi(n_1|k) - \frac{2}{1+ab}\Pi(n_2|k)\right), \tag{2.2.24}$$

[1] Likewise one can calculate the second "B-cycle" period, and it will arise when solving the Picard-Fuchs equations at strong coupling in Section 2.3.2.

where

$$k^2 = 1 - \left(\frac{a+b}{1+ab}\right)^2, \qquad n_1 = \frac{1-a^2}{1+ab}, \qquad n_2 = \frac{b(a^2-1)}{a(1+ab)}\,. \tag{2.2.25}$$

Likewise for the period integral in (2.2.22) we find

$$\begin{aligned} \frac{\partial \mathcal{I}}{\partial \zeta} &= -2\,\frac{\sqrt{ab}}{1+ab}\,K(k'), \\ \frac{\partial \mathcal{I}}{\partial \beta} &= 4\,\frac{\sqrt{ab}}{1+ab}\left(K(k') + \frac{2a(1-b^2)}{(1+ab)(a+b)}(\Pi(n_1'|k') - \Pi(n_2'|k'))\right), \end{aligned} \tag{2.2.26}$$

where

$$k' = \frac{a+b}{1+ab}, \qquad n_1' = \frac{a+b}{b(1+ab)}, \qquad n_2' = \frac{b(a+b)}{1+ab}. \tag{2.2.27}$$

We can now use the dictionary between the lens space matrix model and the ABJM matrix model given by (2.2.11) to get the planar solution of the latter model. In particular, the natural 't Hooft parameters in the ABJM theory

$$\lambda_j = \frac{N_j}{k} \tag{2.2.28}$$

are obtained from the planar solution of the lens space matrix model by the replacement

$$t_1 = 2\pi \mathrm{i} \lambda_1, \qquad t_2 = -2\pi \mathrm{i} \lambda_2. \tag{2.2.29}$$

Since in the ABJM theory the couplings $\lambda_{1,2}$ are real, the matrix model couplings $t_{1,2}$ are pure imaginary. Thanks to (2.2.20) we know that β is of the form

$$\beta = \mathrm{e}^{2\pi \mathrm{i}(\lambda_1 - \lambda_2)} \tag{2.2.30}$$

i.e., it must be a phase.

For later convenience we introduce yet another parameterization of the couplings in terms of B and κ

$$B = \lambda_1 - \lambda_2 + \frac{1}{2}, \qquad \kappa = \mathrm{e}^{-\pi \mathrm{i} B}\zeta\,. \tag{2.2.31}$$

B is identified as the B-field in the dual type IIA background [78]. Notice that it has a shift by $-1/2$ as compared to the original prescription in [3]. Clearly, all calculations in the matrix model are periodic under $B \to B+1$, up to possible monodromies (see (2.5.25) below). As we shall see later, the parameter κ is real for physical values of $\lambda_{1,2}$.

2.2.3 Wilson loops

A family of Wilson loops in this theory has been constructed in [103, 104, 105], with the structure

$$W_R^{1/6} = g_s \mathrm{Tr}_R\, \mathcal{P} \,\exp \int \left(\mathrm{i} A_\mu \dot{x}^\mu + \frac{2\pi}{k}|\dot{x}| M_J^I C_I \bar{C}^J\right) \mathrm{d}s \tag{2.2.32}$$

where A_μ is the gauge connection in the $U(N_1)_k$ gauge group, $x^\mu(s)$ is the parametrization of the loop, and M_I^J is a matrix determined by supersymmetry. It can be chosen so that, if the geometry of the loop is a line or a circle, four real supercharges are preserved. Therefore, we will call (2.2.32) the 1/6 BPS Wilson loop. A similar construction exists for a loop based on the other gauge group,

$$\widehat{W}_R^{1/6} = g_s \mathrm{Tr}_R\, \mathcal{P} \,\exp \int \left(\mathrm{i} \widehat{A}_\mu \dot{x}^\mu + \frac{2\pi}{k}|\dot{x}| M_J^I \bar{C}_I C^J\right) \mathrm{d}s \tag{2.2.33}$$

where A_μ is the $U(N_2)_{-k}$ gauge connection. The planar limit of the vev of (2.2.32) was computed in [103, 104, 105], for $N_1 = N_2 = N$, in the fundamental representation $R = \Box$, and in the weak coupling regime $\lambda \ll 1$, where

$$\lambda = \frac{N}{k} \tag{2.2.34}$$

is the 't Hooft parameter. The result is

$$\langle W_\Box \rangle = 2\pi i \lambda \left(1 + \frac{5\pi^2}{6}\lambda^2 + \mathcal{O}\left(\lambda^3\right)\right). \tag{2.2.35}$$

On the other hand, in the strong coupling regime $\lambda \gg 1$, the Wilson loop vev can be calculated by using the large N string dual, i.e. type IIA theory on $\mathrm{AdS}_4 \times \mathbb{P}^3$ [103, 104, 105]. This gives the prediction

$$\langle W_{\square} \rangle \sim \mathrm{e}^{\pi\sqrt{2\lambda}}. \tag{2.2.36}$$

As in the case of the 1/2 BPS Wilson loop in $\mathcal{N}=4$ Yang–Mills theory, the exact answer for the planar limit of this vev should interpolate between the weak coupling behavior (2.2.35) and the strong coupling prediction of the large N string dual, (2.2.36).

One of the main results of [2] is that the VEV of the 1/6 BPS Wilson loop in ABJM theory, labelled by a representation R or $U(N_1)$, can be obtained by calculating the VEV of the matrix e^{μ_i} in the matrix model (2.2.3), *i.e.*,

$$\langle W_R^{1/6} \rangle = g_s \left\langle \mathrm{Tr}_R \left(\mathrm{e}^{\mu_i}\right) \right\rangle_{\text{ABJM MM}}, \tag{2.2.37}$$

A 1/2 BPS loop $W_{\mathcal{R}}^{1/2}$ was constructed in [99] , where $\mathcal{R}$ is a representation of the supergroup $U(N_1|N_2)$. in [99] it was also shown that it localizes to the matrix model correlator in the ABJM matrix model

$$\langle W_{\mathcal{R}}^{1/2} \rangle = g_s \left\langle \mathrm{Str}_{\mathcal{R}}\, U \right\rangle_{\text{ABJM MM}}, \tag{2.2.38}$$

with the same U as in (2.2.16). Though at first sight the minus sign on the lower block of U, may look surprising, it can be attributed to the fact that the ν_j eigenvalues are shifted by $\pi\mathrm{i}$ from the real line. Due to the relation between the ABJM matrix model and the lens space matrix model, these correlators can be computed in the lens space matrix model as follows:

$$\begin{aligned} \langle W_R^{1/6} \rangle &= g_s \left\langle \mathrm{Tr}_R \left(\mathrm{e}^{\mu_i}\right) \right\rangle_{L(2,1)} \Big|_{N_2 \to -N_2}, \\ \langle W_{\mathcal{R}}^{1/2} \rangle &= g_s \left\langle \mathrm{Tr}_{\mathcal{R}}\, U \right\rangle_{L(2,1)} \Big|_{N_2 \to -N_2}, \end{aligned} \tag{2.2.39}$$

where the super-representation $\mathcal{R}$ is regarded as a representation of $U(N_1+N_2)$.

To evaluate the Wilson loop one uses the resolvent, or equivalently, the eigenvalue densities

$$\begin{aligned} \rho^{(1)}(Z)\mathrm{d}Z &= -\frac{1}{4\pi\mathrm{i}t_1}\frac{\mathrm{d}Z}{Z}\left(\omega(Z+\mathrm{i}\epsilon)-\omega(Z-\mathrm{i}\epsilon)\right), \qquad Z \in \mathcal{C}_1, \\ \rho^{(2)}(Z)\mathrm{d}Z &= \frac{1}{4\pi\mathrm{i}t_2}\frac{\mathrm{d}Z}{Z}\left(\omega(Z+\mathrm{i}\epsilon)-\omega(Z-\mathrm{i}\epsilon)\right), \qquad Z \in \mathcal{C}_2. \end{aligned} \tag{2.2.40}$$

which are each normalized in the planar approximation to unity

$$\int_{\mathcal{C}_i} \rho_0^{(i)}\, \mathrm{d}Z = 1. \tag{2.2.41}$$

For the 1/6 BPS Wilson loop in the fundamental representation one needs to integrate $\mathrm{e}^z = Z$ over the first cut

$$\left\langle W_{\square}^{1/6} \right\rangle = t_1 \int_{\mathcal{C}_1} \rho^{(1)}(Z) Z \mathrm{d}Z = \oint_{\mathcal{C}_1} \frac{\mathrm{d}Z}{4\pi\mathrm{i}}\, \omega(Z). \tag{2.2.42}$$

The correlator relevant for the 1/2 BPS Wilson loop (again in the fundamental representation) is much easier, since

$$\left\langle W_{\square}^{1/2} \right\rangle = t_1 \int_{\mathcal{C}_1} \rho^{(1)}(Z) Z \mathrm{d}Z - t_2 \int_{\mathcal{C}_2} \rho^{(2)}(Z) Z \mathrm{d}Z = \oint_{\infty} \frac{\mathrm{d}Z}{4\pi\mathrm{i}}\, \omega(Z) \tag{2.2.43}$$

and it can be obtained by expanding $\omega(Z)$ around $Z \to \infty$.

The comparison to the case of the 1/2 BPS Wilson loop in $\mathcal{N}=4$ SYM in 4d is straight-forward. In that case the matrix model is Gaussian and the eigenvalue density in the planar approximation follows Wigner's semi-circle law. Doing the integral with the insertion of e^z gives a modified Bessel function [106]

$$\rho_0(z) = \frac{2}{\pi\lambda}\sqrt{\lambda - z^2} \quad \Rightarrow \quad \langle W^{1/2}_{\text{4d } \mathcal{N}=4} \rangle_{\text{planar}} = \int_{-\sqrt{\lambda}}^{\sqrt{\lambda}} \rho_0(z)\, \mathrm{e}^z\, \mathrm{d}z = \frac{2}{\sqrt{\lambda}} I_1(\sqrt{\lambda}). \tag{2.2.44}$$

For the ABJM matrix model all the expressions are more complicated.

The densities $\rho^{(1)}(\mu)$ and $\rho^{(2)}(\nu)$ can be explicitly calculated from (2.2.40) and (2.2.17). We find,

$$\begin{aligned} \rho^{(1)}(X)\mathrm{d}X &= \frac{1}{\pi t_1}\tan^{-1}\left[\sqrt{\frac{\alpha X - 1 - X^2}{\delta X + 1 + X^2}}\right]\frac{\mathrm{d}X}{X}, \\ \rho^{(2)}(Y)\mathrm{d}Y &= -\frac{1}{\pi t_2}\tan^{-1}\left[\sqrt{\frac{\delta Y + 1 + Y^2}{\alpha Y - 1 - Y^2}}\right]\frac{\mathrm{d}Y}{Y}. \end{aligned} \tag{2.2.45}$$

In terms of the variable $x = \log X$ we have

$$\rho^{(1)}(x) = \frac{1}{\pi t_1}\tan^{-1}\left[\sqrt{\frac{\alpha - 2\cosh x}{\delta + 2\cosh x}}\right], \tag{2.2.46}$$

and a similar expression for $\rho^{(2)}(y)$. Notice that, if $t_2 = 0$, one has $\delta = 2$, $\alpha = 4\mathrm{e}^t - 2$, and $\rho^{(1)}(x)$ becomes the density of eigenvalues for the matrix model of Chern–Simons theory on $\mathbb{S}^3$ [90]

$$\rho^{(1)}(x) = \frac{1}{\pi t}\tan^{-1}\left[\frac{\sqrt{\mathrm{e}^t - \cosh^2\left(\frac{x}{2}\right)}}{\cosh\left(\frac{x}{2}\right)}\right]. \tag{2.2.47}$$

The integral (2.2.42) is then given by

$$\langle W^{1/2}_{\square}\rangle = \frac{1}{\pi}\int_{1/a}^{a}\tan^{-1}\left[\sqrt{\frac{\alpha X - 1 - X^2}{\beta X + 1 + X^2}}\right]\mathrm{d}X. \tag{2.2.48}$$

This integral is not easy to calculate in closed form, but its derivatives w.r.t. ζ and β can be written in closed form, like the integrals (2.2.23) and (2.2.24)

$$\begin{aligned} \partial_\zeta\langle W^{1/6}_{\square}\rangle &= -\frac{1}{\pi}\frac{1}{\sqrt{ab}(1+ab)}\left(a\,K(k) - (a+b)\,\Pi(n_2|k)\right) \\ \partial_\beta\langle W^{1/6}_{\square}\rangle &= -\frac{2}{\pi}\frac{\sqrt{ab}}{a+b}\,E(k)\,. \end{aligned} \tag{2.2.49}$$

For the 1/2 BPS Wilson loop of [99] the situation is much simpler and in the planar approximation one needs only the large Z behavior of ω_0 (2.2.17)

$$\omega_0 = t + \frac{\zeta}{Z} + \frac{\zeta^2 + 2\beta^2 - 2}{2Z^2} + \frac{\zeta(\zeta^2 + 6\beta^2 - 3)}{3Z^3} + \mathcal{O}(Z^{-4}). \tag{2.2.50}$$

One finds

$$\langle W^{1/2}_{\square}\rangle_{\text{planar}} = \frac{\zeta}{2}, \tag{2.2.51}$$

which can then be expanded in different regimes. We will elaborate on the weak and strong coupling expansions of the above expression in the next sections and will also turn to the non-planar corrections in Section 2.7.

As a simple generalization, by the replacement $Z \to Z^l$ on the right hand side of (2.2.43), the higher order terms in the expansion (2.2.50) give the expectation values of multiply wrapped 1/2 BPS Wilson loops where $U \to U^l$ in (2.2.38). For even winding the sign in the lower block of the matrix U (2.2.16) is absent. This is consistent with the gauge theory calculation [99], where this sign arose from the requirement of supersymmetry invariance in the presence of the fermionic couplings which are antiperiodic, as should be the case for a singly-wound contractible cycle (see also the discussion in [107]).

The normalization of the Wilson loop as given by (2.2.42) and (2.2.43) is not the same as in the 4d $\mathcal{N} = 4$ case (2.2.44). For the 1/6 BPS loop, the leading term at weak coupling is $t_1 = 2\pi\mathrm{i}N_1/k$. This means that the trace in the fundamental is normalized by a factor of g_s. For the 1/2 BPS loop the leading term is $t_1 \pm t_2 = g_s(N_1 \mp N_2)$, where the sign depends on the winding number. We will comment more about this normalization in Section 2.5.3.

2.3 Moduli space, Picard–Fuchs equations and periods

In this section we present the tools for solving the lens space matrix model using special geometry. We present three special points in the moduli space of the theory and write explicit expressions for the four periods of ω_0 at the vicinity of these points.

In [88] it was shown that the lens space Chern–Simons matrix model is the large N dual of topological string theory on a certain class of local Calabi–Yau geometries, providing in this way a nontrivial generalization of the Gopakumar–Vafa duality [108]. In particular, the $L(2,1)$ lens space matrix model is equivalent to topological string theory on local $\mathbb{F}_0 = \mathbb{P}^1 \times \mathbb{P}^1$. The $1/N$ expansions of the free energy and of the 1/2 BPS Wilson loop VEV are the genus expansions of closed and open topological string amplitudes. The planar content of the theory is encoded in the periods of the mirror geometry described by the family of elliptic curves Σ in $\mathbb{C}^* \times \mathbb{C}^*$, which can be written as

$$Y + \frac{Z^2}{Y} - \sqrt{\frac{z_1}{z_2}}\left(Z^2 - \frac{1}{\sqrt{z_1}}Z + 1\right) = 0, \tag{2.3.1}$$

Here, z_1, z_2 parametrize the moduli space of complex structures, which is the mirror to the enlarged Kähler moduli space of local $\mathbb{F}_0$. This moduli space has a very rich structure first uncovered in [88] and further studied in, for example, [102, 109] by using the standard techniques of mirror symmetry.

Notice that the mirror geometry (2.3.1) is closely related to the resolvent $\omega_0(Z)$. Indeed, one finds that

$$\log Y = \omega_0(Z) \tag{2.3.2}$$

and

$$\zeta = \frac{1}{\sqrt{z_1}}, \qquad \beta = \sqrt{\frac{z_2}{z_1}}. \tag{2.3.3}$$

This can also be expressed as (2.2.31)

$$z_1 = \frac{\mathrm{e}^{-2\pi \mathrm{i} B}}{\kappa^2}, \qquad z_2 = \frac{\mathrm{e}^{2\pi \mathrm{i} B}}{\kappa^2}. \tag{2.3.4}$$

Let us now discuss in some detail the moduli space of (2.3.1), since it will play a fundamental role in the following. It has complex dimension two, corresponding to the two complexified Kähler parameters of local $\mathbb{F}_0$. The coordinates z_1, z_2 (or ζ, β) are global coordinates in this moduli space. Another way of parametrizing it is to use the periods of the meromorphic one-form

$$\omega = \log Y \frac{\mathrm{d}Z}{Z} \tag{2.3.5}$$

As it is well-known, these periods are annihilated by a pair of differential operators called Picard–Fuchs operators. In terms of z_1, z_2, the operators are

$$\begin{aligned} \mathcal{L}_1 &= z_2(1-4z_2)\xi_2^2 - 4z_1^2\xi_1^2 - 8z_1z_2\xi_1\xi_2 - 6z_1\xi_1 + (1-6z_2)\xi_2, \\ \mathcal{L}_2 &= z_1(1-4z_1)\xi_1^2 - 4z_2^2\xi_2^2 - 8z_1z_2\xi_1\xi_2 - 6z_2\xi_2 + (1-6z_1)\xi_1, \end{aligned} \tag{2.3.6}$$

where

$$\xi_i = \frac{\partial}{\partial z_i}. \tag{2.3.7}$$

These operators lead to a system of differential equations known as *Picard–Fuchs* (PF) equations. An important property of the moduli space is the existence of special points, generalizing the regular singular points of ODEs on $\mathbb{C}$. The PF system can be solved around these points, and the solutions give a basis for the periods of the meromorphic one-form. We can use two of the solutions to parametrize the moduli space near a singular point, and the resulting local coordinates, given by periods, are usually called *flat coordinates.*

2.3.1 Orbifold point, or weak coupling

There are three types of special points in the moduli space. The first one is the *orbifold* point discovered in [88], which is the relevant one in order to make contact with the matrix model. To study this point one has to use the global variables

$$x_1 = 1 - \frac{z_1}{z_2}, \qquad x_2 = \frac{1}{\sqrt{z_2}\left(1 - \frac{z_1}{z_2}\right)}. \tag{2.3.8}$$

The orbifold point is then defined as $x_1 = x_2 = 0$, and in terms of these variables the Picard–Fuchs system is given by the two operators

$$\begin{aligned}\mathcal{L}_1 =& \frac{1}{4}(8-8x_1+x_1^2)x_2\partial_{x_2} - \frac{1}{4}\left(4-(2-x_1)^2x_2^2\right)\partial_{x_2}^2 - x_1(2-3x_1+x_1^2)x_2\partial_{x_1}\partial_{x_2}\\ &-(1-x_1)x_1^2\partial_{x_1}+(1-x_1)^2x_1^2\partial_{x_1}^2,\\ \mathcal{L}_2 =& (2-x_1)x_2\partial_{x_2}-(1-(1-x_1)x_2^2)\partial_{x_2}^2 - x_1^2\partial_{x_1} - 2(1-x_1)x_1x_2\partial_{x_1}\partial_{x_2}+(1-x_1)x_1^2\partial_{x_1}^2.\end{aligned} \tag{2.3.9}$$

A basis of periods near the orbifold point was found in [88]. It reads,

$$\begin{aligned}\sigma_1 &= -\log(1-x_1),\\ \sigma_2 &= \sum_{m,n} c_{m,n}x_1^m x_2^n,\\ \mathcal{F}_{\sigma_2} &= \sigma_2\log x_1 + \sum_{m,n} d_{m,n}x_1^m x_2^n\ ,\end{aligned} \tag{2.3.10}$$

where the coefficients $c_{m,n}$ and $d_{m,n}$ vanish for non-positive n or m as well as for all even n. They satisfy the following recursion relations with the seed values $c_{1,1} = 1$, $d_{1,1} = 0$ and $d_{1,3} = -1/6$:

$$\begin{aligned}c_{m,n} &= \frac{(n+2-2m)^2}{4(m-n)(m-1)}c_{m-1,n},\\ c_{m,n} &= \frac{(n-2)^2(m-n+2)(m-n+1)}{n(n-1)(2m-n)^2}c_{m,n-2},\\ d_{m,n} &= \frac{(n+2-2m)^3 d_{m-1,n}+4(n^2-n-2m+2)c_{m,n}}{4(m-1)(m-n)(n+2-2m)},\\ d_{m,n} &= \frac{(n-2)^2(m-n+1)(m-n+2)}{n(n-1)(2m-n)^2}d_{m,n-2}+\left(\frac{1}{m-n+2}+\frac{1}{m-n+1}+\frac{4}{n-2m}\right)c_{m,n}.\end{aligned} \tag{2.3.11}$$

The 't Hooft parameters of the matrix model are period integrals of the meromorphic one-form, therefore they must be linear combinations of the periods above, and one finds [88]

$$t_1 = \frac{1}{4}(\sigma_1+\sigma_2), \qquad t_2 = \frac{1}{4}(\sigma_1-\sigma_2). \tag{2.3.12}$$

An expansion around the orbifold point leads to a regime in which t_1, t_2 are very small. In view of (2.2.29) this corresponds, in the ABJM model, to the *weakly coupled theory*

$$\lambda_1,\ \lambda_2 \ll 1. \tag{2.3.13}$$

The remaining period in (2.3.10) might be used to compute the genus zero free energy of the matrix model. Using the normalization appropriate for the ABJM matrix model, we find

$$\mathcal{I} = 4\frac{\partial F_0}{\partial\sigma_2} - \frac{\pi i t}{2} = \frac{1}{2}\mathcal{F}_{\sigma_2} - \log(4)\sigma_2 - \frac{\pi i}{2}\sigma_1. \tag{2.3.14}$$

2.3.2 Large radius, or strong coupling

The second point that we will be interested in is the so-called *large radius point* corresponding to $z_1 = z_2 = 0$. This is the point where the Calabi–Yau manifold is in its geometric phase, and the expansion of the genus zero free energy near that point leads to the counting of holomorphic curves with Gromov–Witten invariants. The solutions to the Picard–Fuchs equations (2.3.6) near this point can be obtained in a systematic way by considering the so-called fundamental period

$$\varpi_0(z_1,z_2;\rho_1,\rho_2) = \sum_{k,l\geq 0}\frac{\Gamma(2k+2l+2\rho_1+2\rho_2)\Gamma(1+\rho_1)^2\,\Gamma(1+\rho_2)^2}{\Gamma(2\rho_1+2\rho_2)\Gamma(1+k+\rho_1)^2\,\Gamma(1+l+\rho_1)^2}z_1^{k+\rho_1}z_2^{l+\rho_2}. \tag{2.3.15}$$

As reviewed in for example [110], a basis of solutions to the PF equations can be obtained by acting on the fundamental period with the following differential operators

$$D_i^{(1)} = \partial_{\rho_i}, \qquad D_i^{(2)} = \frac{1}{2}\kappa_{ijk}\partial_{\rho_j}\partial_{\rho_k}. \tag{2.3.16}$$

Here κ_{ijk} are the classical triple intersection numbers of the Calabi–Yau. This leads to the periods

$$\begin{aligned} T_i(z_1,z_2) &= -D_i^{(1)}\varpi_0(z_1,z_2;\rho_1,\rho_2)\Big|_{\rho_1=\rho_2=0}, \\ F_i(z_1,z_2) &= D_i^{(2)}\varpi_0(z_1,z_2;\rho_1,\rho_2)\Big|_{\rho_1=\rho_2=0}. \end{aligned} \tag{2.3.17}$$

These periods should be linearly related to those defined in the matrix model in equations (2.2.21) and (2.2.22). We present now some explicit expressions for them that we will use in Sections 2.5.2 and 2.5.4 to solve for these relations (see equations (2.5.16) and (2.5.38)).

In general, one normalizes these periods and divides them by the fundamental period evaluated at $\rho_1 = \rho_2 = 0$. But in local mirror symmetry we have [111]

$$\varpi_0(z_1,z_2;\rho_1,\rho_2)\Big|_{\rho_1=\rho_2=0} = 1. \tag{2.3.18}$$

The T_i are single-logarithm solutions, and they are identified in standard mirror symmetry with the complexified Kähler parameters, while the F_i are double-logarithm solutions and they are identified with the derivatives of the large radius genus zero free energy w.r.t. the T_i. In our case, we find the explicit expressions

$$\begin{aligned} -T_1 &= \log z_1 + \omega^{(1)}(z_1,z_2), \\ -T_2 &= \log z_2 + \omega^{(1)}(z_1,z_2), \end{aligned} \tag{2.3.19}$$

where

$$\omega^{(1)}(z_1,z_2) = 2\sum_{\substack{k,l\geq 0,\\ (k,l)\neq(0,0)}} \frac{\Gamma(2k+2l)}{\Gamma(1+k)^2\Gamma(1+l)^2} z_1^k z_2^l = 2z_1 + 2z_2 + 3z_1^2 + 12z_1z_2 + 3z_2^2 + \cdots \tag{2.3.20}$$

In order to obtain the F_i we have to compute the double derivatives w.r.t. the parameters ρ_1, ρ_2. We find

$$\partial_{\rho_1}^2\varpi_0(z_1,z_2;\rho_1,\rho_2)\Big|_{\rho_1=\rho_2=0} = \log^2 z_1 + 2\log z_1\,\omega^{(1)}(z_1,z_2) + \omega_1^{(2)}(z_1,z_2), \tag{2.3.21}$$

where

$$\omega_1^{(2)}(z_1,z_2) = 8\sum_{\substack{k,l\geq 0,\\ (k,l)\neq(0,0)}} \frac{\Gamma(2k+2l)}{\Gamma(1+k)^2\Gamma(1+l)^2}\left(\psi(2k+2l)-\psi(1+k)\right) z_1^k z_2^l. \tag{2.3.22}$$

Similarly,

$$\partial_{\rho_2}^2\varpi_0(z_1,z_2;\rho_1,\rho_2)\Big|_{\rho_1=\rho_2=0} = \log^2 z_2 + 2\log z_2\,\omega^{(1)}(z_1,z_2) + \omega_2^{(2)}(z_1,z_2) \tag{2.3.23}$$

where

$$\omega_2^{(2)}(z_1,z_2) = 8\sum_{\substack{k,l\geq 0,\\ (k,l)\neq(0,0)}} \frac{\Gamma(2k+2l)}{\Gamma(1+k)^2\Gamma(1+l)^2}\left(\psi(2k+2l)-\psi(1+l)\right) z_1^k z_2^l = \omega_1^{(2)}(z_2,z_1). \tag{2.3.24}$$

Finally,

$$\begin{aligned} \partial_{\rho_1}\partial_{\rho_2}\varpi_0(z_1,z_2;\rho_1,\rho_2)\Big|_{\rho_1=\rho_2=0} &= \log z_1\log z_2 + (\log z_1 + \log z_2)\,\omega^{(1)}(z_1,z_2) \\ &\quad + \frac{1}{2}\left(\omega_1^{(2)}(z_1,z_2) + \omega_2^{(2)}(z_1,z_2)\right). \end{aligned} \tag{2.3.25}$$

The double log periods are obtained as linear combinations of the above, by using the explicit expressions for the classical intersection numbers that can be found in for example [109]

$$\kappa_{111} = \frac{1}{4}, \qquad \kappa_{112} = -\frac{1}{4}, \qquad \kappa_{122} = -\frac{1}{4}, \qquad \kappa_{222} = \frac{1}{4}. \tag{2.3.26}$$

We find:

$$\begin{aligned} F_1(z_1,z_2) &= -\frac{1}{8}\left(D_{\rho_1}^2\omega_0 - 2D_{\rho_1\rho_2}\omega_0 - D_{\rho_1}^2\omega_0\right) \\ &= -\frac{1}{8}\left(\log^2 z_1 - 2\log z_1\log z_2 - \log^2 z_2\right) + \frac{1}{4}\log z_2\,\omega^{(1)}(z_1,z_2) + \frac{1}{8}\omega_2^{(2)}(z_1,z_2), \\ F_2(z_1,z_2) &= -\frac{1}{8}\left(D_{\rho_1}^2\omega_0 - 2D_{\rho_1\rho_2}\omega_0 - D_{\rho_1}^2\omega_0\right) \\ &= -\frac{1}{8}\left(-\log^2 z_1 - 2\log z_1\log z_2 + \log^2 z_2\right) + \frac{1}{4}\log z_1\,\omega^{(1)}(z_1,z_2) + \frac{1}{8}\omega_1^{(2)}(z_1,z_2). \end{aligned} \tag{2.3.27}$$

They satisfy the symmetry property

$$F_1(z_1, z_2) = F_2(z_2, z_1). \tag{2.3.28}$$

The reason why we are interested in the large radius point is because it describes the structure of the ABJM theory at *strong coupling*. In the region where z_2 is small, x_2 is large and the periods $t_{1,2}$ grow. In general, the expansions of the periods around the special points have a finite radius of convergence, but they can be analytically continued to the other "patches". Since their analytic continuation satisfies the PF equation, we know for example that the analytic continuation of the orbifold periods to the large radius patch must be linear combinations of the periods at large radius. This provides an easy way to perform the analytic continuation which will be carried out in detail in the Section 2.5, where we will verify that indeed the region near the large radius point corresponds to

$$\lambda_1, \lambda_2 \gg 1. \tag{2.3.29}$$

2.3.3 Conifold locus

Finally, the third set of special points is the *conifold locus*. This is defined by $\Delta = 0$, where

$$\Delta = 1 - 8(z_1 + z_2) + 16(z_1 - z_2)^2. \tag{2.3.30}$$

In terms of the variables ζ, β, this locus corresponds to the four lines

$$\zeta = -2\beta \pm 2, \qquad \zeta = 2\beta \pm 2. \tag{2.3.31}$$

The conifold locus is the place where cycles in the geometry collapse to zero size. The first two lines correspond to $a = \pm 1$, *i.e.*, the collapse of the $\mathcal{C}_1$ cycle, while the second set of lines corresponds to $b = \mp 1$, *i.e.*, to the collapse of the $\mathcal{C}_2$ cycle. In principle we can solve the PF system near any point in the conifold locus, but in practice it is useful to focus on the point

$$z_1 = z_2 = \frac{1}{16} \tag{2.3.32}$$

which has been studied in [109]. We will call it the symmetric conifold point. Appropriate global coordinates around this point are[2]

$$y_1 = 1 - \frac{z_1}{z_2}, \qquad y_2 = 1 - \frac{1}{16 z_1}. \tag{2.3.33}$$

In terms of these coordinates, the PF system reads

$$\begin{aligned}
\mathcal{L}_1 =& \partial_{y_2} - 2(1-y_2)\partial_{y_2}^2 - 8(1-y_1)^2\partial_{y_1} + 8(1-y_1)^3\partial_{y_1}^2, \\
\mathcal{L}_2 =& -(7-8y_2)\partial_{y_2} + 2(3-7y_2+4y_2^2)\partial_{y_2}^2 - 8(1-y_1)\partial_{y_1} \\
& - 16(1-y_1)(1-y_2)\partial_{y_1}\partial_{y_2} + 8(1-y_1)^2\partial_{y_1}^2.
\end{aligned} \tag{2.3.34}$$

Notice that, strictly speaking, the orbifold point does not belong to the conifold locus, once the moduli space is compactified and resolved [88]. A generic point in the conifold locus has then $t_1 = 0$ or $t_2 = 0$, but not both, and expanding around the conifold locus means, in the ABJM theory, an expansion in the region

$$\lambda_1 \ll 1, \qquad \lambda_2 \sim 1, \tag{2.3.35}$$

or in the region with λ_2 exchanged with λ_1. This regime of the ABJM theory has been considered in [112].

It was observed in [113] that the moduli space of the local $\mathbb{F}_0$ surface can be mapped to a well-known moduli space, namely the Seiberg–Witten (SW) u-plane [10]. This plane is parametrized by a single complex variable u. The relation between the moduli is

$$u = \frac{1}{2}\left(\beta + \beta^{-1}\right) - \frac{\zeta^2}{8\beta}. \tag{2.3.36}$$

[2] These are slightly different from the ones used in [109].

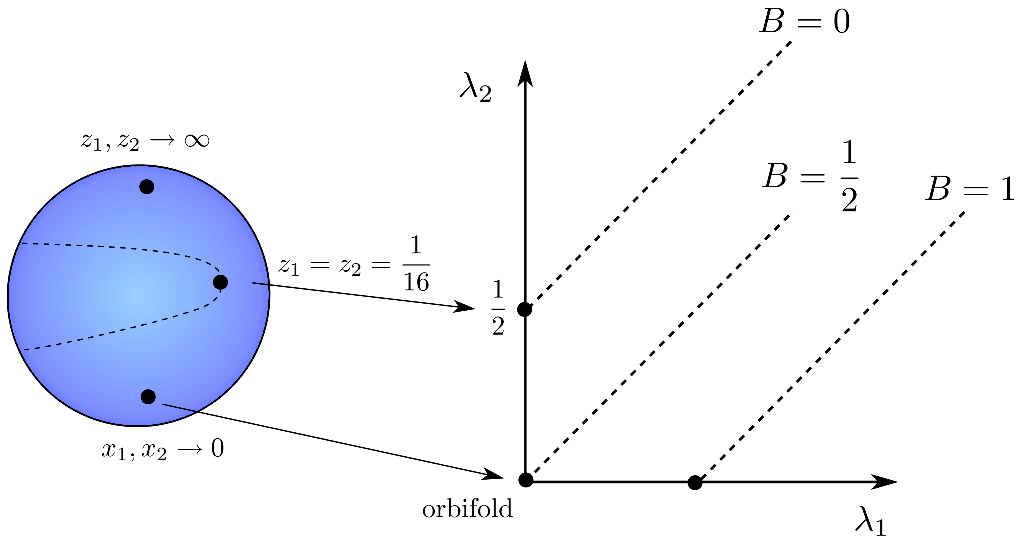

Figure 2.4: The moduli space of the ABJM theory, describing the possible values of the 't Hooft couplings $\lambda_{1,2}$, can be parametrized by a real submanifold of the moduli space of local $\mathbb{F}_0$, here depicted as a sphere. The orbifold point maps to the origin, while the conifold locus (which is represented by a dashed line) maps to the two axes.

The three singular points that we have discussed (large radius, orbifold, and symmetric conifold) map to the points $u = \infty, +1, -1$. These are the semiclassical, monopole and dyon points of SW theory. As we will see, they can be identified with interesting points in ABJM theory.

An important set of quantities in the study of moduli spaces of CY threefolds are the three-point couplings or Yukawa couplings, $C_{z_i z_j z_k}$. These are the components of a completely symmetric degree three covariant tensor on the moduli space. When expressed in terms of flat coordinates they give the third derivatives of the genus zero free energy. In terms of the coordinates z_1, z_2, the Yukawa couplings are given by [88, 109]

$$\begin{aligned} C_{111} &= \frac{(1-4z_2)^2 - 16z_1(1+z_1)}{4z_1^3\Delta}, \\ C_{112} &= \frac{16z_1^2 - (1-4z_2)^2}{4z_1^2 z_2\Delta}, \\ C_{122} &= \frac{16z_2^2 - (1-4z_1)^2}{4z_1 z_2^2\Delta}, \\ C_{222} &= \frac{(1-4z_1)^2 - 16z_2(1+z_2)}{4z_2^3\Delta}. \end{aligned} \tag{2.3.37}$$

2.3.4 The moduli space of the ABJM theory

The matrix model of ABJM is closely related to the lens space matrix model, and therefore so are also the moduli spaces of the theories. Some of the explicit relations needed for this identification will be presented only in the following sections, but we would still like to present here the main points on the moduli space.

We can think about the *moduli space of the planar ABJM theory* as the space of admissible values of the 't Hooft parameters λ_1, λ_2. We will assume for simplicity that $k > 0$. The theory with negative values of k can be obtained from this one by a parity transformation. In the gauge theory $\lambda_{1,2}$ must be rational and non-negative (for $k > 0$). Moreover, according to [3], any value of $\lambda_{1,2}$ is admissible as long as

$$|\lambda_1 - \lambda_2| \leq 1. \tag{2.3.38}$$

This moduli space can be parametrized by the B field and κ, which from the explicit expressions derived below (2.4.1) and (2.5.22) has to be real and positive. It can be identified as a real submanifold of the moduli space of local $\mathbb{F}_0$. Moreover, we can identify the singular points of this moduli space with natural limits of ABJM theory (see Fig. 2.4):

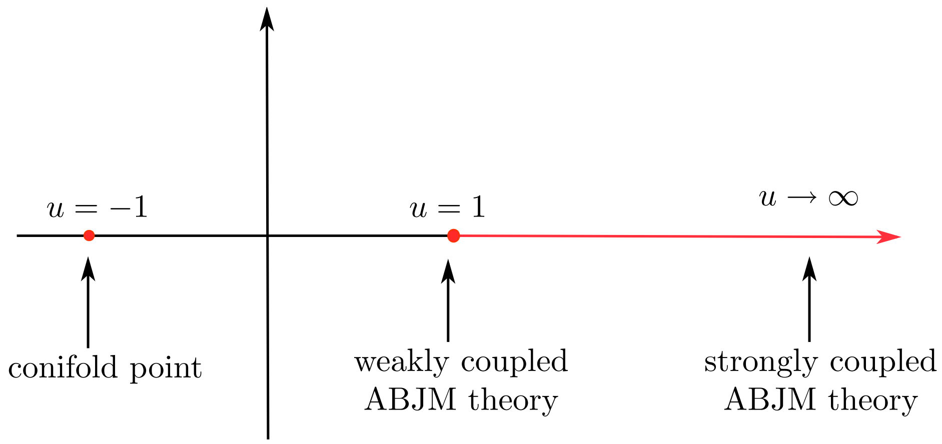

Figure 2.5: The moduli space of the ABJM theory for $B = 1/2$ can be mapped to the line $[1, \infty)$ in the u plane of Seiberg–Witten theory, which is here shown in red. The monopole point corresponds to the weakly coupled ABJM theory, while the semiclassical limit corresponds to the strongly coupled theory.

1. The *weak coupling regime* $\lambda_{1,2} \to 0$ corresponds to the orbifold point of the local $\mathbb{F}_0$ geometry $\kappa = 0$, $B = 1/2$. In terms of type IIA theory, this is also an orbifold geometry with a small radius but a nonzero value for the B field.

2. The *strong coupling regime* $\lambda_{1,2} \to \infty$ (where also $\kappa \to \infty$) corresponds to the large radius limit of the local $\mathbb{F}_0$ geometry.

3. Out of the four lines (2.3.31) in the conifold locus $\Delta = 0$, only two lead to $\kappa \in \mathbb{R}$. They are the curves in the (κ, B) plane with $\kappa = \pm 4 \cos \pi B$, which correspond respectively to $a = 1$ and $b = 1$, therefore to $\lambda_1 = 0$ or $\lambda_2 = 0$. Hence, the boundary of the ABJM moduli space given by $\min(\lambda_1, \lambda_2) = 0$ corresponds to

$$\kappa(B) = \begin{cases} -4\cos \pi B\,, & B > 1/2 \\ 4 \cos \pi B\,, & B < 1/2 \end{cases} \tag{2.3.39}$$

   In particular, the symmetric conifold point $z_1 = z_2 = 1/16$ corresponds to $B = n \in \mathbb{Z}$, $\kappa = \pm 4$. Along the curve (2.3.39), one of the two gauge groups of the ABJM theory is absent, so the theory reduces to topological CS theory. We examine this regime in Section 2.6.

Given a fixed value of the B field, we can describe the real one-dimensional moduli space of the ABJM theory as a real submanifold of the u-plane of Seiberg–Witten theory, by using (2.3.36) in the form

$$u = -\cos(2\pi B) + \frac{\kappa^2}{8}. \tag{2.3.40}$$

Singular points in moduli space become then the well-known singularities of SW theory. For example, when $B = 1/2$, the moduli space, described by $\kappa \in [0, \infty)$, becomes the region $u \in [1, \infty)$. The orbifold point (weakly coupled ABJM theory) maps to the monopole point, while the large radius point (strongly coupled ABJM theory) corresponds to the semi-classical region (see Fig. 2.5). Notice that the conifold point would map to the dyon point of Seiberg–Witten theory, but this does not belong to the moduli space of ABJM theory with $B = 1/2$. We can however realize it by making an analytic continuation of the 't Hooft coupling to complex values. The dyon point corresponds then to the point $\kappa^2 = -16$, which leads by (2.5.9) to an imaginary value

$$\lambda = -\frac{2\mathrm{i}\mathrm{K}}{\pi^2}, \tag{2.3.41}$$

where K is Catalan's number.

As usual, string dualities lead to a full complexification of the moduli space of 't Hooft parameters. In the case of ABJM theory, the complexified moduli space for the variables $\lambda_{1,2}$ is simply the moduli space of the parameters β, ζ, which is a $\mathbb{Z}_2 \times \mathbb{Z}_2$ covering of the moduli space parametrized by $z_{1,2}$.

2.4 Weak coupling

In principle, to study the matrix model at weak coupling one does not need the sophisticated tools presented in the previous section. One can do perturbative calculations directly in the integral expressions (2.2.3) or (2.2.4) for the matrix model. A calculation of the 1/6 BPS Wilson loop to three loop order was indeed done in this way in the original paper [2].

Still, the explicit expressions for the periods $\sigma_{1,2}$ (2.3.10) and their relation to $t_{1,2}$ (2.3.12) gives a much more efficient way to obtain perturbative, planar expansions. Inverting these relations we find the weak coupling expression for κ (2.2.31)

$$\begin{aligned}\kappa = & -2\mathrm{i}(t_1 - t_2) - \frac{\mathrm{i}}{12}\left(t_1^3 + 3t_1^2 t_2 - 3t_1 t_2^2 - t_2^3\right) \\ & - \frac{\mathrm{i}}{960}\left(t_1^5 + 5t_1^4 t_2 - 10t_1^3 t_2^2 + 10t_1^2 t_2^3 - 5t_1 t_2^4 - t_2^5\right) + \mathcal{O}(t^7).\end{aligned} \tag{2.4.1}$$

This agrees with the weak coupling expansion of the inverse of the exact mirror map (2.5.9).

Using the dictionary relating the 't Hooft couplings (2.2.29) we immediately get the result for the 1/2 BPS Wilson loop in the planar approximation (2.2.51)

$$\begin{aligned}\langle W_{\square}^{1/2}\rangle = \mathrm{e}^{\pi \mathrm{i} B}\frac{\kappa}{2} = \mathrm{e}^{\pi\mathrm{i}(\lambda_1-\lambda_2)}\, 2\pi\mathrm{i}(\lambda_1+\lambda_2)\bigg[1 - \frac{\pi^2}{6}\left(\lambda_1^2 - 4\lambda_1\lambda_2 + \lambda_2^2\right) \\ + \frac{\pi^4}{120}\left(\lambda_1^4 - 6\lambda_1^3\lambda_2 - 4\lambda_1^2\lambda_2^2 - 6\lambda_1\lambda_2^3 + \lambda_2^4\right) + \mathcal{O}(\lambda^6)\bigg].\end{aligned} \tag{2.4.2}$$

In this expression we factored out the term $2\pi\mathrm{i}(\lambda_1 + \lambda_2)$, which depends on the overall normalization of the Wilson loop, as mentioned after (2.2.51). There is also the extra phase factor, which appears also at strong coupling and can be attributed to framing. Note that so far this expansion has not been reproduced directly in the gauge theory, as even the two-loop graphs are quite subtle.

For the 1/6 BPS Wilson loop, using the explicit expression (2.2.49) and expanding at low orders one finds

$$\left\langle W_{\square}^{1/6}\right\rangle = \mathrm{e}^{\pi\mathrm{i}\lambda_1} 2\pi\mathrm{i}\lambda_1\left(1 - \frac{\pi^2}{6}\lambda_1(\lambda_1 - 6\lambda_2) - \frac{\pi^3\mathrm{i}}{2}\lambda_1\lambda_2^2 + \frac{\pi^4}{120}\lambda_1\left(\lambda_1^3 - 10\lambda_1^2\lambda_2 - 20\lambda_2^3\right) + \mathcal{O}(\lambda^5)\right). \tag{2.4.3}$$

Again the exponent is a framing factor and the factor of $2\pi\mathrm{i}\lambda_1$ is due to the normalization chosen in (2.2.42). This expression agrees with the 2-loop calculations in [103, 104, 105]. Note that the 3-loop analysis in [105], done for $\lambda_1 = \lambda_2$, misses the next term, due to a projection which essentially removes all terms at odd orders in perturbation theory.

Next we turn to the free energy. Here we notice that the period in (2.2.22) gives only the derivative of the free energy. Indeed, within the formalism of special geometry developed above, the planar free energy of the matrix model is only determined up to quadratic terms in the 't Hooft couplings. These have to be fixed by direct calculation in the matrix model

$$F = \frac{N_1^2}{2}\log\left(\frac{2\pi N_1}{k}\right) + \frac{N_2^2}{2}\log\left(\frac{2\pi N_2}{k}\right) - \frac{3}{4}(N_1^2 + N_2^2) - \log(4)N_1 N_2 + \cdots \tag{2.4.4}$$

The last term comes from the normalization of the cosh term in (2.2.3), while the remaining terms are just the free energies for two Gaussian matrix models with couplings $\pm 2\pi\mathrm{i}/k$. Notice that the above free energy has an imaginary piece given by

$$\frac{\pi\mathrm{i}}{6k}(N_1 - N_2)((N_1 - N_2)^2 - 1). \tag{2.4.5}$$

Using the identification of the periods at weak coupling (2.3.14) we write down the next term in the perturbative expansion

$$\frac{\pi^2}{72k^2}\left(N_1^4 - 6N_1^3 N_2 + 18N_1^2 N_2^2 - 6N_1 N_2^3 + N_2^4\right). \tag{2.4.6}$$

It would be interesting to try to reproduce these expressions directly from studying perturbative ABJM theory on $\mathbb{S}^3$.

2.5 Strong coupling expansion and the AdS dual

We turn now to the strong coupling limit of the matrix model, where we have to find the analytic continuation of the 't Hooft parameters to the strong coupling region, as functions of the global parameters of moduli space. We will see how the shift of the charges discussed in [77, 78] emerges naturally from our computation. We will also evaluate the free energy in this regime and compare with the classical action of the vacuum AdS dual, deriving in this way the $N^{3/2}$ behavior of the degrees of freedom.

2.5.1 ABJM slice

In the original ABJM theory with $N_1 = N_2 = N$ (the case $N_1 \neq N_2$ was considered in [3]) we should look at the slice

$$t_1 = -t_2 = 2\pi \mathrm{i}\lambda, \quad \lambda = \frac{N}{k} \tag{2.5.1}$$

in the moduli space of the dual topological string. From the point of view of the periods σ_1, σ_2 in (2.3.10) this means that we should set

$$\sigma_1 = 0, \tag{2.5.2}$$

therefore $x_1 = 0$. In order to have a nontrivial σ_2, we must consider the double-scaling limit

$$x_1 \to 0, \qquad x_1 x_2 = \zeta \quad \text{fixed.} \tag{2.5.3}$$

The one-dimensional subspace (2.5.1) corresponds, in terms of the variables ζ, ξ, to $\xi = 2$. As in [30], we can find simplified expressions for the periods in this subspace. It is easy to see from the structure of σ_2 that, in the limit (2.5.3), one has

$$\sigma_2 = \sum_{m=0}^{\infty} a_m \zeta^{2m+1}, \quad a_m = c_{2m+1,2m+1}, \tag{2.5.4}$$

and from the recursion relation (2.3.11) we find

$$a_m = \frac{2^{-4m}\Gamma\left(m+\frac{1}{2}\right)^2}{\pi(2m+1)\Gamma\left(m+1\right)^2}. \tag{2.5.5}$$

We then obtain

$$\frac{\mathrm{d}\sigma_2}{\mathrm{d}\zeta} = \frac{2}{\pi} K\left(\frac{\zeta}{4}\right), \tag{2.5.6}$$

which is in fact a particular case of (2.2.23), as it can be easily seen by using the transformation properties of the elliptic integral $K(k)$. The period t_1 itself can be written as a generalized hypergeometric function:

$$t_1(\zeta) = \frac{\zeta}{4} \, {}_3F_2\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}; 1, \frac{3}{2}; \frac{\zeta^2}{16}\right). \tag{2.5.7}$$

In the physical ABJM theory, t_1 is purely imaginary. This means that ζ is purely imaginary as well, so we set

$$\zeta = \mathrm{i}\kappa \tag{2.5.8}$$

and we finally obtain

$$\lambda(\kappa) = \frac{\kappa}{8\pi} \, {}_3F_2\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}; 1, \frac{3}{2}; -\frac{\kappa^2}{16}\right). \tag{2.5.9}$$

As a check, we can perform a weak coupling expansion. The weakly coupled region corresponds to

$$\kappa \ll 1, \qquad \lambda \ll 1, \tag{2.5.10}$$

and in this region the variables are related as

$$\frac{\kappa}{8\pi} = \lambda + \frac{\pi^2\lambda^3}{3} - \frac{7\pi^4\lambda^5}{60} + \frac{173\pi^6\lambda^7}{1260} - \frac{37927\pi^8\lambda^9}{181440} + \mathcal{O}\left(\lambda^{10}\right), \tag{2.5.11}$$

which is obtained from the inversion of (2.5.9). This is in agreement with (2.4.1).

Of course, the main advantage of having analytic expressions is that one can perform a weak-strong coupling interpolation easily. The strong coupling region is

$$\kappa \gg 1, \qquad \lambda \gg 1 \tag{2.5.12}$$

and (2.5.9) leads to the asymptotic expansion

$$\lambda(\kappa) = \frac{\log^2(\kappa)}{2\pi^2} + \frac{1}{24} + \mathcal{O}\left(\frac{1}{\kappa^2}\right). \tag{2.5.13}$$

2.5.2 Analytic continuation and shifted charges

In order to perform the analytic continuation of the 't Hooft parameters, we use the explicit representation of the periods in terms of integrals given in (2.2.21) as well as their derivatives (2.2.23)-(2.2.24). Let us start by discussing t_1. We study its behavior at large ζ but fixed β, which is the large radius region. We find

$$\frac{\partial t_1}{\partial \zeta} = \frac{\mathrm{i}}{\pi\zeta} \log\left(-\frac{\zeta^2}{\beta}\right) + o(\zeta^{-1}), \qquad \frac{\partial t_1}{\partial \beta} = -\frac{\mathrm{i}}{2\pi\beta}\left(\log(-\zeta^2) + \pi \mathrm{i}\right) + o(1), \tag{2.5.14}$$

and this gives the leading behavior

$$t_1 = -\frac{\mathrm{i}}{2\pi}\left(\log(-\zeta^2) + \pi\mathrm{i}\right)\log\frac{\beta}{\zeta} + \cdots \tag{2.5.15}$$

In the physical theory t_1 should be imaginary and β a phase. By examining (2.5.15), this implies that κ is real. From (2.3.4) we then see that $z_1 = \bar{z}_2$ and henceforth we label it $z_1 = z$.

We know also that t_1 must be a linear combination of the periods at large radius. Using that $z_1 = 1/\zeta^2$ and $z_2 = (\beta/\zeta)^2$, and comparing (2.5.15) to the behavior of the periods (2.3.19) and (2.3.27), we find

$$\begin{aligned} t_1 &= \frac{\mathrm{i}}{2\pi}(F_1 + F_2) - \frac{1}{2}T_2 - \frac{\pi \mathrm{i}}{6}, \\ t_2 &= -\frac{\mathrm{i}}{2\pi}(F_1 + F_2) + \frac{1}{2}T_1 + \frac{\pi\mathrm{i}}{6}. \end{aligned} \tag{2.5.16}$$

The constants $\pm\pi\mathrm{i}/6$ cannot be fixed by using the above information, but they can be fixed by specializing to the ABJM slice $z_1 = z_2$, as we will see in a moment.

A simple calculation leads to the following explicit expression

$$\lambda_1(\kappa, B) = \frac{1}{2}\left(B^2 - \frac{1}{4}\right) + \frac{1}{24} + \frac{\log^2\kappa}{2\pi^2} - \frac{\log\kappa}{2\pi^2}\omega^{(1)}(z,\bar{z}) + \frac{1}{16\pi^2}\left(\omega_1^{(2)} + \omega_2^{(2)}\right)(z,\bar{z}). \tag{2.5.17}$$

This expansion is valid in the region $\kappa \to +\infty$. Notice that it is manifestly real when κ is real and positive.

As a check of the above expression, we can particularize to the ABJM slice $\lambda_1 = \lambda_2 = \lambda$, ($B = 1/2$), which corresponds in the gauge theory, to having identical gauge groups in the two nodes of the quiver, *i.e.*, $N_1 = N_2$. The mirror map for this case was obtained in the previous section and its strong coupling expansion (2.5.13) in agreement with (2.5.17). This also fixes the constants in (2.5.16).

The observables of the model are naturally functions of ζ, β (alternatively κ, B), and we have to re-express them in terms of $\lambda_{1,2}$. Equation (2.5.17) shows that the natural variable at strong coupling is not λ_1, but rather

$$\hat{\lambda} = \lambda_1 - \frac{1}{2}\left(B^2 - \frac{1}{4}\right) - \frac{1}{24} = \frac{1}{2}(\lambda_1 + \lambda_2) - \frac{1}{2}(\lambda_1 - \lambda_2)^2 - \frac{1}{24}. \tag{2.5.18}$$

In particular, it is only when expressed in terms of this variable that κ is a periodic function of $\hat{\lambda}, B$.

Remarkably, the above shift is precisely the one found in [78]. In the type IIA realization of the ABJ theory $U(M_2)_k \times U(M_2 + M_4)_{-k}$, where M_2 corresponds to the number of D2 branes and M_4 to the number of D4 branes, the Maxwell charge of the D2 branes is not M_2, but rather

$$Q_2 = M_2 - \frac{k}{2}\left(B^2 - \frac{1}{4}\right) - \frac{1}{24}\left(k - \frac{1}{k}\right), \tag{2.5.19}$$

where

$$B = -\frac{M_4}{k} + \frac{1}{2}. \tag{2.5.20}$$

After dividing by k and taking the large k limit, we recover (2.5.18) with

$$\hat{\lambda} = \frac{Q_2}{k}. \tag{2.5.21}$$

The relation between $\hat{\lambda}$ and κ can be inverted at strong coupling and it is of the form

$$\kappa(\hat{\lambda}, B) = \mathrm{e}^{\pi\sqrt{2\hat{\lambda}}} \left(1 + \sum_{\ell \geq 1} c_\ell \left(\frac{1}{\pi\sqrt{2\hat{\lambda}}}, \beta\right) \mathrm{e}^{-2\ell\pi\sqrt{2\hat{\lambda}}}\right) \tag{2.5.22}$$

where

$$c_\ell(x, \beta) = \sum_{k=0}^{2\ell-1} c_k^{(\ell)}(\beta) x^k. \tag{2.5.23}$$

The coefficients $c_k^{(\ell)}(\beta)$ are Laurent polynomials in β, β^{-1}, of degree ℓ, and symmetric under the exchange $\beta \leftrightarrow \beta^{-1}$. In other words, they can be written as polynomials in $\cos(2\pi m B)$, so they are periodic in B, with period 1. We find, for example,

$$\begin{aligned} c_1(x,\beta) &= -\left(\beta + \beta^{-1}\right)\left(1 - \frac{x}{2}\right), \\ c_2(x,\beta) &= 3 + \frac{x}{8}\left(3\beta^2 - 8 + 3\beta^{-2}\right) - \frac{3x^2}{8}\left(\beta + \beta^{-1}\right)^2 - \frac{x^3}{8}\left(\beta + \beta^{-1}\right)^2. \end{aligned} \tag{2.5.24}$$

The fact that $c_\ell(x,\beta)$ are polynomials in x of degree $2\ell - 1$, rather than power series, comes out from an explicit calculation of the first few cases, and we have not established it.

From the explicit expression (2.5.17) we can implement the symmetries of the model as a function of κ and B (or equivalently, z_1 and z_2). For example, the transformation

$$N_1 \to 2N_1 + k - N_2, \qquad N_2 \to N_1 \tag{2.5.25}$$

simply corresponds to periodicity in the B field

$$B \to B + 1 \tag{2.5.26}$$

while κ remains unchanged. From the point of view of the $z_{1,2}$ variables, this is simply a monodromy transformation $z_{1,2} \to \mathrm{e}^{\mp 2\pi\mathrm{i}} z_{1,2}$. Notice that not all the values of κ lead to admissible values of $\lambda_{1,2}$, since $\min(\lambda_1, \lambda_2) \geq 0$. This means that the boundary of moduli space is the conifold locus (2.3.39).

2.5.3 Wilson loops at strong coupling and semi–classical strings

On the ABJM slice the planar 1/6-BPS Wilson loop is determined by the single equation

$$\frac{\mathrm{d}}{\mathrm{d}\kappa}\langle W_\square^{1/6}\rangle = -\frac{\mathrm{i}}{\pi}\frac{1}{\sqrt{ab}(1+ab)}\left(a\,K(k) - (a+b)\,\Pi(n_2|k)\right). \tag{2.5.27}$$

Then it is easy to check that

$$\frac{\mathrm{d}\langle W_\square^{1/6}\rangle}{\mathrm{d}\kappa} = -\frac{1}{2\pi}\log\kappa + \frac{\mathrm{i}}{4} + \mathcal{O}\left(\frac{1}{\kappa^2}\right) \Rightarrow \langle W_\square^{1/6}\rangle = -\frac{1}{2\pi}\,\kappa\,\log\kappa + \left(\frac{1}{2\pi} + \frac{\mathrm{i}}{4}\right)\kappa + \mathcal{O}\left(\frac{1}{\kappa}\right) \tag{2.5.28}$$

It follows that

$$\langle W_{\square^{1/6}}\rangle \sim -\frac{\sqrt{2\lambda}}{2}\mathrm{e}^{\pi\sqrt{2\lambda}}, \qquad \lambda \gg 1. \tag{2.5.29}$$

The leading exponential is in perfect agreement with the AdS prediction (2.2.36), and the exact answer interpolates between the weak and the strong coupling behaviors.

As an application of the explicit expression for κ (2.5.22), we can use (2.2.31) to immediately obtain the VEV of the 1/2 BPS Wilson loop (2.2.51) at strong coupling

$$\langle W_\square^{1/2}\rangle_{g=0} = \frac{1}{2}\,\mathrm{e}^{\pi\mathrm{i}B}\kappa(\hat{\lambda}, B). \tag{2.5.30}$$

Note that this is a real function of $\hat{\lambda}, B$, up to the overall phase involving the B field. This is the same phase that appears also in the weak-coupling result (2.4.2) and arises also in field theory calculations as a framing-dependant term [46, 147, 148]. The matrix model always gives the answer for framing=1.

At strong coupling we have,

$$\langle W^{1/2}_{\Box}\rangle_{g=0} \sim -\frac{1}{2}\,\mathrm{e}^{\pi \mathrm{i} B}\mathrm{e}^{\pi\sqrt{2\lambda}}. \tag{2.5.31}$$

which displays the same leading exponential behavior predicted by the large N dual.

The computations at strong coupling can be easily extended to the case $N_1 \neq N_2$, but they depend on the direction in which we take the limit in the space of 't Hooft parameters. For ξ fixed and ζ large (therefore $\lambda_1 \sim \lambda_2$), we find the same exponential behavior

$$\langle W_{\Box}\rangle \sim \mathrm{e}^{\pi\sqrt{2\lambda_1}} \sim \mathrm{e}^{\pi\sqrt{\lambda_1+\lambda_2}}. \tag{2.5.32}$$

We would like to comment about the normalization of the operators. As mentioned after (2.2.51), the normalization chosen there is such that the trace of the identity in the fundamental of $U(N_1)$ gives $t_1 = 2\pi \mathrm{i} N_1/k$ and for the fundamental of $U(N_1|N_2)$ (with a minus sign as in (2.2.16), it gives $t_1 - t_2 = 2\pi\mathrm{i}(N_1 + N_2)/k$. In CS theory these normalizations are quite common, but they may be not the most natural ones in the ABJM theory.

An alternative normalization is to divide by this term, such that at weak coupling the expansion of the Wilson loop will be $\langle W\rangle \sim 1 + \cdots$. This is the normalization chosen in [5], and hence the slight differences in the preceding equations from that reference. Note, though, that with such a normalization, one would have to divide the doubly-wound 1/2 BPS Wilson loop in the fundamental representation by the super-trace of the identity, which is $2\pi\mathrm{i}(N_1 - N_2)/k$ and is singular for $N_1 = N_2$.

There should be a natural choice of normalization that would reproduce the correct normalization fo the one-loop partition function of the classical string in $\mathrm{AdS}_4 \times \mathbb{CP}^3$. To this day, though, a fully satisfactory calculation for the analog string in $\mathrm{AdS}_5 \times \mathbb{S}^5$ giving the factor of $\lambda^{-3/4}$ derived from the the Gaussian matrix model does not exist. One argument, based on world-sheet arguments was given in [202], but it is not clear why this argument would be modified for ABJM theory. Direct calculations of the determinant [149, 200] were not conclusive. A possible trick to derive it was proposed in [152] by considering a 1/4 BPS generalization of the circular Wilson loop, where three zero modes of the the Wilson loop of [153] are explicitly broken and the integral over them gives this factor. It would be interesting to construct such generalization to the Wilson loop of [99] and see if a similar argument can be derived from that.

Regardless of the overall normalization, one can compare those of the 1/2 BPS loop and the 1/4 BPS loop. Ignoring numerical constants and the framing factor, the ratio is

$$\frac{\langle W^{1/6}_{\Box}\rangle_{g=0}}{\langle W^{1/2}_{\Box}\rangle_{g=0}} \approx \sqrt{\lambda}, \tag{2.5.33}$$

which is proportional to the volume of a $\mathbb{CP}^1$ inside $\mathbb{CP}^3$. Indeed, it was argued in [103, 105] that the string description of the 1/6 BPS Wilson loop should be in terms of a string smeared over such a cycle.

2.5.4 The planar free energy and a derivation of the $N^{3/2}$ behaviour

In this section we study the free energy at strong coupling. We derive the $N^{3/2}$ behavior characteristic of M2 branes [4], and we match the exact coefficient with a gravity calculation in type IIA superstring on $\mathrm{AdS}_4 \times \mathbb{CP}^3$.

The free energy of the matrix model has a large N expansion of the form

$$F = \log Z = \sum_{g=0}^{\infty} g_s^{2g-2} F_g(\lambda_1, \lambda_2). \tag{2.5.34}$$

This is the way the genus expansion is typically expressed in topological string theory. To compare with the gauge theory and the AdS dual one may choose to rewrite this series as an expansion in powers of $1/N$ by absorbing factors of λ into F_g.

As mentioned in Section 2.4, the formalism of special geometry determines the planar free energy only up to quadratic terms in the 't Hooft couplings, and these have to be fixed from the explicit weak coupling calculation in the matrix model (2.4.4).

Let us now consider the derivative of the genus zero free energy (2.2.22), and study its analytic continuation to strong coupling as we have done for t_i at the top of Section 2.5.2. Expanding (2.2.26) for large κ we find

$$\frac{\partial \mathcal{I}}{\partial \zeta} = -\frac{\pi \mathrm{i}}{\zeta} + \mathcal{O}(\zeta^{-2}), \qquad \frac{\partial \mathcal{I}}{\partial \beta} = \mathcal{O}(\zeta^{-1}), \tag{2.5.35}$$

so

$$\mathcal{I} = -\pi \mathrm{i} \log \zeta + \mathcal{O}(\zeta^0) = -\pi \mathrm{i} \log \kappa + \pi^2 B + \mathcal{O}(\kappa^0, B^0), \qquad \kappa \to \infty, \tag{2.5.36}$$

From this leading large κ behavior we have that in the ABJM slice

$$\frac{\partial F_0}{\partial \lambda} \approx 2\pi^3 \sqrt{2\lambda}, \tag{2.5.37}$$

which can be integrated to give the leading term in (2.5.49) and the match with the supergravity calculation presented below.

But to get the full series of corrections we should proceed more carefully. We know that the result of the continuation should be a linear combination of periods, and comparing to (2.3.19) we see that we can express the period as

$$\mathcal{I} + \frac{\pi \mathrm{i} t}{2} = \frac{\partial F_0}{\partial t_1} - \frac{\partial F_0}{\partial t_2} = -\frac{\pi \mathrm{i}}{4}\left(T_1 + T_2 + 2\pi \mathrm{i}\right). \tag{2.5.38}$$

The constant term can be fixed by looking at the solution on the ABJM slice $N_1 = N_2$, which can be obtained as follows. Since on the slice we effectively have a one-parameter model, there is only one Yukawa coupling, which we can integrate to obtain F_0. From (2.3.37) we easily obtain

$$\partial_\lambda^3 F_0(\lambda) = \frac{1}{4} C_{\lambda\lambda\lambda}\Big|_{\lambda_1=-\lambda_2} = -\frac{128\pi^6}{\kappa(\kappa^2+16)} \frac{1}{K\left(\frac{\mathrm{i}\kappa}{4}\right)^3} \tag{2.5.39}$$

where the factor of 4 is introduced to match the normalization of the matrix model, and we used that

$$\frac{\mathrm{d}\lambda}{\mathrm{d}\kappa} = \frac{1}{4\pi^2} K\left(\frac{\mathrm{i}\kappa}{4}\right). \tag{2.5.40}$$

Integrating once, we find

$$\partial_\lambda^2 F_0(\lambda) = 4\pi^3 \frac{K'\left(\frac{\mathrm{i}\kappa}{4}\right)}{K\left(\frac{\mathrm{i}\kappa}{4}\right)} + a_1, \tag{2.5.41}$$

where a_1 is an integration constant and we have used the Legendre relation

$$E'K + EK' - KK' = \frac{\pi}{2}. \tag{2.5.42}$$

A further integration leads to the following expression in terms of a Meijer function

$$\partial_\lambda F_0(\lambda) = \frac{\kappa}{4} G^{2,3}_{3,3}\left(\begin{matrix} \frac{1}{2}, & \frac{1}{2}, & \frac{1}{2} \\ 0, & 0, & -\frac{1}{2} \end{matrix} \,\middle|\, -\frac{\kappa^2}{16}\right) + a_1 \lambda + a_2. \tag{2.5.43}$$

Comparison with the matrix model free energy at weak coupling (2.4.4) fixes $a_1 = 4\pi^3 \mathrm{i}$, $a_2 = 0$, so we can write

$$\partial_\lambda F_0(\lambda) = \frac{\kappa}{4} G^{2,3}_{3,3}\left(\begin{matrix} \frac{1}{2}, & \frac{1}{2}, & \frac{1}{2} \\ 0, & 0, & -\frac{1}{2} \end{matrix} \,\middle|\, -\frac{\kappa^2}{16}\right) + \frac{\pi^2 \mathrm{i} \kappa}{2}\, {}_3F_2\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}; 1, \frac{3}{2}; -\frac{\kappa^2.}{16}\right). \tag{2.5.44}$$

If we integrate this expression with the following choice of integration constant,

$$F_0(\lambda) = \int_0^\lambda \mathrm{d}\lambda' \, \partial_{\lambda'} F_0(\lambda') \tag{2.5.45}$$

we obtain the correct weak coupling expansion.

We can now analytically continue the r.h.s. of (2.5.44) to $\kappa = \infty$, and we obtain

$$\partial_\lambda F_0(\lambda) = 2\pi^2 \log \kappa + \frac{4\pi^2}{\kappa^2}\, {}_4F_3\left(1, 1, \frac{3}{2}, \frac{3}{2}; 2, 2, 2; -\frac{16}{\kappa^2}\right) \tag{2.5.46}$$

This agrees with (2.5.38) on the ABJM slice. To see this, one notices that

$$\begin{aligned}\omega^{(1)}(z,z) &= 2\sum_{n=1}^{\infty}\sum_{k+l=n}\frac{(2k+2l-1)!}{(k!)^2(l!)^2}z^n = 2\sum_{n=1}^{\infty}\frac{4^n(2n-1)!\Gamma\left(n+\frac{1}{2}\right)}{\sqrt{\pi}\Gamma(n+1)^3}z^n \\ &= 4z\,{}_4F_3\left(1,1,\frac{3}{2},\frac{3}{2};2,2,2;16z\right)\end{aligned} \tag{2.5.47}$$

is precisely the generalized hypergeometric function appearing in (2.5.46).

We are now ready to discuss the calculation of the planar free energy at strong coupling. We have,

$$\partial_{\hat{\lambda}}F_0(\lambda_1,\lambda_2) = 2\pi^2\log\kappa - \pi^2\omega^{(1)}(z,\bar{z}). \tag{2.5.48}$$

After plugging the value of κ in terms of $\hat{\lambda}$ given by the series expansion (2.5.22), and integrating w.r.t. $\hat{\lambda}$, we obtain

$$F_0(\hat{\lambda},B) = \frac{4\pi^3\sqrt{2}}{3}\hat{\lambda}^{3/2} + \frac{\zeta(3)}{2} + \sum_{\ell\geq 1}\mathrm{e}^{-2\pi\ell\sqrt{2\hat{\lambda}}}f_\ell\left(\frac{1}{\pi\sqrt{2\hat{\lambda}}},\beta\right) - \frac{2\pi^3\mathrm{i}}{3}\left(B-\frac{1}{2}\right)^3, \tag{2.5.49}$$

where $f_\ell(x)$ is a polynomial in x of the form

$$f_\ell(x,\beta) = \sum_{k=0}^{2\ell-3} f_k^{(\ell)}(\beta)x^k, \qquad \ell\geq 2. \tag{2.5.50}$$

The coefficients $f_k^{(\ell)}(\beta)$ are Laurent polynomials in β of degree ℓ, and symmetric under the exchange $\beta\leftrightarrow\beta^{-1}$. We have, for the very first cases,

$$\begin{aligned}f_1(x,\beta) &= -\frac{1}{2}\left(\beta+\beta^{-1}\right), \\ f_2(x,\beta) &= \frac{1}{16}\left(\beta^2+16+\beta^{-2}\right) + \frac{x}{4}\left(\beta+\beta^{-1}\right)^2.\end{aligned} \tag{2.5.51}$$

In going from (2.5.48) to (2.5.49) an integration constant $\zeta(3)/2$ appears. Its presence can be checked by comparing (2.5.49) with a numerical calculation of the integral (2.5.45) at intermediate coupling[3]. This constant is nothing but the well-known constant map contribution to the prepotential, first found in [151].

The free energy in the planar approximation is given by rescaling (2.5.49) by the string coupling $F = g_s^{-2}F_0 + \mathcal{O}(g_s^0)$. This expression displays many interesting features. First, note that on the ABJM slice $N_1 = N_2$ the leading term

$$-\frac{\pi\sqrt{2}}{3}k^2\hat{\lambda}^{3/2} \tag{2.5.52}$$

displays the "anomalous" scaling $N^{3/2}$ in the number of degrees of freedom for a theory of M2 branes, as was first derived from a supergravity calculation in [4]. The above calculation is a first principles derivation of this behaviour at strong coupling in the gauge theory. Usually, this behaviour is associated to the thermal free energy on $\mathbb{R}^3$, while (2.5.52) gives rather the free energy of the ABJM theory on $\mathbb{S}^3$ at strong coupling. However, as we have seen in Section 1.4.3 a supergravity calculation of this free energy also leads to the $N^{3/2}$ behavior with exactly the same coefficient. We will show this now, and in particular we will match the numerical coefficient in (2.5.52).

In Fig. 2.6 we show the exact result for the planar limit of $\partial_\lambda F_0(\lambda)$ in the case $N_1 = N_2$, as a function of $\lambda = N/k$, and we compare it to the behavior of the supergravity prediction

$$\partial_\lambda F_0(\lambda) \approx 2\pi^3\sqrt{2(\lambda-1/24)}, \qquad \lambda\to\infty. \tag{2.5.53}$$

We see that the strong coupling behavior gets triggered for values of the coupling $\lambda\approx 0.2$. For $\lambda\to 0$, the behavior of the prepotential is dominated by the Gaussian, weakly coupled result (2.4.4)

$$\partial_\lambda F_0(\lambda) \approx -8\pi^2\lambda\left(\log\left(\frac{\pi\lambda}{2}\right)-1\right), \qquad \lambda\to 0. \tag{2.5.54}$$

[3]This integration constant was incorrectly set to zero in the first version of the orginal paper [6]. It was determined numerically in [150].

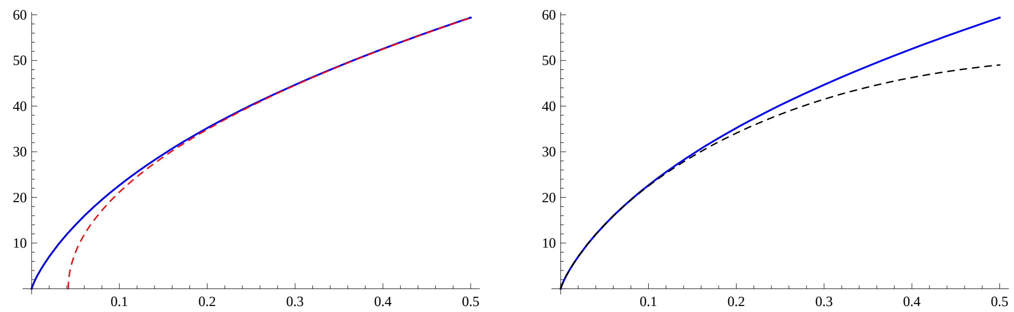

Figure 2.6: Comparison of the exact result for $\partial_\lambda F_0(\lambda)$ given in (2.5.44), plotted as a solid blue line, and the weakly coupled and strongly coupled results. In the figure on the left, the red dashed line is the supergravity result (2.5.53), while in the figure on the right, the black dashed line is the Gaussian result (2.5.54).

A second aspect to notice is that the supergravity result (2.5.49) has corrections which are exponentially suppressed. The exponential is of the form

$$\mathrm{e}^{-\ell A(\mathbb{CP}^1)} \tag{2.5.55}$$

where

$$A(\mathbb{CP}^1) = 2\pi\sqrt{2\hat{\lambda}} \tag{2.5.56}$$

is the area of the $\mathbb{CP}^1$ two-cycle in $\mathbb{CP}^3$. Also, notice that each of these exponential corrections multiplies (at each order in $\hat{\lambda}^{-1/2}$) the polynomial $f_k^{(\ell)}(\beta)$ in β, β^{-1}. Therefore, we have contributions schematically of the form

$$\sum_{n_+ + n_- = \ell} c_{n_+,n_-} \mathrm{e}^{-n_+(A(\mathbb{CP}^1)+2\pi \mathrm{i} B) - n_-(A(\mathbb{CP}^1)-2\pi\mathrm{i}B)} \tag{2.5.57}$$

This is precisely what one should expect for a gas of n_+ instantons and n_- anti–instantons in a σ model on $\mathbb{CP}^3$, where the (anti)instantons wrap the $\mathbb{CP}^1$ cycle. Notice that this kind of corrections are made possible by the non-trivial topology of two cycles in $\mathbb{CP}^3$, *i.e.*, by the fact that $b_2(\mathbb{CP}^3) = 1$, and as such they are absent in $\mathrm{AdS}_5 \times \mathbb{S}^5$. Some aspects of these string instantons have been studied in [206]. It would be interesting to test in detail the possible connection between these string instantons and the exponentially suppressed corrections to the planar free energy.

These instanton corrections are also present in the Wilson loop result (2.5.30), again with an infinite series of corrections. This can be compared with the case of $\mathcal{N} = 4$ SYM in 4d, where the asymptotic large coupling expansion of the Gaussian matrix model (2.2.44) has a single instanton correction which can be explicitly identified with a second saddle point solution in $\mathrm{AdS}_5 \times S^5$ [152, 154].

Finally, we note that when $N_1 \neq N_2$, the planar free energy (2.5.49) includes an imaginary term proportional to $(B - 1/2)^3$, which is derived by the weak coupling calculation (2.4.5). In CS theory such a term is related to framing [46]. It would be very interesting to derive this phase in type IIA string theory.

2.6 Conifold expansion

The expansion around the conifold locus corresponds to a region in the moduli space of the ABJM model where one of the gauge groups has finite coupling, while the other one is weakly coupled. In the lens space matrix model this corresponds to one 't Hooft parameter being small, and the other of order 1. In this section we will study this regime from three different points of view: the exact planar solution in terms of periods and Picard–Fuchs equations, the matrix model, and the gauge theory.

2.6.1 Expansion from the exact planar solution

We can use the exact planar solution to calculate various physical quantities near the conifold locus. For concreteness, we will expand around $t_2 = 0$ but with t_1 arbitrary. The first ingredient we need is an expansion of the global coordinates of moduli space. It turns out that the most convenient method is based on the expressions for the periods (2.2.21). The locus where $t_2 = 0$ is the line

$$\zeta = 2\beta - 2, \tag{2.6.1}$$

where the cut $(-b, -1/b)$ collapses to the point $Z = -1$. The derivative of t_2 w.r.t. ζ can then be computed in terms of residues at this point by expanding the expression in (2.2.23):

$$-\frac{\partial t_2}{\partial \zeta} = \sum_{k\geqslant 0} \frac{1}{4\pi \mathrm{i}} \oint_{-1} \mathrm{d}Z\, \frac{H_k(Z,\beta)\,(\zeta - 2\beta + 2)^k}{(Z+1)^{2k+1}}, \tag{2.6.2}$$

where $H_k(Z,\beta)$ are regular at $Z = -1$. This gives a series for t_2 in powers of $\zeta - 2\beta + 2$,

$$-t_2 = \frac{1}{4\sqrt{\beta}}(\zeta - 2\beta + 2) - \frac{1-\beta}{128\,\beta^{3/2}}(\zeta - 2\beta + 2)^2 + \frac{9 - 2\beta + 9\beta^2}{12288\,\beta^{5/2}}(\zeta - 2\beta + 2)^3 + \mathcal{O}((\zeta - 2\beta + 2)^4) \tag{2.6.3}$$

which can be easily inverted to

$$\zeta = 2\beta - 2 - 4\sqrt{\beta}\, t_2 + \frac{1}{2}(1-\beta)\, t_2^2 + \frac{3 + 10\beta + 3\beta^2}{48\sqrt{\beta}}\, t_2^3 + \mathcal{O}(t_2^4). \tag{2.6.4}$$

As a nice application of this expansion, we can compute the VEV of the 1/2 BPS Wilson loop around the conifold point, which is given in (2.2.51). Using the dictionary (2.2.30), (2.2.31), we find

$$\begin{aligned} \mathrm{e}^{-\pi \mathrm{i} B} \langle W_{\square}^{1/2} \rangle_{g=0} = \; & 2\sin(\pi\lambda_1) + 2\pi\lambda_2\,(2 - \cos(\pi\lambda_1)) + \pi^2\lambda_2^2 \sin(\pi\lambda_1) \\ & + \frac{1}{3}\pi^3\lambda_2^3\left(1 - 5\cos(\pi\lambda_1) + 3\cos^2(\pi\lambda_1)\right) + \mathcal{O}(\lambda_2^4). \end{aligned} \tag{2.6.5}$$

As $\lambda_2 \to 0$, we recover the result for a Wilson loop VEV in $U(N_1)$ CS theory. In the conifold expansion we are then regarding the ABJM theory as a perturbation of $U(N_1)$ CS theory at strong coupling.

The above result can be also obtained by solving the Picard–Fuchs equation around a point in the conifold locus. Let us choose for example the symmetric conifold point (2.3.32), with $B = 1$ and $\kappa = 4$. This corresponds to the point in the conifold locus with

$$\lambda_1 = \frac{1}{2}, \qquad \lambda_2 = 0. \tag{2.6.6}$$

The appropriate global coordinates near this point are (2.3.33). We find that λ_2 is a period solving the PF system (2.3.34) and with leading behavior

$$\lambda_2 = -\frac{1}{4\pi}\,(y_2 + y_1/2) + \mathcal{O}(y^2). \tag{2.6.7}$$

One finds the expansion

$$\begin{aligned} \lambda_2 = \; & \frac{\pi}{4}(B-1)^2 - \frac{5\pi^3}{96}(B-1)^4 + \left(\frac{1}{8\pi} - \frac{\pi}{32}(B-1)^2 + \frac{43\pi^3}{1536}(B-1)^4\right)(\kappa - 4) \\ & + \left(-\frac{1}{128\pi} + \frac{9\pi}{1024}(B-1)^2 - \frac{99\pi^3}{8192}(B-1)^4\right)(\kappa-4)^2 + \mathcal{O}\big((B-1)^6\big) + \mathcal{O}\big((\kappa-4)^3\big), \end{aligned} \tag{2.6.8}$$

which is inverted to

$$\begin{aligned} \kappa = \; & 4 - 2\pi^2\left(\lambda_1 - \frac{1}{2}\right)^2 + \frac{\pi^4}{6}\left(\lambda_1 - \frac{1}{2}\right)^4 + \pi\lambda_2\left(8 + 4\pi\left(\lambda_1 - \frac{1}{2}\right) - \frac{2\pi^3}{3}\left(\lambda_1 - \frac{1}{2}\right)^3\right) \\ & + \mathcal{O}\big(\lambda_2^2\big) + \mathcal{O}\big((\lambda_1 - 1/2)^5\big). \end{aligned} \tag{2.6.9}$$

This is indeed the expansion around $\lambda_1 = 1/2$ of (twice) the series in the r.h.s. of (2.6.5).

Once we know the expansion of the global coordinates, we can consider other quantities in the model, like the genus g free energies. The conifold expansion of $F_g(t_1, t_2)$ has the form

$$F_g(\lambda_1, \lambda_2) = F_g^{\mathrm{G}}(\lambda_2) + \sum_{n\geq 0} F_g^{(n)}(\lambda_1)\lambda_2^n, \tag{2.6.10}$$

where $F_g^{\rm G}(\lambda_2)$ is the free energy of the $U(N_2)$ Gaussian matrix model, and each coefficient $F_g^{(n)}(\lambda_1)$ can be obtained as an exact function of λ_1. Of course,

$$F_g^{(0)}(\lambda_1) = F_g^{\mathbb{S}^3}(\lambda_1) \tag{2.6.11}$$

is the genus g free energy of the CS theory on $\mathbb{S}^3$. When $g=0$, the expansion (2.6.10) can be computed from the exact planar solution in various ways. One can for example use the Yukawa couplings (2.3.37) expanded around the conifold locus in order to compute the third derivatives of F_0, or use the modularity properties of the solution discussed in [109, 113]. In any case, for the first few functions one finds the following results:

$$\begin{aligned} F_0^{(1)}(\lambda_1) &= 2\pi {\rm i}\left(\pi^2\lambda_1^2 + 2{\rm Li}_2\big(-e^{\pi {\rm i}\lambda_1}\big) - 2{\rm Li}_2\big(-e^{-\pi{\rm i}\lambda_1}\big)\right), \\ F_0^{(2)}(\lambda_1) &= -2\pi^3{\rm i}\lambda_1 + 8\pi^2\log\left(\cos\left(\frac{\pi\lambda_1}{2}\right)\right), \\ F_0^{(3)}(\lambda_1) &= \frac{2\pi^3{\rm i}}{3} + \frac{\pi^3}{3}\big(3\cos(\pi\lambda_1)-5\big)\tan\left(\frac{\pi\lambda_1}{2}\right). \end{aligned} \tag{2.6.12}$$

2.6.2 Conifold expansion from the matrix model

It is easy to implement the conifold expansion directly in the lens space matrix model. To do that, we notice that it can be written as two interacting Chern–Simons matrix models on $\mathbb{S}^3$. We recall that the CS matrix model on $\mathbb{S}^3$, first considered in [85], is defined by the partition function

$$Z_{\mathbb{S}^3}(N,g_s) = \frac{1}{N!}\int\prod_{i=1}^N\frac{{\rm d}\mu_i}{2\pi}\prod_{i<j}\left(2\sinh\left(\frac{\mu_i-\mu_j}{2}\right)\right)^2 {\rm e}^{-\frac{1}{2g_s}\sum_i\mu_i^2}. \tag{2.6.13}$$

This is a one-cut matrix model [90]. It can be obtained from the lens space matrix model when one of the two cuts collapses to zero size. In the Z plane the endpoints of the cut are given by a and a^{-1}, where

$$a = 2{\rm e}^t - 1 - 2{\rm e}^{t/2}\sqrt{{\rm e}^t-1}. \tag{2.6.14}$$

Let us consider the following operator in this model:

$$\mathcal{W}(\nu_j) = 2\sum_{i,j}\log\left(2\cosh\left(\frac{\mu_i-\nu_j}{2}\right)\right). \tag{2.6.15}$$

The lens space partition function (2.2.4) can be calculated in two steps. In the first step, we compute

$$Z_1(\nu_j) = \left\langle {\rm e}^{\mathcal{W}(\nu_j)}\right\rangle_{N_1} \tag{2.6.16}$$

where the subindex N_1 indicates that this is an unnormalized VEV in the $\mathbb{S}^3$ CS matrix model with gauge group $U(N_1)$. In a second step, we calculate

$$Z_{L(2,1)} = \langle Z_1(\nu_j)\rangle_{N_2} \tag{2.6.17}$$

in the CS matrix model with gauge group $U(N_2)$. To obtain the conifold expansion, we calculate $Z_1(\nu_j)$ and we expand it in g_s and around $\nu_j=0$. Each term in this expansion can be computed exactly as a function of the Kähler parameter t_1, since the CS matrix model can be solved exactly in the $1/N$ expansion. The resulting double series in g_s and ν_j is then regarded as an operator in the CS matrix model with group $U(N_2)$, which we expand around the Gaussian point as in [85, 88], *i.e.*, we expand the sinh measure around $\nu_j=0$. The partition function $Z_{L(2,1)}$ is then computed as a VEV in the Gaussian matrix model. This procedure gives a method to compute the expansion (2.6.10) directly in the matrix model.

To illustrate this procedure, let us calculate $F_0(t_1,t_2)$ at first order in t_2. In this computation we will denote

$$U_1 = {\rm diag}({\rm e}^{\mu_i}), \qquad U_2 = {\rm diag}({\rm e}^{\nu_j}). \tag{2.6.18}$$

The expansion around $\nu_j=0$ of the operator $\mathcal{W}(\nu_j)$ reads

$$\mathcal{W}(\nu_j) = 2N_2\sum_{i=1}^{N_1}\log\left[2\cosh\left(\frac{\mu_i}{2}\right)\right] - \sum_{j=1}^{N_2}\nu_j\sum_{i=1}^{N_1}\tanh\left(\frac{\mu_i}{2}\right) + \mathcal{O}(\nu_j^2). \tag{2.6.19}$$

The average of the second term in the $U(N_2)$ matrix model vanishes (since it is odd in ν_j), while higher order terms are at least of order t_2^2. The first term can be written as

$$2\sum_{i=1}^{N_1}\log\left[2\cosh\left(\frac{\mu_i}{2}\right)\right]=2\,\mathrm{Tr}\,\log(1+U_1)-\sum_{i=1}^{N_1}\mu_i. \tag{2.6.20}$$

Therefore, in the planar limit and neglecting terms which contribute at order t_2^2, we have

$$\log Z_1(\nu_j)\approx\frac{2t_2}{g_s}\left\langle \mathrm{Tr}\,\log(1+U_1)\right\rangle_{N_1} \tag{2.6.21}$$

since the second term in (2.6.20) is odd in μ_i and its VEV vanishes. We then find,

$$F_0(t_1,t_2)=F_0^{\mathbb{S}^3}(t_1)+2t_2g_s\left\langle \mathrm{Tr}\,\log(1+U_1)\right\rangle+\mathcal{O}(t_2^2). \tag{2.6.22}$$

The VEV in (2.6.22), which is now normalized, can be computed in terms of the resolvent of the CS matrix model, and similar computations appear in [30, 155] in the context of large N instanton corrections. In fact, it follows from (2.7.28) and (2.7.30) that the VEV in (2.6.22) is given by $-g(-1)$, where $g(Y)$ is computed in (D.2). The final result for the linear correction in t_2 is

$$\frac{\pi^2}{3}+\frac{t_1^2}{2}+\mathrm{Li}_2(\mathrm{e}^{-t_1})-2\mathrm{Li}_2(\mathrm{e}^{-t_1/2})+2\mathrm{Li}_2(-\mathrm{e}^{-t_1/2}). \tag{2.6.23}$$

Using dilogarithm identities, this agrees with $\frac{\lambda_2}{t_2}F_0^{(1)}(\lambda_1)$ in (2.6.12). It is interesting to point out that, in the context of CS theory on the lens space $L(2,1)$, this function is essentially the action of the large N instanton corresponding to the flat connection

$$U(N)\to U(N_1)\times U(N_2),\qquad N_2\ll N_1, \tag{2.6.24}$$

as shown in [30]. In the matrix model, this action is obtained by tunneling N_2 eigenvalues from the first cut to the second cut.

We can also calculate the conifold expansion for the VEV of 1/6 and 1/2 BPS Wilson loops directly in the matrix model. We want to compute

$$\langle W^{1/6}_{\square}\rangle=g_s\langle \mathrm{Tr}\,U_1\rangle_{L(2,1)}. \tag{2.6.25}$$

We will again perform this computation in the planar approximation and at linear order in t_2. At this order we can compute instead the normalized average of the operator

$$\frac{\left\langle \mathrm{Tr}\,U_1\,\mathrm{e}^{\mathcal{W}(\nu_j)}\right\rangle_{N_1}}{\left\langle \mathrm{e}^{\mathcal{W}(\nu_j)}\right\rangle_{N_1}}=\langle \mathrm{Tr}\,U_1\rangle+\langle \mathrm{Tr}\,U_1\,\mathcal{W}(\nu_j)\rangle^{(c)}+\cdots \tag{2.6.26}$$

in a Gaussian matrix model for the ν_j. In the last line, all VEVs are normalized VEVs in the $\mathbb{S}^3$ CS matrix model. By completing the square of the Gaussian weight we derive

$$\left\langle \mathrm{Tr}\,U_1\left(\sum_{i=1}^{N_1}\mu_i\right)\right\rangle=\frac{\partial}{\partial j}\left\langle \mathrm{Tr}\,U_1\,\mathrm{e}^{j\sum_{i=1}^{N_1}\mu_i}\right\rangle\Bigg|_{j=0}=g_s\left\langle \mathrm{Tr}\,U_1\right\rangle. \tag{2.6.27}$$

We then find, at this order,

$$\langle W^{1/6}_{\square}\rangle_{g=0}=g_s\langle \mathrm{Tr}\,U_1\rangle+t_2\Big(2\langle \mathrm{Tr}\,U_1\,\mathrm{Tr}\log(1+U_1)\rangle^{(c)}-g_s\langle \mathrm{Tr}\,U_1\rangle\Big)+\mathcal{O}(t_2^2). \tag{2.6.28}$$

The connected correlator

$$\langle \mathrm{Tr}\,U_1\,\mathrm{Tr}\log(1+U_1)\rangle^{(c)}=-\sum_{\ell=1}^{\infty}\frac{(-1)^{\ell}}{\ell}\langle \mathrm{Tr}\,U_1\,\mathrm{Tr}\,U_1^{\ell}\rangle^{(c)} \tag{2.6.29}$$

can be computed by considering the (partially) integrated two-point function (see for example [170])

$$\int \mathrm{d}p\,W_0(p,q)=-\sum_{n,m}\frac{1}{np^nq^{m+1}}\langle \mathrm{Tr}\,U_1^n\,\mathrm{Tr}\,U_1^m\rangle^{(c)} \tag{2.6.30}$$

and extracting the coefficient of q^{-2}. We have,

$$\int \mathrm{d}p\, W_0(p,q) = \frac{1}{2(p-q)}\left(1-\sqrt{\frac{(p-a)(p-a^{-1})}{(q-a)(q-a^{-1})}}\right)+\frac{1}{2\sqrt{(q-a)(q-a^{-1})}}, \tag{2.6.31}$$

which includes the appropriate integration constant. We find, after changing $p \to -p$,

$$-\sum_{\ell=1}^{\infty}\frac{(-1)^\ell}{\ell p^\ell}\langle \mathrm{Tr}\, U_1\, \mathrm{Tr}\, U_1^\ell\rangle^{(c)} = \frac{1}{4}\left(a+a^{-1}+2p-2\sqrt{(p+a)(p+a^{-1})}\right). \tag{2.6.32}$$

When $p=1$ this gives

$$-\sum_{\ell=1}^{\infty}\frac{(-1)^\ell}{\ell}\langle \mathrm{Tr}\, U_1\, \mathrm{Tr}\, U_1^\ell\rangle^{(c)} = \mathrm{e}^{t_1}-\mathrm{e}^{t_1/2}. \tag{2.6.33}$$

Notice that this is an infinite sum of correlators in the CS matrix model. Since

$$\langle \mathrm{Tr}\, U_1\rangle = \frac{\mathrm{e}^{t_1}-1}{g_s}, \tag{2.6.34}$$

we finally obtain,

$$\begin{aligned}\langle W_{\Box}^{1/6}\rangle_{g=0} &= \mathrm{e}^{t_1}-1+t_2\left(\mathrm{e}^{t_1/2}-1\right)^2+\mathcal{O}(t_2^2)\\ &= \mathrm{e}^{t_1/2}\left(2\sinh\frac{t_1}{2}+t_2\left(-2+2\cosh\frac{t_1}{2}\right)+\mathcal{O}(t_2^2)\right).\end{aligned} \tag{2.6.35}$$

Since this is a Wilson loop only in the first group, the framing prefactor depends only on the first 't Hooft coupling.

The 1/2 BPS Wilson loop is obtained by subtracting

$$\langle \mathrm{Tr}\, U_2\rangle_{L(2,1)} = N_2+\mathcal{O}(t_2^2) = \frac{t_2}{g_s}+\mathcal{O}(t_2^2). \tag{2.6.36}$$

We find,

$$\mathrm{e}^{-(t_1+t_2)/2}\langle W_{\Box}^{1/2}\rangle_{g=0} = 2\sinh\left(\frac{t_1}{2}\right)+t_2\left(-2+\cosh\left(\frac{t_1}{2}\right)\right)+\mathcal{O}(t_2^2). \tag{2.6.37}$$

This is the result (2.6.5) obtained from the conifold expansion after using the dictionary (2.2.29).

2.7 More exact results on Wilson loops

In this section we present more exact results on Wilson loops.

2.7.1 $1/N$ corrections

The higher genus corrections to the VEV of 1/2 and 1/6 BPS Wilson loops can be computed in terms of the higher genus corrections to the resolvent of the matrix model. The resolvent has a genus expansion of the form

$$\omega(z) = \sum_{g=0}^{\infty} g_s^{2g}\omega_g(z). \tag{2.7.1}$$

In the same way, the density of eigenvalues has a large N expansion of the form

$$\rho(z) = \sum_{g=0}^{\infty} g_s^{2g}\rho_g(z), \qquad \rho(z) = \rho^{(1)}(z)+\rho^{(2)}(z). \tag{2.7.2}$$

The $\rho_g^{(i)}(z)$ (with $i=1,2$) have their support on the intervals $\mathcal{C}_i$, and they can be obtained by the discontinuity of ω_g at the cuts as in (2.2.40).

The genus expansion of the expectation value of the 1/6 BPS and 1/2 BPS Wilson loops follows the expressions in (2.2.42) and (2.2.43) with the appropriate term in the expansion of $\rho^{(i)}(Z)$ and $\omega(Z)$.

The first step is therefore to compute $\omega_g(p)$. This calculation can be done with the recursive techniques developed in the matrix model literature starting with [170] and culminating with [175]. We will perform an explicit computation for $g = 1$. Calculations for $g \geq 2$ are in principle doable, but they become complicated.

A convenient formula for $\omega_1(p)$ for an algebraic resolvent was found in [173]. To write this formula, we write the discontinuity of the resolvent (also called *spectral curve* in the matrix model literature) as

$$y(p) = M(p)\sqrt{\sigma(p)}, \qquad \sigma(p) = (p - x_1)(p - x_2)(p - x_3)(p - x_4). \tag{2.7.3}$$

$M(p)$ is sometimes called the moment function. Then, one has

$$\omega_1(p) = \frac{4}{\sqrt{\sigma(p)}} \sum_{i=1}^{4} \left(\frac{A_i}{(p - x_i)^2} + \frac{B_i}{p - x_i} + C_i \right), \tag{2.7.4}$$

where

$$\begin{aligned} A_i &= \frac{1}{16} \frac{1}{M(x_i)}, \\ B_i &= -\frac{1}{16} \frac{M'(x_i)}{M^2(x_i)} + \frac{1}{8\, M(x_i)} \left(2\alpha_i - \sum_{j \neq i} \frac{1}{x_i - x_j} \right), \\ C_i &= -\frac{1}{48} \frac{1}{M(x_i)} \sum_{j \neq i} \frac{\alpha_j - \alpha_i}{x_j - x_i} - \frac{1}{16} \frac{M'(x_i)}{M^2(x_i)} \alpha_i + \frac{\alpha_i}{8\, M(x_i)} \left(2\alpha_i - \sum_{j \neq i} \frac{1}{x_i - x_j} \right), \end{aligned} \tag{2.7.5}$$

and the α_i are given by

$$\begin{aligned} \alpha_1 &= \frac{1}{(x_1 - x_2)} \left[1 - \frac{(x_4 - x_2)}{(x_4 - x_1)} \frac{E(k)}{K(k)} \right], \\ \alpha_2 &= \frac{1}{(x_2 - x_1)} \left[1 - \frac{(x_3 - x_1)}{(x_3 - x_2)} \frac{E(k)}{K(k)} \right], \\ \alpha_3 &= \frac{1}{(x_3 - x_4)} \left[1 - \frac{(x_4 - x_2)}{(x_3 - x_2)} \frac{E(k)}{K(k)} \right], \\ \alpha_4 &= \frac{1}{(x_4 - x_3)} \left[1 - \frac{(x_3 - x_1)}{(x_4 - x_1)} \frac{E(k)}{K(k)} \right], \end{aligned} \tag{2.7.6}$$

where the modulus of the elliptic functions is

$$k^2 = \frac{(x_1 - x_2)(x_3 - x_4)}{(x_1 - x_3)(x_2 - x_4)}. \tag{2.7.7}$$

These expressions differ from the ones in [173] in a permutation of the roots, as explained in [158]. The overall factor of 4 in (2.7.4) is due to the fact that our resolvent has a different normalization than the one in [173].

Although the resolvent of the lens space matrix model (2.2.17) is not algebraic, its discontinuity can be written in the form (2.7.3) with

$$\sigma(p) = f(p)^2 - 4\beta^2 p^2, \qquad f(p) = p^2 - \zeta p + 1 \tag{2.7.8}$$

and

$$M(p) = \frac{2}{p\sqrt{\sigma(p)}} \tanh^{-1} \frac{\sqrt{\sigma(p)}}{f(p)}. \tag{2.7.9}$$

This form of the spectral curve is typical of the mirrors of toric geometries [172, 201]. The branch points are

$$x_1 = -b, \quad x_2 = -\frac{1}{b}, \quad x_3 = \frac{1}{a}, \quad x_4 = a. \tag{2.7.10}$$

Using these expressions, it is possible to compute the integral

$$\langle W_{\square}^{1/6} \rangle_{g=1} = \frac{1}{4\pi \mathrm{i}} \oint_{\mathcal{C}_1} \omega_1(Z) Z \mathrm{d}Z \tag{2.7.11}$$

in closed form, in terms of elliptic functions E, K and the elliptic integral of the third kind $\Pi(n,k)$, with

$$n = \frac{(a^2-1)b}{(1+ab)a} \quad . \tag{2.7.12}$$

One finds the rather complicated expression

$$\begin{aligned}\langle W^{1/6}_{\square}\rangle_{g=1} &= \frac{1}{12\pi\sqrt{a}\, b^{3/2}(1+ab)(a^2-1)^2(b^2-1)K}\Big[-3(b-2a+a^2b)\,(1+ab)^4E^2 + \Big[a(1+a^4) \\ &- b + a^2(4+4a^2-a^4)b - 4a(1-3a^2+a^4)b^2 - a^2(1+a^2)b^3(2+b^2) + a(1-8a^2+a^4)b^4\Big]K^2 \\ &+ \big(b^3(1+6a^2+a^4) + 4a(1+a^2)(b^2-1) + b(3-14a^2+3a^4)\big)\,(1+ab)^2E\,K\Big] \\ &+ \frac{(ab-1)(a^2-b^2)}{12\pi\,(ab)^{3/2}(1+ab)k^4K^2}\Big[-6E^2 + 4(2-k^2)E\,K - (2-2k^2+k^4)\,K^2\Big]\Pi.\end{aligned} \tag{2.7.13}$$

To check this formula, we expand it around the weakly coupled point $\lambda_1 = \lambda_2 = 0$. After using the inverse mirror map given by (2.4.1) we find

$$\begin{aligned}\langle W^{1/6}_{\square}\rangle_{g=1} = &-\frac{\pi \mathrm{i}}{12}\lambda_1 + \frac{\pi^2}{12}\lambda_1^2 + \frac{\pi^2}{4}\lambda_1\lambda_2 + \frac{\pi^3\mathrm{i}}{18}\lambda_1^3 + \frac{\pi^3\mathrm{i}}{24}\lambda_1^2\lambda_2 - \frac{\pi^3\mathrm{i}}{4}\lambda_1\lambda_2^2 \\ &- \frac{\pi^4}{36}\lambda_1^4 + \frac{\pi^4}{24}\lambda_1^3\lambda_2 + \frac{5\pi^4}{24}\lambda_1^2\lambda_2^2 - \frac{\pi^4}{6}\lambda_1\lambda_2^3 + \mathcal{O}(\lambda^5).\end{aligned} \tag{2.7.14}$$

We can test this expansion with a perturbative calculation in the ABJM matrix model. At order $\mathcal{O}(g_s^4)$ we have found,

$$\begin{aligned}\frac{\mathrm{e}^{-g_sN_1/2}}{2\pi \mathrm{i}\lambda_1}\langle W^{1/6}_{\square}\rangle = 1 &- \left(\frac{1}{24}N_1^2 - \frac{1}{4}N_1N_2 + \frac{1}{24}\right)g_s^2 + \left(\frac{1}{16}N_1^2N_2 - \frac{1}{16}N_2\right)g_s^3 \\ &+ \left(\frac{3}{5760}N_1^4 - \frac{10}{1920}N_1^3N_2 - \frac{20}{1920}N_1N_2^3 - \frac{10}{5760}N_1^2 + \frac{5}{192}N_1N_2 + \frac{1}{32}N_2^2 + \frac{7}{5760}\right)g_s^4 + \cdots\end{aligned} \tag{2.7.15}$$

It is straightforward to see that this agrees with (2.7.14).

The $1/N$ correction to the 1/2 BPS Wilson loop is much easier to obtain, since it can be computed as a residue at infinity. We have that

$$\omega_1(Z) = \frac{4}{Z^2}\sum_{i=1}^{4} C_i + \mathcal{O}(Z^{-3}), \tag{2.7.16}$$

where the C_i are given in (2.7.5). We find, at weak coupling,

$$\begin{aligned}\langle W^{1/2}_{\square}\rangle_{g=1} = &-\frac{\pi \mathrm{i}}{12}(\lambda_1+\lambda_2) + \frac{\pi^2}{12}(\lambda_1^2-\lambda_2^2) + \frac{\pi^3\mathrm{i}}{18}(\lambda_1^3+\lambda_2^3) - \frac{5\pi^3\mathrm{i}}{24}\lambda_1\lambda_2(\lambda_1-\lambda_2) \\ &- \frac{\pi^4}{36}(\lambda_1^4-\lambda_2^4) + \frac{5\pi^4}{24}\lambda_1\lambda_2(\lambda_1^2-\lambda_2^2) + \mathcal{O}(\lambda^5).\end{aligned} \tag{2.7.17}$$

At strong coupling we find (we consider for simplicity the ABJM slice)

$$\langle W^{1/2}_{\square}\rangle_{g=1} = \frac{1}{24\mathrm{i}}\,\frac{3+2\log^2\kappa - 4\log\kappa}{\log^2\kappa}\kappa + \mathcal{O}(1) \tag{2.7.18}$$

The leading exponent is exactly as at genus zero (2.5.30), representing the same minimal surface with an extra degenerate handle attached. Its effect is to modify the one-loop determinant, which (with our normalization and ignoring instantons) can be written as

$$\langle W^{1/2}_{\square}\rangle_{g=1} = -\mathrm{i}\left(\frac{1}{12} - \frac{1}{6\pi\sqrt{2\lambda}} + \frac{1}{16\pi^2\lambda}\right)\mathrm{e}^{\pi\sqrt{2\lambda}}, \qquad \lambda\to\infty. \tag{2.7.19}$$

2.7.2 Giant Wilson loops

It has been argued in [145, 146, 159, 203] that a D-brane probe in $AdS_5 \times \mathbb{S}^5$ represents an insertion of a Wilson loop in the dual 4d $\mathcal{N}=4$ SYM with a large symmetric or antisymmetric representation (in the case of D3 branes and D5 branes, respectively). These "giant Wilson loops" are characterized by a representation with n boxes, and one considers the limit

$$n,\, N \to \infty, \qquad \frac{n}{N} \text{ fixed.} \tag{2.7.20}$$

In terms of the Gaussian matrix model of the Wilson loops in that theory, the giant Wilson loop in the symmetric representation is represented by an additional eigenvalue outside the cut and the antisymmetric representation by a "hole" in the original cut.

Let us review now the known D-brane solutions which could be relevant for ABJM theory. The usual 1/2 BPS Wilson loop in the fundamental representation is described by a string with world-volume $AdS_2 \subset AdS_4$. In M-theory it is an M2-brane wrapping also the orbifold cycle on $\mathbb{S}^7/\mathbb{Z}_k$. When considering $k/2$ coincident M2-branes (or k, when it is odd) the M2-brane solution develops an extra branch, where the circle becomes a linear combination of the orbifold direction and a contractible circle in AdS_4 [160]. In type IIA these configurations are D2-branes with world-volume $AdS_2 \times \mathbb{S}^1 \subset AdS_4$, where the radius of the $\mathbb{S}^1$ is a free modulus. From the M-theory point of view these are continuous deformations of the system of $k/2$ coincident M2-branes describing a Wilson loop in a $k/2$ dimensional representation. In the field theory they are the vortex loop operators of [114], which have a description as semi-classical field configurations and carry the same charge as $k/2$ Wilson loops.

These solutions have further moduli associated to rotations away from the orbifold cycle inside $\mathbb{S}^7/\mathbb{Z}_k$. Such M2-brane configurations preserve 8 supercharges (1/3 BPS) [103, 114].

There is also a known family of D6-brane solutions which were argued in [103] to represent the 1/6 BPS Wilson loops in anti-symmetric representations. The action for this D-branes is (for $N_1 = N_2$)

$$S_{\rm D6} = -\pi\sqrt{2\lambda}\,\frac{n(N-n)}{N}\,, \tag{2.7.21}$$

which matches that of n strings for small n and has the $n \to N-n$ symmetry of the antisymmetric representation. In the matrix model these D6-branes should correspond to creating a "hole" in one of the two cuts, splitting it in two.

We turn now to the lens space matrix model and try to find the appropriate description for these objects, and in particular the 1/2 BPS vortex loop operators. As pointed out in [161], the calculation of Wilson loops in the matrix model in this limit can be done in a saddle-point approximation. We will now reformulate the arguments of [161] and adapt them to the lens space matrix model.

We will focus on the case of 1/2 BPS Wilson loops, where we want to calculate

$$W_n^\eta = \langle \mathrm{Tr}_{\mathcal{R}_n^\eta} U\rangle, \qquad \eta = \pm 1, \tag{2.7.22}$$

where U is the same matrix as in (2.2.16) and $\mathcal{R}_n^{\pm 1} = S_n,\, A_n$ are respectively the totally symmetric and the totally antisymmetric representations of $U(N_1+N_2)$ with n boxes. It will turn out that the relevant limit in this theory is slightly different from (2.7.20) and is given by fixing

$$\nu = \eta\,\frac{n}{k} = \frac{\eta g_s\, n}{2\pi \mathrm{i}}. \tag{2.7.23}$$

Positive ν will correspond to symmetric representations and negative ν to antisymmetric ones. In the 't Hooft limit, for fixed N/k, the two scalings are clearly equivalent.

The calculation of (2.7.22) is very similar to the calculation of partition functions of n bosons or fermions in the *canonical* ensemble, where n is fixed and large. But at large n, in the thermodynamic limit, this calculation can be done as well in the *grand canonical* ensemble. We then introduce the fugacity z and consider the grand-canonical partition function, using the expression for the determinant as the generating function of the characters

$$\Xi_\eta(z) = \sum_{n\geq 0} z^n W_n^\eta = \left\langle \det\left(1-\eta z\, U\right)^{-\eta}\right\rangle = \left\langle \exp\left(\sum_{\ell\geq 1}\frac{\mathrm{Tr}\, U^\ell}{\ell}\eta^{\ell-1} z^\ell\right)\right\rangle. \tag{2.7.24}$$

The average value of n in this ensemble is given by (we remove the average notation here, as is standard in the grand canonical formalism)

$$n = z\frac{\partial}{\partial z}\log \Xi_\eta. \tag{2.7.25}$$

This is inverted to determine the fugacity as a function of the number of particles

$$z_* = z_*(n), \tag{2.7.26}$$

and then the original VEV can be calculated, in a saddle point approximation, as

$$W_n^\eta \approx z_*^{-n}\Xi_\eta(z_*) = \left\langle \exp\left(-n\log z_* + \sum_{\ell\geq 1} \frac{\operatorname{Tr} U^\ell}{\ell}\eta^{\ell-1} z_*^\ell \right)\right\rangle. \tag{2.7.27}$$

For convenience, let us henceforth absorb $Y = \eta z$. It can be seen that, at leading order in large N, the grand-canonical partition function (2.7.24) is given by disconnected planar graphs. Therefore

$$\Xi_\eta(Y) \approx \exp\left(\frac{\eta}{g_s}\, g(Y)\right), \qquad g(Y) = g_s \sum_{\ell\geq 1} \frac{\langle \operatorname{Tr} U^\ell\rangle_0}{\ell}\, Y^\ell, \tag{2.7.28}$$

where the subscript 0 refers to the planar part. We now observe that the function $g(Y)$ is related to the planar resolvent in the lens space matrix model (2.2.15) and (2.2.17) by

$$\begin{aligned} Y\frac{\partial}{\partial Y} g(Y) =& \frac{1}{2}\left(\omega_0(Y^{-1}) - t\right) \\ =& -\log\left(\frac{1}{2}\left[\sqrt{(Y+b)(Y+1/b)} + \sqrt{(Y-a)(Y-1/a)}\right]\right). \end{aligned} \tag{2.7.29}$$

Note that compared to ω_0 in (2.2.17), the sign between the two square roots is reversed. Integrating this equation we get

$$g(Y) = -\int_0^Y \frac{\mathrm{d}Y'}{Y'} \log\left(\frac{1}{2}\left[\sqrt{(Y'+b)(Y'+1/b)} + \sqrt{(Y'-a)(Y'-1/a)}\right]\right). \tag{2.7.30}$$

The initial point of integration is chosen to be $Y = 0$, since around that point the integrand approaches a constant $\zeta/2 + \mathcal{O}(Y)$. This guaranties that for small Y the result of the integration will be proportional to the 1/2 BPS Wilson loop (2.2.51).

The saddle point equation (2.7.25) determining the mean value of n is then given by

$$\nu = \frac{1}{2\pi\mathrm{i}}\, Y \frac{\partial}{\partial Y} g(Y). \tag{2.7.31}$$

i.e., (2.7.29)

$$\mathrm{e}^{-2\pi\mathrm{i}\nu} = \frac{1}{2}\left[\sqrt{(Y_*+b)(Y_*+1/b)} + \sqrt{(Y_*-a)(Y_*-1/a)}\right], \qquad Y_* = \eta z_*. \tag{2.7.32}$$

This can be solved explicitly in terms of β, ζ or alternatively in terms of B and κ. The solution reads

$$Y_* = \frac{\mathrm{i}\kappa\, \mathrm{e}^{-\pi\mathrm{i}(2\nu+B)}}{4\sin(2\pi(\nu+B))}\left(1 - \sqrt{1 - \frac{16\sin(2\pi\nu)\sin(2\pi(\nu+B))}{\kappa^2}}\right). \tag{2.7.33}$$

The choice of sign is such that $Y_* = 0$ when $\nu = 0$. We will write

$$W_n^\eta \approx \exp\left(A_\eta/g_s\right) \tag{2.7.34}$$

where A_η, which is identified with the action of a brane probe in the large N string/M-theory dual, is given by

$$A_\eta = -2\pi\mathrm{i}\eta\nu\log(\eta Y_*) + \eta g(Y_*). \tag{2.7.35}$$

In the original variables, in terms of ω_0, the integral (2.7.30) is from infinity to a finite position Y_*^{-1}, and represents the effect of adding a single eigenvalue to the system. This fits with the standard dictionary [36] identifying a brane with a single eigenvalue.

This integral gives an expression for the action of the giant Wilson loop, in the limit (2.7.20) which is exact as a function of the 't Hooft couplings. The derivatives of this integral with respect to β and ζ can be evaluated in closed form, as in (2.2.23), in terms of *incomplete* elliptic integrals. The resulting expression can then be studied at the different limits of the ABJM theory as done for other observables in earlier sections.

If we go to the conifold limit, setting $\lambda_2 = 0$, we get an expression for the giant Wilson loop in Chern–Simons theory on $\mathbb{S}^3$. In that case there exists an exact expression for the Wilson loop for all n. As we show in Appendix D, the above derivation in this limit indeed reproduces the CS answer.

We will now discuss the expansion of the result for the giant Wilson loop for large κ, since this is the strong coupling limit in which one makes contact with the AdS geometry [203]. In terms of B and κ, the integral (2.7.30) reads

$$g(Y_*) = -\int_0^{Y_*} \frac{\mathrm{d}Y'}{Y'} \log\left(\frac{1}{2}\left[\sqrt{(1+Y')^2 - \mathrm{e}^{\pi\mathrm{i}B}Y'(\kappa - 4\mathrm{i}\sin(\pi B))}\right.\right. \\ \left.\left. + \sqrt{(1-Y')^2 - \mathrm{e}^{\pi\mathrm{i}B}Y'(\kappa + 4\mathrm{i}\sin(\pi B))}\right]\right) \tag{2.7.36}$$

where Y_* is given in (2.7.33).

Expanding Y_* at leading order at large κ we get

$$Y_* = 2\mathrm{i}\,\mathrm{e}^{-\pi\mathrm{i}(2\nu+B)}\frac{\sin(2\pi\nu)}{\kappa} + \mathcal{O}(\kappa^{-2}) = \frac{1-\mathrm{e}^{-4\pi\mathrm{i}\nu}}{\kappa}\,\mathrm{e}^{-\pi\mathrm{i}B} + \mathcal{O}(\kappa^{-2}) \tag{2.7.37}$$

This suggests rescaling Y in the integral (2.7.36) by κ, which allows for a systematic expansion in powers of κ^{-1}. At leading order the integral becomes

$$g(Y_*) = -\int_0^{Y_*} \frac{\mathrm{d}Y'}{Y'}\left(\log\sqrt{1-\mathrm{e}^{\pi\mathrm{i}B}\kappa Y'} + \mathcal{O}(\kappa^{-1})\right). \tag{2.7.38}$$

This yields

$$g(Y_*) = \frac{1}{2}\mathrm{Li}_2\left(\mathrm{e}^{\pi\mathrm{i}B}\kappa Y_*\right) + \mathcal{O}(\kappa^{-2}) = \frac{1}{2}\mathrm{Li}_2\left(1-\mathrm{e}^{-4\pi\mathrm{i}\nu}\right) + \mathcal{O}(\kappa^{-2}) \tag{2.7.39}$$

Another way to get this estimate is to notice that the highest powers of ζ in the series expansion in y of $g(y)$ are captured by

$$g(y) = \frac{1}{2}\mathrm{Li}_2(\zeta y) + \cdots. \tag{2.7.40}$$

Using the dilogarithem identity (D.3) we conclude that the action (2.7.35), written in terms of the original variable n, is

$$\frac{1}{g_s}A_\eta = n\pi\sqrt{2\hat\lambda} + \frac{n\pi\mathrm{i}}{2}(2B-1+\eta) + \frac{\eta k}{4\pi\mathrm{i}}\left(\frac{\pi^2}{6} - \mathrm{Li}_2\left(\mathrm{e}^{-4\pi\mathrm{i}n/k}\right)\right) + \mathcal{O}(\hat\lambda^{-1/2}, \mathrm{e}^{-2\pi\sqrt{2\hat\lambda}}). \tag{2.7.41}$$

Notice that this formula does not display the exchange symmetry $n \leftrightarrow N-n$ for the antisymmetric case $\eta = -1$. This is because this symmetry is not present for the antisymmetric super-representation, as pointed out in [208].

The leading order in λ in (2.7.41) is as expected, *i.e.*, n times the action of the fundamental string (and n times an extra framing factor). The non-trivial dependence on ν only appears at subleading order in λ, and therefore will not be visible in the supergravity approximation. As mentioned above, there are no known 1/2 BPS brane solutions carrying less than $k/2$ units of electric charge other than fundamental strings. So we expect that the above action describes the interaction of these coincident strings.

For n a multiple of $k/2$ (or of k, if it is odd), we see from (2.7.33) that $Y_* = 0$ and the integral (2.7.30) is over a full cycle. The argument of the dilogarithm in (2.7.41) is unity, canceling the $\pi^2/6$ term. Since Y^* passed through one of the cuts $\mathcal{C}_1$ or $\mathcal{C}_2$, it is now on a different sheet, and exactly at the branch point of the logarithm in $\omega_0(Y^{-1})$. This happens exactly for the value of n where the strings describing the Wilson loop can be replaced by D2-branes, which are the string theory incarnation of the vortex loop operators [114]. This suggests that the vortex loop operators are related to eigenvalues along the logarithmic branch-cut. It is possible to use our formalism to calculate the perturbative and instanton corrections to the these configurations and it would be interesting to understand further their significance in the matrix model.

2.8 Flavored theory and its gravity dual

It is possible to flavor the ABJM theory by adding matter hypermultiplets in the fundamental representation [17, 129, 130]. More precisely, one adds $N_f^{(i)}$ multiplets $(Q_i, \widetilde{Q}_i)$, with $i = 1, 2$. The fields Q_i, $i = 1, 2$ are in the representations $(N_1, 1)$ and $(1, N_2)$, respectively, while $\widetilde{Q}_i$ are in the conjugate representations $(\overline{N}_1, 1)$ and $(1, \overline{N}_2)$, respectively. This matter content breaks the $\mathcal{N} = 6$ supersymmetry of the ABJM theory down to $\mathcal{N} = 3$. Notice that the ABJM theory can be obtained, formally, as the limit

$$N_f^{(i)} \to 0 \tag{2.8.1}$$

of the flavored theory. We will denote by

$$N_f = N_f^{(1)} + N_f^{(2)} \tag{2.8.2}$$

the total number of flavours.

According to the general localization procedure described in the eprevious chapter the inclusion of extra matter hypermultiplets just leads to the insertion of determinant-type operators in the matrix integral (2.2.3):

$$\begin{aligned} &Z_{\mathcal{N}=3}(N_1, N_2, N_f^{(1)}, N_f^{(2)}, g_s) \\ &= \frac{1}{N_1! N_2!} \int \prod_{i=1}^{N_1} \frac{\mathrm{d}\mu_i}{2\pi} \prod_{j=1}^{N_2} \frac{\mathrm{d}\nu_j}{2\pi} \frac{\prod_{i<j} \left(2 \sinh\left(\frac{\mu_i - \mu_j}{2}\right)\right)^2 \left(2 \sinh\left(\frac{\nu_i - \nu_j}{2}\right)\right)^2}{\prod_{i,j} \left(2 \cosh\left(\frac{\mu_i - \nu_j}{2}\right)\right)^2} \\ &\times \prod_{i=1}^{N_1} \left(2 \cosh \frac{\mu_i}{2}\right)^{-N_f^{(1)}} \prod_{i=1}^{N_2} \left(2 \cosh \frac{\nu_i}{2}\right)^{-N_f^{(2)}} \mathrm{e}^{-\frac{1}{2g_s}\left(\sum_i \mu_i^2 - \sum_j \nu_j^2\right)}. \end{aligned} \tag{2.8.3}$$

As was previously discussed the large N dual of the ABJM Chern–Simons–matter theory is given by type IIA string theory on $\mathrm{AdS}_4 \times \mathbb{CP}^3$, which lifts to M-theory on $\mathrm{AdS}_4 \times \mathbb{S}^7/\mathbb{Z}_k$. It was conjectured in [17, 129, 130] that, when $N_1 = N_2 = N$, the $\mathcal{N} = 3$ theory with flavor has a type IIA large N dual where N_f D6 branes wrap the $\mathbb{RP}^3$ cycle inside $\mathbb{CP}^3$. This is the four-dimensional counterpart of the original construction of [164], which adds flavor to AdS_5 by wrapping D7 branes around an $\mathbb{S}^3$ inside $\mathbb{S}^5$. The flavored $\mathcal{N} = 3$ theory also describes N M2 branes probing an eight-dimensional hyperKähler cone $\mathcal{M}_8$ with Sp(2) holonomy. The base of this cone is a tri-Sasakian space X_7. The space $\mathcal{M}_8$ is a particular member of a family of hyperKähler cones $\mathcal{M}_8(\mathbf{t})$ labeled by three natural numbers $\mathbf{t} = (t_1, t_2, t_3)$. These cones can be constructed as hyperKähler quotients

$$\mathbb{H}^3///U(1), \tag{2.8.4}$$

where the $U(1)$ action is characterized by the three charges $\mathbf{t}$. The bases $X_7(\mathbf{t})$ of these cones give an infinite family of tri-Sasakian manifolds known as Eschenburg spaces, see [18] for a detailed study and references to the relevant literature. The dual to the $\mathcal{N} = 3$ Chern–Simons–matter theory with a total number of N_f fundamentals has charges

$$\mathbf{t} = (N_f, N_f, k)\,. \tag{2.8.5}$$

In the following, the eight-dimensional cone corresponding to this charge will be simply denoted by $\mathcal{M}_8$.

At large N the above theory of N M2 branes is described by M-theory on the manifold

$$\mathrm{AdS}_4 \times X_7, \tag{2.8.6}$$

where X_7 is the tri-Sasakian seven manifold corresponding to (2.8.5). This background is a particular example of backgrounds considered in Section 1.4.3. The volume of X_7 (with unit radius) is given by [18]

$$\mathrm{vol}(X_7) = \frac{\mathrm{vol}(\mathbb{S}^7)}{k\, \xi^2(\mu)}, \tag{2.8.7}$$

where

$$\xi(\mu) = \frac{1 + \mu}{\sqrt{1 + \mu/2}}, \qquad \mu = \frac{N_f}{k}. \tag{2.8.8}$$

Then from (1.4.57) it follows that the free energy of the flavoured theory has the following strong coupling behaviour.

$$-F_{\mathcal{N}=3}(\mathbb{S}^3)=\frac{\pi\sqrt{2}}{3}N^{3/2}k^{1/2}\xi\left(\mu\right) \tag{2.8.9}$$

Another quantity that we are interested in is the VEV of supersymmetric Wilson loops. As usual in the AdS/CFT correspondence, this can be calculated by evaluating the regularized area of a fundamental string in the type IIA reduction of the above M-theory background. The resulting geometry, which includes the full backreaction of the D6 branes, is a warped compactification and we have not performed such a calculation. However, it was pointed out in [17] that

$$R^2_{\rm str}\sim\frac{1}{4}\frac{R^3_{X_7}}{N_f+k}=\frac{2\pi N^{1/2}}{\sqrt{N_f+2k}} \tag{2.8.10}$$

and we then expect

$$\langle W_\Box\rangle\sim\exp\left(\frac{2\pi N^{1/2}}{\sqrt{N_f+2k}}\right) \tag{2.8.11}$$

for both the 1/2 and 1/6 BPS Wilson loops. This VEV incorporates the screening effect on Wilson loops due to unquenched flavor. Indeed, we see from (2.8.11) that, as the number of flavors grows, the exponent decreases. This might be interpreted as a conformal avatar of the screening effect. Also notice that, when $N_f\to 0$, one recovers in (2.8.11) the right value for the ABJM limit. Notice that, in the computation leading to (2.8.11), a possible contribution to the vev of strings ending on the D6 branes has not been taken into account. We will see however in this paper that a gauge theory computation at strong coupling leads to a result in agreement with (2.8.11), thus indicating that such contributions are absent or subleading.

2.9 Strong coupling limit and tropical geometry

In order to make contact with the AdS dual, one has to calculate the gauge theory/matrix model quantities at strong coupling. For ABJM theory considered in the previous sections this was done essentially by computing exact interpolating functions at all couplings and then going to the strong coupling regime. However, the calculation of interpolating functions might become hard, specially in more complicated generalizations of ABJM theory like the theories with matter considered in [17, 129, 130]. In particular, one would like to have a computational framework to do calculations directly at strong coupling, without going through the determination of exact interpolating functions. We will now propose such a framework, and we will illustrate it by considering the ABJM theory.

As was explained earlier the strong coupling limit of the ABJM theory corresponds to the large radius limit of this Calabi–Yau moduli space. This is the limit where

$$z_1,\ z_2\to 0. \tag{2.9.1}$$

In terms of the periods T_i defined in (2.3.19) this limit can be defined as

$$\mathrm{Re}\,(T_i)\to\infty,\qquad i=1,2, \tag{2.9.2}$$

and we can write, up to exponentially suppressed corrections,

$$z_1\approx \mathrm{e}^{-T_1},\quad z_2\approx\mathrm{e}^{-T_2}\,. \tag{2.9.3}$$

The 't Hooft parameters of the ABJM model are large in this limit, and they behave like [5, 6]

$$\lambda_i\sim\frac{T_1T_2}{8\pi^2},\qquad i=1,2. \tag{2.9.4}$$

We now ask the following question: what is the behavior of the planar resolvent, or equivalently the spectral curve (2.3.1), in this limit? Notice that the coefficients of (2.3.1), regarded as an equation for an algebraic curve, become exponentially large or small (or they remain constant). This kind of behavior has been very much studied recently in the mathematical literature and it is known as the *tropical limit* of the algebraic curve (or the *ultradiscretization* of the algebraic curve), see for example [165, 166]. This limit is only non-trivial if we scale z,y in the same way, where

$$z=\log Z,\qquad y=\log Y, \tag{2.9.5}$$

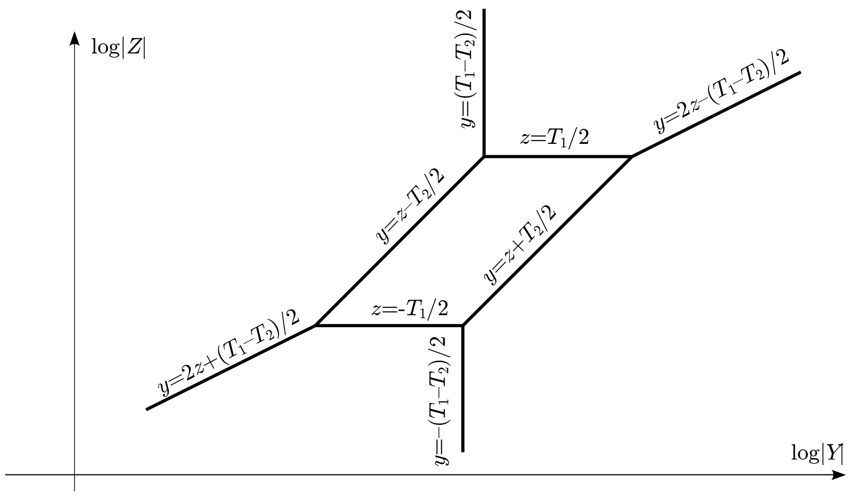

Figure 2.7: The strong coupling limit of the curve (2.3.1) can be represented as a set of segments where the relation between z and y is linear. This limit is nothing but the "ultradiscretization" or "tropicalization" of the spectral curve.

i.e. we have to consider the limit in which

$$\mathrm{Re}(z),\ \mathrm{Re}(y) \to \infty \tag{2.9.6}$$

as well. For generic values of z, y in this regime there is only one dominating term in (2.3.1), and the equation cannot be satisfied. To have a nontrivial equation we need at least two dominating terms which cancel each other. This gives us a set of linear equations on $\mathrm{Re}(z)$ and $\mathrm{Re}(y)$. Therefore, the "ultradiscrete" limit of the curve can be represented as a collection of segments in the real plane. On each of these segments there is a linear relation between z and y. It is an easy exercise to show that, for our particular example (2.3.1), the resulting diagram can be represented as in Fig. 2.7[4]. This diagram is called a *tropical curve*.

This two-dimensional plane can be understood as the base of the fibration

$$\begin{aligned} &\mathbb{C}^* \times \mathbb{C}^* \to \mathbb{R}^2 \\ &\quad (Z, Y) \mapsto (\log|Z|, \log|Y|). \end{aligned} \tag{2.9.7}$$

The fiber is $\mathbb{S}^1 \times \mathbb{S}^1$ and it is parametrized by the imaginary parts of (z, y). A linear relation of the form $mz = ny + c,\ m, n \in \mathbb{Z}$ gives a line in $\mathbb{R}^2$ with a fiber $\mathbb{S}^1 \subset \mathbb{S}^1 \times \mathbb{S}^1$ with winding number (n, m). Thus the lines in the picture correspond, in the original curve, to thin tubes connected at the vertices. This type of picture is familiar from local mirror symmetry: as emphasized in [167], the mirror curve of a toric manifold, like (2.3.1), can be regarded as the thickening of the toric diagram in which lines become cylinders or tubes. In the strong coupling or large radius limit, the tubes become thinner and we get back the toric skeleton, which is now interpreted as a tropical curve.

In order to do calculations at strong coupling we have to understand what happens to the period integrals in the regime (2.9.2). It can be shown rigorously (see for example [166]) that, in the limit (2.9.2), the periods of differentials on the original curve can be computed directly on the tropical curve, and they reduce to simple contour integrals along the two-dimensional diagram in Fig. 2.7. We then have to determine what is the tropical limit of the contours. To do this, we first note that in the limit (2.9.2) the endpoints of the cuts behave like

$$\begin{aligned} &A \approx \log\zeta \approx T_1/2 \approx T_2/2, \\ &a \approx -b \approx \zeta \approx \mathrm{e}^{A}, \\ &\frac{1}{a} \approx -\frac{1}{b} \approx \mathrm{e}^{-A}. \end{aligned} \tag{2.9.8}$$

[4]In writing the linear equations for the segments, we have neglected constant imaginary parts, which are small in the tropical limit.

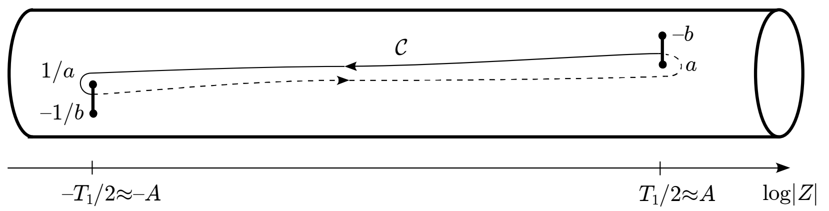

Figure 2.8: The tropical limit of the cuts of the Z-plane, represented here as a cylinder.

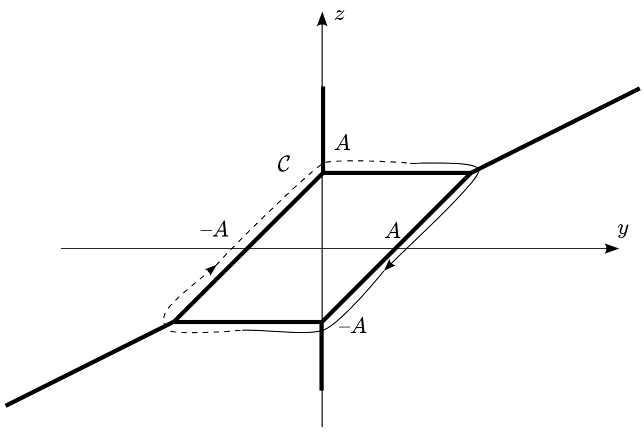

Figure 2.9: The contour $\mathcal{C}$ around $[1/a, a]$ becomes a parallelogram around the tropical curve.

Let us now represent the $\mathbb{C}^*$ domain of the variable Z as an infinite cylinder, as in Fig. 2.8. This picture also shows the contour $\mathcal{C} \equiv \mathcal{C}_1$ around the cut $[1/a, a]$ in the tropical limit. Since our curve (2.3.1) is a double covering of this cylinder, we can build it out of two copies of $\mathbb{C}^*$ glued along the cuts shown in Fig. 2.8. To see this in detail, let us look at the diagram shown in Fig. 2.7 and let us thicken it in order to reconstruct the original curve (2.3.1). If we remove the two horizontal segments, the thickening gives two infinite tubes which can be parametrized by z. Each of these tubes can be identified in turn with a copy of $\mathbb{C}^*$. In order to recover the full curve, we have to add the thickened horizontal segments. They give two horizontal tubes connecting the two copies of $\mathbb{C}^*$ at $z \approx -A$ and $z \approx A$. Notice that these locations are the positions of the small cuts drawn in Fig. 2.8. Since the solid and dashed pieces of $\mathcal{C}$ depicted in Fig. 2.8 lie on different copies of $\mathbb{C}^*$, we conclude that the contour $\mathcal{C}$ around the cut $[1/a, a]$ becomes, in the tropical limit, the two-dimensional contour around the parallelogram shown in Fig. 2.9.

Let us now consider the contour $\mathcal{D}$ in Fig. 2.3, which encircles the cut $[-1/b, 1/a]$. This cut corresponds to the horizontal tube at $z \approx -A$, therefore the contour becomes a non-trivial cycle around the tube. In the tropical limit it can be schematically drawn as in Fig. 2.10.

We can now use this formalism to compute some interesting physical quantities at strong coupling. For simplicity we will restrict ourselves to the ABJM slice $N_1 = N_2$. The resolvent is

$$\omega_0(z) = y(z)\,\mathrm{d}z. \tag{2.9.9}$$

We first determine the relation between the 't Hooft parameter and the modulus A as follows:

$$2\pi\mathrm{i}\lambda = t_1 = \frac{1}{4\pi\mathrm{i}} \oint_{\mathcal{C}} y(z)\,\mathrm{d}z \approx -\frac{A^2}{\pi\mathrm{i}}, \tag{2.9.10}$$

which leads to

$$A \approx \pi\sqrt{2\lambda} \tag{2.9.11}$$

in agreement with the result (2.9.4) from [5, 6]. The 1/6 BPS Wilson loop is given by a integral around

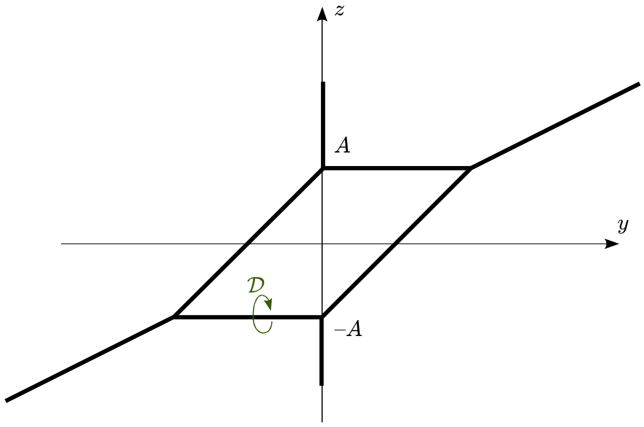

Figure 2.10: The contour $\mathcal{D}$ around $[-1/b, 1/a]$ becomes a non-trivial cycle around the shrinking tube.

the contour $\mathcal{C}$:

$$\left\langle W_{\square}^{1/6} \right\rangle_0 = \oint_{\mathcal{C}} \frac{\mathrm{d}z}{4\pi \mathrm{i}} \mathrm{e}^z y(z). \tag{2.9.12}$$

In the tropical limit, this becomes an elementary integral around the parallelogram shown in Fig. 2.9. We then find,

$$\left\langle W_{\square}^{1/6} \right\rangle_0 \approx \frac{1}{4\pi \mathrm{i}} \left(\int_{-A}^{A} (z - A)\, \mathrm{e}^z \, \mathrm{d}z - \int_{-A}^{A} (z + A)\, \mathrm{e}^z \, \mathrm{d}z \right) \approx \frac{\mathrm{i}A}{2\pi} \mathrm{e}^A. \tag{2.9.13}$$

Using (2.9.11) we find

$$\left\langle W_{\square}^{1/6} \right\rangle_0 \approx \frac{\mathrm{i}}{2} \sqrt{2\lambda} \mathrm{e}^{\pi \sqrt{2\lambda}} \tag{2.9.14}$$

which is the result obtained in section 2.5 from the exact interpolating function, up to an overall phase (this is due to the fact that we neglected constant subleading imaginary pieces in the equations for the segments of the tropical curve).

Another quantity that we can compute is the free energy at strong coupling. From (2.2.22) we find,

$$\frac{\partial F_0}{\partial \lambda} = -\pi \mathrm{i} \oint_{\mathcal{D}} y \, \mathrm{d}z = \pi \mathrm{i} \oint_{\mathcal{D}} z \, \mathrm{d}y \approx -\pi \mathrm{i} A \, \mathrm{mon}_{\mathcal{D}}\, y, \tag{2.9.15}$$

where $\mathrm{mon}_{\mathcal{D}}$ denotes the monodromy along the cycle $\mathcal{D}$. Since $a \sim b \to \infty$ and $Z \sim 1/a \sim 1/b$, we have

$$y \approx 2 \log \left\{ \sqrt{Z - 1/a} - \sqrt{Z + 1/b} \right\} + \text{const.}, \tag{2.9.16}$$

and

$$\mathrm{mon}_{\mathcal{D}}\, y = 2\pi \mathrm{i}. \tag{2.9.17}$$

We conclude that

$$\frac{\partial F_0}{\partial \lambda} \approx 2\pi^2 A \approx 2\pi^3 \sqrt{2\lambda} \tag{2.9.18}$$

which is the result obtained in section 2.5. Of course, the interest of this tropical formalism is the generalization to more complicated situations. This we will do in the next section, where we will consider the ABJM theory with fundamental matter introduced in section 2.

2.10 Quenched flavor in Chern–Simons–matter theories

2.10.1 The quenched approximation in the matrix model

In studying theories with fundamental matter multiplets in the context of the AdS/CFT correspondence, there have been essentially two approaches. In the first one, called the *quenched* or the *probe* approximation, one assumes that the number of flavors is much smaller than the number of colors. Since the flavor multiplets are usually obtained by adding branes to the original theory, the quenched approximation is equivalent to treating these branes as probes, and one assumes that they do not backreact on the background (see [168] for a review and a list of references for this approach). One can go beyond the quenched approximation and consider *unquenched* flavor, where the full backreaction of the branes is taken into account, see [169] for a recent review with references. It turns out that these two approaches have counterparts in the study of the matrix model (2.8.3) including flavor multiplets. We will first set up the matrix model analogue of the quenched approximation, and we will consider the full unquenched theory in the next section.

In the matrix model (2.8.3), the inclusion of fundamental flavors leads to the insertion of two determinant-like operators

$$\begin{aligned}&\prod_{i=1}^{N_1}\left(2\cosh\frac{\mu_i}{2}\right)^{-N_f^{(1)}}\prod_{j=1}^{N_2}\left(2\cosh\frac{\nu_j}{2}\right)^{-N_f^{(2)}}\\&=\exp\left[-N_f^{(1)}\sum_{i=1}^{N_1}\log\left(2\cosh\frac{\mu_i}{2}\right)-N_f^{(2)}\sum_{j=1}^{N_2}\log\left(2\cosh\frac{\nu_j}{2}\right)\right].\end{aligned} \tag{2.10.1}$$

We can treat these insertions as operators which perturb the partition function without changing the spectral curve or resolvent of the ABJM theory. To see how this works in practice, we write the partition function (2.8.3) as a normalized vev in the ABJM theory,

$$Z_{\mathcal{N}=3}(N_1,N_2,N_f^{(1)},N_f^{(2)},g_s)=\langle \mathrm{e}^{-\mathcal{W}}\rangle_{\mathrm{ABJM}}Z_{\mathrm{ABJM}}(N_1,N_2,g_s) \tag{2.10.2}$$

where

$$\mathcal{W}=N_f^{(1)}\mathcal{W}_1+N_f^{(2)}\mathcal{W}_2 \tag{2.10.3}$$

and

$$\mathcal{W}_1=\sum_{i=1}^{N_1}\log\left[2\cosh\frac{\mu_i}{2}\right],\qquad \mathcal{W}_2=\sum_{j=1}^{N_2}\log\left[2\cosh\frac{\nu_i}{2}\right]. \tag{2.10.4}$$

We can then calculate the free energy of the $\mathcal{N}=3$ theory as a cumulant expansion,

$$F_{\mathcal{N}=3}(N_1,N_2,N_f^{(1)},N_f^{(2)},g_s)=F_{\mathrm{ABJM}}(N_1,N_2,g_s)+\sum_{k=1}^{\infty}\frac{(-1)^k}{k!}\langle\mathcal{W}^k\rangle^{(c)}_{\mathrm{ABJM}} \tag{2.10.5}$$

where (c) denotes as usual the connected vev. Since $\mathcal{W}^k$ is a polynomial of degree k in the number of flavours $N_f^{(i)}$, the above cumulant expansion is an expansion around $N_f^{(i)}=0$. Equivalently, we can introduce the Veneziano parameters [163]

$$t_f^{(i)}=g_sN_f^{(i)}. \tag{2.10.6}$$

The perturbative series (2.10.5) is an expansion in the Veneziano parameters around $t_f^{(i)}=0$, which is valid for

$$t_f^{(i)}\ll 1, \tag{2.10.7}$$

or equivalently

$$N_f^{(i)}\ll \min(N_1,N_2), \tag{2.10.8}$$

which corresponds indeed to a quenched approximation. Each term in this series is given by an integrated correlator in the ABJM theory, which is computed with the master field described by the resolvent (2.2.17). Since the spectral curve is not changed, this is equivalent to neglecting the backreaction of the D-branes on the original geometry. Diagramatically, the genus g correction to $\langle\mathcal{W}^k\rangle^{(c)}_{\mathrm{ABJM}}$ gives the contribution of k "quark" loops to the genus g free energy, where all gluon diagrams of genus g have been resummed.

A similar perturbative scheme can be constructed for the calculation of operators $\mathcal{O}$ in the matrix model (like for example Wilson loops):

$$\langle\mathcal{O}\rangle_{\mathcal{N}=3}=\frac{\langle\mathcal{O}\mathrm{e}^{-\mathcal{W}}\rangle_{\mathrm{ABJM}}}{\langle\mathrm{e}^{-\mathcal{W}}\rangle_{\mathrm{ABJM}}}=\langle\mathcal{O}\rangle_{\mathrm{ABJM}}+\sum_{k=1}^{\infty}\frac{(-1)^k}{k!}\langle\mathcal{O}\mathcal{W}^k\rangle^{(c)}_{\mathrm{ABJM}} \tag{2.10.9}$$

The operator vevs appearing in (2.10.5) and (2.10.9) can be computed by using the connected correlation functions of the ABJM model. These are defined by

$$W(Z_1,\cdots,Z_h)=\left\langle \mathrm{Tr}\frac{1}{Z_1-U}\cdots\mathrm{Tr}\frac{1}{Z_h-U}\right\rangle^{(c)} \tag{2.10.10}$$

where U is given in (2.2.16). These correlators have a genus expansion

$$W(Z_1,\cdots,Z_h)=\sum_{g=0}^{\infty} g_s^{2g-2+h} W_g(Z_1,\cdots,Z_h) \tag{2.10.11}$$

which can be computed systematically with the techniques started in [170] and culminated in [24, 171] (for the Chern–Simons matrix models analyzed in this paper, one has to consider the slightly modified version of these techniques considered in [172, 201]). Let us consider for example the operators

$$\mathcal{O}_a(X)=\mathrm{Tr}\, f_a(X),\quad \widehat{\mathcal{O}}_b(Y)=\mathrm{Tr}\, g_b(Y),\qquad a=1,\cdots,h_1,\quad b=1,\cdots,h_2, \tag{2.10.12}$$

where

$$X=\mathrm{diag}\left(\mathrm{e}^{\mu_i}\right),\qquad Y=\mathrm{diag}\left(\mathrm{e}^{\nu_i}\right). \tag{2.10.13}$$

In this notation, the operators (2.10.4) are written as

$$\mathcal{W}_1(X)=\mathrm{Tr}\log\left(X^{\frac{1}{2}}+X^{-\frac{1}{2}}\right),\qquad \mathcal{W}_2(Y)=\mathrm{Tr}\log\left(Y^{\frac{1}{2}}+Y^{-\frac{1}{2}}\right). \tag{2.10.14}$$

We have then the following result for the connected correlators of these operators,

$$\begin{aligned}&\langle \mathcal{O}_1(X)\cdots\mathcal{O}_{h_1}(X)\widehat{\mathcal{O}}_1(Y)\cdots\widehat{\mathcal{O}}_{h_2}(Y)\rangle^{(c)}=\\ &\quad\oint_{\mathcal{C}_1}\frac{\mathrm{d}X_1}{2\pi\mathrm{i}}\cdots\oint_{\mathcal{C}_1}\frac{\mathrm{d}X_{h_1}}{2\pi\mathrm{i}}\oint_{\mathcal{C}2}\frac{\mathrm{d}Y_1}{2\pi\mathrm{i}}\cdots\oint_{\mathcal{C}2}\frac{\mathrm{d}Y_{h_2}}{2\pi\mathrm{i}}\,W(X_1,\cdots,X_{h_1},Y_1,\cdots,Y_{h_2})\\ &\quad\cdot f_1(X_1)\cdots f_{h_1}(X_{h_1})g_1(Y_1)\cdots g_{h_2}(Y_{h_2}).\end{aligned} \tag{2.10.15}$$

This leads to a systematic $1/N$ expansion by using (2.10.11). The planar limit of the one-point functions is given by the equivalent expressions

$$\langle\mathcal{O}(X)\rangle_0=t_1\int_{1/a}^{a}\rho_1(\mu)f(\mu)\mathrm{d}\mu,\qquad \langle\mathcal{O}(Y)\rangle_0=t_2\int_{-b}^{-1/b}\rho_2(\nu)g(\nu)\mathrm{d}\nu. \tag{2.10.16}$$

2.10.2 Quenched expansion at weak coupling

We will now present some concrete examples of the quenched approximation, calculated at weak coupling. The results can be tested with perturbative calculations in the matrix model. For simplicity, we will set $N_f^{(2)}=0$, and we will focus on the free energy. The first order correction in $N_f^{(1)}$ to the planar free energy is given by

$$-\langle\mathcal{W}_1\rangle_0=-\oint_{\mathcal{C}_1}\frac{\mathrm{d}Z}{2\pi\mathrm{i}}\log\left(Z^{\frac{1}{2}}+Z^{-\frac{1}{2}}\right)W_0(Z), \tag{2.10.17}$$

where

$$W_0(Z)=\frac{1}{2Z}\omega_0(Z). \tag{2.10.18}$$

The second order planar correction is

$$\frac{1}{2!}\langle\mathcal{W}_1^2\rangle_0^{(c)}=\frac{1}{2!}\oint_{\mathcal{C}_1}\frac{\mathrm{d}X_1}{2\pi\mathrm{i}}\oint_{\mathcal{C}_1}\frac{\mathrm{d}X_2}{2\pi\mathrm{i}}\log\left(X_1^{\frac{1}{2}}+X_1^{-\frac{1}{2}}\right)\log\left(X_2^{\frac{1}{2}}+X_2^{-\frac{1}{2}}\right)W_0(X_1,X_2) \tag{2.10.19}$$

and $W_0(X_1,X_2)$ is the two-cut, two-point planar correlator of the matrix model. It is related to the Bergmann kernel of the spectral curve $B(X_1,X_2)$ by [171]

$$W_0(X_1,X_2)=B(X_1,X_2)-\frac{1}{(X_1-X_2)^2} \tag{2.10.20}$$

and it was first calculated by Akemann [173] in the useful form:

$$\begin{aligned} W_0(X_1, X_2) = & \frac{1}{4(X_1 - X_2)^2} \left(\sqrt{\frac{(X_1 - a)(X_1 - 1/a)(X_2 + b)(X_2 + 1/b)}{(X_1 + b)(X_1 + 1/b)(X_2 - a)(X_2 - 1/a)}} \right. \\ & \left. + \sqrt{\frac{(X_1 + b)(X_1 + 1/b)(X_2 - a)(X_2 - 1/a)}{(X_1 - a)(X_1 - 1/a)(X_2 + b)(X_2 + b)}} \right) \\ & + \frac{(a + 1/b)(b + 1/a)}{4\sqrt{\sigma(X_1)\sigma(X_2)}} \frac{E(k)}{K(k)} - \frac{1}{2(X_1 - X_2)^2}, \end{aligned} \tag{2.10.21}$$

where $\sigma(Z)$ is given in (2.7.8).

An efficient way to calculate the above integrals at weak coupling is to perform the change of variables

$$X = \frac{a - a^{-1}}{2} y + \frac{a + a^{-1}}{2}, \tag{2.10.22}$$

and expand the integrand in series in t_i around $t_i = 0$. The coefficients of the resulting series are relatively simple integrals, which can be computed by deforming the contour in terms of residues at infinity. The result one obtains is

$$\begin{aligned} \langle \mathcal{W}_1 \rangle_0 &= \frac{t_1^2}{8} + \frac{1}{96} t_1^2 (t_1 + 6t_2) + \frac{1}{64} t_1^2 t_2^2 + \frac{1}{3072} t_1^2 t_2^2 \left(t_1^2 - 12 t_1 t_2 + t_2^2 \right) + \cdots, \\ \frac{1}{2!} \langle \mathcal{W}_1^2 \rangle_0^{(c)} &= \frac{t_1^2}{64} + \frac{1}{64} t_1^2 t_2 - \frac{1}{6144} t_1^2 \left(t_1^2 + 24 t_1 t_2 - 48 t_2^2 \right) + \cdots \end{aligned} \tag{2.10.23}$$

This can be explicitly checked against a direct calculation of the matrix integral. Indeed, we find in matrix model perturbation theory

$$\begin{aligned} F_{\mathcal{N}=3} = F_{\text{ABJM}} - N_f^{(1)} & \left[g_s \frac{N_1^2}{8} + g_s^2 \left(\frac{N_1^3}{96} - \frac{N_1^2 N_2}{16} - \frac{5N_1}{192} \right) + g_s^3 \left(\frac{N_1^2 N_2^2}{64} + \frac{N_1^2}{192} \right) + \cdots \right] \\ + (N_f^{(1)})^2 & \left[g_s^2 \frac{N_1^2}{64} - g_s^3 \left(\frac{N_1^2 N_2}{64} + \frac{N_1}{128} \right) \right. \\ & \left. - g_s^4 \left(\frac{N_1^4}{6144} - \frac{N_1^3 N_2}{256} - \frac{N_1^2 N_2^2}{128} - \frac{11 N_1^2}{248} - \frac{3 N_1 N_2}{512} \right) + \cdots \right] \end{aligned} \tag{2.10.24}$$

whose planar part agrees with (2.10.23).

2.10.3 Quenched expansion at strong coupling

Since the correlation functions (2.10.10) are given by contour integrals of meromorphic differentials, we can compute them with the tropical techniques that we introduced in the last section. We will focus on the free energy on $\mathbb{S}^3$, and we will assume for simplicity that $N_f^{(2)} = 0$ so that $N_f^{(1)} = N_f$. We can write

$$-\langle \mathcal{W}_1 \rangle_0 = -\frac{1}{4\pi i g_s} \oint_{\mathcal{C}} \omega_0(z) f(z) \mathrm{d}z \tag{2.10.25}$$

where

$$f(z) = \log \left(2 \cosh \frac{z}{2} \right). \tag{2.10.26}$$

In the tropical limit in which z is large the function f simplifies as

$$f(z) \approx \frac{|z|}{2}. \tag{2.10.27}$$

Then, the contour integral (2.10.25) becomes

$$-\oint_{\mathcal{C}} \omega_0(z) f(z) \mathrm{d}z \approx \int_{-A}^{A} \frac{|z|}{2} (z + A)\, \mathrm{d}z - \int_{-A}^{A} \frac{|z|}{2} (z - A)\, \mathrm{d}z = A^3 \approx \left(2\pi^2 \lambda \right)^{3/2} \tag{2.10.28}$$

and the first correction of order $\mathcal{O}(N_f)$ to the free energy is

$$-\frac{\pi}{4}N_f N\sqrt{2\lambda}\,. \tag{2.10.29}$$

The next order $\mathcal{O}(N_f^2)$ is much harder to compute with this technique, but it still can be done. Since this correction involves the two-point correlation function, which is essentially equal to the Bergmann kernel of the curve, what we have to do is to find the tropical limit of this kernel. To do this, we first discuss the tropical limit of holomorphic forms.

In tropical geometry a tropical holomorphic 1-form is a locally constant real 1-form with a "conservation" condition in the vertices, and which is zero on the external legs (for basic notions of tropical geometry see e.g. [165]). The dimension of the space of holomorphic 1-forms is obviously equal to the number of independent cycles of the graph, which coincides with the genus of the complex curve. In our case this space is a one-dimensional space with a basis h such that

$$h = \pm \mathrm{d}z \tag{2.10.30}$$

on the left and right sides of the parallelogram in Fig. 2.7, respectively, and

$$h = \pm \mathrm{d}y \tag{2.10.31}$$

on the upper and lower sides. We also have $h = 0$ on the external legs.

One can realize a tropical holomorphic 1-form as a limit of a complex holomorphic 1-form: as we have seen, for each edge of the tropical curve we have an integer direction vector (n, m). Then one can associate the integral of a complex holomorphic 1-form around the corresponding tube with the value of the tropical one form on this vector. In this way, the "conservation" condition in the vertices is a trivial consequence of holomorphicity. Since the external legs of the graph correspond to marked points on the complex curve, the absence of poles of the complex holomorphic 1-form at these points corresponds to the condition that the tropical holomorphic 1-form is zero on the external legs. In our particular case one can show explicitly that, in the tropical limit,

$$\frac{\mathrm{d}Z}{\sqrt{\sigma(Z)}} \approx \mathrm{e}^{-A}h\,. \tag{2.10.32}$$

The general notions of tropical Jacobian, Abel-Jacobi map and theta function were introduced in [174]. In our case the tropical Jacobian of our tropical curve C_{trop} is just

$$J_{\text{trop}} = \mathbb{R}/L\mathbb{Z} \cong \mathbb{S}^1 \tag{2.10.33}$$

where

$$L = \oint_{\mathcal{C}} h = 8A \tag{2.10.34}$$

is the perimeter[5] of the parallelogram. The tropical version of the Abel-Jacobi map is

$$\begin{aligned} u_{\text{trop}} &: C_{\text{trop}} \to J_{\text{trop}} \\ p &\mapsto u_{\text{trop}}(p) = \int_{p_0}^{p} h \mod L\mathbb{Z} \end{aligned} \tag{2.10.35}$$

which equals the length of a path between the points p and p_0. It can be obtained as a tropical limit of the ordinary Abel-Jacobi map:

$$u(p) \approx \frac{\mathrm{i}}{2\pi}u_{\text{trop}}(p). \tag{2.10.36}$$

The tropical theta function (with an odd characteristic) is

$$\theta_{\text{trop}}(v) = \max_{n\in\mathbb{Z}}\left\{-\frac{1}{2}Ln^2 + n(v - L/2)\right\}\,. \tag{2.10.37}$$

One can easily show that

$$\theta'_{\text{trop}}(v) = \left[\frac{v}{L}\right] \tag{2.10.38}$$

[5] As usual in tropical geometry, the length of the edge is a "geometric" length with an extra weight $(n^2 + m^2)^{-1/2}$.

where $[\cdot]$ denotes the floor function. Thus

$$\theta''_{\rm trop}(v) = \delta(v \bmod L\mathbb{Z})\,. \tag{2.10.39}$$

This tropical theta function can be obtained as a limit of the ordinary theta function with an odd characteristic

$$\Theta(z) = \sum_{n\in\mathbb{Z}} {\rm e}^{\pi{\rm i}\tau n^2 + 2\pi{\rm i}(z+\tau/2)}\,. \tag{2.10.40}$$

In the tropical limit we have

$$\tau \approx \frac{{\rm i}L}{2\pi} \to {\rm i}\infty, \tag{2.10.41}$$

so one exponential will dominate the others in the sum. Thus one can deduce that

$$\log\Theta\left(\frac{v}{2\pi{\rm i}}\right) \approx \theta_{\rm trop}(v). \tag{2.10.42}$$

Any Bergmann kernel can be written as

$$B = B_{\rm sing} + B_{\rm hol}\,, \tag{2.10.43}$$

where $B_{\rm sing}$ is given by (see e.g. [175])

$$B_{\rm sing}(p_1,p_2) = {\rm d}_1{\rm d}_2 \log\Theta(u(p_1)-u(p_2)) \tag{2.10.44}$$

and $B_{\rm hol}$ is a holomorphic part. In our case it should be chosen such that

$$\oint_{\mathcal{C}} B = 0\,. \tag{2.10.45}$$

In the tropical limit

$$B_{\rm hol}(p_1,p_2) \approx {\rm const}\cdot h(p_1)h(p_2)\,, \tag{2.10.46}$$

while

$$\begin{aligned} B_{\rm sing}(p_1,p_2) &\approx {\rm d}_1{\rm d}_2\theta_{\rm trop}(u_{\rm trop}(p_2)-u_{\rm trop}(p_1)) = -\delta(u_{\rm trop}(p_2)-u_{\rm trop}(p_1))h(p_1)h(p_2) \\ &\equiv -h_{\rm diag}(p_1,p_2)\,, \end{aligned} \tag{2.10.47}$$

where $h_{\rm diag}$ is supported on the diagonal and has the property

$$\oint_{\mathcal{C}\times\mathcal{C}} h_{\rm diag}(p_1,p_2)f(p_1,p_2) = \oint_{\mathcal{C}} h(p)f(p,p). \tag{2.10.48}$$

Imposing the condition (2.10.45) we get

$$B(p_1,p_2) \approx B_{\rm trop}(p_1,p_2) = -h_{\rm diag}(p_1,p_2) + \frac{h(p_1)h(p_2)}{L}\,. \tag{2.10.49}$$

We can now compute the second order correction at order $\mathcal{O}(N_f^2)$, (2.10.19), by using tropical techniques. It is given by

$$\frac{N_f^2}{2!}\frac{1}{(2\pi{\rm i})^2}\oint_{\mathcal{C}\times\mathcal{C}} B(p_1,p_2)\,\log\left(2\cosh\frac{z_1}{2}\right)\,\log\left(2\cosh\frac{z_2}{2}\right), \tag{2.10.50}$$

since the double pole subtracted in (2.10.20) does not contribute to the double contour integral. In the tropical limit this integral reads

$$\begin{aligned} \oint_{\mathcal{C}\times\mathcal{C}} B_{\rm trop}\left|\frac{z_1}{2}\right|\left|\frac{z_2}{2}\right| &= \frac{1}{2^5 A}\left(\int_{\mathcal{C}} h|z|\right)^2 - \frac{1}{2^2}\int_{\mathcal{C}} h|z|^2 = \frac{1}{2^2}\left\{\frac{\left(6A^2\right)^2}{8A} - \frac{16A^3}{3}\right\} \\ &= -\frac{5}{3\cdot 2^3}\,A^3. \end{aligned} \tag{2.10.51}$$

Using (2.9.11) we obtain that the correction of order $\mathcal{O}(N_f^2)$ is

$$\frac{5\,N_f^2\pi\sqrt{2}}{96}\lambda^{3/2}. \tag{2.10.52}$$

On the other hand, the AdS prediction for the free energy is given by (2.8.9). The quenched approximation is obtained by expanding this quantity for small N_f. Since

$$\xi(\mu) = 1 - \sum_{k=1}^{\infty} \frac{(1+2k)(2k-3)!!}{4^k\,k!}(-\mu)^k = 1 + \frac{\mu}{4} - \frac{5\mu^2}{32} + \cdots \tag{2.10.53}$$

we find

$$F_{\mathcal{N}=3}(\mathbb{S}^3) = -\frac{\pi\sqrt{2}}{3}N^{3/2}k^{1/2} - \frac{\pi}{4}N_f N\sqrt{2\lambda} + \frac{5\pi\sqrt{2}}{96}N_f^2\lambda^{3/2} + \mathcal{O}(N_f^3). \tag{2.10.54}$$

We then see that the tropical computations (2.10.29), (2.10.52) reproduce correctly the first two terms in this expansion.

One can try to compute the next corrections by calculating the tropical limit of the connected correlators (2.10.10) for $h \geq 3$, but as we will see in the next section it is possible to solve the planar theory at all values of $N_f^{(i)}$ (i.e. in the Veneziano limit) and calculate the tropical limit directly.

2.11 Unquenched flavor in Chern–Simons–matter theories

We now solve the matrix model (2.8.3) in the planar limit, but for all values of $N_f^{(i)}$, by using the techniques of [100, 162].

2.11.1 Exact resolvent in the Veneziano limit

The starting point in the calculation of the resolvent are the saddle-point equations

$$\begin{aligned}
\frac{\mu_i}{g_s} + \frac{N_f^{(1)}}{2}\tanh\left(\frac{\mu_i}{2}\right) &= \sum_{j\neq i}^{N_1}\coth\frac{\mu_i-\mu_j}{2} - \sum_{a=1}^{N_2}\tanh\frac{\mu_i-\nu_a}{2},\\
-\frac{\nu_a}{g_s} + \frac{N_f^{(2)}}{2}\tanh\left(\frac{\nu_i}{2}\right) &= \sum_{b\neq a}^{N_2}\coth\frac{\nu_a-\nu_b}{2} - \sum_{i=1}^{N_1}\tanh\frac{\nu_a-\mu_i}{2}.
\end{aligned} \tag{2.11.1}$$

We will solve instead the problem

$$\begin{aligned}
\mu_i + \frac{t_f^{(1)}}{2}\tanh\left(\frac{\mu_i}{2}\right) &= \frac{t_1}{N_1}\sum_{j\neq i}^{N_1}\coth\frac{\mu_i-\mu_j}{2} + \frac{t_2}{N_2}\sum_{a=1}^{N_2}\tanh\frac{\mu_i-\nu_a}{2},\\
\nu_a - \frac{t_f^{(2)}}{2}\tanh\left(\frac{\nu_i}{2}\right) &= \frac{t_2}{N_2}\sum_{b\neq a}^{N_2}\coth\frac{\nu_a-\nu_b}{2} + \frac{t_1}{N_1}\sum_{i=1}^{N_1}\tanh\frac{\nu_a-\mu_i}{2},
\end{aligned} \tag{2.11.2}$$

analytically in the parameters $t_{1,2}$ and $t_f^{(1,2)}$, and then we will perform the analytic continuation

$$t_2 \to -t_2. \tag{2.11.3}$$

The procedure to solve this type of equations is as in [88, 100, 162]. We first introduce exponentiated variables

$$Z_i = \mathrm{e}^{\mu_i}, \qquad W_a = \mathrm{e}^{\nu_a}. \tag{2.11.4}$$

In terms of these variables the saddle–point equations read

$$\begin{aligned}
\log Z_i + \frac{t_f^{(1)}}{2}\frac{Z_i-1}{Z_i+1} &= t_1\frac{N_1-1}{N_1} + t_2 + \frac{2t_1}{N_1}\sum_{j\neq i}^{N_1}\frac{Z_j}{Z_i-Z_j} - \frac{2t_2}{N_2}\sum_{a=1}^{N_2}\frac{W_a}{Z_i+W_a},\\
\log W_a - \frac{t_f^{(2)}}{2}\frac{W_a-1}{W_a+1} &= t_1 + t_2\frac{N_2-1}{N_2} + \frac{2t_2}{N_2}\sum_{b\neq a}^{N_2}\frac{W_b}{W_a-W_b} - \frac{2t_1}{N_1}\sum_{i=1}^{N_1}\frac{Z_i}{W_a+Z_i}.
\end{aligned} \tag{2.11.5}$$

The resolvent $\omega_0(Z)$ is defined as in (2.2.15), and it will have two cuts corresponding to the set of eigenvalues. Let $[a,b],[c,d]\subset\mathbb{R}$ be the cuts corresponding to $-W_a$ and Z_i, respectively. In terms of the planar resolvent we have,

$$\begin{aligned}\log Z+\frac{t_f^{(1)}}{2}\frac{Z-1}{Z+1}&=\frac{1}{2}\left(\omega_0(Z+\mathrm{i}0)+\omega_0(Z-\mathrm{i}0)\right),\\ \log(-W)-\frac{t_f^{(2)}}{2}\frac{W+1}{W-1}&=\frac{1}{2}\left(\omega_0(W+\mathrm{i}0)+\omega_0(W-\mathrm{i}0)\right).\end{aligned}\tag{2.11.6}$$

As in [162] we now define the functions,

$$\begin{aligned}F(Z)&=\sqrt{\sigma(Z)}\int_c^d \mathrm{d}X\,\frac{f(X)}{Z-X},\\ G(Z)&=\sqrt{\sigma(Z)}\int_a^b \mathrm{d}X\,\frac{g(X)}{Z-X},\end{aligned}\tag{2.11.7}$$

where

$$\begin{aligned}f(x)&=\frac{1}{\sqrt{|\sigma(x))|}}\left(\log x+\frac{t_f^{(1)}}{2}\frac{x-1}{x+1}\right),\\ g(x)&=-\frac{1}{\sqrt{|\sigma(x))|}}\left(\log(-x)-\frac{t_f^{(2)}}{2}\frac{x+1}{x-1}\right).\end{aligned}\tag{2.11.8}$$

It is then easy to show that the planar resolvent, defined again by the VEV (2.2.15), is given by

$$\omega_0(Z)=\frac{1}{\pi}(F(Z)+G(Z)).\tag{2.11.9}$$

One can check that, as in the ABJM matrix model,

$$ab=1,\quad cd=1.\tag{2.11.10}$$

This follows from the symmetry of the saddle point equations under $Z\to Z^{-1}$, $W\to W^{-1}$, together with the conditions on the endpoints imposed by the asymptotic behavior of $\omega_0(z)$ at infinity,

$$\omega_0(z)\sim t,\qquad z\to\infty.\tag{2.11.11}$$

From now on we will denote the two independent endpoints as a,b, as in the ABJM model, so that the cuts are $[1/a,a]$ and $[-b,-1/b]$. The two equations

$$t_1=\oint_{\mathcal{C}_1}\frac{\mathrm{d}Z}{4\pi\mathrm{i}}\frac{\omega_0(Z)}{Z},\qquad t_2=\oint_{\mathcal{C}_2}\frac{\mathrm{d}Z}{4\pi\mathrm{i}}\frac{\omega_0(Z)}{Z}\tag{2.11.12}$$

determine them a,b as a function of t_1,t_2. Here, $\mathcal{C}_i$ encircle $[1/a,a]$ and $[-b,-1/b]$, respectively.

We can now calculate the planar resolvent explicitly. It is given by two pieces. The first one is,

$$\int_{1/a}^{a}\frac{\mathrm{d}X}{Z-X}\frac{\log X}{\sqrt{|\sigma(X)|}}-\int_{-b}^{-1/b}\frac{\mathrm{d}X}{Z-X}\frac{\log(-X)}{\sqrt{|\sigma(X)|}},\tag{2.11.13}$$

which is simply the resolvent of the lens space matrix model (2.2.17). The second piece is

$$\frac{t_f^{(1)}}{2}\int_{1/a}^{a}\frac{\mathrm{d}X}{Z-X}\frac{X-1}{X+1}\frac{1}{\sqrt{|\sigma(X)|}}+\frac{t_f^{(2)}}{2}\int_{-b}^{-1/b}\frac{\mathrm{d}X}{Z-X}\frac{X+1}{X-1}\frac{1}{\sqrt{|\sigma(X)|}}.\tag{2.11.14}$$

These integrals can be expressed in terms of elliptic functions. In order to do so we use the results

$$\begin{aligned}\int_{1/a}^{a}\frac{\mathrm{d}X}{(Z-X)\sqrt{|\sigma(X)|}}&=\frac{2\sqrt{ab}}{1+ab}\frac{1}{(Z-a)(Z+b)}\left((a+b)\Pi(n^+(Z),k)+(Z-a)K(k)\right),\\ \int_{-b}^{-1/b}\frac{\mathrm{d}X}{(Z-X)\sqrt{|\sigma(X)|}}&=\frac{2\sqrt{ab}}{1+ab}\frac{1}{(Z-a)(Z+b)}\left(-(a+b)\Pi(n^-(Z),k)+(Z+b)K(k)\right),\end{aligned}\tag{2.11.15}$$

where

$$k^2 = \frac{(a^2-1)(b^2-1)}{(1+ab)^2}, \quad n^+(Z) = -\frac{a^2-1}{1+ab}\frac{Z+b}{Z-a}, \quad n^-(Z) = -\frac{b^2-1}{1+ab}\frac{Z-a}{Z+b}. \tag{2.11.16}$$

Defining the auxiliary function

$$\begin{aligned} &J(a,b,Z,s) = \\ &\frac{2\sqrt{ab}}{1+ab}\frac{1}{Z+s}\left(\left((a+b)\frac{Z-s}{(Z-a)(Z+b)}\Pi(n^+(Z),k) + \frac{Z-s}{Z+b}K(k)\right) - (Z\to -s)\right), \end{aligned} \tag{2.11.17}$$

we finally obtain

$$\omega_0(Z) = \omega_0^{\rm ABJM}(Z) + \frac{\sqrt{\sigma(Z)}}{2\pi}\left(t_f^{(1)} J(a,b,Z,1) + t_f^{(2)} J(-b,-a,Z,-1)\right), \tag{2.11.18}$$

where $\omega_0^{\rm ABJM}(Z)$ is the resolvent in the theory without matter, and it is given in (2.2.17). The asymptotic behavior (2.11.11) determines

$$\begin{aligned} &t = \log(\beta) \\ &-\frac{\sqrt{ab}}{\pi(1+ab)}\left(t_f^{(1)}\left((b+1)K(k) - (a+b)\Pi(n_a,k)\right) + t_f^{(2)}\left(-(a+1)K(k) + (a+b)\Pi(n_b,k)\right)\right), \end{aligned} \tag{2.11.19}$$

where we have used the notation

$$n_a = \frac{1-a^2}{1+ab}, \quad n_b = \frac{1-b^2}{1+ab}. \tag{2.11.20}$$

The relation (2.11.19) reduces to (2.2.20) when both $t_f^{(1)}$ and $t_f^{(2)}$ go to zero.

2.11.2 Weak coupling limit in the unquenched theory

In order to test the above expressions, we can compute the expansion of the resolvent at weak 't Hooft coupling (i.e. around $t_i = 0$) but for arbitrary $t_f^{(i)}$. To do this, the first step is to express the endpoints of the cuts in terms of the 't Hooft parameters. The period integrals for t_i can be expanded around $a = 1$, $b = 1$, and these series expansions can be inverted. At the first few orders in t_i we find

$$\begin{aligned} a &= 1 + \frac{1}{\sqrt{T_f^{(1)}}}2\sqrt{t_1} + \frac{1}{T_f^{(1)}}2t_1 + \frac{1}{6\left(T_f^{(1)}\right)^{3/2}}\left(7 - \frac{1}{T_f^{(1)}}\right)\frac{3}{2}t_1^{3/2} + \frac{1}{\left(T_f^{(1)}\right)^{3/2}}\frac{1}{2}\sqrt{t_1}t_2 \\ &\quad + \frac{1}{2\left(T_f^{(1)}\right)^2}\left(3 - \frac{1}{T_f^{(1)}}\right)t_1^2 + \frac{1}{\left(T_f^{(1)}\right)^2}t_1t_2 + \cdots, \\ b &= 1 + \frac{1}{\sqrt{T_f^{(2)}}}2\sqrt{t_2} + \frac{1}{T_f^{(2)}}2t_2 + \frac{1}{6\left(T_f^{(2)}\right)^{3/2}}\left(7 - \frac{1}{T_f^{(2)}}\right)\frac{3}{2}t_2^{3/2} + \frac{1}{\left(T_f^{(2)}\right)^{3/2}}\frac{1}{2}\sqrt{t_2}t_1 \\ &\quad + \frac{1}{2\left(T_f^{(2)}\right)^2}\left(3 - \frac{1}{T_f^{(2)}}\right)t_2^2 + \frac{1}{\left(T_f^{(2)}\right)^2}t_2t_1 + \cdots. \end{aligned} \tag{2.11.21}$$

In these equations,

$$T_f^{(1)} = 1 + \frac{t_f^{(1)}}{4}, \qquad T_f^{(2)} = 1 - \frac{t_f^{(2)}}{4}. \tag{2.11.22}$$

When $t_f^{(i)} = 0$ we recover the mirror map at the orbifold point of [88]. As a test of these results, we can calculate the coefficient of $1/Z$ in the resolvent. After the analytic continuation (2.11.3), this coefficient computes the planar VEV of the supertrace of U in the matrix model (2.8.3), as in (2.2.38):

$$\begin{aligned} &2g_s\left\langle {\rm Str}_\square\, U\right\rangle_0 = \zeta \\ &+ \frac{t_f^{(1)}}{2\pi\sqrt{ab}(1+ab)}\left((1+ab)^2E(k) + (b-a+a^2b+3ab^2)K(k) - 4ab(a+b)\Pi(n_a,k)\right) \\ &+ \frac{t_f^{(2)}}{2\pi\sqrt{ab}(1+ab)}\left((1+ab)^2E(k) + (a-b+ab^2+3a^2b)K(k) - 4ab(a+b)\Pi(n_b,k)\right), \end{aligned} \tag{2.11.23}$$

where (2.11.3) must be implemented at the end of the calculation. At weak coupling we find,

$$\begin{aligned} g_s \langle \mathrm{Str}_\square U \rangle_0 &= t_1 + t_2 + \frac{1}{2T_f^{(1)}} t_1^2 - \frac{1}{2T_f^{(2)}} t_2^2 \\ &+ \frac{1}{4\left(T_f^{(1)}\right)^2}\left(1 - \frac{1}{3T_f^{(1)}}\right) t_1^3 - \frac{t_1^2 t_2}{4\left(T_f^{(1)}\right)^2} - \frac{t_1 t_2^2}{4\left(T_f^{(2)}\right)^2} + \frac{1}{4\left(T_f^{(2)}\right)^2}\left(1 - \frac{1}{3T_f^{(2)}}\right) t_2^3 + \cdots \end{aligned} \tag{2.11.24}$$

The r.h.s. has the expected symmetries of the matrix model. Indeed, it is odd under

$$t_1 \leftrightarrow -t_2, \qquad t_f^{(1)} \leftrightarrow -t_f^{(2)}. \tag{2.11.25}$$

Notice that each term in the expansion (2.11.24) is a rational function of the Veneziano parameters. From the diagrammatic point of view, each of these terms corresponds to a fixed planar "gluon" diagram (with a boundary associated to the insertion of U) in which we have summed over all the "quark" loops, i.e. the "gluons" are quenched and the "quarks" are dynamical. One can actually check the first few terms written down in (2.11.24) againts an explicit perturbative calculation in the matrix model.

Based on [99], one should expect that the VEV (2.11.24) computes (twice) the VEV of the 1/2 BPS Wilson loop operator. However, in order to assert this one should first check that the construction of [99] of this operator extends to the flavored theory that we are considering here. In any case, the formulae (2.2.22), (2.2.42) remain valid in the flavored theory, since the 1/6 BPS Wilson loop operator can be constructed for all Chern–Simons–matter theories considered in [2]. We will now evaluate these formulae in the unquenched theory at strong coupling, by using tropical techniques.

2.11.3 Strong coupling limit in the unquenched theory

For simplicity we will set $N_f^{(2)} = 0$, $N_f = N_f^{(1)}$. We will write (2.11.18) as

$$\omega_0(z) = y(z)\,\mathrm{d}z, \quad y(z) = y_\mathrm{p}(z) + \mu\, y_\mathrm{m}(z). \tag{2.11.26}$$

In this equation, μ is defined in (2.8.8),

$$y_\mathrm{p}(z) = \omega_0^{\mathrm{ABJM}}(z) \tag{2.11.27}$$

is the equation of the spectral curve (2.2.17) in the ABJM model, and

$$y_\mathrm{m}(z) = \frac{1}{2}\oint_{\mathcal{C}} \frac{\mathrm{d}X}{2\pi\mathrm{i}} \frac{1}{Z-X} \frac{X-1}{X+1} \frac{\sqrt{\sigma(Z)}}{\sqrt{\sigma(X)}}. \tag{2.11.28}$$

In the tropical limit $\zeta \approx \mathrm{e}^A$ is large, and we will set

$$\beta = \mathrm{e}^K \tag{2.11.29}$$

where K is a parameter to be determined. In the ABJM model with $N_1 = N_2$ one has $K = 0$. We will shortly determine the value of K in the theory with unquenched flavor, in the tropical limit. We will assume that $A > 0$, $|K| < A$, which will be justified *a posteriori.*

The integral (2.11.28) can be evaluated "tropically." To do this we use the tropical limit (2.10.32) of the holomorphic one-form, as well as the limits

$$\frac{Z}{Z-X} \approx \Phi(x,z) \equiv \begin{cases} 1, & \text{if } x < z \\ -\mathrm{e}^{-(x-z)}, & \text{if } x > z \end{cases} \tag{2.11.30}$$

and

$$\frac{\mathrm{e}^{-A}\sqrt{\sigma(Z)}}{Z} \approx \pm\Xi(z) \equiv \pm \begin{cases} -\mathrm{e}^{-(x-z)}, & \text{if } z < -A \\ 1, & \text{if } -A < z < A \\ -\mathrm{e}^{z-A}, & \text{if } z > A \end{cases} \tag{2.11.31}$$

Here the $\pm$ sign corresponds to the two determinations of the square root. One finally obtains

$$y_\mathrm{m}(z) \approx \pm\frac{1}{2}\,\Xi(z) \oint_{\mathcal{C}} h_x \Phi(x,z)\,\mathrm{sign}\,x, \tag{2.11.32}$$

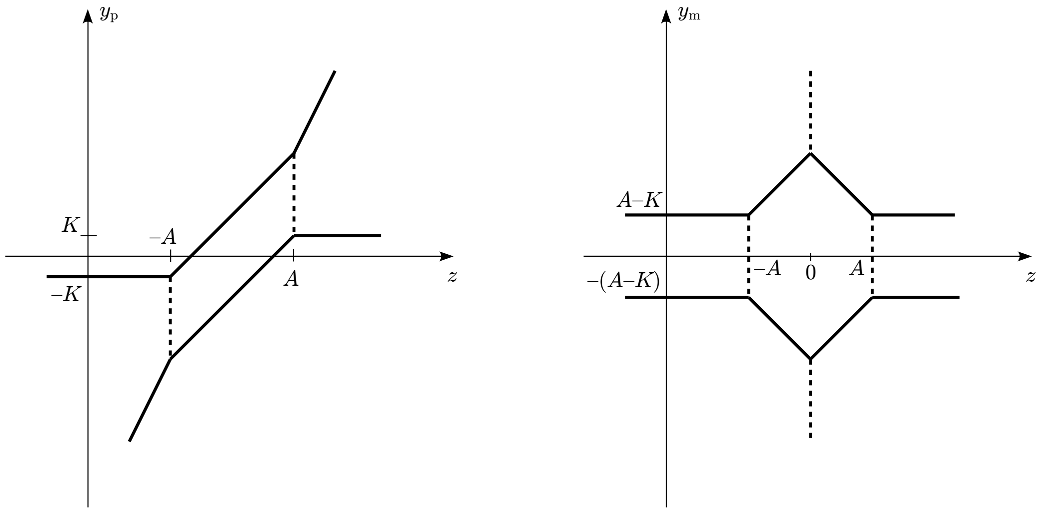

Figure 2.11: The two-dimensional graphs representing the tropical limits of $y_{\rm p}(z)$ (left) and $y_{\rm m}(z)$ (right).

where h_x is defined in (2.10.32) but with z replaced by x. In this result, $\operatorname{sign} x/2$ can be interpreted as the derivative of $|x|/2$, which is the tropical limit of the potential deformation, see (2.10.27), and in principle we can generalized it to other deformations. The tropical limit of $y_{\rm m}(z)$ can be rewritten as

$$y_{\rm m}(z) \approx \pm \frac{1}{2} \int_{\mathcal{C}_z} h_x \operatorname{sign} x = \pm \begin{cases} A - K + (A - |z|), & \text{if } -A < z < A, \\ A - K, & \text{otherwise,} \end{cases} \tag{2.11.33}$$

where $\mathcal{C}_z$ is a line connecting two different points on the curve with the same value of z. Equivalently one can obtain (2.11.33) by taking the tropical limit in the explicit expression (2.11.18). The tropical limit of the two-valued functions $y_{\rm p}(z)$, $y_{\rm m}(z)$ can be represented by the two-dimensional graphs shown in Fig. 2.11. Of course, the diagram for $y_{\rm p}(z)$ is nothing but the tropical curve represented in Fig. 2.7.

We now want to find the relation between A and K. To do this, we will impose for simplicity that $N_1 = N_2 = N$ in the $\mathcal{N} = 3$ Chern–Simons–matter theory. This means that the total 't Hooft parameter $t = t_1 + t_2$ vanishes. It follows from (2.11.12) that this sum can be evaluated by deforming the sum of the contours $\mathcal{C}_1$ and $\mathcal{C}_2$ to infinity and the origin, so we obtain

$$\operatorname{res}_{Z=\infty} \omega_0(Z) - \operatorname{res}_{Z=0} \omega_0(Z) = 0, \tag{2.11.34}$$

which leads to

$$K = \frac{\mu}{1+\mu} A. \tag{2.11.35}$$

Notice that in the limit $\mu = 0$ we correctly reproduce $K = 0$. With this relation at hand we can already add the two graphs to obtain the tropical curve representing the new resolvent $\omega_0(z)$, which is shown in Fig. 2.12. The calculation of the different periods reduces, like before, to trivial line integrals on the plane. We first have to relate A to the 't Hooft parameter. We have, for the period (2.11.12),

$$2\pi{\rm i}\lambda = t_1 = \frac{1}{4\pi{\rm i}} \oint_{\mathcal{C}} y(z)\, {\rm d}z \approx -\frac{(1+\mu/2)A^2}{\pi{\rm i}} \tag{2.11.36}$$

and we find

$$A \approx \frac{\pi\sqrt{2\lambda}}{\sqrt{1+\mu/2}}, \tag{2.11.37}$$

which is the deformation of the relationship (2.9.11) in the ABJM theory. For the planar free energy we have, as in the period integral (2.9.15) for the ABJM theory,

$$\frac{\partial F_0}{\partial \lambda} = -\pi{\rm i} A \operatorname{mon}_{\mathcal{D}} y. \tag{2.11.38}$$

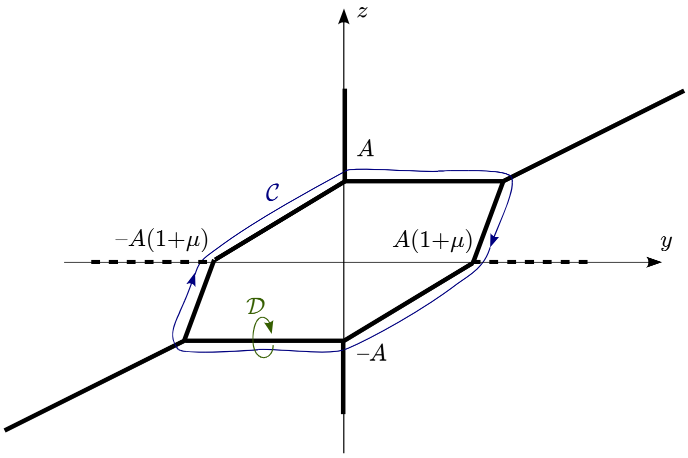

Figure 2.12: The tropical curve representing the tropical limit of the resolvent (2.11.18) of the $\mathcal{N} = 3$ theory.

The monodromy of $y_{\rm m}(z)$ can be computed as (2.9.16). Indeed, the main contribution to the integral for $y_{\rm m}$ is given by the vicinity of $1/a$, and we have

$$\begin{aligned} y_{\rm m} &\approx \int_{1/a}^{\infty} \frac{{\rm d}X}{X-Z} \frac{\sqrt{(Z-1/a)(Z+1/b)}}{\sqrt{(X-1/a)(X+1/b)}} \\ &= \log\left\{\sqrt{Z-1/a}-\sqrt{Z+1/b}\right\} - \log\left\{\sqrt{Z-1/a}+\sqrt{Z+1/b}\right\}. \end{aligned} \tag{2.11.39}$$

Then the monodromy around the cut $[-1/b, 1/a]$ is

$$\text{mon}_{\mathcal{D}} y_{\rm m} \approx 2\pi {\rm i}\,. \tag{2.11.40}$$

We conclude that

$$\frac{\partial F_0}{\partial \lambda} \approx 2\pi^2 A(1+\mu), \tag{2.11.41}$$

or equivalently,

$$F_0(\lambda) \approx \frac{1}{3}\pi^3 \sqrt{2\lambda^3}\, \frac{1+\mu}{\sqrt{1+\mu/2}}. \tag{2.11.42}$$

This is in perfect agreement with the AdS prediction (2.8.9). Notice in particular that we have been able to reconstruct the full nontrivial function $\xi(\mu)$ involved in the volume of the tri-Sasakian target space (2.8.7).

Finally, we can calculate the vev of the 1/6 BPS Wilson loop, which is given again by (2.9.12) but now with the new resolvent. We obtain,

$$\begin{aligned} \left\langle W^{1/6}_{\square} \right\rangle_0 &\approx \frac{1}{4\pi i}\left(\int_{-A}^{A} {\rm e}^z \left(z - A - \mu(A-|z|)\right) {\rm d}z - \int_{-A}^{A} {\rm e}^z \left(z + A + \mu(A-|z|)\right) {\rm d}z \right) \\ &\approx \frac{\rm i}{2}\sqrt{\frac{2\lambda}{1+\mu/2}} \exp\left(\pi\sqrt{\frac{2\lambda}{1+\mu/2}}\right). \end{aligned} \tag{2.11.43}$$

With some more work one can show that the would-be 1/2 BPS Wilson loop has the same leading, exponential dependence. This is in perfect agreement with the AdS prediction (2.8.11), and the exponent should be equal to the regularized area of a fundamental string in the corresponding type IIA background.

2.12 The all genus free energy

In this section we present and extend the results for the all genus free energy of ABJM theory and we give an M-theory/string theory interpretation for them.

2.12.1 Genus expansion in the matrix model

In this section we provide an efficient, recursive method to compute the $1/N$ corrections to the free energy in the case $N_1 = N_2 = N$. This is based on the modular properties of the solution and the technique of direct integration of the holomorphic anomaly equations. The method determines *a priori* the full $1/N$ expansion. In practice it is quite efficient and it makes possible to calculate the F_g corrections for high genera. This is then used to estimate non-perturbative effects in the large N expansion.

As noted in [113], we can use the relation between the local $\mathbb{F}_0$ theory and Seiberg–Witten theory to write all the quantities in the model in terms of modular forms. This representation becomes particularly useful when we restrict ourselves to a one-parameter model, as it was shown in a different context in [180]. When $N_1 = N_2$, $\beta = 1$ and the modulus u becomes simply

$$u = 1 + \frac{\kappa^2}{8}. \tag{2.12.1}$$

In Seiberg–Witten theory, u is related to the modular parameter τ of the Seiberg–Witten curve by

$$u = \frac{\vartheta_4^4 - \vartheta_2^4}{\vartheta_3^4}(\tau) = 1 - 32q^{1/2} + 256q + \cdots \tag{2.12.2}$$

where $q = \mathrm{e}^{2\pi \mathrm{i}\tau}$. This formula can be inverted to

$$\tau = \mathrm{i}\frac{K'\left(\frac{\mathrm{i}\kappa}{4}\right)}{K\left(\frac{\mathrm{i}\kappa}{4}\right)}, \tag{2.12.3}$$

therefore we see that the modular parameter τ is related to the specific heat of the theory through (2.5.41). Let us now introduce the quantity

$$\xi = \frac{2}{\vartheta_2^2(\tau)\vartheta_4^4(\tau)}. \tag{2.12.4}$$

This is proportional to the third derivative of the genus zero free energy, therefore to the Yukawa coupling $C_{\lambda\lambda\lambda}$. More precisely, we have

$$\partial_\lambda^3 F_0(\lambda) = -8\pi^3 \mathrm{i}\,\xi. \tag{2.12.5}$$

Therefore, the planar content of the theory can be elegantly encoded in terms of modular forms on the Seiberg–Witten curve.

One powerful application of the modular properties of the ABJM theory is the determination of the higher genus corrections to the free energy, $F_g(\lambda)$. These can be obtained in principle from the matrix model (2.2.3), or equivalently from the formalism of [24] (appropriately modified as in [172, 201]). However, as emphasized in for example [129, 156, 180], this formalism is not very convenient to do calculations at higher genus. One should rather use the fact that the F_g are quasi-modular forms that can be promoted to non-holomorphic modular forms. The resulting non-holomorphic objects satisfy the holomorphic anomaly equations of [144], as shown in [129, 157], and these can be in turn solved with the technique of direct integration developed in [109, 129, 143, 180] for local CY manifolds and matrix models.

The basic strategy of direct integration is the following. First, we assume an ansatz for F_g of the form

$$F_g(\tau) = \xi^{2g-2} f_g(\tau) \tag{2.12.6}$$

where

$$f_g(\tau) = \sum_{k=0}^{3g-3} E_2^k(\tau) c_k^{(g)}(\tau)\,, \qquad g \geq 2, \tag{2.12.7}$$

is an almost modular form of weight $6g - 6$, with respect to a monodromy group $\Gamma \subset SL(2, \mathbb{Z})$. $F_g(\tau)$ can be promoted to a non-holomorphic modular form $F_g(\tau, \bar{\tau})$ by changing

$$E_2(\tau) \to \widehat{E}_2(\tau, \bar{\tau}) = E_2(\tau) - \frac{3}{\pi\,\mathrm{Im}(\tau)}. \tag{2.12.8}$$

The resulting $F_g(\tau, \bar{\tau})$ satisfies the holomorphic anomaly equations of [144], which govern their anti–holomorphic dependence. Since this dependence is contained in $\widehat{E}_2(\tau, \bar{\tau})$, these equations govern the E_2 content of F_g. This means that the coefficients $c_k^{(g)}(\tau)$, which are modular forms of weight $6g-6-2k$, can be obtained recursively for $k > 0$ if one knows the lower F_g. In order to write down the recursive equation, it is useful to introduce a covariant derivative d_ξ taking a form of weight k into a form of weight $k+2$:

$$\mathrm{d}_\xi = \partial_\tau + \frac{k}{3} \frac{\partial_\tau \xi}{\xi} \tag{2.12.9}$$

Then, the holomorphic anomaly equations lead to

$$\frac{\mathrm{d} f_g}{\mathrm{d} E_2} = -\frac{1}{3} \left\{ \mathrm{d}_\xi^2 f_{g-1} + \frac{1}{3} \frac{\partial_\tau \xi}{\xi} \mathrm{d}_\xi f_{g-1} + \sum_{r=1}^{g-1} \mathrm{d}_\xi f_r \, \mathrm{d}_\xi f_{g-r} \right\}, \qquad g \geq 2. \tag{2.12.10}$$

If $F_{g'}$ are known, with $g' < g$, the above equation determines all the coefficients $c_k^{(g)}(\tau)$ in f_g, with the exception of $c_0^{(g)}(\tau)$, which plays the rôle of an integration constant. This coefficient is a holomorphic form of weight $6g-6$ and it is called the *holomorphic ambiguity*.

In order to fix the holomorphic ambiguity we need two pieces of information. The first one concerns its functional dependence. Since $c_0^{(g)}(\tau)$ is a modular form w.r.t. some monodromy subgroup, it belongs to a finitely generated ring. This means that it is determined by a finite number of coefficients, which typically grows with g. The second piece of information comes from boundary conditions at singular points in moduli space. A very powerful boundary condition for matrix models and local Calabi–Yau manifolds is the so-called *gap condition*, discovered in [142] and further used in [109, 180] to fix the holomorphic ambiguity. According to the gap condition, near certain points p_i in moduli space, parametrized by a flat coordinate t_i, the genus g free energy behaves as

$$F_g^{(i)} = \frac{a_g}{t_i^{2g-2}} + \mathcal{O}(1). \tag{2.12.11}$$

The superscript (i) means that the genus g free energy has to be transformed to the duality frame which is appropriate for the i-th singularity, as it is well-known in special geometry. The “gap" refers to the absence of singular terms t^{-k} with $0 < k < 2g-2$ in the local expansion near $t_i = 0$. The vanishing of these terms provides boundary conditions for $c_0^{(g)}(\tau)$, and in some cases it fixes them completely.

In our case, the relevant ring is that of Γ_2 modular forms which is generated by the theta functions

$$b = \vartheta_2^4(\tau) \; , \quad c = \vartheta_3^4(\tau) \; , \quad d = \vartheta_4^4(\tau). \tag{2.12.12}$$

Since $c = b + d$, only two of them are independent, and we will choose b and d. Using standard formulae in the theory of modular forms, one finds

$$\frac{\partial_\tau \xi}{\xi} = \frac{b - E_2}{4}, \tag{2.12.13}$$

as well as

$$\mathrm{d}_\xi b = \frac{b^2 + bd}{3}, \qquad \mathrm{d}_\xi (bd) = \frac{(bd)b}{6}, \qquad \mathrm{d}_\xi E_2 = \frac{1}{12} \left(-E_2^2 + 2bE_2 - E_4 \right). \tag{2.12.14}$$

The modular expression for the genus one free energy is known [113] and reads

$$F_1 = -\log \eta(\tau), \tag{2.12.15}$$

therefore we have

$$\mathrm{d}_\xi f_1 = -\frac{E_2}{24}. \tag{2.12.16}$$

These are all the ingredients needed for the recursion. The holomorphic ambiguity can be written as

$$c_0^{(g)}(\tau) = \sum_{j=0}^{3g-3} \alpha_j^{(g)} b^j d^{3g-3-j} \tag{2.12.17}$$

and it involves $3g-2$ unknowns. Let us see how we can fix these by looking at the behavior near the three singular points of moduli space.

At the orbifold point, the F_g are the genus g amplitudes of the super-matrix model (2.2.3) with $N_1 = N_2$. Their leading behavior near $\lambda = 0$ is governed by two copies of the Gaussian matrix model, therefore they behave as

$$F_g^{(\mathrm{o})}(\lambda) = \frac{B_{2g}}{g(2g-2)}(2\pi\mathrm{i}\lambda)^{2-2g} + \mathcal{O}(\lambda^2). \tag{2.12.18}$$

This gives g conditions, since the ansatz (2.12.17) for the holomorphic ambiguity only involves even powers of λ.

The symmetric conifold point $z_1 = z_2 = 1/16$ is related to the orbifold point through an S-duality transformation. The appropriate global coordinates near this point are given in (2.3.33). In the ABJM slice one has

$$y_1 = 0, \qquad y_2 = y = 1 - \frac{\zeta^2}{16}. \tag{2.12.19}$$

The following period is a good local, flat coordinate near the symmetric conifold point:

$$t = \sum_{n=0}^{\infty} \frac{a_n}{(n+1)\,2^{4n}} y^{n+1}, \tag{2.12.20}$$

where

$$a_n = \frac{1}{\binom{2n}{n}} \sum_{k=0}^{n} \binom{2k}{k}\binom{4k}{2k}\binom{2n-2k}{n-k}\binom{4n-4k}{2n-2k}. \tag{2.12.21}$$

It was noticed in [109] that the genus g amplitude at the conifold point behaves like

$$F_g^{(\mathrm{c})}(t) = \frac{B_{2g}}{2g(2g-2)}\left(\frac{t}{2\mathrm{i}}\right)^{2-2g} + \mathcal{O}(1). \tag{2.12.22}$$

This fixes $2g-2$ conditions. Together with the g conditions coming from the orbifold point, this completely fixes the $3g-2$ unknowns in the holomorphic ambiguity.

The result can be verified by looking at the radius point, which is related to the orbifold point by an STS transformation. The genus g free energy at this point is the generating function of Gromov–Witten invariants of the local $\mathbb{F}_0$ geometry in the slice $T_1 = T_2 = T$. More precisely, one has

$$F_g^{(\mathrm{GW})}(Q) = (-4)^{g-1}\left\{\frac{(-1)^g |B_{2g}B_{2g-2}|}{g(2g-2)(2g-2)!} + \sum_{d\geq 1} N_{d,g} Q^d\right\}, \qquad Q = \mathrm{e}^{-T} \tag{2.12.23}$$

where

$$N_{d,g} = \sum_{d_1+d_2=d} N_{d_1,d_2,g} \tag{2.12.24}$$

is a sum of Gromov–Witten invariants at genus g, $N_{d_1,d_2,g}$, of local $\mathbb{F}_0$ (the degrees d_1, d_2 correspond to the two Kähler classes of this geometry). The constant term in (2.12.23) is the well-known constant map contribution to the higher genus free energy [144] for a manifold with “effective" Euler characteristic $\chi = 4$. It can be checked that the higher genus free energies obtained from the integration of the holomorphic anomaly equation with the above boundary conditions reproduce the well-known large radius free energies (2.12.23).

Let us see how this works in some detail when $g = 2$. The integration of the holomorphic anomaly equation gives,

$$f_2 = \frac{1}{3}\cdot\frac{1}{24^2}\left(-\frac{5}{3}E_2^3 + 3bE_2^2 - 2E_4E_2\right) + c_0^{(2)}(\tau), \tag{2.12.25}$$

where $c_0^{(2)}(\tau)$ is of the form (2.12.17). The expansion around the orbifold and conifold points read, respectively,

$$\begin{aligned} F_2^{(\mathrm{o})}(\lambda) &= \frac{1}{432(2\pi\mathrm{i}\lambda)^2}\left(-\frac{11}{3} + 1728\alpha_0^{(2)}\right) + 4\left(\frac{2}{3}\left(\alpha_0^{(2)} - \frac{11}{5184}\right) - \frac{3\alpha_0^{(2)}}{2} + \alpha_1^{(2)} + \frac{1}{576}\right) \\ &\quad + \mathcal{O}(\lambda^2), \\ F_2^{(\mathrm{c})}(t) &= -\frac{5 + 1296\alpha_3^{(2)}}{1296t^2} + \frac{-1 - 864(12\alpha_2^{(2)} + 15\alpha_3^{(2)})}{10368t} + \mathcal{O}(1), \end{aligned} \tag{2.12.26}$$

Imposing the conditions (2.12.18), (2.12.22) and (2.12.23) we fix

$$\alpha_0^{(2)} = \frac{1}{25920}, \quad \alpha_1^{(2)} = -\frac{1}{3456}, \quad \alpha_2^{(2)} = \frac{1}{3456}, \quad \alpha_3^{(2)} = \frac{1}{3240}. \tag{2.12.27}$$

We finally obtain

$$F_2^{(\mathrm{o})} = \frac{1}{432bd^2}\left(-\frac{5}{3}E_2^3 + 3bE_2^2 - 2E_4E_2\right) + \frac{16b^3 + 15db^2 - 15d^2b + 2d^3}{12960bd^2}, \tag{2.12.28}$$

which gives at large radius the expansion Since τ depends on λ through (2.12.3) and (2.5.9), this gives the exact expression for the genus two free energy on $\mathbb{S}^3$ in the ABJM model, for any value of the 't Hooft coupling.

Notice that the modular ring appearing here and parametrizing the holomorphic ambiguity is different from the one appearing in Seiberg–Witten theory [129, 143] or in the cubic matrix model [180]. This is due to the fact that, although the curves are the same, the meromorphic forms defining the theory are different.

Using this method, we have computed the free energies up to high genus. The genus g free energies $F_g(\lambda) = F_2^{(\mathrm{o})}(\lambda)$ obtained in this way are exact interpolating functions of the 't Hooft parameter, and they can be studied in various regimes. In the strong coupling regime $\lambda \to \infty$ it is more convenient to use the shifted variable

$$\hat{\lambda} = \lambda - \frac{1}{24} = \frac{\log^2 \kappa}{2\pi^2} + \mathcal{O}(\kappa^{-2}), \qquad \kappa \gg 1. \tag{2.12.29}$$

One finds the following structure. For F_0 and F_1 one has, at strong coupling,

$$\begin{aligned} F_0 &= \frac{4\pi^3\sqrt{2}}{3}\hat{\lambda}^{3/2} + \mathcal{O}\left(\mathrm{e}^{-2\pi\sqrt{2\hat{\lambda}}}\right), \\ F_1 &= \frac{1}{6}\log\kappa - \frac{1}{2}\log\left[\frac{2\log\kappa}{\pi}\right] + \mathcal{O}\left(\frac{1}{\kappa^2}\right), \end{aligned} \tag{2.12.30}$$

while for $g \geq 2$ one has

$$F_g = f_g\left(\frac{1}{\log\,\kappa}\right) + \mathcal{O}\left(\frac{1}{\kappa^2}\right), \tag{2.12.31}$$

where

$$f_g(x) = \sum_{j=0}^{g} c_j^{(g)} x^{2g-3+j} \tag{2.12.32}$$

is a polynomial. One finds, for the very first genera,

$$\begin{aligned} f_2(x) &= \frac{15x^3 - 6x^2 + x}{144}, \\ f_3(x) &= \frac{405x^6 - 135x^5 + 18x^4 - x^3}{5184}, \\ f_4(x) &= \frac{9945x^9 - 3240x^8 + 450x^7 - 32x^6 + x^5}{82944}, \\ f_5(x) &= \frac{274590x^{12} - 89505x^{11} + 12960x^{10} - 1050x^9 + 48x^8 - x^7}{995328}, \end{aligned} \tag{2.12.33}$$

and the leading, strong coupling behavior is given by

$$F_g(\lambda) \sim \lambda^{\frac{3}{2}-g}, \qquad \lambda \to \infty, \quad g \geq 0. \tag{2.12.34}$$

2.12.2 Expansion in the type IIA and M-theory duals

It is possible to translate the all-genus expansion of the matrix model into expansions in the type IIA and the M-theory duals. In the type IIA dual, the genus expansion of the matrix model becomes the genus expansion of superstring theory. In M-theory, the genus expansion becomes an expansion in the Planck length (or, equivalently, in Newton's constant). In order to translate the matrix model results in a string/M-theory result we need a precise dictionary relating gauge theory quantities to gravity quantities. In particular, one has to take into account the anomalous shifts relating the rank of the gauge group N

to the Maxwell charge Q, which in turn determines the compactification radius L [77, 78]. The relation is

$$Q = N - \frac{1}{24}\left(k - \frac{1}{k}\right). \tag{2.12.35}$$

The charge Q determines the compactification radius in M-theory according to

$$\left(\frac{L}{\ell_p}\right)^6 = 32\pi^2 Qk. \tag{2.12.36}$$

This means that the shifted variable $\hat{\lambda}$ introduced in (2.12.29) is given, in M-theory variables, by

$$\hat{\lambda} = \frac{1}{32\pi^2 k^2}\left(\frac{L}{\ell_p}\right)^6 \left(1 - \frac{4\pi^2}{3}\left(\frac{\ell_p}{L}\right)^6\right). \tag{2.12.37}$$

When considering the type IIA expansion, we have to trade k for the string coupling constant $g_{\rm st}$, and the Planck length by the string lenght ℓ_s. In the end we find

$$\begin{aligned} k^2 &= g_{\rm st}^{-2}\left(\frac{L}{\ell_s}\right)^2, \\ \hat{\lambda} &= \frac{1}{32\pi^2}\left(\frac{L}{\ell_s}\right)^4 \left(1 - \frac{4\pi^2 g_{\rm st}^2}{3}\left(\frac{\ell_s}{L}\right)^6\right). \end{aligned} \tag{2.12.38}$$

The exponentially small corrections (2.12.31) should correspond, in the type IIA superstring, to worldsheet instantons wrapping the $\mathbb{CP}^1$ inside $\mathbb{CP}^3$, and in M-theory to membrane instantons [120] wrapping the $\mathbb{S}^3 \subset \mathbb{S}^7$. In the following we will drop these nonperturbative corrections, although they can be of course computed to any given order in the exponentiated worldsheet instanton/membrane action.

Let us first write down the type IIA superstring expansion. Using the dictionary (2.12.38) we find

$$F = -\frac{g_{\rm st}^{-2}}{384\pi^2}\left(\frac{L}{\ell_s}\right)^8 + \frac{3}{64}\left(\frac{L}{\ell_s}\right)^2 + \frac{1}{2}\log\left[2\pi\left(\frac{\ell_s}{L}\right)^2\right] + \sum_{g=2}^{\infty} r_g\left(\left(\frac{\ell_s}{L}\right)^2\right) g_{\rm st}^{2g-2} \tag{2.12.39}$$

where

$$r_g(x) = \sum_{k=3g-4}^{4(g-1)} r_{g,k} x^k, \qquad g \geq 2, \tag{2.12.40}$$

is a polynomial. One finds, for the very first genera,

$$\begin{aligned} r_2(x) &= -\frac{1}{192}\pi^2 x^4 \left(5120x^4 - 576x^2 + 27\right), \\ r_3(x) &= \frac{1}{32}\pi^4 x^{10}\left(163840x^6 - 15360x^4 + 576x^2 - 9\right), \\ r_4(x) &= -\frac{1}{576}\pi^6 x^{16}\left(1158676480x^8 - 106168320x^6 + 4147200x^4 - 82944x^2 + 729\right), \\ r_5(x) &= \frac{1}{32}\pi^8 x^{22}\big(37916508160x^{10} - 3476029440x^8 + 141557760x^6 \\ &\quad - 3225600x^4 + 41472x^2 - 243\big). \end{aligned} \tag{2.12.41}$$

Notice that [181] predicts, for general Sasaki-Einstein manifolds in M-theory, a correction for F_1 scaling as $\lambda^{1/2}$, like the second term in (2.12.39). It would be interesting to see if the precise numerical coefficient also agrees with theirs.

We can now work out the M-theory expansion. If we use again the dictionary (2.12.37), we see that the M-theory free energy on $\mathrm{AdS}_4 \times \mathbb{S}^7/\mathbb{Z}_k$ has the structure

$$F = -\frac{1}{384\pi^2 k}\left(\frac{L}{\ell_p}\right)^9 + \frac{3}{64k}\left(\frac{L}{\ell_p}\right)^3 + \frac{1}{2}\log\left[2\pi k\left(\frac{\ell_p}{L}\right)^3\right] + \frac{1}{k}\sum_{n=1}^{\infty} p_n(k)\left(\frac{\ell_p}{L}\right)^{3n}, \tag{2.12.42}$$

where $p_n(k)$ is a *polynomial* in k of degree at most $[(n+3)/3]$. At each order n only a finite number of terms in the original genus expansion contribute, and the maximal genus contributing is

$$g = \left[\frac{n+3}{2}\right]. \tag{2.12.43}$$

The polynomials $p_n(k)$ are given, for the first few orders, by

$$\begin{aligned} p_1(k) &= -\frac{9\pi^2}{64}, \\ p_2(k) &= 3\pi^2, \\ p_3(k) &= -\frac{80\pi^2}{3}k^2 - \frac{9\pi^4}{32}. \end{aligned} \tag{2.12.44}$$

Since each coefficient in the series (2.12.42) is a polynomial in k, one can compute from the genus expansion in the matrix model the free energy of M-theory in the large radius expansion, at a given order in $(\ell_p/L)^3$, and for *any* value of k.

It turns out that the expansion (2.12.42) has a remarkable hidden structure. As we see, the natural parameter in the power series is

$$\left(\frac{\ell_p}{L}\right)^3 \tag{2.12.45}$$

as expected in a generic M-theory expansion. However, if we introduce the following "renormalized" parameter

$$\frac{\widehat{\ell}_p}{L} = \frac{\ell_p/L}{\left[1 - 12\pi^2\left(\ell_p/L\right)^6\right]^{1/6}}, \tag{2.12.46}$$

it turns out that the expansion can be resummed in the following way,

$$F = -\frac{1}{384\pi^2 k}\left(\frac{L}{\widehat{\ell}_p}\right)^9 + \frac{1}{6}\log\left[8\pi^3 k^3 \left(\frac{\widehat{\ell}_p}{L}\right)^9\right] + \sum_{n=1}^{\infty} d_{n+1}\pi^{2n}k^n \left(\frac{\widehat{\ell}_p}{L}\right)^{9n}, \tag{2.12.47}$$

where the coefficients d_n are just rational numbers:

$$d_2 = -\frac{80}{3}, \qquad d_3 = 5120, \qquad d_4 = -\frac{18104320}{9}, \qquad d_5 = 1184890880, \qquad \cdots \tag{2.12.48}$$

This resummation is based on a highly non-trivial property of the polynomials (2.12.33) which is not at all manifest from their matrix model origin, and is begging for an interpretation in the context of M-theory/string theory. A similar simplification can be obtained in the type IIA expansion by introducing a "renormalized" parameter ℓ_s/L, which depends also on $g_{\rm st}$.

What is the interpretation of the M-theory expansion (2.12.42) and its resummation (2.12.47)? In other M-theory expansions (like the two-graviton potential in M(atrix) theory), the terms which go like $(\ell_p/L)^9$ are interpreted like classical supergravity interactions, since they correspond to integral powers of the eleven-dimensional Newton's coupling constant. The other terms, with powers which are not multiples of 9, are usually interpreted as "quantum gravity" corrections (see for example the discussion in [182], IV.A.5). The resummation (2.12.47) suggest that in this case these quantum gravity corrections can be rewritten in terms of a classical expansion, but involving the "renormalized" coupling (2.12.46).

2.13 Instantons and the genus expansion

2.13.1 Instantons in matrix models

In this subsection we review some results on instantons in matrix models, following the work of [178, 179, 180], which contains much more details and references.

The study of instantons in matrix models has been pursued in many works, starting with the pioneering papers of David [183]. An important insight, first developed in relation to matrix models of two-dimensional gravity, is that instantons are obtained by *eigenvalue tunneling*. In order to make the discussion simpler, let us consider the cubic matrix model, where the effective potential has two critical points. In the one-cut phase of this model, all eigenvalues are located in the neighborhood of one critical point. The ℓ-instanton configuration in this phase is simply obtained by removing ℓ eigenvalues from this cut and tunneling them to the other critical point, as shown in Fig. 2.13. The instanton action for the one-cut phase admits a beautiful geometric interpretation in terms of the spectral curve $y(x)$ describing the planar limit of the matrix model, and it is given by the integral

$$A_B = \int_a^{x_0} y(x)\mathrm{d}x, \tag{2.13.1}$$

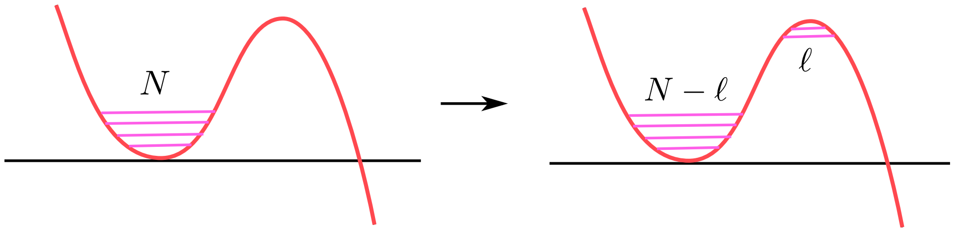

Figure 2.13: An ℓ-instanton can be obtained by tunneling ℓ eigenvalues from one critical point to another one.

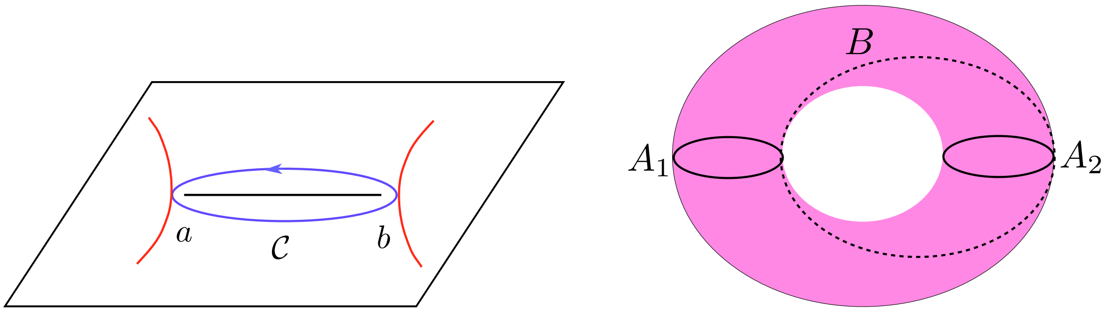

Figure 2.14: The left hand side shows the spectral curve in the one-cut phase of the cubic matrix model. The instanton action relevant in the double-scaling limit is obtained by calculating the B-period of the one-form $y(x)\mathrm{d}x$, which goes from the filled cut A_1 to the pinched point. The two-cut phase, in which the pinched point becomes a filled interval, is shown on the right hand side. The instanton action is still given by the B-period integral.

where a is the endpoint of the filled cut, and x_0 is the location of the critical point which corresponds to an empty cut (x_0 is actually a singular point where the spectral curve has a nodal singularity). More geometrically, we can write this as a period integral of the natural meromorphic form $y(x)\mathrm{d}x$, corresponding to a B cycle which goes from the filled cut to the pinched point [184, 185]:

$$A_B = \frac{1}{2}\oint_B y(x)\mathrm{d}x. \tag{2.13.2}$$

In Fig. 2.13.1 (left) we show the pinched curve, where the A_1 cycle corresponds to the filled cut, and the B cycle goes from A_1 to the pinched cycle. This picture extends to one-cut matrix models with generic polynomial potentials: instantons are given by eigenvalue tunneling, and their actions are B-type period, going from the filled cut to other critical points.

Based on the connection between instantons and the large order behavior of perturbation theory [186], we should expect these instantons to control the behavior of the genus g amplitudes F_g of one-cut matrix models at large g. Indeed, one can verify in examples [178] that

$$F_g(t) \sim (2g)!(A(t))^{-2g}, \qquad g \gg 1, \tag{2.13.3}$$

where $A(t)$ is a period of $y(x)\mathrm{d}x$. Notice that (2.13.3) is just the leading behavior of the full asymptotics at large g, which involves a series of corrections in $1/g$ (see for example [178] for more details). The relevant period $A(t)$ appearing in the leading asymptotics (2.13.3) depends on the value of t. For small t, the behavior of the free energy is dominated by its Gaussian part,

$$F_g(t) \approx \frac{B_{2g}}{2g(2g-2)t^{2g}}, \qquad t \to 0, \tag{2.13.4}$$

and the action $A(t)$ in (2.13.3) is in fact the A_1-period going around the filled cut, which is just proportional to the 't Hooft parameter:

$$A_{A_1}(t) = 2\pi\mathrm{i}t = \frac{1}{2}\oint_{A_1} y(x)\mathrm{d}x. \tag{2.13.5}$$

Notice that this period vanishes at the origin $t = 0$. In other regions of the t-plane, the large genus behavior will be controlled by B-periods $A_B(t)$ of the form (2.13.1). In general, the action controlling the large order behavior at a given point t is the smallest period (in absolute value). Notice that the B-type periods $A_B(t)$ vanish at critical values of the 't Hooft parameter and the other couplings, so in both cases the instanton action is given by a vanishing cycle in moduli space. An equivalent way of formulating the rôle of instantons is that their actions give the location of the singularities for the Borel transform of the asymptotic series (0.2.2).

This result can be generalized to the two-cut phase of the cubic matrix model, where the pinched point is now resolved into a second cut A_2. The instanton action is still given by the B-cycle integral, now going from the first cycle A_1 to the second cycle A_2, and it controls the large order behavior of $F_g(t_1, t_2)$ in the appropriate regions of moduli space [180]. The above analysis of instanton configurations seems to apply to matrix models with generic polynomial potentials. However, there are important matrix models, like the Chern–Simons matrix model [85], which display a more subtle structure. It was found in [187], for example, that due to the multivaluedness of the effective potential, the instanton actions in the Chern–Simons matrix model are given by

$$2\pi \mathrm{i}\,(t + 2\pi \mathrm{i} n)\,, \qquad n \in \mathbb{Z}. \tag{2.13.6}$$

For $n = 0$ one recovers the action governing the Gaussian behavior. The instantons with $n = \pm 1$ can be detected through the large order behavior of the genus g free energies, once the Gaussian part is subtracted.

It is then natural to ask if there is a common structure describing the instantons of general matrix models. All the models we have in mind are characterized by the fact that their planar limit is described by special geometry on a local Calabi–Yau manifold, and it is then desirable to describe their instantons in that language as well. This is precisely what we will do now.

2.13.2 Instantons and special geometry

We will suppose that we are given a local Calabi–Yau manifold, whose geometry is encoded in a spectral curve $y(x)$. This curve can be an algebraic curve in $\mathbb{C} \times \mathbb{C}$, like the curves arising in polynomial matrix models, or a curve in $\mathbb{C}^* \times \mathbb{C}^*$, like the ones arising in Chern–Simons matrix models and in the mirrors of toric Calabi–Yau threefolds. We will denote by $\mathcal{M}$ the moduli space associated to the geometry described by $y(x)$. In order to write down the genus g amplitudes F_g, one has to choose first a symplectic frame. In order to make this choice manifest we will write $F_g^{(f)}$, where f specifies the choice of frame. The different $F_g^{(f)}$ are related by symplectic transformations and they transform as quasi-modular forms [113].

Usually, $\mathcal{M}$ has special points corresponding to physical singularities, and near each of these points there are preferred frames. A famous example is Seiberg–Witten theory [10], where $\mathcal{M}$ is the moduli space of the Seiberg–Witten elliptic curve and there are three special points corresponding to the semiclassical regime, the monopole point and the dyon point. The corresponding frames are usually called electric, magnetic and dyonic, respectively. In the most relevant example for this paper, the ABJM matrix model, the moduli space $\mathcal{M}$ has three critical points usually called large radius, orbifold and conifold points, so there will be three preferred frames.

Our main proposal, based on the results reviewed above, is that instanton actions are always given by complex linear combinations of the periods of special geometry. More precisely, we propose

$$A^{(f)}\left(t_i; a_i, b_i, c\right) = \sum_{i=1}^{s} \left(a_i t_i^{(f)} + b_i \frac{\partial F_0^{(f)}}{\partial t_i} \right) + c, \tag{2.13.7}$$

where $s = \dim(\mathcal{M})$ and a_i, b_i, c are complex numbers. The first term is the sum gives the contribution of the A-cyles, while the second term gives the contribution of the B-cycles. Notice that this is also the structure of central charges in special geometry. In particular, we propose that the large genus behavior of the $F_g^{(f)}$ at generic points of the moduli space is governed by an instanton action of this form. A particular rôle is played by the instanton actions which govern the large order behavior of the $F_g^{(f)}$ near the singular points of moduli space. We will denote these actions by

$$A_p^{(f)}(t_i) \tag{2.13.8}$$

where p labels the singular points in $\mathcal{M}$. According to our proposal, these actions are given by (2.13.7), for a specific choice of the constants a_i, b_i, c which depends on the point p.

Of course, the main problem is to determine the values of the constants a_i, b_i, c which describe the possible instantons in the problem at hand. Unfortunately we don't have a general principle to do this.

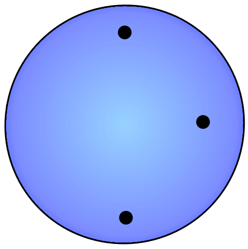

Figure 2.15: The moduli space of the ABJM theory has three special points.

However, when the singular point p is in the interior of $\mathcal{M}$ (i.e. for singular points of the conifold or orbifold type), we expect the $A_g^{(f)}$ to be given by *vanishing periods*. One way to motivate this is to notice that, near the singular points, the $F_g^{(f)}(t_i)$ diverge for all g. The instanton action controlling their large order behavior should then vanish at those points. The identification of vanishing periods makes it possible to fix the constants a_i, b_i and c in many situations and leads to a determination of the large genus behavior near orbifold and conifold points. At generic points there will be a competition between the different instanton actions, and the dominant contribution to the large order behavior will be given by the instanton action which is smaller in absolute value (or, equivalently, by the instanton action which is closest to the origin in the Borel plane).

Of course, the proposal above recovers and generalizes the known description of instantons in matrix models, where the instanton actions are given by A or B periods, as we have already discussed. In the remaining of this section we will analyze in detail the ABJM matrix model following these general principles, and we will present ample evidence that in this model the relevant instanton actions describing the large genus behavior are indeed of the form (2.13.7).

2.13.3 Instantons in the ABJM model

The moduli space $\mathcal{M}$ of the ABJM matrix model was studied in detail in section 2.3, and it is shown schematically in Fig. 2.13.3. It can be parametrized by λ (which is a period), or equivalently by the global modulus κ. Notice that, although in the original ABJM theory λ is a rational number, in the planar solution it is naturally promoted to a complex variable, and $\mathcal{M}$ will be regarded here as a complex one-dimensional space. There are three singular points in this moduli space: the orbifold, large radius and conifold points. The first two points correspond respectively to the weak coupling limit and the strong coupling limits of the ABJM theory. The conifold point, which occurs for $\kappa = -4\mathrm{i}$, or equivalently, at

$$\lambda_c = -\frac{2\mathrm{i}K}{\pi^2}, \tag{2.13.9}$$

where K is Catalan's constant, has no obvious interpretation in the gauge theory (although we will comment on this later on).

The frame in which the genus g free energies F_g give the $1/N$ expansion of the matrix model is the orbifold or weak coupling frame, as first discovered in [88]. We will study this frame first, and we will determine the relevant instanton actions near the singular points. In this frame, which we will denote by w, the appropriate period coordinate is λ, and the orbifold singularity occurs at $\lambda = 0$. Near this singularity the relevant instanton action is simply

$$A_w^{(w)}(\kappa) = -4\pi^2 \lambda(\kappa). \tag{2.13.10}$$

Since the 't Hooft coupling in the matrix model is $t = 2\pi\mathrm{i}\lambda$, this action is the standard one (2.13.5) controlling the Gaussian point. It vanishes of course at $\lambda = 0$. On the other hand, it is easy to find the vanishing period at the conifold point, and this leads to the instanton action

$$A_c^{(w)}(\kappa) = \frac{\mathrm{i}}{\pi}\frac{\partial F_0^{(w)}}{\partial \lambda} + 4\pi^2\lambda - \pi^2 = \frac{\mathrm{i}\kappa}{4\pi} G^{2,3}_{3,3}\left(\begin{matrix} \frac{1}{2}, & \frac{1}{2}, & \frac{1}{2} \\ 0, & 0, & -\frac{1}{2} \end{matrix} \middle| -\frac{\kappa^2}{16}\right) - \pi^2. \tag{2.13.11}$$

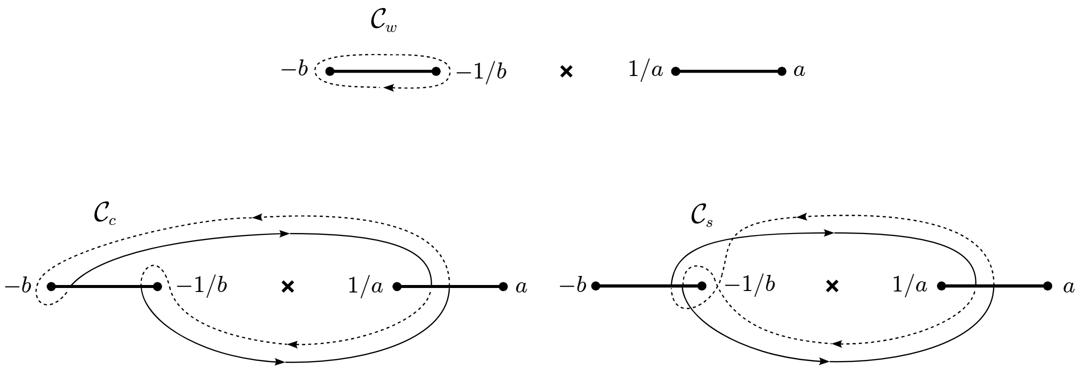

Figure 2.16: The topology of the contours $\mathcal{C}_w$, $\mathcal{C}_c$, $\mathcal{C}_s$ for a vicinity of the orbifold point in the moduli space.

Finally, we have to consider the large radius, or strong coupling, point. The relevant action turns out to be

$$\begin{aligned} A_s^{(w)}(\kappa) =& \frac{\mathrm{i}}{\pi}\frac{\partial F_0^{(w)}}{\partial \lambda} - \pi^2 \\ =& \frac{\mathrm{i}\kappa}{4\pi} G^{2,3}_{3,3}\left(\begin{array}{ccc} \frac{1}{2}, & \frac{1}{2}, & \frac{1}{2} \\ 0, & 0, & -\frac{1}{2} \end{array} \middle| -\frac{\kappa^2}{16}\right) - \frac{\pi\kappa}{2}\, {}_3F_2\left(\frac{1}{2},\frac{1}{2},\frac{1}{2};1,\frac{3}{2};-\frac{\kappa^2}{16}\right) - \pi^2. \end{aligned} \tag{2.13.12}$$

For this action, the coefficients appearing in (2.13.7) cannot be determined by requiring it to be a vanishing period, but it has a simple structure, since it is just given by

$$A_s^{(w)}(\kappa) = A_w^{(w)}(\kappa) + A_c^{(w)}(\kappa). \tag{2.13.13}$$

One can verify numerically that it is the right action in the sense that it controls the large order behavior of $F_g^{(w)}$ in the region where it dominates the asymptotics, as we will show in the next subsection. It is tempting to conjecture that all instanton actions appearing in the theory are just integer linear combinations of $A_w^{(w)}(\kappa)$ and $A_c^{(w)}(\kappa)$. This is in fact what we would expect if these instantons could be identified with Euclidean D-branes of the string dual, as we will argue later.

The actions (2.13.10), (2.13.11) and (2.13.12) can be written as period integrals on the spectral curve of the ABJM matrix model. In terms of the variables

$$Y = \mathrm{e}^y, \; X = \mathrm{e}^x \tag{2.13.14}$$

the spectral curve is given by the equation [5, 6, 88, 100]

$$Y + \frac{X^2}{Y} - X^2 + \mathrm{i}\kappa\, X - 1 = 0\,. \tag{2.13.15}$$

The Riemann surface of (2.13.15) can be represented by two X-planes glued along the cuts $[1/a, a]$ and $[-b, -1/b]$. The position of the endpoints can be determined from

$$a + \frac{1}{a} + b + \frac{1}{b} = 4, \qquad a + \frac{1}{a} - b - \frac{1}{b} = 2\mathrm{i}\kappa\,. \tag{2.13.16}$$

Let us note that a, $b \to 1$ at the orbifold (weak coupling) point, and that y has a logarithmic singularity at the origin (and at infinity) on one of the two X-sheets. The actions describing the large g behavior can be represented as

$$A_p = \frac{1}{2}\oint_{\mathcal{C}_p} y(x)\mathrm{d}x \tag{2.13.17}$$

where the contours $\mathcal{C}_p$ are depicted in Fig. 2.16.

As a last remark, notice that these actions appear in pairs $A_p^{(w)}$, $-A_p^{(w)}$, and this is reflected in the fact that the genus expansion that they govern involves only even powers of the string coupling constant. Equivalently, there are singularities in the Borel plane of the g_s coupling constant at the points $\pm A_p^{(w)}$.

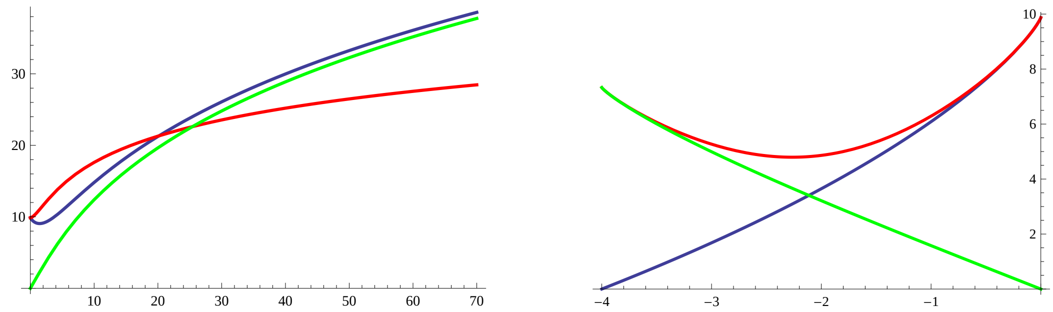

Figure 2.17: In this figure we depict the absolute value of the three instanton actions in the orbifold or weakly coupled frame. On the left side, the horizontal axis represents the positive real axis of the κ variable. The curve in green, which vanishes at the origin, is $|A_w^{(w)}(\kappa)|$, while the blue and red lines represent $|A_c^{(w)}(\kappa)|$ and $|A_s^{(w)}(\kappa)|$, respectively. Notice that, when κ is large (i.e. the strong coupling region), the smallest action in absolute value is $A_s^{(w)}(\kappa)$. On the right side, the horizontal axis represents the imaginary axis of the κ variable. The conifold action $A_c^{(w)}(\kappa)$ vanishes at $\kappa_c = -4\mathrm{i}$, and therefore dominates the large order behavior near that point.

This is also the case in simpler cases related to noncritical string theory, like the Painlevé I equation (see for example [188]).

As explained above, at each point in moduli space we expect the large order behavior to be dominated by the smallest action in absolute value. In Fig. 2.17 we show the absolute value of the instanton actions in the weakly coupled frame, and along two different directions in the complex moduli space parametrized by κ: the real axis (left) and the imaginary axis (right). For real $\kappa \gg 1$ (which corresponds to the strong coupling regime $\lambda \gg 1$), the smallest instanton action is $A_s^{(w)}$, while near the origin the smallest action is $A_w^{(w)}$. For λ imaginary and near the conifold point λ_c, the smallest instanton action is clearly $A_c^{(w)}$. For generic points in the moduli space there is a competition between the different actions. For example, for imaginary κ, there is a point κ_* where

$$\left|A_c^{(w)}(\kappa_*)\right| = \left|A_w^{(w)}(\kappa_*)\right|. \tag{2.13.18}$$

This is the point where the two lines cross in the graphic on the right side of Fig. 2.17. For $|\kappa| > |\kappa_*|$ we should expect the large order behavior to be controlled by the conifold action $A_c^{(w)}$, while for $|\kappa| < |\kappa_*|$ it should be controlled by the weak coupling action $A_w^{(w)}$. We will present explicit checks of these expectations in a moment.

So far we have made the analysis in the weakly coupled frame, but we can do the analysis in the other preferred frames. It turns out that the relevant instanton actions near the singular points are just given by the analytic continuations of the instanton actions in other frames. This is not surprising, since for example vanishing periods near a singular point are uniquely defined, independently of the frame. This in particular means that the large genus behavior of the $F_g^{(f)}$ in different frames will be governed by the same instanton action.

Let us consider for example the conifold frame. An appropriate flat coordinate in this frame is given by [6, 109]

$$\lambda^{(c)} = \frac{1}{4\pi}\sum_{n=0}^{\infty}\frac{a_n}{(n+1)\,2^{4n}}y^{n+1}, \tag{2.13.19}$$

where

$$y = 1 + \frac{\kappa^2}{16}, \qquad a_n = \frac{1}{\binom{2n}{n}}\sum_{k=0}^{n}\binom{2k}{k}\binom{4k}{2k}\binom{2n-2k}{n-k}\binom{4n-4k}{2n-2k}. \tag{2.13.20}$$

This coordinate vanishes at the conifold point $y = 0$. The conifold free energies near this point behave as [109]

$$F_g^{(c)} \sim \frac{B_{2g}}{2g(2g-2)}\left(2\pi\mathrm{i}\lambda^{(c)}\right)^{2-2g}, \qquad g \geq 2, \tag{2.13.21}$$

and we would expect the appropriate instanton action in this frame to be

$$A_c^{(c)}(\kappa) = -4\pi^2 \lambda^{(c)}. \tag{2.13.22}$$

Indeed, one can verify that this is just the analytic continuation of (2.13.11) to $\kappa = -4\mathrm{i}$. Similar considerations apply to the other instanton actions in the conifold and strong coupling frame.

2.13.4 Large order behavior

We now provide some numerical evidence that the actions we have found control indeed the large order behavior of the genus expansion. We will only consider the behavior in the weak coupling frame, but similar considerations and tests can be made for the other frames. For simplicity of notation, in this subsection we will remove the superscript (w) in our expressions. Our numerical analysis is done for the original sequence F_g coming from the matrix model. It can be easily shown that the redefinition of the F_gs which occurs when we use the type IIA parameters, as explained in (2.12.38), does not change the leading asymptotics (2.13.3), and it only affects the subleading $1/g$ corrections.

Generically, the instanton actions we have found are complex, and we will write them as

$$A_p(\lambda) = |A_p(\lambda)|\, \mathrm{e}^{\mathrm{i}\theta_p(\lambda)}. \tag{2.13.23}$$

If the genus g amplitudes are real (as it happens for example for λ and k real), complex instantons governing the large order behavior must appear in complex conjugate pairs (this was pointed out already in [177], in ordinary quantum mechanics). This means that, for real λ, the large order behavior of $F_g(\lambda)$ must be given by

$$\begin{aligned} F_g(\lambda) &\sim \Gamma(2g-1)\left\{ C_p(\lambda)\,(A_p(\lambda))^{-2g} + \overline{C}_p(\lambda)\left(\overline{A}_p(\lambda)\right)^{-2g} \right\} \\ &\sim \Gamma(2g-1)\,|A_p(\lambda)|^{-2g} \cos\left(2g\theta_p(\lambda) + \delta_p(\lambda)\right), \qquad g \gg 1. \end{aligned} \tag{2.13.24}$$

In this equation, $C_p(\lambda)$ is the next correction to the asymptotics, which in some simple matrix models can be obtained by a one-loop calculation in the background of an instanton [178, 183], and

$$\delta_p(\lambda) = \arg\left(C_p(\lambda)\right). \tag{2.13.25}$$

The choice of instanton action here depends on the value of λ, as explained above. If both $(A_p(\lambda))^2$ and the $F_g(\lambda)$ are real, the large genus behavior is given simply by

$$F_g(\lambda) \sim \Gamma(2g-1)\,(A_p(\lambda))^{-2g}, \qquad g \gg 1. \tag{2.13.26}$$

This is what happens for example for λ imaginary and negative, near the conifold point λ_c.

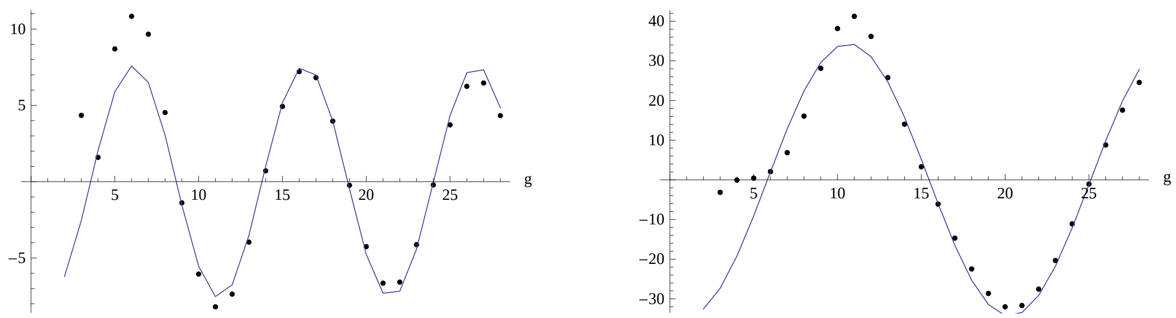

Figure 2.18: In these figures, the dots represent the sequence (2.13.27) for values of λ in the strong coupling region: $\lambda \approx 1.2838$ (left) and $\lambda \approx 4.6687$ (right). The action is then $A_s(\lambda)$, given in (2.13.12). The continuous line represents the oscillatory behavior in the r.h.s. of (2.13.28), where the angle is the one associated to the strong coupling action $\theta_s(\lambda)$.

When the instanton action is complex, the asymptotics is much harder to study numerically, since the standard techniques of acceleration of convergence (like Richardson extrapolation) do not apply to the oscillatory behavior (2.13.24), and in addition the phase $\delta_p(\lambda)$ is not known. In these cases the sequence

$$R_g^p = (-1)^{g+1} \frac{\pi F_g}{\Gamma(2g-1)\,|A_p(\lambda)|^{-2g+1}} \tag{2.13.27}$$

should behave as

$$R^p_g \sim \cos\left(2g\theta_p(\lambda) + g\pi + \delta_p(\lambda)\right), \tag{2.13.28}$$

i.e. it should lead to an oscillatory behavior in g, with (unknown) constant amplitude but with a known frequency given by $\theta_p(\lambda)$. The factor $(-1)^{g+1}$ in (2.13.27) has been introduced for convenience, in view of the forthcoming discussion on Borel summability, and it leads to the shift by $g\pi$ in (2.13.28).

When the action is real, we can actually extract the value of the instanton action from the sequence

$$Q^p_g = \frac{4g^2 F_g(\lambda)}{F_{g+1}(\lambda)\left(A_p(\lambda)\right)^2} \tag{2.13.29}$$

which as $g \to \infty$ should asymptote 1. Standard acceleration methods like Richardson extrapolation can be used to test this behavior to high precision, as in [178].

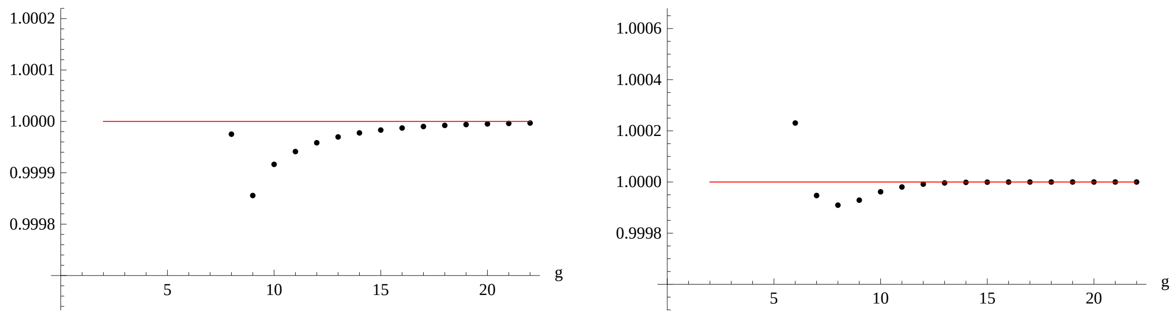

Figure 2.19: In these figures, the dots represent the fifth Richardson transform of the sequence (2.13.29) for values of λ along the negative imaginary axis $\lambda \approx -0.1386\,\mathrm{i}$ (left) and $\lambda \approx -0.0620\,\mathrm{i}$ (right), and for the conifold and the weak coupling action, respectively. They converge quite rapidly to unity, verifying in this way that the proposed instanton actions control the large order behavior of the genus g free energies.

We now present two tests of the large order behavior of the genus g free energies F_g, as predicted by the instanton analysis of the previous subsection.

For λ real and large, we expect the large order behavior to be controlled by the action (2.13.12). This action is complex, and should lead to an oscillatory behavior in F_g. We can then compare the sequence R^s_g, for $g = 2, \cdots, 28$, as computed numerically in (2.13.27), to the expected behavior (2.13.28). This is done in Fig. 2.18 for two values of λ in the strong coupling region. The agreement is rather good. In order to plot the continuous line in these figures, we have taken $\delta_s(\lambda) = -2\theta_s(\lambda)$, which leads to a good matching.

When λ (or κ) is on the negative imaginary axis, the relevant instanton actions are the conifold and the weak coupling actions, as shown in the figure on the right in Fig. 2.17. These are real and pure imaginary, respectively. Therefore, we can use the sequence (2.13.29) and its Richardson transforms to test the expected large order behavior. In this region there is a competition between the conifold and weak coupling instanton actions, and we should pass from a regime dominated by the weak coupling action near $\lambda = 0$, to a regime dominated by the conifold action near $\lambda = \lambda_c$. This is precisely what the numerical analysis shows. As an example, we show in Fig. 2.19 the fifth Richardson transform of the sequence (2.13.29) for two different values of λ and two different instanton actions: on the left, we consider $\lambda \approx -0.1386\,\mathrm{i}$ and the conifold action, while on the right we consider $\lambda \approx -0.0620\,\mathrm{i}$ and the weak coupling action. As we see, the expected asymptotic value (unity) is reached quite accurately.

2.13.5 Borel summability

In the physical ABJM theory, λ is real and g_s is purely imaginary. The expansion (0.2.2) should be written in terms of the real coupling constant $2\pi/k$, i.e. as

$$F(\lambda, k) = \sum_{g=0}^{\infty} \left(\frac{2\pi}{k}\right)^{g-2} (-1)^{g-1} F_g(\lambda). \tag{2.13.30}$$

We get an extra $(-1)^{g-1}$ sign at each genus, and this is what motivated the introduction of this sign in (2.13.27). Equivalently, this leads to an extra $-\mathrm{i}$ factor in the instanton actions computed above. We

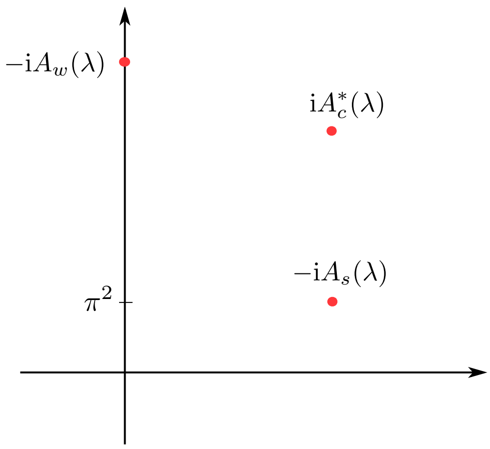

Figure 2.20: Singularities in the first quadrant of the Borel plane, for $\lambda \sim 1$.

can now ask whether this factorially divergent series is Borel summable or not. At strong coupling, the behavior of the genus g free energy $(-1)^{g-1}F_g(\lambda)$ is oscillatory for generic λ. This is because the strong coupling action (2.13.12), which controls the asymptotics in this regime, is complex:

$$\mathrm{Im}\,(-\mathrm{i}A_s(\lambda)) = \pi^2. \tag{2.13.31}$$

In fact, for large λ we have,

$$-\mathrm{i}A_s(\lambda) = 2\pi^2\sqrt{2\lambda} + \pi^2\mathrm{i} + \mathcal{O}\left(\mathrm{e}^{-2\pi\sqrt{2\lambda}}\right), \qquad \lambda \gg 1. \tag{2.13.32}$$

This suggests that the $1/N$ expansion is Borel summable for generic values of λ in the strong coupling region, as it happens in simple quantum-mechanical examples [177]. More precisely, Borel summability requires that there are no instantons with positive real action, i.e. that there are no singularities along the positive real axis in the Borel plane of the coupling constant $2\pi/k$. For $\lambda > 1/4$ none of the actions $A_p(\lambda)$ lies on the positive real axis. This is illustrated in Fig. 2.20, which shows the singularities in the first quadrant of the Borel plane for $\lambda \sim 1$. They are associated to the instanton actions $-\mathrm{i}A_s(\lambda)$, $-\mathrm{i}A_w(\lambda)$, and to the conjugate action $(-\mathrm{i}A_c(\lambda))^*$. Notice that there are other singularities in the other quadrants, related to the ones shown in Fig. 2.20 by flipping the sign and by conjugation.

Since

$$\mathrm{Im}\,(-\mathrm{i}A_c(\kappa)) = \pi^2\,(1 - 4\lambda(\kappa))\,, \tag{2.13.33}$$

the conifold action is real for $\lambda = 1/4$, and we have in principle an obstruction to Borel summability. It might happen that there are other instantons in the theory which we have not identified and lead to singularities in the real axis, even at strong coupling. However, if all the instantons are integer linear combinations of (2.13.10) and (2.13.11), as we conjectured before, there will be only a countable set of values of λ for which this happens.

Notice that, at large λ, the imaginary part of the dominant instanton action is subleading, and the action is approximately real. This peculiar behavior, in which Borel summability is lost in some limit of the parameter space of the model, has been found before in much more conventional models. Indeed, as shown in [189], the classical $O(N)$ one-dimensional spin chain has a Borel summable $1/N$ expansion, where each term in the expansion is itself a function of the temperature. However, in the low temperature limit the imaginary part of the instanton action is subleading. Correspondingly, when each term in the $1/N$ series is truncated to its low-temperature limit, the resulting expansion is not Borel summable anymore. Nevertheless, it should be kept in mind that in general the asymptotics does not commute with taking limits in parameter space. In the case of ABJM, for example, the asymptotics of the strongly coupled F_g (where we neglect worldsheet instanton corrections) is not governed by the strong coupling limit of the instanton action (2.13.32).

We conclude that for generic, real values of λ in the strong coupling region, the genus expansion is very likely Borel summable. This means that all the information about the partition function of the dual superstring theory can be retrieved from the perturbative genus expansion, by Borel resummation. This is in contrast with the genus expansion of unitary models coupled to gravity, which is not Borel summable [190]. Of course, the lack of Borel summability is not a problem if we have an unambiguous non-perturbative definition, as it is the case here. It just means that we have to add explicit non-perturbative effects in the theory in a careful way, as illustrated for example in [191]. But it is interesting that type IIA superstring on this AdS_4 background leads to a Borel resummable string expansion, since this means that it represents a stable background with respect to quantum-mechanical tunnelling effects in the string coupling constant.

Let us now consider the analytic continuation of the ABJM theory to an *imaginary* value of the Chern–Simons level,

$$k = \mathrm{i}\gamma, \qquad \gamma \in \mathbb{R}, \tag{2.13.34}$$

so that λ is also imaginary. In this case g_s is real, and the action controlling large order behavior near $\lambda = \lambda_c$ is the conifold action, which is also real. The free energies F_g have all the same sign now and the expansion in $1/\gamma$ is *not* Borel summable.

2.13.6 Double scaling limit near the conifold point

One interesting aspect of the ABJM free energy is the existence of the critical point (2.13.9) for an imaginary value of the coupling, which corresponds in the Calabi–Yau language to a conifold point. The genus g free energies $F_g(\lambda)$ are singular at this point, and their behavior near $\lambda = \lambda_c$ is given by

$$F_g \sim C_g \left(\frac{2\pi^2 \mathrm{i}\,(\lambda - \lambda_c)}{\log(\lambda - \lambda_c)} \right)^{2-2g}, \qquad g \geq 2, \tag{2.13.35}$$

where

$$C_g = \frac{B_{2g}}{2g(2g-2)}. \tag{2.13.36}$$

This is of course the critical behavior associated to the $c = 1$ string at the self-dual radius (see for example [117]). The scaling variable is

$$\mu \sim \frac{\lambda - \lambda_c}{\log(\lambda - \lambda_c)}. \tag{2.13.37}$$

This $c = 1$ behavior is expected from the Calabi–Yau point of view [192], but it is more surprising from the point of view of ABJM theory and its string dual.

The scenario one finds for ABJM theory near λ_c (i.e. a non-trivial critical point at imaginary 't Hooft coupling, a non-trivial double-scaling limit, and the lack of Borel summability which signals an instability) has been advocated in [193] in order to analytically continue AdS theories to de Sitter space. It would be interesting to understand this better.

2.14 Instanton at strong coupling

Based on the AdS/CFT correspondence, we would expect that the instanton configurations that we have found in the matrix model/gauge theory context should correspond to instanton configurations in the string theory dual. A natural source for such instanton effects are Euclidean D-branes wrapped around submanifolds in the target $AdS_4 \times \mathbb{CP}^3$. In this section, we want to interpret the strong coupling instanton A_s (which is the dominant configuration in the strongly coupled region) as a D-brane configuration, and we will find a D2-brane whose action coincides with the action A_s at large λ. Notice that, after including the coupling constant, the action (2.13.32) becomes, at strong coupling,

$$\frac{k}{2\pi \mathrm{i}} A_s(\lambda) \approx k\pi\sqrt{2\lambda} + \frac{\pi \mathrm{i} k}{2}. \tag{2.14.1}$$

In terms of the string coupling constant this can be also written as $A_{\mathrm{st}}/g_{\mathrm{st}}$, where

$$A_{\mathrm{st}} \approx \frac{1}{4} \left(\frac{L}{\ell_s} \right)^3 \left(1 + 2\pi \mathrm{i} \frac{\ell_s^2}{L^2} \right). \tag{2.14.2}$$

The leading part of this action has the appropriate form for a extended object in three dimensions, and it is natural to identify it with an Euclidean D2 brane. In fact, it can be written as

$$T_{\mathrm{D2}} \mathrm{vol}(\mathbb{RP}^3) \tag{2.14.3}$$

and seems to correspond to an Euclidean D2 brane wrapping a $\mathbb{RP}^3$ inside $\mathbb{CP}^3$. We will now make this more precise by an explicit calculation. Note from (2.14.2) that the imaginary part of the instanton action, which makes the $1/N$ expansion Borel summable, is in fact an α' correction. This means that Borel summability is in this case a stringy effect, and it is invisible in the supergravity limit.

2.14.1 D2-brane instantons

We work in the coordinate system for $\mathbb{CP}^3$ given in Appendix F. The metric has the form (F.7). We consider a D2-brane wrapping the submanifold of fixed α with $\vartheta_1 = \vartheta_2 = \vartheta$ and $\varphi_1 = -\varphi_2 = \varphi$. The metric is that of a warped $\mathbb{RP}^3$ (note that the period of χ is 2π)

$$\mathrm{d}s^2 = \frac{L^3}{4k}\left[\mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2 + \sin^2\alpha\,(\mathrm{d}\chi + \cos\vartheta\,\mathrm{d}\varphi)^2\right], \tag{2.14.4}$$

and in addition in the world-volume we include a field strength $F_{\vartheta\varphi} = E\sin\vartheta$.

The classical action including the Dirac-Born-Infeld (DBI) and Chern–Simons (CS) terms is

$$\mathcal{S}_{\mathrm{D2}} = T_{\mathrm{D2}}\int \mathrm{e}^{-\Phi}\sqrt{\det(g + 2\pi\alpha' F)} + T_{\mathrm{D2}}\int \pi\mathrm{i}\alpha' P[C_1]\wedge F\,, \tag{2.14.5}$$

where $P[C_1]$ is the pullback to the world-volume of the one-form (F.9), which in our subspace is

$$C_1 = \frac{k}{2}\Big[\cos\alpha(\mathrm{d}\chi + \cos\vartheta\,\mathrm{d}\varphi) - \mathrm{d}\chi - \mathrm{d}\varphi\Big]. \tag{2.14.6}$$

The extra $\mathrm{d}\chi$ and $\mathrm{d}\varphi$ terms, which are exact, make the expression regular at $\alpha = \vartheta = 0$. Similar terms with opposite signs will be regular at $\alpha, \vartheta = \pi$.

Plugging our ansatz in we find

$$\mathcal{S}_{\mathrm{D2}} = \frac{T_{\mathrm{D2}}L^3}{8}\int \mathrm{d}\chi\,\mathrm{d}\vartheta\,\mathrm{d}\varphi\,\sin\vartheta\left[\sin\alpha\sqrt{1+\beta^2E^2} + \mathrm{i}\beta E(\cos\alpha - 1)\right], \tag{2.14.7}$$

with $\beta = 8\pi k/L^3 = \sqrt{2/\lambda}$ (setting $\alpha' = 1$). Note that we are using conventions where the D2-brane tension is $T_{\mathrm{D2}} = 1/4\pi^2$.

The equation of motion for α gives the relation

$$\beta E = -\mathrm{i}\cos\alpha\,. \tag{2.14.8}$$

Then the electric flux density is the conserved momentum dual to the electric field

$$p = \mathrm{i}\,\frac{\delta\mathcal{L}}{\delta E} = \beta\sin\vartheta\,. \tag{2.14.9}$$

The classical action should be expressed in terms of p, rather than E, and the Legendre transform gives

$$\mathcal{S}_{\mathrm{D2}}^{\mathrm{classical}} = \frac{T_{\mathrm{D2}}L^3}{8}\int \mathrm{d}\chi\,\mathrm{d}\vartheta\,\mathrm{d}\varphi\,\mathrm{i}pE + \mathcal{S}_{\mathrm{D2}} = \frac{L^3}{4} = \pi k\sqrt{2\lambda}\,, \tag{2.14.10}$$

In precise agreement with the leading, real part of the action of the matrix model instanton (2.14.1). This is the same as the action of $k/2$ string instantons.

Above we wrote the DBI action suppressing fluctuations in the three orthogonal directions in $\mathbb{CP}^3$. It is easy to include them and one finds that this D2-brane is a classical solution which is unstable to fluctuations in these three directions.

2.14.2 Supersymmetry

The D-brane we have found should be a BPS state since $\mathbb{RP}^3$ is a generalized Lagrangian submanifold in $\mathbb{CP}^3$ (this has been established in a related context in [142]). We will now confirm this by direct calculation, see [194, 195, 196] for similar considerations.

Our choice of frame and the corresponding expression for the Killing spinors are given in Appendix G. For our ansatz (with all the other fields set to zero) they are

$$\epsilon = e^{\frac{\alpha}{4}(\hat\gamma\gamma_4 - \gamma_{7\natural})}e^{\frac{\vartheta}{4}(\hat\gamma\gamma_5 - \gamma_{8\natural} + \gamma_{79} + \gamma_{46})}e^{-\frac{\xi_1}{2}(\hat\gamma\gamma_\natural - \gamma_{47}) - \frac{\xi_2}{2}(\gamma_{58} - \gamma_{69})}\epsilon_0 = \mathcal{M}\epsilon_0\,, \tag{2.14.11}$$

ϵ_0 is a constant 32-component spinor and the Dirac matrices satisfy $\gamma_{0123456789\natural} = 1$.

The angles ξ_i are the phases from (F.4)

$$\xi_1 = \frac{\chi + \varphi}{2}, \qquad \xi_2 = \frac{\chi - \varphi}{2}. \tag{2.14.12}$$

The supersymmetries preserved by a D2-brane are determined by solving the following equation on the D2-brane solution

$$\Gamma\,\epsilon = \epsilon\,, \tag{2.14.13}$$

where Γ for our D2-brane solution is given by (see e.g. [195])

$$\Gamma = \frac{\mathrm{i}}{\mathcal{L}_{DBI}} \left(\Gamma^{(3)} + 2\pi\alpha' F_{\vartheta\varphi}\,\Gamma^{(1)}\gamma_\natural \right). \tag{2.14.14}$$

Here

$$\Gamma^{(3)} = \Gamma_{\mu_1\mu_2\mu_3} \frac{\partial x^{\mu_1}}{\partial\sigma^1} \frac{\partial x^{\mu_2}}{\partial\sigma^2} \frac{\partial x^{\mu_3}}{\partial\sigma^3}, \qquad \Gamma^{(1)} = \Gamma_\chi\,, \tag{2.14.15}$$

are the pullback of the curved space-time Dirac matrices in the world-volume directions (with and without the directions of the field strength $F_{\vartheta\varphi}$). Plugging in our choice of coordinates and the details of the solution we find

$$\begin{aligned} \Gamma^{(3)} &= \frac{1}{8}\sin\alpha\sin\vartheta\,\gamma_{758}\,e^{\frac{\alpha}{2}(\gamma_{56}-\gamma_{89})}\,, \\ 2\pi\alpha' F_{\vartheta\varphi}\Gamma^{(1)} &= -\frac{\mathrm{i}}{8}\cos\alpha\sin\alpha\sin\vartheta\,\gamma_7\,, \\ \mathcal{L}_{DBI} &= \frac{1}{8}\sin^2\alpha\sin\vartheta\,. \end{aligned} \tag{2.14.16}$$

And we therefore find that (2.14.13) reads

$$\left(\mathrm{i}\,\gamma_{758}\,e^{\frac{\alpha}{2}(\gamma_{56}-\gamma_{89})} + \cos\alpha\,\gamma_{7\natural}\right)\epsilon = \sin\alpha\,\epsilon\,. \tag{2.14.17}$$

Simple manipulations allow to write this equation as

$$-\,\mathrm{i}\,\gamma_{58\natural}\,e^{\frac{\alpha}{2}(2\gamma_{7\natural}+\gamma_{56}-\gamma_{89})}\,\epsilon = \epsilon\,. \tag{2.14.18}$$

Next we need to commute this operator through $\mathcal{M}$ in (2.14.11). As it turns out, only the α dependent term in $\mathcal{M}$ does not commute through and we find the equation

$$-\,\mathrm{i}\,\mathcal{M}^{-1}\gamma_{58\natural}\,e^{\frac{\alpha}{2}(2\gamma_{7\natural}+\gamma_{56}-\gamma_{89})}\,\epsilon = -\mathrm{i}\,\gamma_{58\natural}\,e^{\frac{\alpha}{2}(\hat\gamma\gamma_4+\gamma_{7\natural}+\gamma_{56}-\gamma_{89})}\,\epsilon_0 = \epsilon_0\,. \tag{2.14.19}$$

It is easy to see that the operator appearing in this equation squares to unity, and half its eigenvalues are $+1$ and half -1. Since it does not commute with the S_i operators in (G.6), the D2-brane is 1/2 BPS.

Note in particular that for $\alpha = 0$ we find the equation $\mathrm{i}\,\gamma_{58\natural}\epsilon_0 = \epsilon_0$, which is the projector equation for a fundamental string wrapping the ϑ_1, φ_1 sphere. In this limit the D2-brane instanton indeed degenerates to $k/2$ regular string instantons. While the supercharges at different values of α are not the same, it is possible to choose the orientation of the D2-branes in $\mathbb{CP}^3$ such that their supersymmetry is compatible with the supercharge used for localization and with world-sheet instantons (of certain orientation). Therefore, these D2-branes have the right structure to be responsible for the non-perturbative effects we have found in the matrix model.

Chapter 3

Fermi gas approach

In this chapter we develop an alternative method to study Chern–Simons–matter models based on a relation to a certain Fermi gas system. This chapter is organized as follows. In section 3.1 we discuss how the perturbative parts of all genus free energies of ABJM theory can be resummed into a closed simple expression. In section 3.2 we show that the matrix integral of a general class of $\mathcal{N} \geq 3$ CSM theories (necklace quivers with fundamental matter) can be written as the partition function of an ideal Fermi gas with a non-trivial one-particle Hamiltonian. In sections 3.3 and 3.4 we present the tools to analyze the Fermi gas, and we illustrate them in ABJM theory. More precisely, in section 3.3 we study the Fermi gas in the thermodynamic limit, by passing to the grand canonical ensemble. This makes it possible to derive the leading $N^{3/2}$ behavior of the free energy of ABJM theory, by using elementary tools in Statistical Mechanics. In section 3.4 we study the quantum corrections to the grand canonical potential. In section 3.5 we extend our techniques to more general CSM theories, including necklace quivers and theories with fundamental matter. We show that, when the free energy on the three-sphere is real, the $1/N$ expansion at fixed k gets resummed by an Airy function, thus proving the property 0.2.15 for this family of examples. We also consider the "massive" theory of [176], where a different $N^{5/3}$ scaling has been found for the free energy, and we rederive it with our techniques.

3.1 The ABJM matrix model in the 't Hooft expansion

3.1.1 $1/N$ expansion and non-perturbative effects

As we have seen in section 2.12.1, when expanded at strong coupling, the genus g free energies have the structure

$$F_g(\hat{\lambda}) = F_g^{\rm p}(\hat{\lambda}) + F_g^{\rm np}(\hat{\lambda}). \tag{3.1.1}$$

The first term represents the perturbative contribution in α', while the second term is non-perturbative in α',

$$F_g^{\rm np}(\hat{\lambda}) \sim \mathcal{O}\left(\mathrm{e}^{-2\pi\sqrt{2\hat{\lambda}}}\right) \tag{3.1.2}$$

and it was interpreted as the contribution of worldsheet instantons in the type IIA dual.

Besides the non-perturbative effects in α' there are also non-perturbative effects in the string coupling constant of the form

$$\exp\left(-k\pi\sqrt{2\lambda}\right) \tag{3.1.3}$$

at large λ. These instanton effects were deduced in the previous chapter by analysis of large-order behavior of perturbation theory These were interpreted as D2-branes wrapped around generalized Lagrangian cycles of the target geometry. We will refer to these non-perturbative effects as membrane instanton effects, since they can be interpreted as M2 instantons in M-theory [120] but they are invisible in ordinary string perturbation theory.

It was shown in [23] that the genus expansion of the perturbative free energies can be resummed. In order to do that, one has to use the variable

$$\lambda_{\rm ren} = \lambda - \frac{1}{24} - \frac{1}{3k^2}. \tag{3.1.4}$$

If we define the perturbative partition function as

$$Z^{\rm p}_{\rm ABJM} = \exp\left[\sum_{g=0}^{\infty} F_g^{\rm p}(\hat\lambda) g_s^{2g-2}\right] \tag{3.1.5}$$

then

$$Z^{\rm p}_{\rm ABJM} \propto {\rm Ai}\left[\left(\frac{\pi^2 k^4}{2}\right)^{1/3} \lambda_{\rm ren}\right], \tag{3.1.6}$$

where Ai is the Airy function. This can be also written in terms of N as

$$Z^{\rm p}_{\rm ABJM} \propto {\rm Ai}\left[C^{-1/3}\left(N - \frac{k}{24} - \frac{1}{3k}\right)\right], \tag{3.1.7}$$

where

$$C = \frac{2}{\pi^2 k}. \tag{3.1.8}$$

The expansion resummed in (3.1.7) makes perfect sense for finite k. Therefore, even if (3.1.7) was obtained from a calculation in the 't Hooft expansion, it should be part of the M-theory answer. Indeed, one of our goals in this paper is to verify this by computing $Z_{\rm ABJM}$ directly in the M-theory expansion.

The Airy function appearing in (3.1.7) gives an exact resummation of the long-distance expansion in M-theory. The shift (3.1.4) was interpreted in (2.12.46) as a renormalization of the expansion parameter ℓ_p/L. Then the argument of the Airy function (3.1.7) is given by

$$\left(256\, k\, \pi^2\right)^{-2/3} \left(\frac{L}{\widehat{\ell}_p}\right)^6. \tag{3.1.9}$$

The $1/N$ expansion of the ABJM matrix model was derived in the previous chapter by using the holomorphic anomaly equations [144] of topological string theory. The result (3.1.7) was obtained in [23] by looking at the recursive structure of these equations. There is however a much simpler method to obtain (3.1.7) which exploits the wavefunction behavior of the topological string partition function. Our derivation of (3.1.7) in this paper does not depend at all on ideas from topological string theory, but since it is formally very similar, we will now present this simpler argument. Readers who are not familiar with topological string theory can skip the rest of this section and proceed to the next one.

As shown in [121], it follows from the holomorphic anomaly equations that the topological string partition function is a wavefunction on moduli space. In particular, its transformation from one symplectic frame to the other is given by a Fourier transform. This property was spelled out in detail and exploited in [113]. The main result is summarized as follows. Let

$$\Gamma = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in {\rm SL}(2, \mathbb{Z}) \tag{3.1.10}$$

be a symplectic transformation relating two different frames (we assume for simplicity that there is a single modulus in the problem). This means that the periods $(\partial_a F_0, a)$ transform as

$$\begin{pmatrix} \partial_{a^\Gamma} F_0^\Gamma \\ a^\Gamma \end{pmatrix} = \Gamma \begin{pmatrix} \partial_a F_0 \\ a \end{pmatrix}. \tag{3.1.11}$$

Then, the full topological string partition function

$$Z(a) = \exp\left[\sum_{g=0}^{\infty} F_g(a) g_s^{2g-2}\right] \tag{3.1.12}$$

transforms as

$$Z^\Gamma(a^\Gamma) = \int {\rm d}a\, {\rm e}^{-S(a, a^\Gamma)/g_s^2} Z(a), \tag{3.1.13}$$

where

$$S(a, a^\Gamma) = -\frac{1}{2}\delta\gamma^{-1} a^2 + \gamma^{-1} a a^\Gamma - \frac{1}{2}\alpha\gamma^{-1}\left(a^\Gamma\right)^2. \tag{3.1.14}$$

In the context of ABJM theory, as explained in detail in [5, 6], the relevant quantities correspond to topological string computations in the so-called orbifold frame, where the natural periods are λ (the 't Hooft coupling of the gauge theory) and the derivative $\partial_\lambda F_0$. On the other hand, the most familiar frame in topological string theory is the large radius or Gromov–Witten frame, where the natural periods are T (the Kähler modulus) and the derivative $\partial_T F_0^{\rm GW}$. The genus g free energies in the large radius frame are given by the standard formulae,

$$\begin{aligned} F_0^{\rm GW} &= \frac{T^3}{6} + \sum_{k>0} N_{0,k} {\rm e}^{-kT}, \\ F_1^{\rm GW} &= \frac{T}{12} + \sum_{k>0} N_{1,k} {\rm e}^{-kT}, \\ F_{g>1}^{\rm GW} &= \sum_{k>0} N_{g,k} {\rm e}^{-kT}, \end{aligned} \tag{3.1.15}$$

where $N_{g,k}$ are Gromov–Witten invariants in the local $\mathbb{P}^1 \times \mathbb{P}^1$ geometry (there is no constant term contribution at higher genus). The fact that the total free energy is at most cubic in T, up to exponentially small corrections, is a well known fact in topological string theory.

In the previous chapter, the periods in the orbifold frame were written in terms of periods in the large radius frame in order to perform analytic continuations to strong coupling. By general principles, this relation must be a symplectic transformation like (3.1.10). In fact, it is easy to see that the results of the previous chaper relating the periods can be written as the following symplectic transformation:

$$\begin{pmatrix} \partial_{\tilde\lambda} \widetilde{F}_0 \\ \tilde\lambda \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} \partial_{\widetilde{T}} \widetilde{F}_0^{\rm GW} \\ \widetilde{T} \end{pmatrix} \tag{3.1.16}$$

where

$$\tilde\lambda = \frac{4\pi^2}{c}\lambda, \quad \widetilde{T} = \frac{\pi {\rm i}}{2c} T, \qquad c^2 = 2\pi {\rm i}, \tag{3.1.17}$$

and

$$\begin{aligned} \widetilde{F}_0 &= F_0 - \pi^3 {\rm i} \lambda, \\ \widetilde{F}_g^{\rm GW} &= (-4)^{g-1} \left(F_g^{\rm GW} - \delta_{g,0} \frac{\pi^2 T}{3} \right). \end{aligned} \tag{3.1.18}$$

Then, according to (3.1.13), (3.1.14), the total partition functions are related by the following formula:

$$\exp\left[F(\lambda) - \pi^3 {\rm i} \lambda / g_s^2\right] \propto \int {\rm d}\widetilde{T} \exp\left[-\widetilde{T}^2/g_s^2 + \widetilde{T}\tilde\lambda/g_s^2 + \widetilde{F}^{\rm GW}(\widetilde{T}) \right]. \tag{3.1.19}$$

Notice that, up to nonperturbative terms in T, this is the integral of the exponential of a cubic polynomial, therefore we will indeed get an Airy function. Let us introduce the new variable μ through

$$T = \frac{4\mu}{k} - \pi {\rm i}. \tag{3.1.20}$$

Then, one finds the expression

$$\begin{aligned} \exp F(\lambda) &\propto \int {\rm d}\mu \exp\left\{ \frac{2\mu^3}{3k\pi^2} - \mu N + \frac{k}{24}\mu + \frac{1}{3k}\mu + \mathcal{O}\left({\rm e}^{-\frac{4\mu}{k}}\right) \right\} \\ &\propto {\rm Ai}\left[C^{-1/3}(N - B) \right] \left(1 + \mathcal{O}({\rm e}^{-2\pi\sqrt{2\lambda}}) \right), \end{aligned} \tag{3.1.21}$$

where we used the following integral representation of the Airy function,

$${\rm Ai}(z) = \frac{1}{2\pi {\rm i}} \int_{\mathcal{C}} {\rm d}t \, \exp\left(\frac{t^3}{3} - zt \right), \tag{3.1.22}$$

and $\mathcal{C}$ is a contour in the complex plane from ${\rm e}^{-{\rm i}\pi/3}\infty$ to ${\rm e}^{{\rm i}\pi/3}\infty$. In (3.1.21), C is given in (3.1.8) and

$$B = \frac{k}{24} + \frac{1}{3k}. \tag{3.1.23}$$

The result of (3.1.21) is of course the expression obtained in (3.1.7). Notice that the first term in the shift B comes from $F_0^{\rm GW}$ in (3.1.18), while the second term is due to the first, perturbative term in $F_1^{\rm GW}$. The exponentially small corrections in N in (3.1.21), which are due to the worldsheet instantons at large radius of the topological string, become, after Fourier transform, the worldsheet instantons (3.1.2) of the type IIA superstring.

This derivation is nice, but it seems difficult to generalize it in its current form to other Chern–Simons–matter theories, and prove in this way the property (0.2.15) for other cases. In this paper we will find a completely different approach to the derivation of the Airy function which turns out to formally equivalent to the one based on topological string theory. However, this approach can be extended to many $\mathcal{N} \geq 3$ CSM theories and makes it possible to verify the conjecture 0.2.15 for many of them.

3.2 Chern–Simons–matter theories as Fermi gases

3.2.1 ABJM theory as a Fermi gas

Our Fermi gas approach is based on the following observation. The interaction term between the eigenvalues in (2.2.3) can be written in a different way by using the Cauchy identity:

$$\begin{aligned}\frac{\prod_{i<j}\left[2\sinh\left(\frac{\mu_i-\mu_j}{2}\right)\right]\left[2\sinh\left(\frac{\nu_i-\nu_j}{2}\right)\right]}{\prod_{i,j}2\cosh\left(\frac{\mu_i-\nu_j}{2}\right)} &= \det_{ij}\frac{1}{2\cosh\left(\frac{\mu_i-\nu_j}{2}\right)} \\ &= \sum_{\sigma\in S_N}(-1)^{\epsilon(\sigma)}\prod_i\frac{1}{2\cosh\left(\frac{\mu_i-\nu_{\sigma(i)}}{2}\right)}.\end{aligned} \tag{3.2.1}$$

In this equation, S_N is the permutation group of N elements, and $\epsilon(\sigma)$ is the signature of the permutation σ. This identity has been used in other matrix models in [31, 118, 119] in order to study them in the grand canonical ensemble, as we will do here. In the context of ABJM theory, it was used in [204] in order to prove the equivalence of (2.2.3) and the matrix integral for $\mathcal{N}=8$ super Yang–Mills theory in three dimensions, when $k=1$. The manipulations in [204] can be easily generalized to arbitrary k, and one obtains the following expression for the ABJM matrix model,

$$Z(N)=\frac{1}{N!}\sum_{\sigma\in S_N}(-1)^{\epsilon(\sigma)}\int\frac{{\rm d}^N x}{(2\pi k)^N}\frac{1}{\prod_i 2\cosh\left(\frac{x_i}{2}\right)2\cosh\left(\frac{x_i-x_{\sigma(i)}}{2k}\right)}. \tag{3.2.2}$$

We will derive this expression below with a different technique, which can be used for more general Chern–Simons–matter theories. The main property of (3.2.2) is that it makes contact with the standard formalism to study partition functions of ideal Fermi gases. Indeed, let us introduce the function

$$\rho(x_1,x_2)=\frac{1}{2\pi k}\frac{1}{\left(2\cosh\frac{x_1}{2}\right)^{1/2}}\frac{1}{\left(2\cosh\frac{x_2}{2}\right)^{1/2}}\frac{1}{2\cosh\left(\frac{x_1-x_2}{2k}\right)}. \tag{3.2.3}$$

If we interpret it as a one-particle density matrix in the position representation

$$\rho(x_1,x_2)=\langle x_1|\hat\rho|x_2\rangle, \tag{3.2.4}$$

the matrix integral (2.2.3) can be written as the partition function of an ideal Fermi gas with N particles

$$Z(N)=\frac{1}{N!}\sum_{\sigma\in S_N}(-1)^{\epsilon(\sigma)}\int{\rm d}^N x\prod_i\rho(x_i,x_{\sigma(i)}). \tag{3.2.5}$$

It is well-known that the sum over permutations appearing in the canonical free energy of an ideal quantum gas can be written as a sum over conjugacy classes of the permutation group (see for example [122]). A conjugacy class is specified by a set of integers $\{m_\ell\}$ satisfying

$$\sum_\ell \ell m_\ell = N. \tag{3.2.6}$$

Let us define

$$Z_\ell=\int{\rm d}x_1\cdots{\rm d}x_\ell\,\rho(x_1,x_2)\rho(x_2,x_3)\cdots\rho(x_{\ell-1},x_\ell)\rho(x_\ell,x_1). \tag{3.2.7}$$

Then, the partition function is given by,

$$Z(N) = \sum_{\{m_\ell\}}{}' \prod_\ell \frac{\eta^{(\ell-1)m_\ell} Z_\ell^{m_\ell}}{m_\ell! \ell^{m_\ell}} \tag{3.2.8}$$

where the $'$ means that we only sum over the integers satisfying the constraint (3.2.6).

Due to the constrained sum, the canonical partition function is not easy to handle for large N. As usual, the remedy is to consider the grand partition function

$$\Xi = 1 + \sum_{N=1}^{\infty} Z(N) z^N, \tag{3.2.9}$$

where

$$z = \mathrm{e}^{\mu} \tag{3.2.10}$$

plays the rôle of the fugacity and μ is the chemical potential. The grand-canonical potential is

$$J(\mu) = \log \Xi. \tag{3.2.11}$$

Notice that this potential (like the free energy) has the opposite sign to the usual conventions in Statistical Mechanics. A standard argument (presented for example in [122]) tells us that the sum over conjugacy classes in (3.2.8) can be written as

$$J(\mu) = -\sum_{\ell \geq 1} Z_\ell \frac{(-z)^\ell}{\ell}. \tag{3.2.12}$$

The canonical partition function is recovered from the grand-canonical potential as

$$Z(N) = \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} \frac{\Xi}{z^{N+1}}. \tag{3.2.13}$$

At large N, this integral can be computed by applying the saddle-point method to

$$Z(N) = \frac{1}{2\pi \mathrm{i}} \int \mathrm{d}\mu \, \exp\left[J(\mu) - \mu N\right]. \tag{3.2.14}$$

The saddle point occurs at

$$N = \frac{\partial J}{\partial \mu} = -\sum_{\ell \geq 1} Z_\ell (-z)^\ell, \tag{3.2.15}$$

and defines a function $\mu_*(N)$. The free energy is given, at leading order as $N \to \infty$, by

$$F(N) = J(\mu_*) - \mu_* N. \tag{3.2.16}$$

However, it is possible to compute the $1/N$ corrections to this relation by simply computing the corrections to the full integral in (3.2.14). This is what we will eventually do. Notice the similarity between the traditional inverse transform (3.2.14) and the Fourier transform (3.1.21) in topological string theory.

We have then shown that the original ABJM matrix integral can be computed as the canonical partition function of a system of N non-interacting fermions, where the one-particle density matrix is given by (3.2.3). We just have to solve the corresponding one-body problem in order to compute the relevant thermodynamic quantities of the system. Equivalently, one should compute the quantity Z_ℓ introduced in (3.2.7). This quantity can be regarded as the partition function of a classical lattice gas with ℓ particles in a periodic lattice with nearest-neighbour interactions, as shown in Fig. 3.1. The density matrix ρ plays the rôle of the classical transfer matrix of the system (see for example chapter 12 of [123]). It defines a symmetric kernel

$$\langle x | \hat{\rho} | \phi \rangle = \int \mathrm{d}x' \, \rho(x, x') \phi(x'), \tag{3.2.17}$$

so that

$$Z_\ell = \mathrm{Tr}\, \hat{\rho}^\ell. \tag{3.2.18}$$

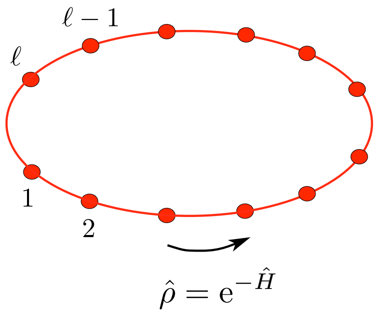

Figure 3.1: A one-dimensional periodic lattice with ℓ sites. The transfer matrix $\hat{\rho}$ can be regarded as the quantum propagator for a single particle in Euclidean, discretized time with Hamiltonian $\hat{H}$.

It is easy to see that this kernel is a non-negative Hilbert–Schmidt operator, therefore it has a discrete, positive spectrum

$$\hat{\rho}|\phi_n\rangle = \lambda_n|\phi_n\rangle, \qquad n = 0, 1, \cdots, \tag{3.2.19}$$

where $|\phi_n\rangle$ are orthonormal eigenfunctions and we assume that

$$\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \cdots. \tag{3.2.20}$$

We can then write the density matrix as

$$\hat{\rho} = \sum_{n\geq 0} \lambda_n |\phi_n\rangle\langle\phi_n|. \tag{3.2.21}$$

In terms of these eigenvalues we have,

$$Z_\ell = \sum_{n\geq 0} \lambda_n^\ell. \tag{3.2.22}$$

When ℓ is large, this sum is dominated by the largest eigenvalue λ_0,

$$Z_\ell \approx \lambda_0^\ell, \qquad \ell \gg 1. \tag{3.2.23}$$

It also follows from this representation that the grand-canonical partition function is given by a Fredholm determinant,

$$\Xi = \det(1 + z\hat{\rho}) = \prod_{n\geq 0} (1 + z\lambda_n). \tag{3.2.24}$$

Instead of using the formulation of the lattice problem in terms of the density matrix operator, we can introduce a quantum Hamiltonian in the standard way,

$$\hat{\rho} = \mathrm{e}^{-\hat{H}}. \tag{3.2.25}$$

This leads to the well-known equivalence between the partition function of a classical lattice gas (3.2.18) and the propagator of a quantum particle in ℓ units of discretized time (see for example [123, 124]). We can then write

$$Z_\ell = \mathrm{Tr}\, \mathrm{e}^{-\ell \hat{H}}. \tag{3.2.26}$$

To find the Hamiltonian corresponding to the ABJM matrix model, we first write the density matrix (3.2.3) as

$$\hat{\rho} = \mathrm{e}^{-\frac{1}{2}U(\hat{q})}\mathrm{e}^{-T(\hat{p})}\mathrm{e}^{-\frac{1}{2}U(\hat{q})}. \tag{3.2.27}$$

In this equation, $\hat{q}, \hat{p}$ are canonically conjugate operators,

$$[\hat{q}, \hat{p}] = \mathrm{i}\hbar, \tag{3.2.28}$$

and

$$\hbar = 2\pi k. \tag{3.2.29}$$

This is a key aspect of this formalism: $\hbar$ is the inverse coupling constant of the gauge theory/string theory, therefore semiclassical or WKB expansions in the Fermi gas correspond to strong coupling expansions in gauge theory/string theory. The potential $U(q)$ in (3.2.27) is given by

$$U(q) = \log\left(2\cosh\frac{q}{2}\right), \tag{3.2.30}$$

and the kinetic term $T(p)$ is given by the same function,

$$T(p) = \log\left(2\cosh\frac{p}{2}\right). \tag{3.2.31}$$

The peculiar kinetic term (3.2.31) can be regarded as a non-trivial dispersion relation interpolating between the quadratic behavior of a non-relativistic particle at small p,

$$T(p) \sim \log(2) + \frac{p^2}{8}, \qquad p \to 0, \tag{3.2.32}$$

and the linear behavior of an ultra-relativistic particle at large p,

$$T(p) \sim \frac{|p|}{2}, \qquad |p| \to \infty. \tag{3.2.33}$$

Notice that, as it is standard for Hamiltonians defined by transfer matrices at finite lattice spacing [123, 124], the quantum operator $\hat{H}$ defined by (3.2.25) and (3.2.27) differs from

$$T(\hat{p}) + U(\hat{q}) \tag{3.2.34}$$

in $\hbar$ corrections. There is a very elegant method to obtain these corrections based on the phase-space or Wigner approach to quantization. This method will be also extremely useful in setting the semiclassical or WKB expansion of our thermodynamic problem. We first recall that the Wigner transform of an operator $\hat{A}$ is given by (see [125] for a detailed exposition of phase-space quantization)

$$A_{\rm W}(q,p) = \int {\rm d}q' \left\langle q - \frac{q'}{2}\right| \hat{A} \left| q + \frac{q'}{2}\right\rangle {\rm e}^{{\rm i}pq'/\hbar}. \tag{3.2.35}$$

The Wigner transform of a product is given by the $\star$-product of their Wigner transforms,

$$\left(\hat{A}\hat{B}\right)_{\rm W} = A_{\rm W} \star B_W \tag{3.2.36}$$

where the star operator is given as usual by

$$\star = \exp\left[\frac{{\rm i}\hbar}{2}\left(\overleftarrow{\partial}_q \overrightarrow{\partial}_p - \overleftarrow{\partial}_p \overrightarrow{\partial}_q\right)\right], \tag{3.2.37}$$

and is invariant under linear canonical transformations. Another useful property is that

$${\rm Tr}\,\hat{A} = \int \frac{{\rm d}p{\rm d}q}{2\pi\hbar} A_{\rm W}(q,p). \tag{3.2.38}$$

In order to calculate the $\hbar$ corrections to the Hamiltonian, we consider the Wigner transform of the density matrix (3.2.27). By using (3.2.36) we find,

$$\rho_{\rm W} = {\rm e}^{-\frac{1}{2}U(q)} \star {\rm e}^{-T(p)} \star {\rm e}^{-\frac{1}{2}U(q)}. \tag{3.2.39}$$

Let us note that the partition function depends only on the eigenvalues λ_n of $\hat{\rho}$ (or, equivalently, on the traces $Z_\ell = {\rm Tr}\,\hat{\rho}^\ell$). Therefore there is the following freedom in the choice of $\hat{\rho}$:

$$\hat{\rho} \to \hat{V}\hat{\rho}\hat{V}^{-1} \tag{3.2.40}$$

which translates into

$$\rho_{\rm W}(q,p) \to V_{\rm W}(q,p) \star \rho_{\rm W}(q,p) \star (V^{-1})_{\rm W}(q,p) \tag{3.2.41}$$

after the Wigner transform. Equation (3.2.39) defines the Wigner transform of our Hamiltonian through

$$\rho_{\rm W} = {\rm e}_\star^{-H_{\rm W}}, \tag{3.2.42}$$

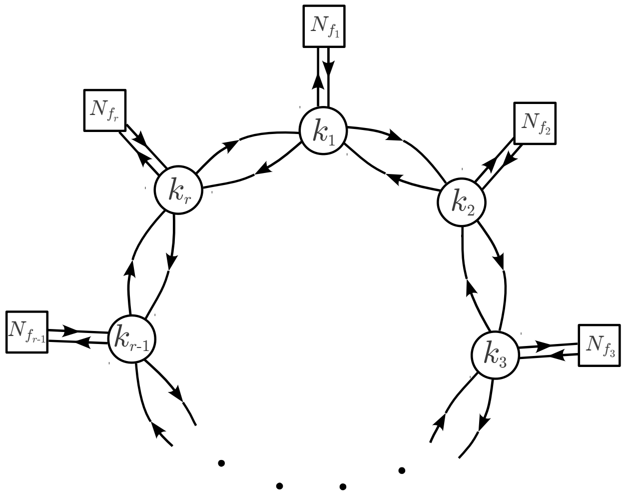

Figure 3.2: A quiver with r nodes forming a necklace.

where the $\star$-exponential is defined by

$$\exp_\star(A) = 1 + A + \frac{1}{2} A \star A + \cdots . \tag{3.2.43}$$

The quantum Hamiltonian $H_{\rm W}$ can be computed by using the Baker–Campbell–Hausdorff formula, as applied to the $\star$-product. One finds,

$$\begin{aligned} H_{\rm W}(q,p) &= T + U + \frac{1}{12} [T, [T, U]_\star]_\star + \frac{1}{24} [U, [T, U]_\star]_\star + \ldots \\ &= T(p) + U(q) - \frac{\hbar^2}{12} (T'(p))^2 U''(q) + \frac{\hbar^2}{24} (U'(q))^2 T''(p) + \mathcal{O}(\hbar^4), \end{aligned} \tag{3.2.44}$$

where we have used the fact that, at leading order in $\hbar$, the Moyal bracket is the Poisson bracket

$$[A, B]_\star \equiv A \star B - B \star A = \mathrm{i}\hbar \{A, B\} + \mathcal{O}(\hbar^2). \tag{3.2.45}$$

Further corrections to (3.2.44) can be computed to any desired order, see (B.1) for the result at order $\mathcal{O}(\hbar^4)$.

3.2.2 More general Chern–Simons–matter theories

The identification of the matrix model of ABJM theory as the partition function of a Fermi gas can be also made for more general $\mathcal{N} \geq 3$ Chern–Simons–matter theories. We will set up the formalism for the necklace quivers with r nodes considered in [126, 127], and with fundamental matter in each node (see Fig. 3.2). These theories are given by a

$$U(N)_{k_1} \times U(N)_{k_2} \times \cdots U(N)_{k_r} \tag{3.2.46}$$

Chern–Simons quiver. Each node will be labelled with the letter $a = 1, \cdots, r$. There are bifundamental chiral superfields A_{aa+1}, B_{aa-1} connecting adjacent nodes, and in addition we will suppose that there are N_{f_a} matter superfields $(Q_a, \tilde{Q}_a)$ in each node, in the fundamental representation. We will write

$$k_a = n_a k, \tag{3.2.47}$$

and we will assume that

$$\sum_{a=1}^{r} n_a = 0. \tag{3.2.48}$$

According to the general localization computation in [2], the matrix model computing the $\mathbb{S}^3$ partition function of a necklace quiver is given by

$$Z(N) = \frac{1}{(N!)^r} \int \prod_{a,i} \frac{\mathrm{d}\lambda_{a,i}}{2\pi} \frac{\exp\left[\frac{\mathrm{i} n_a k}{4\pi}\lambda_{a,i}^2\right]}{\left(2\cosh\frac{\lambda_{a,i}}{2}\right)^{N_{f_a}}} \prod_{a=1}^{r} \frac{\prod_{i<j}\left[2\sinh\left(\frac{\lambda_{a,i}-\lambda_{a,j}}{2}\right)\right]^2}{\prod_{i,j} 2\cosh\left(\frac{\lambda_{a,i}-\lambda_{a+1,j}}{2}\right)}. \tag{3.2.49}$$

This matrix model is very similar to the $\hat{A}_{r-1}$ models considered in for example [31, 128], and one can use a very similar strategy in order to rewrite them as Fermi gases. First of all, we define a kernel corresponding to a pair of connected nodes (a, b) by,

$$K_{ab}(x', x) = \frac{1}{2\pi k} \frac{\exp\left\{\mathrm{i}\frac{n_b x^2}{4\pi k}\right\}}{2\cosh\left(\frac{x'-x}{2k}\right)} \left[2\cosh\frac{x}{2k}\right]^{-N_{f_b}}, \tag{3.2.50}$$

where we set $x = \lambda/k$. The grand canonical partition function corresponding to the above matrix model is defined as in (3.2.9). Then, if we use the Cauchy identity (3.2.1), a simple generalization of the above arguments makes it possible to write it again as a Fredholm determinant (3.2.24), where now [31]

$$\hat{\rho} = \hat{K}_{r1}\hat{K}_{12}\cdots\hat{K}_{r-1,r} \tag{3.2.51}$$

is the product of the kernels (3.2.50) around the quiver. Therefore, we can apply exactly the same techniques that we used before in ABJM theory. In a sense, we are “integrating out" $r-1$ nodes of the quiver in order to define an effective theory in the r-th node, but with a complicated Hamiltonian which takes into account the other nodes.

This idea can be made very concrete by looking at the Wigner transform of the operator $\hat{\rho}$ in (3.2.51). We first compute the Wigner transform of the kernel (3.2.50),

$$K_{ab}^{\mathrm{W}}(q,p) = \frac{1}{2\cosh\frac{p}{2}} \star \frac{\mathrm{e}^{\frac{\mathrm{i} n_b q^2}{2\hbar}}}{\left[2\cosh\frac{q}{2k}\right]^{N_{f_b}}} \tag{3.2.52}$$

where the $\hbar$ in the $\star$ product is given again by (3.2.29). Let us note that

$$\mathrm{e}^{\frac{\mathrm{i}nq^2}{2\hbar}} \star f(p) \star \mathrm{e}^{-\frac{\mathrm{i}nq^2}{2\hbar}} = f\left(\mathrm{e}^{\frac{\mathrm{i}nq^2}{2\hbar}} \star p \star \mathrm{e}^{-\frac{\mathrm{i}nq^2}{2\hbar}}\right) = f\left(\mathrm{e}^{\mathrm{ad}_\star\left[\frac{\mathrm{i}nq^2}{2\hbar}\right]} p\right) = f(p-nq), \tag{3.2.53}$$

where we used that

$$\left[q^2, p\right]_\star = 2\mathrm{i}\hbar q. \tag{3.2.54}$$

We obtain then, for the Wigner transform of the density operator (3.2.51)

$$\begin{aligned}\rho_{\mathrm{W}}(q,p) = {} & \frac{1}{2\cosh\frac{p}{2}} \star \frac{1}{\left[2\cosh\frac{q}{2k}\right]^{N_{f_1}}} \star \frac{1}{2\cosh\frac{p-n_1 q}{2}} \star \\ & \frac{1}{\left[2\cosh\frac{q}{2k}\right]^{N_{f_2}}} \star \frac{1}{2\cosh\frac{p-(n_1+n_2)q}{2}} \star \frac{1}{\left[2\cosh\frac{q}{2k}\right]^{N_{f_3}}} \star \\ & \cdots \star \frac{1}{2\cosh\frac{p-(n_1+\cdots+n_{r-1})q}{2}} \star \frac{1}{\left[2\cosh\frac{q}{2k}\right]^{N_{f_r}}}\end{aligned} \tag{3.2.55}$$

where we used (3.2.48). For necklace theories without fundamental matter this is simply

$$\rho_{\mathrm{W}}(q,p) = \frac{1}{2\cosh\frac{p}{2}} \star \frac{1}{2\cosh\frac{p-n_1 q}{2}} \star \frac{1}{2\cosh\frac{p-(n_1+n_2)q}{2}} \star \cdots \star \frac{1}{2\cosh\frac{p-(n_1+\cdots+n_{r-1})q}{2}}. \tag{3.2.56}$$

In particular, for the ABJM necklace $(-k, k)$ with fundamental matter $N_{f_1} = N_{f_2} = N_f$ first considered in [17, 129, 130], we have

$$\rho_{\mathrm{W}}(q,p) = \frac{1}{2\cosh\frac{p}{2}} \star \frac{1}{\left[2\cosh\frac{q}{2k}\right]^{N_f}} \star \frac{1}{2\cosh\frac{p+q}{2}} \star \frac{1}{\left[2\cosh\frac{q}{2k}\right]^{N_f}}. \tag{3.2.57}$$

If we perform a canonical transformation

$$p \to -q, \qquad q \to p+q \tag{3.2.58}$$

and we conjugate by $\left[2\cosh\frac{q}{2}\right]^{1/2}$ to obtain a symmetric kernel, we get the equivalent representation,

$$\rho_{\rm W}(q,p) = \frac{1}{\left[2\cosh\frac{q}{2}\right]^{1/2}} \star \frac{1}{\left[2\cosh\frac{p+q}{2k}\right]^{N_f}} \star \frac{1}{2\cosh\frac{p}{2}} \star \frac{1}{\left[2\cosh\frac{p+q}{2k}\right]^{N_f}} \star \frac{1}{\left[2\cosh\frac{q}{2}\right]^{1/2}} \tag{3.2.59}$$

which, for $N_f = 0$, agrees with the result (3.2.39). In this way, we have reduced the general necklace quiver theory to an ideal Fermi gas whose one-particle quantum Hamiltonian is defined by the above density matrices through (3.2.42).

Notice that, in general, the density operators $\hat{\rho}$ are not Hermitian, and correspondingly $H_{\rm W}$ is generally not real. This reflects the fact that the free energy on the three-sphere of these CSM theories is in general complex.

3.3 Thermodynamic limit

It is well-known that the thermodynamic limit of an ideal quantum gas can be evaluated by treating the one-particle problem in the semiclassical or WKB approximation. Moreover, the $1/N$ corrections to the thermodynamic limit can be obtained by studying the quantum corrections to the semiclassical limit. In this section we will present general results about the thermodynamic limit and we will illustrate them in ABJM theory. More general theories will be considered in section 3.5.

3.3.1 The thermodynamic limit of ideal Fermi gases

In the following we will need several standard results in the analysis of ideal quantum gases. The distribution operator at zero temperature is given by,

$$\hat{n}(E) = \theta(E - \hat{H}) \tag{3.3.1}$$

where $\theta(x)$ is the Heaviside step function. The trace of this operator gives the function $n(E)$, counting the number of eigenstates whose energy is less than E:

$$n(E) = {\rm Tr}\, \hat{n}(E) = \sum_n \theta(E - E_n). \tag{3.3.2}$$

Notice that

$$E_n = -\log \lambda_n, \tag{3.3.3}$$

where λ_n are the eigenvalues (3.2.19) of the density matrix. The density of eigenstates is defined by

$$\rho(E) = \frac{{\rm d}n(E)}{{\rm d}E} = \sum_n \delta(E - E_n). \tag{3.3.4}$$

The one-particle canonical partition function is then given by the standard formula,

$$Z_\ell = \int_0^\infty {\rm d}E\, \rho(E)\, {\rm e}^{-\ell E}, \tag{3.3.5}$$

while the grand-canonical potential of the N particle system is given by

$$J(\mu) = \int_0^\infty {\rm d}E\, \rho(E)\, \log\left(1 + z{\rm e}^{-E}\right). \tag{3.3.6}$$

Let us now consider the thermodynamic limit of the system, when $N \to \infty$. In this regime, the behavior of the system is semiclassical and the spectrum of the one-particle Hamiltonian is encoded in the functions $n(E)$, $\rho(E)$. The thermodynamic limit is governed by the behavior of these functions as $E \gg 1$. We notice that, if

$$n(E) \approx CE^s, \qquad E \gg 1, \tag{3.3.7}$$

then the grand-canonical potential is given by

$$J(\mu) \approx sC \int_0^\infty \log\left(1 + z\mathrm{e}^{-E}\right) E^{s-1} \mathrm{d}E = -C\Gamma(s+1)\,\mathrm{Li}_{s+1}(-\mathrm{e}^{\mu}), \tag{3.3.8}$$

where Li_s is the usual polylogarithm function. The number of particles is related to the chemical potential by

$$N(\mu) \approx C\Gamma(s+1)\,\mathrm{Li}_s(-\mathrm{e}^{\mu}), \tag{3.3.9}$$

and large N corresponds to large μ. In this regime, we have

$$J(\mu) \approx \frac{C}{s+1}\mu^{s+1}, \qquad N(\mu) \approx C\mu^s\,. \tag{3.3.10}$$

The second equation defines μ as function of N, and we deduce from (3.2.16) that the canonical free energy is given, as $N \to \infty$, by

$$F(N) \approx -\frac{s}{s+1} C^{-1/s} N^{\frac{s+1}{s}}\,. \tag{3.3.11}$$

These formulae should be familiar from the elementary theory of ideal quantum gases. For example, the textbook ideal Fermi gas in three dimensions has $s = 3/2$.

To determine the value of s for a given system we notice that, in the semiclassical limit, the trace is replaced by an integral over phase space

$$\mathrm{Tr} \to \int \frac{\mathrm{d}q\mathrm{d}p}{2\pi\hbar} \tag{3.3.12}$$

which gives the standard semiclassical formula

$$n(E) \approx \int \frac{\mathrm{d}p\mathrm{d}q}{2\pi k}\theta(E - H(q,p)) = \frac{\mathrm{Vol}(E)}{2\pi\hbar}, \tag{3.3.13}$$

i.e. the number of eigenstates is just given by the volume of phase space. The surface

$$H(q,p) = E \tag{3.3.14}$$

is just the Fermi surface of the system. For a one-dimensional ideal gas whose one-particle Hamiltonian is of the form

$$H \sim A|p|^{\alpha} + B|q|^{\beta} \tag{3.3.15}$$

we have

$$s = \frac{1}{\alpha} + \frac{1}{\beta}. \tag{3.3.16}$$

This will be useful later on.

3.3.2 A simple derivation of the $N^{3/2}$ behavior in ABJM theory

We can now study the thermodynamic limit of the partition function of ABJM theory. In this case, the Hamiltonian appearing in the semiclassical formula (3.3.13) is just given by the classical counterpart of (3.2.34),

$$H_{\mathrm{cl}}(q,p) = T(p) + U(q) = \log\left(2\cosh\frac{p}{2}\right) + \log\left(2\cosh\frac{q}{2}\right). \tag{3.3.17}$$

Here we have neglected the $\hbar$ corrections appearing in H_{W}. It is easy to show that the minimum energy is

$$E_0 = 2\log 2, \tag{3.3.18}$$

which corresponds to the maximal eigenvalue of the density matrix

$$\lambda_0 = \frac{1}{4}. \tag{3.3.19}$$

This is the semiclassical value given by the leading WKB approximation, and it will be corrected quantum-mechanically. In the large E regime, the discrete spectrum "condenses" along a cut in the complex plane, and λ_0 signals the endpoint of the cut.

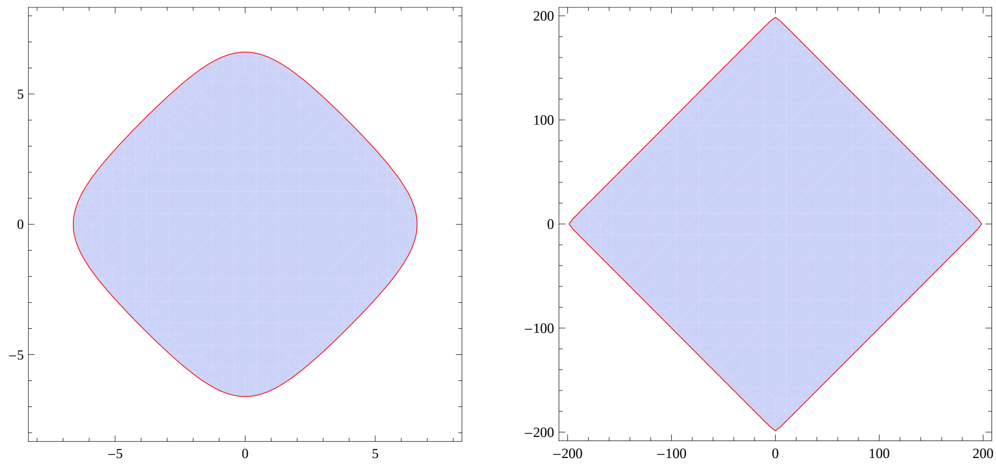

Figure 3.3: The Fermi surface (3.3.20) for ABJM theory in the q-p plane, for $E = 4$ (left) and $E = 100$ (right). When the energy is large, the Fermi surface approaches the polygon (3.3.22).

In order to proceed with the analysis of the thermodynamic limit, we should determine the Fermi surface

$$H_{\rm cl}(q,p) = E \tag{3.3.20}$$

controling the density of eigenvalues. We show the shape of this surface in Fig. 3.3 for $E = 4$ (left) and $E = 100$ (right). It is clear that in the thermodynamic limit, when E is large, the surface can be approximated by considering the values of $U(q)$, $T(p)$ for q, p large. In this regime we have

$$U(q) \approx \frac{|q|}{2}, \quad |q| \to \infty, \qquad T(p) \approx \frac{|p|}{2}, \quad |p| \to \infty, \tag{3.3.21}$$

so that (3.3.20) is approximately given by

$$|q| + |p| = 2E, \tag{3.3.22}$$

as it is manifest in the graphic on the right in Fig. 3.3. From (3.3.15) and (3.3.16) we deduce that

$$s = 2. \tag{3.3.23}$$

Since

$$\text{Vol}(E) \approx 8E^2, \tag{3.3.24}$$

the number of states is given by

$$n(E) \approx \frac{2}{\pi^2 k} E^2. \tag{3.3.25}$$

By comparing with (3.3.7), we find

$$C = \frac{2}{\pi^2 k}. \tag{3.3.26}$$

The equation (3.3.11) gives immediately

$$F(N) \approx -\frac{\pi\sqrt{2k}}{3} N^{3/2}. \tag{3.3.27}$$

This is exactly the result found in the previous chapter using the 't Hooft expansion of the matrix model. The derivation presented here is however completely elementary, and relies on basic notions of quantum Statistical Mechanics: the 3/2 scaling of the number of degrees of freedom is nothing but the scaling of the free energy of an ultrarelativistic gas of one-dimensional fermions in a linearly confining potential. No matrix model techniques are needed. In this sense, our derivation is even simpler than the one presented in [38], which required some detailed analysis of the eigenvalue interaction in the matrix integral.

We would like to emphasize that the above result (3.3.27) provides the right large N behavior of the system at finite k. This is because the true expansion parameter in the semiclassical expansion is $\hbar/E$, which is small for large E even at finite $\hbar$. This can be proved rigorously for some spectral problems defined by kernels of the form (3.2.3) [131], and we will verify it in section 3.4 by a detailed analysis of the WKB expansion.

3.3.3 Large N corrections

One advantage of the statistical-mechanical framework presented here is that it makes it possible to compute corrections to the thermodynamic limit in a systematic way. To start the study of these corrections, we now look at the thermodynamics of the Fermi gas of ABJM theory in the semiclassical approximation, but taking into account the exact value of the volume of phase space (i.e. we go beyond the polygonal approximation in (3.3.22)). As expected, this gives sub-leading and exponentially suppressed corrections at large N.

The computation of the exact volume is equivalent to computing all the Z_ℓ exactly in the semiclassical approximation, and resumming the resulting series (3.2.12). Using that

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}\xi}{\left(2\cosh\frac{\xi}{2}\right)^{\ell}} = \frac{\Gamma^2(\ell/2)}{\Gamma(\ell)} \tag{3.3.28}$$

we find

$$Z_\ell \approx \frac{1}{\hbar} Z_\ell^{(0)}, \tag{3.3.29}$$

where

$$Z_\ell^{(0)} = \int \frac{\mathrm{d}p\mathrm{d}q}{2\pi} \mathrm{e}^{-\ell H_{\rm cl}(q,p)} = \frac{1}{2\pi}\frac{\Gamma^4(\ell/2)}{\Gamma^2(\ell)}\,. \tag{3.3.30}$$

Therefore,

$$J(\mu) \approx \frac{1}{k} J_0(\mu) \tag{3.3.31}$$

where

$$\begin{aligned} J_0(\mu) &= -\sum_{\ell=1}^{\infty} \frac{(-z)^\ell}{4\pi^2}\frac{\Gamma^4(\ell/2)}{\ell\Gamma^2(\ell)} \\ &= \frac{1}{4} z\, {}_3F_2\left(\frac{1}{2},\frac{1}{2},\frac{1}{2};1,\frac{3}{2};\frac{z^2}{16}\right) - \frac{z^2}{8\pi^2}\, {}_4F_3\left(1,1,1,1;\frac{3}{2},\frac{3}{2},2;\frac{z^2}{16}\right). \end{aligned} \tag{3.3.32}$$

This function has a branch cut in the z-plane at $(-\infty,-4]$. This is expected: indeed, from (3.2.24) there should be a cut starting at

$$z = -\lambda_0^{-1} = -4, \tag{3.3.33}$$

indicating the condensation of eigenvalues for the one-particle density matrix. The function (3.3.32) has the following asymptotics for large μ,

$$J_0(\mu) = \frac{2\mu^3}{3\pi^2} + \frac{\mu}{3} + \frac{2\zeta(3)}{\pi^2} + J_0^{\rm np}(\mu). \tag{3.3.34}$$

The leading, cubic term in μ, is the responsible for the behavior (3.3.27). The subleading term in μ gives a correction of order $N^{1/2}$ to the leading behavior (3.3.27). The last, non-perturbative term involves an infinite power series of exponentially small corrections in μ. They have the structure,

$$J_0^{\rm np}(\mu) = \sum_{\ell=1}^{\infty} \left(a_{0,\ell}\mu^2 + b_{0,\ell}\mu + c_{0,\ell}\right)\mathrm{e}^{-2\ell\mu}. \tag{3.3.35}$$

Explicitly, one finds for the very first orders

$$\begin{aligned} J_0^{\rm np}(\mu) = &\frac{2}{3\pi^2}\left(6-\pi^2+6\mu-6\mu^2\right)\mathrm{e}^{-2\mu} + \frac{1}{2\pi^2}\left(25-6\pi^2-66\mu-36\mu^2\right)\mathrm{e}^{-4\mu} \\ &+ \mathcal{O}\left(\mu^2\mathrm{e}^{-6\mu}\right). \end{aligned} \tag{3.3.36}$$

The non-perturbative part of the grand potential leads to exponentially small corrections in N in the canonical free energy. In fact, using (3.3.10) we find that, once evaluated at the saddle-point,

$$\exp(-2\mu) \approx \exp\left(-\sqrt{2}\pi k^{1/2} N^{1/2}\right). \tag{3.3.37}$$

This is precisely the action for membrane instantons (3.1.3) found in the previous chapter as large N instantons of the matrix model in the 't Hooft expansion. We conclude that the exponentially small corrections in μ, which in this approach appear already in the semi-classical approximation, correspond in fact to non-perturbative corrections in the genus expansion, and should be identified as membrane instanton contributions.

As mentioned before, the calculation of these exponentially small corrections to the grand-canonical potential is equivalent to the exact calculation of the volume (3.3.13) of classical phase space. To see this, we notice that we can write this volume as a period of the one-form $p\mathrm{d}q$ along the curve (3.3.14)

$$\mathrm{Vol}(E) = \oint_{H_{\mathrm{cl}}(q,p)=E} p\mathrm{d}q. \tag{3.3.38}$$

This period vanishes at the point $E = E_0$. It turns out that its exact value is given by a Meijer function,

$$\mathrm{Vol}(E) = \frac{\mathrm{e}^E}{\pi} G^{2,3}_{3,3}\left(\frac{\mathrm{e}^{2E}}{16}\,\middle|\, \begin{matrix} \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \\ 0, 0, -\frac{1}{2} \end{matrix}\right) - 4\pi^2. \tag{3.3.39}$$

This leads to the following large E expansion of the number of states,

$$n(E) = \frac{2E^2}{\pi^2 k} - \frac{1}{3k} + \mathcal{O}(E\mathrm{e}^{-2E}) + \mathcal{O}(\hbar), \tag{3.3.40}$$

where the first term agrees of course with the semiclassical calculation at large E done before. One can then check that the expression (3.3.6) for the grand-canonical potential reproduces (3.3.32), once the density obtained from (3.3.39) is used.

3.3.4 Relation to previous results

The semiclassical limit of the one-particle Hamiltonian turns out to be closely related to the planar limit of ABJM theory studied in [5, 6, 8].

First of all, the semiclassical quantization of the one-dimensional problem leads to the Fermi surface

$$T(p) + U(q) = E \tag{3.3.41}$$

which is in fact a *curve* in phase space. Let us now make the following change of variables,

$$x = \frac{q}{2} + \frac{p}{2}, \qquad y = p + \pi\mathrm{i}, \tag{3.3.42}$$

which, up to an overall constant, preserves the form $p\mathrm{d}q$. In terms of the exponentiated variables

$$X = \mathrm{e}^{q/2+p/2}, \qquad Y = -\mathrm{e}^p. \tag{3.3.43}$$

the Fermi surface (3.3.41) reads

$$Y + \frac{X^2}{Y} - X^2 + \mathrm{i}\kappa X - 1 = 0, \tag{3.3.44}$$

where

$$\mathrm{i}\kappa = \mathrm{e}^E. \tag{3.3.45}$$

The curve (3.3.44) is nothing but the spectral curve (2.3.1) of the ABJM matrix model. The minimal energy E_0 given in (3.3.18) corresponds to the conifold point $\kappa = 4\mathrm{i}$ studied in detail in the previous chapter. The volume of phase space, which as we remarked after (3.3.38) is a vanishing period at $E = E_0$, is actually proportional to the conifold period studied in section 2.13. Finally, the large energy limit, in which the Fermi surface becomes a polygon, is nothing but the tropical limit of the spectral curve studied in the previous chapter.

3.4 Quantum corrections

In the previous section we have recalled the semiclassical limit of ideal Fermi gases, and we have studied in detail the case of ABJM theory. We now study the corrections to the semiclassical limit in a systematic and general way. These corrections lead to a power series in $\hbar^2 \propto k^2$ for the grand-canonical partition function. As we will see, only the first $\hbar^2$ correction contributes to the asymptotic series in $1/N$ of the canonical free energy, up to an additive function of k but independent of N. This means that we can compute the *full* series of $1/N$ corrections to the original matrix model partition function, up to an overall, N-independent constant. However, the exponentially small terms in μ appearing in $J(\mu)$ receive corrections to all orders in $\hbar^2$.

3.4.1 Quantum-corrected Hamiltonian and Wigner–Kirkwood expansion

There are two sources of $\hbar$ corrections in the one-body problem appearing in our Fermi gas formulation. The first one appears already in the Hamiltonian $\hat{H}$: when we compute $\hat{H}$ starting from (3.2.27), the non-commutativity of the operators in (3.2.27) leads to $\mathcal{O}(\hbar)$ corrections to (3.2.34). This first source of corrections is nicely encoded in the Wigner transform (3.2.44). Another source of corrections is due to the standard semiclassical expansion of the density of eigenvalues. We now present a formalism to treat in a systematic way both types of corrections. This formalism is a generalization of the standard Wigner–Kirkwood $\hbar$ expansion [132, 133] in quantum statistical mechanics, and it incorporates general, $\hbar$-dependent Hamiltonians. As in section 3.2.1, the formalism is most conveniently formulated in the phase-space approach to quantization, and it has been developed in the context of many-body physics. The most elegant presentation is due to Voros [134, 135] (see also [136]).

Let $\hat{H}$ be the Hamiltonian of a one-particle, one-dimensional quantum system, and let $H_{\rm W}$ be its Wigner transform. We would like to compute systematically the $\hbar$ expansion of the canonical partition function and of the density of states. Following [135] we notice that it is possible to expand any function $f(\hat{H})$ of $\hat{H}$ around $H_{\rm W}(q,p)$, which is a c-number. This gives,

$$f(\hat{H}) = \sum_{r\geq 0} \frac{1}{r!} f^{(r)}(H_{\rm W}) \left(\hat{H} - H_{\rm W}(q,p)\right)^r . \tag{3.4.1}$$

The semiclassical expansion of this object is obtained simply by evaluating its Wigner transform, and we obtain

$$f(\hat{H})_{\rm W} = \sum_{r\geq 0} \frac{1}{r!} f^{(r)}\left(H_{\rm W}\right) \mathcal{G}_r \tag{3.4.2}$$

where

$$\mathcal{G}_r = \left[\left(\hat{H} - H_{\rm W}(q,p)\right)^r\right]_{\rm W} \tag{3.4.3}$$

and the Wigner transform is evaluated at the same point q, p. Of course, one has

$$\mathcal{G}_0 = 1, \qquad \mathcal{G}_1 = 0, \tag{3.4.4}$$

and the quantities $\mathcal{G}_r$ for $r \geq 2$ can be computed again by using (3.2.36). They have an $\hbar$ expansion of the form

$$\mathcal{G}_r = \sum_{n\geq \left[\frac{r+2}{3}\right]} \hbar^{2n} \mathcal{G}_r^{(n)}, \qquad r \geq 2. \tag{3.4.5}$$

This means, in particular, that to any order in $\hbar^2$, only a finite number of $\mathcal{G}_r$'s are involved. One finds, for the very first orders [135, 136],

$$\begin{aligned} \mathcal{G}_2 &= -\frac{\hbar^2}{4}\left[\frac{\partial^2 H_{\rm W}}{\partial q^2}\frac{\partial^2 H_{\rm W}}{\partial p^2} - \left(\frac{\partial^2 H_{\rm W}}{\partial q \partial p}\right)^2\right] + \mathcal{O}(\hbar^4), \\ \mathcal{G}_3 &= -\frac{\hbar^2}{4}\left[\left(\frac{\partial H_{\rm W}}{\partial q}\right)^2 \frac{\partial^2 H_{\rm W}}{\partial p^2} + \left(\frac{\partial H_{\rm W}}{\partial p}\right)^2 \frac{\partial^2 H_{\rm W}}{\partial q^2} - 2\frac{\partial H_{\rm W}}{\partial q}\frac{\partial H_{\rm W}}{\partial p}\frac{\partial^2 H_{\rm W}}{\partial q \partial p}\right] + \mathcal{O}(\hbar^4). \end{aligned} \tag{3.4.6}$$

One can then apply this method to compute the semiclassical expansion of any function of the Hamiltonian operator. For example, when applied to (3.3.2), one finds,

$$\hat{n}(E)_{\rm W} = \theta(E - H_{\rm W}) + \sum_{r=2}^{\infty} \frac{1}{r!} \mathcal{G}_r \delta^{(r-1)}(E - H_{\rm W}), \tag{3.4.7}$$

therefore

$$n(E) = \int_{H_{\rm W}(q,p)\leq E} \frac{{\rm d}q{\rm d}p}{2\pi\hbar} + \sum_{r=2}^{\infty} \frac{1}{r!} \int \frac{{\rm d}q{\rm d}p}{2\pi\hbar} \mathcal{G}_r \delta^{(r-1)}(E - H_{\rm W}). \tag{3.4.8}$$

When applied to the canonical density matrix at inverse temperature β, one finds,

$$\left({\rm e}^{-\beta\hat{H}}\right)_{\rm W} = \left(\sum_{r=0}^{\infty} \frac{(-\beta)^r}{r!} \mathcal{G}_r\right) {\rm e}^{-\beta H_{\rm W}}. \tag{3.4.9}$$

The standard Wigner–Kirkwood $\hbar$ expansion of the canonical partition function [132, 133] is just a particular case of (3.4.9) when the Hamiltonian is

$$\hat{H} = \frac{\hat{p}^2}{2} + U(\hat{q}). \tag{3.4.10}$$

Let us now apply this formalism to our case. First of all, the quantum-corrected Hamiltonian is given by a power series in $\hbar$ of the form,

$$H_{\mathrm{W}} = \sum_{n\geq 0} \hbar^{2n} H_{\mathrm{W}}^{(n)}. \tag{3.4.11}$$

At leading order we find of course the classical Hamiltonian (3.3.17),

$$H_{\mathrm{W}}^{(0)} = T(p) + U(q), \tag{3.4.12}$$

the $\mathcal{O}(\hbar^2)$ term is written down in (3.2.44), and the $\mathcal{O}(\hbar^4)$ term can be found in (B.1). The one-particle canonical partition function can then be computed as a power series in $\hbar$,

$$Z_\ell = \frac{1}{\hbar} \sum_{n=0}^{\infty} Z_\ell^{(n)} \hbar^{2n}, \tag{3.4.13}$$

where

$$Z_\ell^{(0)} = \int \frac{\mathrm{d}q \mathrm{d}p}{2\pi} \mathrm{e}^{-\ell H_{\mathrm{cl}}} \tag{3.4.14}$$

is the classical limit. The expansion is obtained by grouping $\hbar^2$ corrections in the expression

$$Z_\ell = \frac{1}{\hbar} \sum_{r\geq 0} \frac{(-\ell)^r}{r!} \int \frac{\mathrm{d}q\, \mathrm{d}p}{2\pi} \mathcal{G}_r \mathrm{e}^{-\ell H_{\mathrm{W}}}. \tag{3.4.15}$$

The power series in $\hbar$ for Z_ℓ leads to the following power series in k for $J(\mu)$,

$$J(\mu) = \frac{1}{k} \sum_{n=0}^{\infty} J_n(\mu) k^{2n}, \tag{3.4.16}$$

where

$$J_n(\mu) = -(2\pi)^{2n-1} \sum_{\ell=1}^{\infty} \frac{(-z)^\ell}{\ell} Z_\ell^{(n)}. \tag{3.4.17}$$

As an illustration of the above, general considerations, we will now calculate the first, $\hbar^2$ correction to the semiclassical result of ABJM theory obtained in section 3.3.3. Using the formulae above, we find

$$\begin{aligned} Z_\ell^{(1)} &= \int \frac{\mathrm{d}q \mathrm{d}p}{2\pi} \mathrm{e}^{-\ell H_{\mathrm{cl}}} \left\{ -\ell H_{\mathrm{W}}^{(1)} + \frac{\ell^2}{2} \mathcal{G}_2^{(1)} - \frac{\ell^3}{6} \mathcal{G}_3^{(1)} \right\} \\ &= -\ell \int \frac{\mathrm{d}q \mathrm{d}p}{2\pi} \mathrm{e}^{-\ell H_{\mathrm{cl}}} \left[\frac{1}{24} (U'(q))^2 T''(p) - \frac{1}{12} (T'(p))^2 U''(q) \right] \\ &+ \int \frac{\mathrm{d}q \mathrm{d}p}{2\pi} \mathrm{e}^{-\ell H_{\mathrm{cl}}} \left\{ \frac{\ell^3}{24} \left[(U'(q))^2 T''(p) + U''(q) (T'(p))^2 \right] - \frac{\ell^2}{8} U''(q) T''(p) \right\}. \end{aligned} \tag{3.4.18}$$

To evaluate these coefficients, we need the integral appearing in (3.3.28), as well as

$$\int_{-\infty}^{\infty} \mathrm{d}\xi \frac{\tanh^2(\xi/2)}{(2\cosh(\xi/2))^\ell} = \frac{\Gamma^2(\ell/2)}{\Gamma(\ell)} - 4 \frac{\Gamma^2(\ell/2+1)}{\Gamma(\ell+2)}. \tag{3.4.19}$$

We then find,

$$Z_\ell^{(1)} = \frac{\ell}{48\pi} (2\ell^2+1) \left[\frac{\Gamma^2(\ell/2+1)\Gamma^2(\ell/2)}{4\Gamma(\ell+2)\Gamma(\ell)} - \frac{\Gamma^4(\ell/2+1)}{\Gamma^2(\ell+2)} \right] - \frac{\ell^2}{16\pi} \frac{\Gamma^4(\ell/2+1)}{\Gamma^2(\ell+2)}. \tag{3.4.20}$$

From (3.4.20) one can compute $J_1(z)$ in closed form. Let us introduce the function

$$\begin{aligned} f(z) = {}& {}_3F_2\left(1,1,1; \frac{3}{2}, \frac{3}{2}; \frac{z^2}{16}\right) - \frac{z^2}{24} {}_3F_2\left(1,1,2; \frac{3}{2}, \frac{5}{2}; \frac{z^2}{16}\right) \\ &+ \frac{1}{z}\left(-2\pi E\left(\frac{z}{4}\right) - z + \pi^2\right) \end{aligned} \tag{3.4.21}$$

where $E(k)$ is the complete elliptic integral of the second kind with modulus k. Then, one finds

$$J_1(\mu) = \frac{1}{24}\left\{ f(z) - \left(z\frac{\partial}{\partial z}\right)^2 f(z)\right\}. \tag{3.4.22}$$

The asymptotic expansion of the above function at large μ is given by

$$f(z) - \left(z\frac{\partial}{\partial z}\right)^2 f(z) = \mu - 2 + \mathcal{O}\left(\mu^2 e^{-2\mu}\right). \tag{3.4.23}$$

Therefore, we find, at next-to-leading order in k, the following expression for the grand canonical potential of ABJM theory,

$$J_{\rm ABJM}(\mu) \approx \frac{2\mu^3}{3k\pi^2} + \mu\left(\frac{1}{3k} + \frac{k}{24}\right) + \frac{2\zeta(3)}{\pi^2 k} - \frac{k}{12} + \mathcal{O}\left(\mu^2 e^{-2\mu}\right). \tag{3.4.24}$$

Notice that the non-perturbative corrections in μ to (3.4.23) involve only *even* powers of z. This is consistent with their interpretation as membrane instantons.

3.4.2 General structure of quantum corrections

As we mentioned above, we can compute the quantum corrections to $J(\mu)$ either by working out the corrections to the Z_ℓ integrals, or by working out the corrections to the function $n(E)$. In order to understand the general structure of these corrections for a Fermi gas, the second point of view is more convenient. In this section we will analyze this general structure in detail, and we will make a precise connection between the structure of $n(E)$ and the expected Airy function behavior.

First of all, we have to understand more precisely the relationship between the structure of $n(E)$ and the structure of $J(\mu)$. Let us write the density function $n(E)$ in the form,

$$n(E) = CE^2 + n_0 + n_{\rm np}(E), \tag{3.4.25}$$

where the last term has the following asymptotics at infinity,

$$n_{\rm np}(E) = \mathcal{O}(E e^{-E}), \quad E \to \infty. \tag{3.4.26}$$

We know from (3.3.40) that this is indeed the case at leading order in k for ABJM theory and in the next subsection we will show that quantum corrections do not spoil this behavior. Notice that, since all eigenvalues of our Hamiltonian are positive, we must have

$$n(0) = 0, \tag{3.4.27}$$

therefore

$$n_{\rm np}(0) = -n_0. \tag{3.4.28}$$

If we now plug (3.4.25) in (3.3.6) we find,

$$\begin{aligned} J(\mu) &= \int_0^\infty dE\, n'(E) \log(1 + z e^{-E}) \\ &= -2C\,{\rm Li}_3(-z) + \mu \int_0^\infty dE\, n'_{\rm np}(E) - \int_0^\infty dE\, n'_{\rm np}(E) E + \int_0^\infty dE\, n'_{\rm np}(E) \log(1 + e^E/z). \end{aligned} \tag{3.4.29}$$

The second integral gives

$$\int_0^\infty dE\, n'_{\rm np}(E) = -n_{\rm np}(0) = n_0, \tag{3.4.30}$$

where we used (3.4.28). The last term can be calculated as

$$\int_0^\infty dE\, n'_{\rm np}(E) \log(1 + e^E/z) = n_0 \log(1 + 1/z) - \int_0^\infty dE \frac{n_{\rm np}(E)}{1 + z e^{-E}}, \tag{3.4.31}$$

and both terms are non-perturbative in μ. Indeed,

$$\int_0^\infty \mathrm{d}E \frac{n_{\mathrm{np}}(E)}{1+z\mathrm{e}^{-E}} \sim \int_0^\infty \mathrm{d}E \frac{E\mathrm{e}^{-E}}{1+z\mathrm{e}^{-E}} = \mathcal{O}\left(\mu\, \mathrm{e}^{-\mu}\right). \tag{3.4.32}$$

Then, by using the standard asymptotics of the trilogarithm

$$\mathrm{Li}_3(-z) = -\frac{\mu^3}{6} - \frac{\pi^2}{6}\mu + \mathcal{O}\left(\mathrm{e}^{-\mu}\right), \tag{3.4.33}$$

we deduce the following asymptotic expansion of $J(\mu)$ for large μ:

$$J(\mu) = \frac{C}{3}\mu^3 + B\mu + A + J_{\mathrm{np}}(\mu), \tag{3.4.34}$$

where

$$\begin{aligned} B &= n_0 + \frac{\pi^2 C}{3}, \\ A &= -\operatorname{Tr}' \hat{H} \equiv -\int_0^\infty \mathrm{d}E\, E\, n'_{\mathrm{np}}(E), \end{aligned} \tag{3.4.35}$$

and

$$J_{\mathrm{np}}(\mu) = \mathcal{O}\left(\mu\, \mathrm{e}^{-\mu}\right). \tag{3.4.36}$$

Notice that A is a non-trivial function of k, but it doesn't depend on μ. If we now plug this in (3.2.14), we find immediately

$$Z(N) = C^{-1/3}\mathrm{e}^A \operatorname{Ai}\left[C^{-1/3}(N-B)\right] + Z_{\mathrm{np}}(N), \tag{3.4.37}$$

where the last term is non-perturbative in N.

We then see that, if we are able to derive the structural results (3.4.25) and (3.4.26) for the density of states of a given theory, the conjecture 0.2.15 for the M-theory expansion is proved. In fact, so far we have not specified in which regime we are working in k. In practice, we have to work in an expansion in k around $k = 0$. However, we expect that C will only get contributions at leading order in k (i.e. the strict semiclassical limit), and that B will be only corrected at the next-to-leading order in k. We will now verify this in ABJM theory. In contrast, the μ-independent term A gets corrected at all orders in k.

3.4.3 Quantum corrections in ABJM theory

We now study the general structure of quantum corrections in ABJM theory, by using the strategy explained above, i.e. by looking at the number of eigenvalues $n(E)$. Our goal is to show that $n(E)$ has the structure (3.4.25). This involves a somewhat detailed argument. Since not every reader might go through it, we want to emphasize that the physics behind this argument is very simple. The WKB expansion of the density of eigenvalues of a quantum system is in fact an expansion in

$$\left(\hbar \frac{\mathrm{d}}{\mathrm{d}E}\right)^2. \tag{3.4.38}$$

Therefore, if the leading order term in $n(E)$ is of the form CE^2, the first quantum correction gives the constant term in (3.4.25), and further terms in the WKB expansion do not correct the polynomial part of $n(E)$. They can only give exponentially small corrections in E. In the rest of this section, we will verify that this qualitative argument is actually correct in the case of the one-body problem appearing in ABJM theory.

Our starting point in the study of quantum corrections in ABJM theory is (3.4.8). As we know, there are two sources of $\hbar$ corrections in this formula. One is the quantum-corrected Hamiltonian, and the other are the terms $\mathcal{G}_r$ appearing in the generalized Wigner–Kirkwood expansion. We will consider first the quantum corrections coming from H_{W}, i.e. from the first term in (3.4.8). Since we have a symmetry $q \to -q$ and $p \to -p$ in the problem, we can restrict ourselves to the case $q \geqslant 0$ and $p \geqslant 0$. We want to solve the equation

$$H_{\mathrm{W}}(q,p) = E \tag{3.4.39}$$

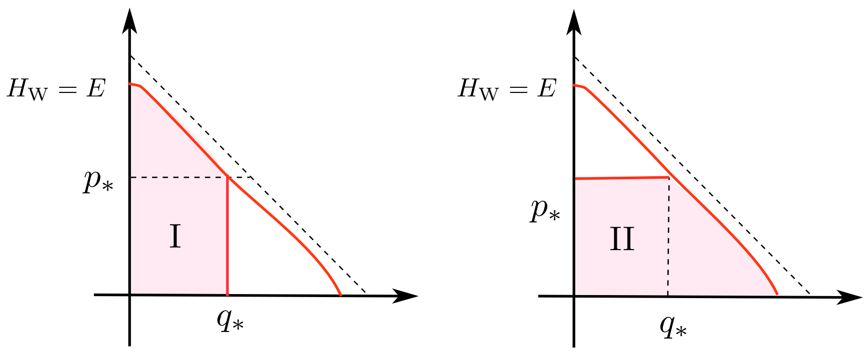

Figure 3.4: The regions I (left) and II (right) under the quantum curve $H_W(q,p) = E$ in the positive quadrant. The diagonal dashed line is the polygonal curve (3.3.22).

in the limit $E \to \infty$. This defines a "quantum curve" or "quantum Fermi surface," including explicit $\hbar$ corrections. At leading order in E the curve is given by (3.3.22), and the corresponding domain (in the positive quadrant) has volume

$$\mathrm{Vol}_0(E) = 2E^2. \tag{3.4.40}$$

One crucial ingredient in what follows is the fact that the function $U(q)$ and its derivatives have the following asymptotics as $q \gg 1$:

$$\begin{aligned} U(q) &= \log 2\cosh\frac{q}{2} = \frac{q}{2} + \sum_{k\geqslant 1}\frac{(-1)^{k+1}}{k}\mathrm{e}^{-kq}, \\ U'(q) &= \frac{1}{2}\tanh\frac{q}{2} = \frac{1}{2} + \sum_{k\geqslant 1}(-1)^k\mathrm{e}^{-kq}, \\ U''(q) &= \frac{1}{4\cosh^2\frac{q}{2}} = \sum_{k\geqslant 1}k(-1)^{k+1}\mathrm{e}^{-kq}. \end{aligned} \tag{3.4.41}$$

The same results hold for $T(p)$. Notice that, if we take a number large enough of derivatives of these functions, they become exponentially suppressed at infinity. This will be eventually the source of the simplifications at large E.

Let us now consider the point (q_*, p_*) in the curve (3.4.39), where

$$p_* = E. \tag{3.4.42}$$

It is easy to see from the explicit form of H_W that

$$q_* = E + \mathcal{O}\left(\mathrm{e}^{-E}\right) \tag{3.4.43}$$

where the exponentially small corrections in E are themselves power series in $\hbar^2$. This point divides the curve (3.4.39) into two segments, and defines two regions for the fully corrected volume, as shown in Fig. 3.4. Region I is defined as the region under the quantum curve with $p \geq p_*$, while region II is defined by $q \geq q_*$. We have

$$\mathrm{Vol}(E) = 4\mathrm{Vol}_{\mathrm{I}}(E) + 4\mathrm{Vol}_{\mathrm{II}}(E), \tag{3.4.44}$$

where

$$\mathrm{Vol}_{\mathrm{I}}(E) = \int_0^{q_*(E)} p(E,q)\,\mathrm{d}q, \qquad \mathrm{Vol}_{\mathrm{II}}(E) = \int_0^{p_*(E)} q(E,p)\mathrm{d}p - p_*q_*, \tag{3.4.45}$$

and $p(E,q)$ and $q(E,p)$ are local solutions of $H_W(q,p) = E$.

Let us first consider the curve bounding region I. Along this curve, $p(q,E) \geq E$, therefore exponential terms in p in H_W are bounded by exponential terms in E. We can then write

$$T(p) = \frac{p}{2} + \mathcal{O}\left(\mathrm{e}^{-E}\right), \qquad T'(p) = \frac{1}{2} + \mathcal{O}\left(\mathrm{e}^{-E}\right), \tag{3.4.46}$$

and

$$T^{(n)}(p) = \mathcal{O}\left(\mathrm{e}^{-E}\right), \quad n \geq 2. \tag{3.4.47}$$

In the quantum corrections to the function $H_{\rm W}$ we will have terms of the form $(T^{(k)}(p))^n$, with $k \geq 1$. Due to (3.4.46) and (3.4.47), and neglecting exponentially small corrections of the form $\mathcal{O}\left(\mathrm{e}^{-E}\right)$, we should keep only the terms $(T'(p))^n$ with $n \geq 1$ (like the third term in the second line of (3.2.44)). But these terms always multiply terms of the form $U^{(2n)}(q)$. We conclude that, on the curve bounding region I,

$$H_{\rm W} = \frac{p}{2} + U(q) - \frac{\hbar^2}{48}U''(q) + \frac{1}{2}\sum_{n>1} \hbar^{2n} c_n U^{(2n)}(q) + \mathcal{O}\left(\mathrm{e}^{-E}\right). \tag{3.4.48}$$

The third term in this expression comes from the third term in the second line of (3.2.44). The fourth term comes from higher quantum corrections (see the first term in the last line of (B.2) for an example of such a term at order $\mathcal{O}(\hbar^4)$). We can now solve for p along this curve,

$$p(E,q) = 2E - q + \Delta p(E,q), \tag{3.4.49}$$

where

$$\Delta p(E,q) = q - 2U(q) + \frac{\hbar^2}{24}U''(q) - \sum_{n>1} \hbar^{2n} c_n U^{(2n)}(q) + \mathcal{O}(\mathrm{e}^{-E}). \tag{3.4.50}$$

We calculate the volume of region I as follows,

$$\mathrm{Vol}_{\rm I} = \mathrm{Vol}_{\rm I}^0 + \Delta\mathrm{Vol}_{\rm I}. \tag{3.4.51}$$

The first term comes from the polygonal limit of the curve,

$$\mathrm{Vol}_{\rm I}^0(E) = \int_0^{q_*(E)} (2E - q)\mathrm{d}q = 2Eq_*(E) - \frac{q_*^2(E)}{2}. \tag{3.4.52}$$

The second term comes from the corrections to the curve, and it is given by

$$\begin{aligned} \Delta\mathrm{Vol}_{\rm I}(E) &= \int_0^{q_*(E)} \Delta p(E,q)\mathrm{d}q \\ &- 2\int_0^{q_*(E)} \left(U(q) - \frac{q}{2}\right)\mathrm{d}q + \frac{\hbar^2}{24}\int_0^{q_*(E)} U''(q)\mathrm{d}q - \sum_{n>1}\hbar^{2n}c_n \int_0^{q_*(E)} U^{(2n)}(q) + \mathcal{O}\left(E\mathrm{e}^{-E}\right) \\ &= -2\int_0^{\infty} \left(U(q) - \frac{q}{2}\right)\mathrm{d}q + \frac{\hbar^2}{24}\int_0^{\infty} U''(q)\mathrm{d}q - \sum_{n>1}\hbar^{2n}c_n \int_0^{\infty} U^{(2n)}(q) + \mathcal{O}\left(E\mathrm{e}^{-E}\right) \\ &= -\frac{\pi^2}{6} + \frac{\hbar^2}{48} + \mathcal{O}\left(E\mathrm{e}^{-E}\right). \end{aligned} \tag{3.4.53}$$

In the last calculation we used that, up to non-perturbative terms in E, we can extend the integration region to infinity, and also that

$$\int_0^{\infty} U^{(2n)}(q)\mathrm{d}q = U^{(2n-1)}(\infty) - U^{(2n-1)}(0) = 0 \quad \text{for } n > 1. \tag{3.4.54}$$

A similar calculation can be done for region II. We obtain, from the polygonal approximation of the curve,

$$\mathrm{Vol}_{\rm II}^0(E) = 2Ep_*(E) - \frac{p_*^2(E)}{2} - p_*(E)q_*(E), \tag{3.4.55}$$

while the corrections give,

$$\Delta\mathrm{Vol}_{\rm II}(E) = -\frac{\pi^2}{6} - \frac{\hbar^2}{96} + \mathcal{O}\left(E\mathrm{e}^{-E}\right). \tag{3.4.56}$$

Using that

$$p_*(E) + q_*(E) = 2E + \mathcal{O}\left(\mathrm{e}^{-E}\right) \tag{3.4.57}$$

we finally get

$$\mathrm{Vol}(E) = 8E^2 - \frac{4\pi^2}{3} + \frac{\hbar^2}{24} + \mathcal{O}\left(E\mathrm{e}^{-E}\right). \tag{3.4.58}$$

We now consider the contribution from the quantum corrections to the density. In fact, these terms only give non-perturbative corrections in E. Using that

$$\delta(E - H_{\rm W}(q,p)) = \delta(p - p(E,q)) \Big/ \frac{\partial H_{\rm W}(q,p)}{\partial p} = \delta(q - q(E,p)) \Big/ \frac{\partial H_{\rm W}(q,p)}{\partial q} \tag{3.4.59}$$

one can always decompose an integral over the phase space as a sum of one-dimensional integrals in regions I and II, as in (3.4.44). For region I one can use again the expression (3.4.48) and the properties (3.4.46), (3.4.47). The only nontrivial term which gives an $\hbar^2$ correction comes from $\mathcal{G}_3$ and gives,

$$\begin{aligned}\frac{\hbar^2}{24}\frac{\partial^2}{\partial E^2}\int_0^{q_*(E)} \mathrm{d}q\, \frac{\partial H_{\rm W}(q,p)}{\partial p}\frac{\partial^2 H_{\rm W}(q,p)}{\partial q^2}\bigg|_{p=p(E,q)} &= \\ \frac{\hbar^2}{24}\frac{\partial^2}{\partial E^2}\int_0^{q_*(E)} \mathrm{d}q \left(\sum_{n\geq 0}\hbar^{2n} c_n U^{(2n+2)}(q) + \mathcal{O}(\mathrm{e}^{-E})\right) &= \mathcal{O}(E\mathrm{e}^{-E}).\end{aligned} \tag{3.4.60}$$

For higher order corrections everything that contains a term with

$$\frac{\partial^r H_{\rm W}(q,p)}{\partial p^r}, \qquad r>1, \tag{3.4.61}$$

or with

$$\frac{\partial^2 H_{\rm W}(q,p)}{\partial p \partial q} \tag{3.4.62}$$

is of order e^{-E}. Since the derivatives ∂_p and ∂_q always come in pairs, the only terms possibly contributing are of the form

$$\prod_i \left(\frac{\partial H_{\rm W}(q,p)}{\partial p}\right)^{n_i} \frac{\partial^{n_i} H_{\rm W}(q,p)}{\partial q^{n_i}} = \prod_i \frac{\partial^{n_i} H_{\rm W}(q,p)}{\partial q^{n_i}} + \mathcal{O}(\mathrm{e}^{-E}) \tag{3.4.63}$$

where $n_i \geq 2, \forall i$. After integrating and applying $\partial^r/\partial E^r$ this gives a correction of order $\mathcal{O}(E\mathrm{e}^{-E})$ by the same reason.

We conclude that, to all orders in the $\hbar$ expansion,

$$n(E) = \frac{\mathrm{Vol}(E)}{2\pi\hbar} + \mathcal{O}(E\mathrm{e}^{-E}) = \frac{2E^2}{\pi k} - \frac{1}{3k} + \frac{k}{24} + \mathcal{O}(E\mathrm{e}^{-E}). \tag{3.4.64}$$

Therefore, by using (3.4.34) and (3.4.35) we find the expression

$$J_{\rm ABJM}(\mu) = \frac{2\mu^3}{3k\pi^2} + \mu\left(\frac{1}{3k} + \frac{k}{24}\right) + A(k) + J_{\rm np}(\mu) \tag{3.4.65}$$

where

$$J_{\rm np}(\mu) = \sum_{\ell,n=1}^{\infty} \left(a_{\ell,n}\mu^2 + b_{\ell,n}\mu + c_{\ell,n}\right) k^{2n-3} \mathrm{e}^{-2\ell\mu}. \tag{3.4.66}$$

It is not manifest from the above results that this series involves only even powers of z^{-1}, but we have verified it to be the case for the first three orders in k, and we believe it is a general feature. Finally, it follows from (3.4.64) and (3.4.37) that

$$Z_{\rm ABJM}(N) = C^{-1/3}\mathrm{e}^{A(k)}\,\mathrm{Ai}\left[C^{-1/3}\left(N - \frac{1}{3k} - \frac{k}{24}\right)\right] + Z_{\rm np}(N), \tag{3.4.67}$$

where C is given in (3.1.8) and $Z_{\rm np}(N)$ are exponentially suppressed corrections at large N. This concludes our derivation of the Airy behavior for ABJM theory. The function $A(k)$ can in principle be determined, order by order in k, by computing the $Z_\ell^{(n)}$, resumming the resulting series, and expanding at $\mu = \infty$, as we did in sections 3.3.3 and 3.4.1. One obtains,

$$A(k) = \frac{2\zeta(3)}{\pi^2 k} - \frac{k}{12} - \frac{\pi^2 k^3}{4320} + \mathcal{O}(k^5). \tag{3.4.68}$$

A sketch of the computation leading to the third term of this expansion can be found in the Appendix B. In principle one can also compute $A(k)$ by using the representation (3.4.35).

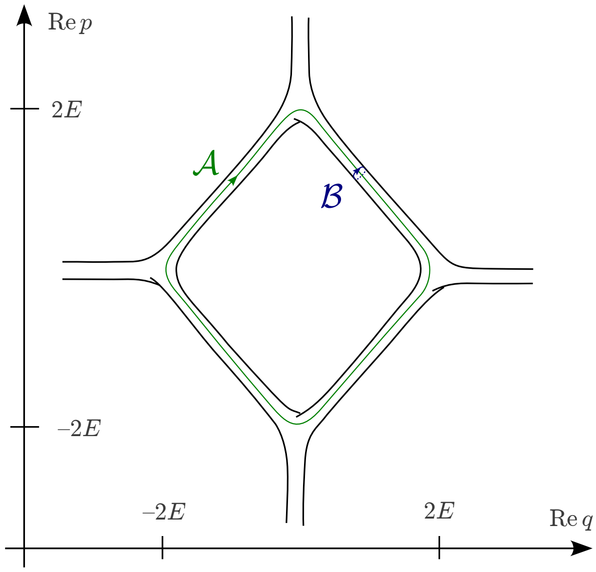

Figure 3.5: The Riemann surface of $H_{\rm cl}(p, q) = E$ for ABJM theory for large E. The four interior tubes form the limiting polygon of the Fermi surface.

3.4.4 Quantum-mechanical instantons as worldsheet instantons

One obvious question that one can ask at this point is the following: where are the worldsheet instantons (3.1.2) that one finds perturbatively in g_s in the 't Hooft expansion? We now give some preliminary evidence that worldsheet instantons correspond to the quantum-mechanical instantons of the Hamiltonian $H_{\rm W}$.

So far our focus has been in the perturbative corrections in $\hbar$, but one should expect generically non-perturbative corrections due to instantons, of order $\exp(-1/\hbar)$. To understand these quantum-mechanical instantons in our problem, with a non-conventional Hamiltonian, we need a general, geometric approach to non-perturbative WKB expansions, like the one proposed in [137, 138]. In this approach, instanton contributions are obtained by looking at the complexified curve

$$H(q, p) = E \tag{3.4.69}$$

where $H(q, p)$ is the Hamiltonian of the model. Perturbative WKB expansions are associated to periods of the above curve around "A-type" cycles, while non-perturbative corrections to the WKB method are associated to "B-type" cycles. In the case of ABJM theory, the complexified curve is identical to the spectral curve (3.3.44), after an appropriate choice of the variables. Its Riemann surface looks as shown in the Fig. 3.5. Let us introduce canonical coordinates Q, P related to the q, p coordinates as

$$Q = q, \qquad P = p + q. \tag{3.4.70}$$

This preserves the symplectic form. The coordinate P is chosen so that it has no monodromy along the contour $\mathcal{B}$. Then in the large E limit

$$n(E) \approx \frac{1}{2\pi\hbar}\oint_{\mathcal{A}} P\mathrm{d}Q \approx \frac{1}{2\pi\hbar}\mathrm{Vol}\left\{(q, p) : |p| + |q| < 2E\right\} = \frac{4E^2}{\pi\hbar}. \tag{3.4.71}$$

The instanton contribution is of order

$$\exp\left[\frac{\mathrm{i}}{\hbar}\oint_{\mathcal{B}} P\mathrm{d}Q\right], \tag{3.4.72}$$

where in the large E limit

$$\oint_{\mathcal{B}} P\mathrm{d}Q = 2E \cdot 4\pi\mathrm{i} + \mathcal{O}(\mathrm{e}^{-cE}). \tag{3.4.73}$$

Here we used that, in the interior of the upper-right tube, $P = p + q = 2E + \mathcal{O}(\mathrm{e}^{-cE})$ for some constant c, and that the monodromy of Q around the tube is $4\pi\mathrm{i}$. The above period can be computed exactly with the results of section 2.13 since the behavior (3.4.73) fixes it completely:

$$\begin{aligned}\oint_{\mathcal{B}} P\mathrm{d}Q &= -2\mathrm{i}\mathrm{e}^E\pi\,{}_3F_2\left(\frac{1}{2},\frac{1}{2},\frac{1}{2};1,\frac{3}{2};\frac{\mathrm{e}^{2E}}{16}\right)-\frac{\mathrm{e}^E}{\pi}G^{2,3}_{3,3}\left(\frac{\mathrm{e}^{2E}}{16}\,\middle|\,\begin{matrix}\frac{1}{2},\frac{1}{2},\frac{1}{2}\\ 0,0,-\frac{1}{2}\end{matrix}\right)+4\pi^2\\ &= 8\mathrm{i}\pi E+\mathcal{O}(\mathrm{e}^{-2E}).\end{aligned} \tag{3.4.74}$$

In fact, after the identification (3.3.45), this period is equal to $-4A_s$, where A_s is the strong coupling instanton action computed in section 2.13. For large energy, (3.4.74) gives a contribution to the density of states of order

$$\exp\left[-4E/k\right] \tag{3.4.75}$$

which becomes a contribution

$$\exp\left[-4\mu/k\right] \tag{3.4.76}$$

to the grand canonical potential, and a contribution

$$\sim \exp\left[-2\pi\sqrt{2N/k}\right] \tag{3.4.77}$$

to the canonical free energy. This is precisely the weight of a worldsheet instanton (3.1.2) in ABJM theory.

Quantum-mechanical instantons are of course invisible in the perturbative $\hbar$ expansion of H_{W} and in the Wigner–Kirkwood expansion, but they appear in the 't Hooft expansion. In fact, the 't Hooft expansion of the canonical free energy

$$F(\lambda,k)=\sum_{g\geq 0}k^{2-2g}F_g(\lambda) \tag{3.4.78}$$

leads to a genus expansion of the grand canonical potential of the form [119]

$$J_{\text{'t Hooft}}(\mu,k)=\sum_{g\geq 0}k^{2-2g}\mathcal{J}_g(\mu/k). \tag{3.4.79}$$

Notice that, in the Fermi gas approach, only the perturbative part in μ of $J(\mu)$ can be written in this form (3.4.79). The membrane instanton contributions and the function $A(k)$ do not have the right functional dependence in μ/k to fit into the 't Hooft expansion, while the weight associated to a quantum-mechanical instanton (3.4.76) is again of the right form. In the case of ABJM theory, we see from (3.4.24) that the Fermi gas approach gives

$$\begin{aligned}\mathcal{J}_0(\zeta)&=\frac{2\zeta^3}{3\pi^2}+\frac{\zeta}{24}+\mathcal{O}\left(\mathrm{e}^{-4\zeta}\right),\\ \mathcal{J}_1(\zeta)&=\frac{\zeta}{3}+\mathcal{O}\left(\mathrm{e}^{-4\zeta}\right),\\ \mathcal{J}_g(\zeta)&=\mathcal{O}\left(\mathrm{e}^{-4\zeta}\right),\qquad g\geq 2,\end{aligned} \tag{3.4.80}$$

where $\zeta=\mu/k$. From the point of view of the topological string, it follows from (3.1.20) that the variable ζ is essentially the period T at large radius, and a perturbative Fermi gas approach makes possible to recover the leading, perturbative genus zero and genus one free energies of the topological string given in (3.1.15).

Finally, we should mention that there is an extra source of worldsheet instanton-like corrections. In general, the exact representation (3.2.13) and the saddle-point integral (3.2.14) are only equivalent up to exponentially small corrections in N. Since we are taking into account such corrections, we have to be more careful here. The expression (3.2.13) is equivalent to

$$Z(N)=\frac{1}{2\pi\mathrm{i}}\int_{\mu_*-\mathrm{i}\pi}^{\mu_*+\mathrm{i}\pi}\mathrm{d}\mu\,\exp\left[J(\mu)-\mu N\right], \tag{3.4.81}$$

where the integration contour is parallel to the imaginary axis, and μ_* is arbitrary. To apply the saddle-point method, one chooses for μ_* the saddle-point of the exponent, and then extends the integration contour to infinity along the imaginary axis (this is what gives the Airy function behavior we have found many times in this paper). As it is well-known, it is in this last step of extending the integration contour

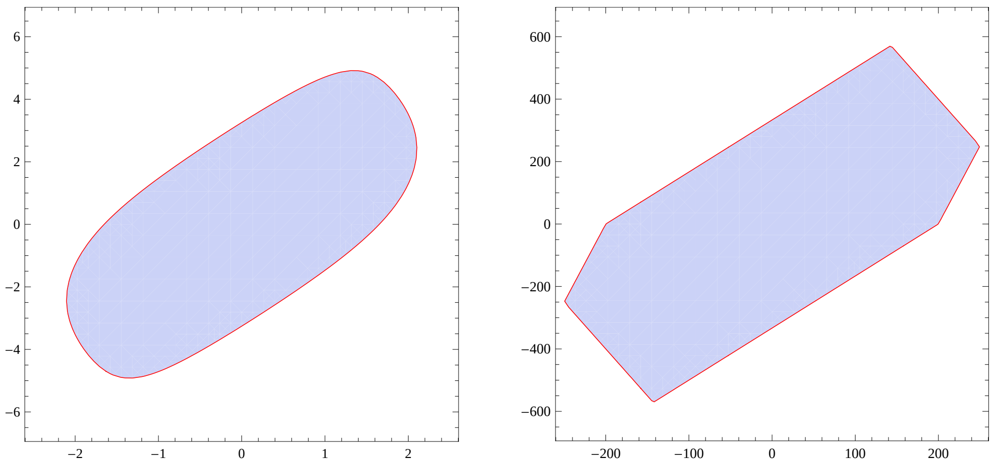

Figure 3.6: The Fermi surface for the three-node quiver with $n_a = (1, 3, -4)$ in the q-p plane, for $E = 5$ (left) and $E = 500$ (right). At large energy it approaches the polygon (3.5.1).

that one introduces exponentially small errors in N. A rough estimate of these errors can be done as follows. The saddle-point expansion involves integrating a Gaussian of the form

$$\exp\left[\frac{1}{2}(\mu-\mu_*)^2 J''(\mu_*)\right]. \tag{3.4.82}$$

The error in going to (3.2.14) can then be estimated by evaluating this Gaussian at the true endpoints in (3.4.81). This gives, by using the leading term in (3.4.24),

$$\sim \exp\left[-2\mu_*/k\right] \tag{3.4.83}$$

which is the square root of (3.4.76). Therefore, these type of corrections should also be taken in account when trying to extract information about worldsheet instantons.

3.5 More general Chern–Simons–matter theories

In this section we consider in detail more general CSM theories. We first study the thermodynamic limit of necklace quivers, and derive a general formula for the large N limit of their free energy which agrees with the result obtained in [115, 116] by analyzing the matrix model. Then we extend the considerations of section 3.4.3 to the general necklace CSM theories considered in section 3.2.2. For technical reasons we restrict ourselves to theories whose Hamiltonian is Hermitian, i.e. whose free energy is real, and we show that, with that assumption, the Airy behavior of the resummed $1/N$ expansion found in [23] is indeed generic. These general considerations are then illustrated in detail in the case of the ABJM theory (i.e. the two-node theory) with fundamental matter. Finally, we consider the massive theories of [176], and we derive the $N^{5/3}$ behavior found in [42, 139] with our techniques.

3.5.1 Thermodynamic limit for general necklace quivers

In this section we study the general necklace quiver considered in subsection 3.2.2. From (3.2.55) it follows that, for large energy, the Fermi surface is defined by the polygonal equation

$$\sum_{j=1}^{r}\left|p-\left(\sum_{i=1}^{j-1}n_i\right)q\right| + \left(\sum_{j=1}^{r}\frac{N_{f_j}}{k}\right)|q| = 2E. \tag{3.5.1}$$

As an example, we show in Fig. 3.6 the classical Fermi surface (3.5.1) at small and large energy, for a three-node quiver with $n_a = (1, 3, -4)$. By the by now familiar argument of the previous sections, the number of eigenstates is given by the semiclassical formula (3.3.13) applied to the domain bounded by (3.5.1). One finds,

$$n(E) \approx CE^2, \tag{3.5.2}$$

where the constant C is given by

$$\pi^2 C = \text{ Vol}\left\{(x,y)\ :\ \sum_{j=1}^{r}\left|y-\left(\sum_{i=1}^{j-1}k_i\right)x\right|+\left(\sum_{j=1}^{r}N_{f_j}\right)|x|<1\right\} \tag{3.5.3}$$

and the variables y, x differ from p, q in (3.5.1) by rescaling. Once this constant has been determined, the large N asymptotics of the free energy is given by (3.3.11)

$$F(N) \approx -\frac{2}{3}C^{-1/2}N^{3/2}. \tag{3.5.4}$$

In order to compute C, we notice that the right hand side of (3.5.3) is the volume of a convex polygon which can be easily calculated. Suppose for simplicity that $N_{f_j} = 0$. Let us introduce new parameters c_j, related to k_i in the following way

$$c_{\sigma(j)} = c + \sum_{i=1}^{j-1} k_i \tag{3.5.5}$$

where c is an auxiliary constant and σ is a permutation chosen so that $c_i \leq c_{i+1}, \forall i$. Then

$$\pi^2 C = \text{Vol}\left\{(x,y)\ :\ \sum_{j=1}^{r}|y-c_j x|<1\right\}. \tag{3.5.6}$$

Let us note that the c_j differ by a permutation from the parameters q_a introduced in [38] (where they were defined in such a way that $k_a = q_{a+1} - q_a$). Notice also that the expression (3.5.6) is explicitly invariant under permutations of the c_j. The inequality $\sum_{j=1}^{r}|y - c_j x| < 1$ defines a convex hull of $2r$ points $(\pm x_s, \pm y_s)$ so that

$$y_s = c_s x_s, \qquad \sum_{j=1}^{r}|y_s - c_j x_s| = 1. \tag{3.5.7}$$

One finds

$$x_s = \frac{1}{\sum_{j=1}^{r}|c_s - c_j|}, \qquad y_s = \frac{c_s}{\sum_{j=1}^{r}|c_s - c_j|}. \tag{3.5.8}$$

Then one can use the standard formula for the area of a convex hull to find

$$\pi^2 C = \sum_{s=1}^{r}\frac{|c_{s+1}-c_s|}{\left(\sum_{j=1}^{r}|c_{s+1}-c_j|\right)\left(\sum_{j=1}^{r}|c_s-c_j|\right)} \tag{3.5.9}$$

where as usual we use the convention $c_{r+1} \equiv c_1$. Let us illustrate this formula by applying it to necklaces with three and four nodes. For the necklace with three nodes (k_1, k_2, k_3) we can assume, without loss of generality, that

$$|c_1 - c_2| = |k_3|, \qquad |c_2 - c_3| = |k_1|, \qquad |c_1 - c_3| = |k_2|. \tag{3.5.10}$$

Then

$$\frac{\pi^2 C}{2} = \frac{|k_1||k_2| + |k_2||k_3| + |k_3||k_1|}{(|k_1|+|k_2|)(|k_2|+|k_3|)(|k_3|+|k_1|)}. \tag{3.5.11}$$

For the quiver with four nodes, let us assume without loss of generality that $\sum_{i=1}^{4} c_i = 0$. Then an easy computation gives

$$\frac{\pi^2 C}{2} = \frac{1}{32}\left(\frac{1}{c_1} - \frac{1}{c_4} + 12\,\frac{1}{c_3+c_4} + 4\,\frac{c_1+c_3}{(c_3+c_4)^2}\right). \tag{3.5.12}$$

These formulae for the three and four-node quivers agree with the results first found in [38] (for the four-node quiver, their formula is obtained by setting $c_1 = q_3$, $c_2 = q_1$, $c_3 = q_2$, $c_4 = q_4$). In fact, the above general result for the free energy of these CSM theories, involving the area of the polygon (3.5.3), has been derived in this form in [115, 116] by refining the analysis of the matrix model done in [38] (where a different, but equivalent general formula for the free energy was proposed). This class of CSM theories is dual to M-theory on $\text{AdS}_4 \times X_7$, where X_7 is an appropriate tri-Sasaki Einstein space [127].

Therefore, the coefficient C should be proportional to the volume of the X_7 manifold, and one should have

$$\frac{\mathrm{Vol}(X_7)}{\mathrm{Vol}(\mathbb{S}^7)} = \frac{\pi^2 C}{2} = \frac{1}{2}\,\mathrm{Vol}\left\{(x,y)\,:\,\sum_{j=1}^{p}\left|y-\left(\sum_{i=1}^{j-1}k_i\right)x\right| + \left(\sum_{j=1}^{p}N_{f_j}\right)|x|<1\right\}. \tag{3.5.13}$$

This is indeed the case, as it was proved in [115].

We then see that the Fermi gas approach allows us to rederive the result for the large N energy obtained in [115], but in a simpler way. The polygon appearing in the matrix model analysis of [115] has here a very simple interpretation: it is the large energy limit of the Fermi surface for the ideal Fermi gas.

3.5.2 Airy function behavior for a class of CSM theories

We now extend the considerations of section 3.4.3 to more general necklace quivers with matter. In the next subsection we apply the general considerations developed here to the case of ABJM theory with matter.

Let us assume that the Wigner transform of the density matrix can be written in the form

$$\rho_{\mathrm{W}}(q,p) \equiv \exp_\star\{-H_W\} = \mathrm{e}^{-\Phi_1(Q_{R_1})}\star\mathrm{e}^{-\Phi_2(Q_{R_2})}\star\cdots\star\mathrm{e}^{-\Phi_m(Q_{R_m})}, \tag{3.5.14}$$

where

$$Q_R = a_R q + b_R p \tag{3.5.15}$$

for suitable a_R, b_R, and the different Q_R are given by linearly independent combinations of q,p. We will also suppose that the functions Φ_i are real valued, even, and that[1]

$$\Phi_i(Q) = \gamma_i|Q| + \mathcal{O}(\mathrm{e}^{-c|Q|}), \qquad Q\to\infty, \quad \gamma_i>0. \tag{3.5.16}$$

These assumptions are obviously true for the general necklace quivers considered in subsection 3.2.2. In addition, we will suppose that $H_{\mathrm{W}}(q,p)$ is real. This corresponds to the case when the quantum operator $\hat{H}$ is Hermitian. It should be possible to treat the general case, when $\hat{H}$ has complex eigenvalues, with similar techniques, but we will not do it here.

As usual, the leading contribution to the number of states is given by the volume of a polygon:

$$n(E) \approx \frac{1}{2\pi\hbar}\mathrm{Vol}\left\{(q,p) : \sum_{i=1}^{m}\gamma_i|a_{R_i}q + b_{R_i}p| < E\right\} = CE^2. \tag{3.5.17}$$

Let

$$\mathcal{C}_E = \{(q,p)\,:\,H_{\mathrm{W}}(p,q)=E\} \tag{3.5.18}$$

be the real curve describing the Fermi surface. One can always decompose $\mathcal{C}_E$ into patches $\mathcal{U}_R$, $\mathcal{U}'_R$ so that both $\mathcal{U}_R$ and $\mathcal{U}'_R$ contain one point where $Q_R=0$, $\mathcal{U}'_R$ is related by to $\mathcal{U}_R$ by reflection through the center of the polygon, and

$$\left.\sum_{i=1}^{m}\gamma_i|a_{R_i}q+b_{R_i}p|\right|_{\partial\mathcal{U}_R,\,\partial\mathcal{U}'_R} = E + \mathcal{O}(\mathrm{e}^{-cE}). \tag{3.5.19}$$

In Fig. 3.7 we depict the general structure of such decomposition. In the case of ABJM theory considered in section 3.4.3, the boundaries of the regions I and II lying on the curve $\mathcal{C}_E$ are halves of the patches $\mathcal{U}_1$ and $\mathcal{U}_2$. A particular example of such a decomposition of the Fermi surface in regions is shown in Fig. 3.10, in the case of ABJM theory with fundamental matter. We will now argue that, for each patch $\mathcal{U}_R$, the structure of corrections is essentially the same as in the ABJM case.

One can always choose

$$P_R = c_R p + d_R q \tag{3.5.20}$$

so that

$$\mathrm{d}Q_R\wedge\mathrm{d}P_R = \mathrm{d}q\wedge\mathrm{d}p \tag{3.5.21}$$

[1] In what follows, the constant $c>0$ may have different values in different formulae.

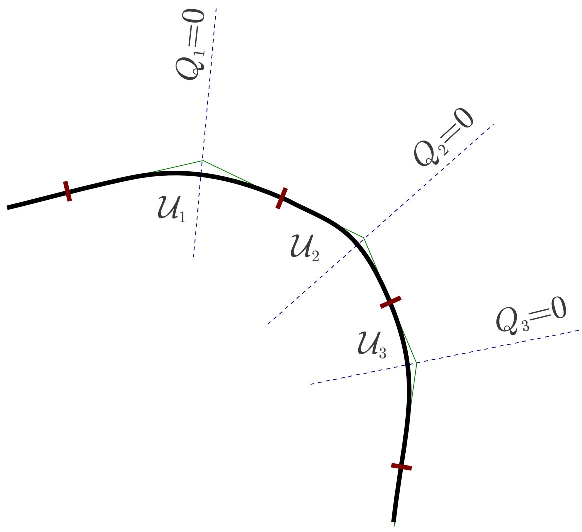

Figure 3.7: The black thick line depicts the Fermi surface (3.5.18). The green thin line depicts the limiting polygon. The red thick dashes mark the boundaries of the patches $\mathcal{U}_R$.

and such that, in the domain $\mathcal{U}_R$, the Hamiltonian can be written as follows:

$$H_{\rm W}(q,p) = \beta_R P_R + \sum_{i|R_i=R} \Phi_i(Q_R) + \sum_{r>0}\sum_{i|R_i=R} \hbar^{2r} c^R_{r,i} \Phi_i^{(2r)}(Q_R) + \mathcal{O}(\mathrm{e}^{-cE}) \qquad (3.5.22)$$

where $\mathcal{O}(\mathrm{e}^{-cE})$ denotes an estimate which is uniform in $\mathcal{U}_R$. Without loss of generality one can assume that $\beta_R > 0$. Then in the domain $\mathcal{U}_R$ the solution to $H_W^{(R)}(P_R,Q_R) \equiv H_W(p,q) = E$ can be written as

$$P_R(E,Q_R) = \frac{1}{\beta_R}\left(E - \sum_{i|R_i=R} \gamma_i |Q_R|\right) + \Delta P_R(E,Q_R). \qquad (3.5.23)$$

In this equation,

$$\Delta P_R(E,Q_R) = \Delta_{\rm p} P_R(Q_R) + \mathcal{O}(\mathrm{e}^{-cE}), \qquad (3.5.24)$$

where $\Delta_{\rm p}$ denotes the perturbative part of the correction. It is given by

$$\Delta_{\rm p} P_R(Q_R) = -\frac{1}{\beta_R}\left(\sum_{i|R_i=R} (\Phi_i(Q_R) - \gamma_i|Q_R|) + \sum_{r>0}\sum_{i|R_i=R} \hbar^{2r} c^R_{r,i}\Phi_i^{(2r)}(Q_R)\right), \qquad (3.5.25)$$

and satisfies the property,

$$\Delta_{\rm p} P_R(Q_R) = \mathcal{O}\left(\mathrm{e}^{-c|Q_R|}\right), \; Q_R \to \infty \qquad (3.5.26)$$

The property (3.5.19) implies that

$$\mathrm{Vol}\left\{H_{\rm W}(q,p) < E\right\} = \mathrm{Vol}\left\{\sum_{i=1}^m \gamma_i |Q_R| < E\right\} + 2\sum_R \Delta\mathrm{Vol}_R(E) + \mathcal{O}(\mathrm{e}^{-cE}) \qquad (3.5.27)$$

where

$$\Delta\mathrm{Vol}_R(E) = \int_{\mathcal{U}_R} \Delta P_R(E,Q_R)\mathrm{d}Q_R. \qquad (3.5.28)$$

From (3.5.24) it follows that

$$\Delta\mathrm{Vol}_R(E) = \int_{\mathcal{U}_R} \Delta_{\rm p} P_R(Q_R)\mathrm{d}Q_R + \mathcal{O}(\mathrm{e}^{-cE}). \qquad (3.5.29)$$

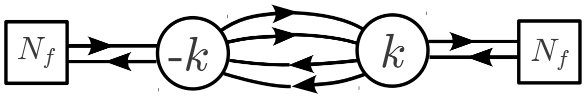

Figure 3.8: Quiver for the two-node theory with fundamental matter.

By using (3.5.26), we can extend the integration region to infinity, up to non-perturbative corrections, and we obtain

$$\Delta\mathrm{Vol}_R(E) = \int_{-\infty}^{\infty} \Delta_{\mathrm{p}} P_R(Q_R)\mathrm{d}Q_R + \mathcal{O}(\mathrm{e}^{-cE}). \tag{3.5.30}$$

Let us denote

$$\Delta_{\mathrm{p}}\mathrm{Vol}_R \equiv \int_{-\infty}^{\infty} \Delta_{\mathrm{p}} P_R(Q_R)\mathrm{d}Q_R = -\frac{1}{\beta_R}\int_{-\infty}^{\infty} \sum_{i|R_i=R} (\Phi_i(Q_R) - \gamma_i|Q_R|)\mathrm{d}Q_R - \frac{2\hbar^2}{\beta_R}\sum_{i|R_i=R} c^R_{1,i}\gamma_i. \tag{3.5.31}$$

Similarly to what happened in ABJM theory, there are no perturbative corrections to $n(E)$ from higher terms of the Wigner–Kirkwood expansion. Therefore we obtain

$$n(E) = CE^2 + n_0 + \mathcal{O}(\mathrm{e}^{-cE}) \tag{3.5.32}$$

with

$$n_0 = \frac{1}{\pi\hbar}\sum_R \Delta_{\mathrm{p}}\mathrm{Vol}_R. \tag{3.5.33}$$

As was shown in subsection 3.4.2, it then follows that the $1/N$ corrections are resummed to an Airy function

$$Z(N) = C^{-1/3}\mathrm{e}^A \,\mathrm{Ai}\left[C^{-1/3}(N-B)\right] + Z_{\mathrm{np}}(N) \tag{3.5.34}$$

where

$$B = n_0 + \frac{C\pi^2}{3}. \tag{3.5.35}$$

The above general argument gives an explicit algorithm to compute the constant (3.5.33). In the next subsection we consider the example of the ABJM theory with matter as an illustration of this argument.

3.5.3 ABJM theory with fundamental matter

The ABJM theory with matter we will consider is described by a two-node quiver with equal number of fundamentals in each node, see Fig. 3.8. The density matrix of this theory is given by (3.2.57):

$$\rho_{\mathrm{W}} \equiv \exp_\star\{-H_{\mathrm{W}}\} = \mathrm{e}^{-U(q)/2} \star \mathrm{e}^{-\Psi(p+q)} \star \mathrm{e}^{-T(p)} \star \mathrm{e}^{-\Psi(p+q)} \star \mathrm{e}^{-U(q)/2} \tag{3.5.36}$$

where $U(q)$, $T(p)$ are given respectively in (3.2.30) and (3.2.31), and

$$\Psi(p+q) = N_f \log 2\cosh\frac{p+q}{2k}. \tag{3.5.37}$$

The Wigner transform of the Hamiltonian has the following $\hbar$ expansion, which can be obtained with the use of the Baker-Campbell-Hausdorff formula:

$$\begin{aligned}
H_{\mathrm{W}}(p,q) &= H_{\mathrm{cl}}(p,q) + \frac{\hbar^2}{24}U'(q)^2\left(2\Psi''(p+q) + T''(p)\right) \\
&- \frac{\hbar^2}{12}\left(-2\Psi'(p+q)^2\left(T''(p) - 2U''(q)\right) + T'(p)^2\left(2\Psi''(p+q) + U''(q)\right) + 4T'(p)U''(q)\Psi'(p+q)\right) \\
&+ \frac{\hbar^2}{6}U'(q)\left(T''(p)\Psi'(p+q) - T'(p)\Psi''(p+q)\right) + \mathcal{O}\left(\hbar^4\right),
\end{aligned} \tag{3.5.38}$$

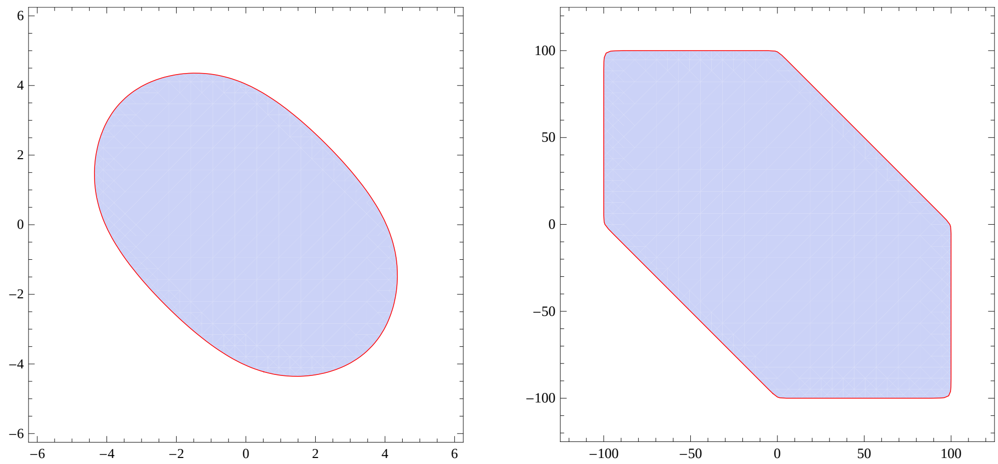

Figure 3.9: The Fermi surface (3.5.18) in the q-p plane for the ABJM theory with fundamental matter and $N_f = 1$, $k = 2$, for $E = 5$ (left) and $E = 100$ (right).

where the first term is the "classical" Hamiltonian[2]:

$$H_{\rm cl}(p,q) = T(p) + U(q) + 2\Psi(p+q). \tag{3.5.39}$$

The function Ψ has an asymptotic behavior similar to that of T and U:

$$\Psi(Q) = \frac{\alpha}{2}|Q| + O(e^{-c|Q|}), \qquad |Q| \gg 1, \tag{3.5.40}$$

where

$$\alpha = \frac{N_f}{k}. \tag{3.5.41}$$

Since $H_{\rm W}$ is real, (3.5.36) is a particular example of the general case (3.5.14) considered in the previous subsection. We have three different local coordinates

$$Q_1 = q, \qquad Q_2 = p, \qquad Q_3 = p+q. \tag{3.5.42}$$

For large energy the Fermi surface $H_{\rm W}(q,p) = E$ approaches a polygon given by

$$\frac{|p|}{2} + \frac{|q|}{2} + \alpha|p+q| = E, \tag{3.5.43}$$

see Fig. 3.9. Therefore, the leading contribution to the number of states is

$$n(E) \approx \frac{1}{2\pi\hbar} \text{Vol}\left\{ \frac{|p|}{2} + \frac{|q|}{2} + \alpha|p+q| < E \right\} = CE^2 \tag{3.5.44}$$

where

$$C = \frac{2(1+\alpha)}{\pi^2 k(1+2\alpha)^2}. \tag{3.5.45}$$

It follows that

$$F(N) \approx -\frac{\sqrt{2}}{3}\pi k \frac{1+2\alpha}{\sqrt{1+\alpha}} N^{3/2} \tag{3.5.46}$$

which reproduces the result of [7, 38]. Notice that, as in the case of ABJM theory, the large energy limit of the Fermi surface is closely related to the tropical limit of the spectral curve obtained in the previous chapter.

Now let us compute the corrections according to the general scheme described in the previous subsection. The regions $\mathcal{U}_R, \mathcal{U}'_R$ for $R = 1,2,3$, as well as the lines $Q_R = 0$, are shown in Fig. 3.10. In the domain $\mathcal{U}_1$ the Hamiltonian can be written as

$$H_{\rm W}(q,p) = p\left(\alpha + \frac{1}{2}\right) + U(q) + q\alpha - \frac{1}{48}\hbar^2(2\alpha+1)^2 U''(q) + \sum_{n>1} \hbar^{2n} c_n^1 U^{(2n)}(q) + \mathcal{O}(\mathrm{e}^{-cE}). \tag{3.5.47}$$

[2]We use quotation marks because Ψ still contains $k = \hbar/(2\pi)$.

Therefore we can take

$$P_1 = p + \frac{\alpha}{1/2+\alpha} q \tag{3.5.48}$$

and

$$\Delta_{\rm p} P_1(Q_1) = -\frac{2}{1+2\alpha}\left\{ U(Q_1) - \frac{|Q_1|}{2} - \frac{1}{48}\hbar^2(2\alpha+1)^2 U''(Q_1) + \sum_{n>1} \hbar^{2n} c_n^1 U^{(2n)}(Q_1)\right\}, \tag{3.5.49}$$

$$\Delta_{\rm p}{\rm Vol}_1 = -\frac{\pi^2}{3(1+2\alpha)} + \frac{\hbar^2(1+2\alpha)}{24}. \tag{3.5.50}$$

In the domain $\mathcal{U}_2$ we have,

$$H_{\rm W}(q,p) = T(p) - p\alpha - q\left(\alpha + \frac{1}{2}\right) + \frac{1}{96}(2\alpha+1)^2\hbar^2 T''(p) + \sum_{n>1}\hbar^{2n} c_n^2 T^{(2n)}(p) + \mathcal{O}({\rm e}^{-cE}), \tag{3.5.51}$$

therefore

$$P_2 = -q - \frac{\alpha}{1/2+\alpha} p \tag{3.5.52}$$

and

$$\Delta_{\rm p} P_2(Q_2) = -\frac{2}{1+2\alpha}\left\{ T(Q_2) - \frac{|Q_2|}{2} + \frac{1}{96}\hbar^2(2\alpha+1)^2 T''(Q_2) + \sum_{n>1} \hbar^{2n} c_n^2 T^{(2n)}(Q_2)\right\}, \tag{3.5.53}$$

$$\Delta_{\rm p}{\rm Vol}_2 = -\frac{\pi^2}{3(1+2\alpha)} - \frac{\hbar^2(1+2\alpha)}{48}. \tag{3.5.54}$$

In the domain $\mathcal{U}_3$,

$$H_{\rm W}(q,p) = p/2 - q/2 + 2\Psi(p+q) + \frac{1}{48}\hbar^2\Psi''(p+q) + \sum_{n>1}\hbar^{2n}c_n^3\Psi^{(2n)}(p+q) + \mathcal{O}({\rm e}^{-cE}). \tag{3.5.55}$$

Therefore

$$P_3 = \frac{p-q}{2} \tag{3.5.56}$$

and

$$\Delta_{\rm p} P_3(Q_3) = -\left(2\Psi(Q_3) - \alpha|Q_3|\right) - \frac{1}{48}\hbar^2\Psi''(Q_3) - \sum_{n>1}\hbar^{2n}c_n^3\Psi^{(2n)}(Q_3), \tag{3.5.57}$$

$$\Delta_{\rm p}{\rm Vol}_3 = -\frac{5}{48}\hbar^2\alpha \tag{3.5.58}$$

Finally, we obtain

$$n_0 = \frac{1}{\pi\hbar}\sum_{R\in\{1,2,3\}}\Delta_{\rm p}{\rm Vol}_R = -\frac{1}{3k(1+2\alpha)} + \frac{k}{24}(1-3\alpha), \tag{3.5.59}$$

and we conclude that the partition function is given by the Airy function

$$Z(N) = C^{-1/3}{\rm e}^A\,{\rm Ai}\left[C^{-1/3}(N-B)\right] + Z_{\rm np}(N) \tag{3.5.60}$$

with

$$B = n_0 + \frac{C\pi^2}{3} = \frac{1}{3k(1+2\alpha)^2} + \frac{k}{24}(1-3\alpha). \tag{3.5.61}$$

The above procedure can be repeated for other quivers with a Hermitian Hamiltonian in order to determine the precise value of B. For example, for the four-node quiver with levels

$$(k, -2k, 2k, -k), \tag{3.5.62}$$

one finds

$$B = -\frac{13}{135k} + \frac{k}{8}. \tag{3.5.63}$$

Clearly, it would be nice to have a close answer for the shift for a more general class of quivers (like for example four node quivers with a Hermitian density matrix). In addition, it would be interesting to compare the shifts (3.5.61), (3.5.63) with a direct calculation from the M-theory/type IIA geometry, as in [77, 78].

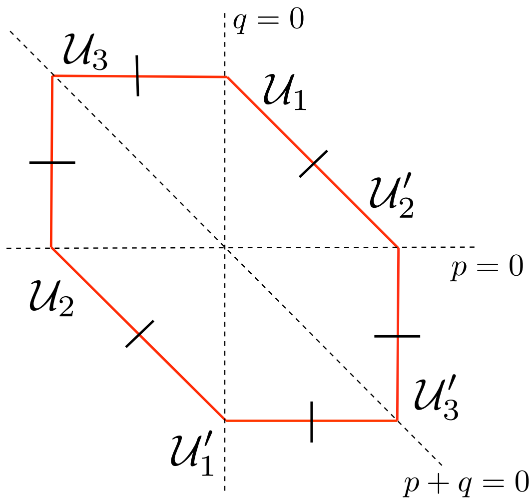

Figure 3.10: The regions $\mathcal{U}_R$, $\mathcal{U}_R'$ defined in section 3.5.2 for the Fermi surface of ABJM theory with matter. The dashed lines are defined by $Q_R=0$, where the different coordinates are given in (3.5.42).

3.5.4 The massive theory

The techniques developed in this paper can be also applied to a variant of ABJM theory in which the Chern–Simons levels k_1, k_2 do not add up to zero [176]. We will denote

$$2\pi \mathrm{i}\theta = -\frac{1}{k_1}-\frac{1}{k_2}, \qquad \frac{4\pi}{\hbar}=\frac{1}{k_1}-\frac{1}{k_2}, \tag{3.5.64}$$

so that the original ABJM theory is recovered when $\theta=0$. Notice that θ is in principle imaginary, but it will be useful to Wick-rotate it to real values (see also [141]). The theory with $k_1+k_2\neq 0$ was studied in [176], where it was argued that a non-zero θ corresponds to a non-zero Romans mass in type IIA supergravity. For this reason, we will call this theory the "massive" theory. The massive theory was further investigated in [139], where it was found that its free energy scales with N as

$$F(N)\approx (k_1+k_2)^{1/3}N^{5/3}. \tag{3.5.65}$$

This scaling was reproduced in [42] from an analysis of the matrix model representing the partition function,

$$Z(N,\theta) = \frac{1}{N!^2}\int \frac{\mathrm{d}^N x}{(2\pi)^N}\frac{\mathrm{d}^N y}{(2\pi)^N}\frac{\prod_{i<j}\left[2\sinh\left(\frac{\mu_i-\mu_j}{2}\right)\right]^2\left[2\sinh\left(\frac{\nu_i-\nu_j}{2}\right)\right]^2}{\prod_{i,j}\left[2\cosh\left(\frac{\mu_i-\nu_j}{2}\right)\right]^2}\exp\left[\frac{\mathrm{i}}{4\pi}\sum_{i=1}^{N}\left(k_1\mu_i^2+k_2\nu_i^2\right)\right]. \tag{3.5.66}$$

The exact planar resolvent of this theory was found, in a somewhat implicit form, in [140]. The scaling (3.5.65) can be also derived from this resolvent by using the techniques of [141].

In order to apply the Fermi gas picture to this theory, we have to find an appropriate density matrix. An elementary computation leads to

$$Z(N,\xi)=\frac{1}{N!}\sum_{\sigma\in S_N}(-1)^{\epsilon(\sigma)}\int \mathrm{d}^N x\prod_i \rho(x_i,x_{\sigma(i)};\theta). \tag{3.5.67}$$

where

$$\rho(x_1,x_2;\theta)=\mathrm{e}^{-\frac{1}{2}U_\theta(x_1)}K(x_1,x_2;\theta)\mathrm{e}^{-\frac{1}{2}U_\theta(x_2)}. \tag{3.5.68}$$

Here, the one-body potential is given by

$$U_\theta(q)=\log\left(2\cosh\frac{q}{2}\right)+\frac{\theta}{2}q^2, \tag{3.5.69}$$

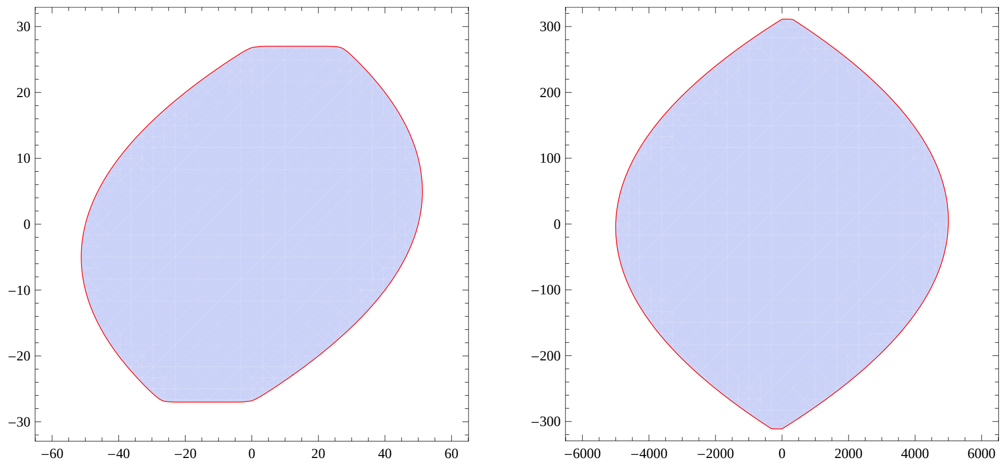

Figure 3.11: The Fermi surface (3.5.81) in the q-P plane, for $\theta = 1/10$, $k = 1$, $E = 50$ (left) and $E = 5000$ (right). When the energy is large, the Fermi surface approaches the surface defined by (3.5.82).

while the function K is given by

$$\begin{aligned} &K(x_1, x_2; \theta) \\ &= \sqrt{1 + \hbar^2\theta^2/4} \int_{-\infty}^{\infty} \frac{\mathrm{d}y}{4\pi\hbar \cosh \frac{y}{2}} \exp\left\{ -\frac{\theta}{2} y^2 - y \left[\frac{\theta}{2}(x_1 + x_2) + \frac{\mathrm{i}}{\hbar}(x_1 - x_2) \right] \right\}. \end{aligned} \tag{3.5.70}$$

Although (3.5.70) is complicated, its Wigner transform is very simple,

$$K_{\mathrm{W}}(q, p; \theta) = \sqrt{1 + \hbar^2\theta^2/4}\, \mathrm{e}^{-T_\theta(q,p)} \tag{3.5.71}$$

where

$$T_\theta(q, p) = \log\left(2 \cosh \frac{p}{2}\right) + \frac{\theta}{2} p^2 + \theta p q, \tag{3.5.72}$$

and of course

$$K_W(q, p; 0) = \mathrm{e}^{-T(p)}. \tag{3.5.73}$$

The Wigner transform of the density matrix is then

$$\rho_{\mathrm{W}}(\theta) = \sqrt{1 + \hbar^2\theta^2/4}\, \mathrm{e}^{-\frac{1}{2}U_\theta(q)} \star \mathrm{e}^{-T_\theta(q,p)} \star \mathrm{e}^{-\frac{1}{2}U_\theta(q)}, \tag{3.5.74}$$

and defines the Hamiltonian of the theory through

$$\rho_{\mathrm{W}}(\theta) = \mathrm{e}_\star^{-H_{\mathrm{W}}(\theta)}. \tag{3.5.75}$$

For $\theta = 0$ we recover the density matrix of ABJM theory (3.2.39).

We can now use the technology developed before to analyze the theory. We will content ourselves with an analysis of the thermodynamic limit, which leads to a nice interpretation of the $N^{5/3}$ behavior found in [42, 139]. We will also assume that

$$|\theta\hbar| \ll 1, \tag{3.5.76}$$

or equivalently, that

$$\left| \frac{k_1 + k_2}{k_1 - k_2} \right| \ll 1. \tag{3.5.77}$$

In this limit we can safely ignore quantum corrections and just look at the classical Hamiltonian

$$H_{\mathrm{cl}}(q, P; \theta) = U(q) + T(P - q) + \frac{\theta}{2} P^2 - \frac{1}{2} \log\left(1 + \frac{\theta^2\hbar^2}{4}\right), \tag{3.5.78}$$

where

$$P = p + q. \tag{3.5.79}$$

This linear change of variables preserves the volume form in phase space,

$$\mathrm{d}q \wedge \mathrm{d}P = \mathrm{d}q \wedge \mathrm{d}p. \tag{3.5.80}$$

At large E the Fermi surface

$$H_{\rm cl}(q, P; \theta) = E \tag{3.5.81}$$

becomes simply

$$\frac{\theta}{2}P^2 + |q| = E, \tag{3.5.82}$$

as we can see in Fig. 3.11. Notice that, once $\theta \neq 0$, the equation defining the Fermi surface at large E has a quadratic term in the new momentum coordinate P which dominates at large E. In other words, the Fermi gas has now a non-relativistic dispersion relation, and this changes the scaling of the free energy. Looking at (3.3.15) we deduce that

$$s = \frac{3}{2}, \tag{3.5.83}$$

therefore the free energy should scale now as $N^{5/3}$, as found in [42, 139][3]. We find

$$n(E) \approx \frac{4}{2\pi\hbar}\int_0^{\sqrt{\frac{2E}{\theta}}} \mathrm{d}P \left(E - \frac{\theta}{2}P^2\right) = \frac{4}{3\pi\hbar}\sqrt{\frac{2}{\theta}}E^{3/2}. \tag{3.5.84}$$

The free energy can now be computed from (3.3.11) and reads,

$$F(N) \approx -\frac{3}{5}\left(\frac{3\sqrt{2}\pi\hbar}{8}\right)^{2/3} \theta^{1/3} N^{5/3}. \tag{3.5.85}$$

If we express this in terms of $k_1 + k_2$ we find,

$$F(N) \approx -\frac{3^{5/3}}{5 \cdot 2^{4/3}}\pi \mathrm{e}^{-\frac{\mathrm{i}\pi}{6}}(k_1 + k_2)^{1/3}\left(1 + \frac{\theta^2\hbar^2}{4}\right)^{1/3} N^{5/3}. \tag{3.5.86}$$

Since we are assuming (3.5.76), our result can be written as

$$F(N) \approx -\frac{3^{5/3}}{5 \cdot 2^{4/3}}\pi \mathrm{e}^{-\frac{\mathrm{i}\pi}{6}}(k_1 + k_2)^{1/3} N^{5/3}, \tag{3.5.87}$$

which is precisely what [42] obtained. Notice that, in [42], this result was derived based on an assumption on the behavior of the eigenvalues of the matrix model at large N, while here we have obtained it directly. When the parameter $\theta^2\hbar^2$ is not small, one has to take into account the quantum corrections to the Hamiltonian, and the equation of the Fermi surface is modified. It would be interesting to study in more detail the different regimes that can occur in this theory as we vary the coupling constants.

[3]The matrix model analyzed in [119] displays the same scaling.

Appendix

A Harmonic analysis on $\mathbb{S}^3$

A.1 Maurer–Cartan forms

We will first introduce some results and conventions for the Lie algebra and the Maurer–Cartan forms. The basis of a Lie algebra $\mathbf{g}$ satisfies

$$[T_a, T_b] = f_{abc} T_c. \tag{A.1}$$

If $g \in G$ is a generic element of G, one defines the *Maurer–Cartan forms* ω_a through the equation

$$g^{-1}\mathrm{d}g = \sum_a T_a \omega_a, \tag{A.2}$$

and they satisfy

$$\mathrm{d}\omega_a + \frac{1}{2} f_{abc} \omega_b \wedge \omega_c = 0. \tag{A.3}$$

This is due to the identity

$$\mathrm{d}\left(g^{-1}\mathrm{d}g\right) + g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g = 0. \tag{A.4}$$

Let us now specialize to $SU(2)$. A basis for the Lie algebra is given by:

$$T_a = \frac{\mathrm{i}}{2}\sigma_a. \tag{A.5}$$

Explicitly

$$T_1 = \frac{\mathrm{i}}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad T_2 = \frac{\mathrm{i}}{2}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad T_3 = \frac{\mathrm{i}}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{A.6}$$

The structure constants are

$$f_{abc} = -\epsilon_{abc}. \tag{A.7}$$

Any element of $SU(2)$ can be written in the form

$$g = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}, \qquad |\alpha|^2 + |\beta|^2 = 1. \tag{A.8}$$

We parametrize this element as (see for example [207])

$$|\alpha| = \cos\frac{t_1}{2}, \qquad |\beta| = \sin\frac{t_1}{2}, \qquad \operatorname{Arg}\alpha = \frac{t_2 + t_3}{2}, \qquad \operatorname{Arg}\beta = \frac{t_2 - t_3 + \pi}{2}, \tag{A.9}$$

where t_i are the Euler angles and span the range

$$0 \le t_1 < \pi, \qquad 0 \le t_2 < 2\pi, \qquad -2\pi \le t_3 < 2\pi. \tag{A.10}$$

The general element of $SU(2)$ will then be given by

$$\begin{aligned} g = u(t_1, t_2.t_3) &= \begin{pmatrix} \cos(t_1/2)\mathrm{e}^{\mathrm{i}(t_2+t_3)/2} & \mathrm{i}\sin(t_1/2)\mathrm{e}^{\mathrm{i}(t_2-t_3)/2} \\ \mathrm{i}\sin(t_1/2)\mathrm{e}^{\mathrm{i}(-t_2+t_3)/2} & \cos(t_1/2)\mathrm{e}^{-\mathrm{i}(t_2+t_3)/2} \end{pmatrix} \\ &= u(t_2, 0, 0)u(0, t_1, 0)u(0, 0, t_3). \end{aligned} \tag{A.11}$$

We then have

$$\Omega = g^{-1}\mathrm{d}g = \frac{\mathrm{i}}{2}\begin{pmatrix} \mathrm{d}t_3 + \cos t_1 \mathrm{d}t_2 & \mathrm{e}^{-\mathrm{i}t_3}(\mathrm{d}t_1 + \mathrm{i}\mathrm{d}t_2 \sin t_1) \\ \mathrm{e}^{\mathrm{i}t_3}(\mathrm{d}t_1 - \mathrm{i}\mathrm{d}t_2 \sin t_1) & -\mathrm{d}t_3 - \cos t_1 \mathrm{d}t_2 \end{pmatrix}. \tag{A.12}$$

Therefore,

$$\begin{aligned}\omega_1 &= \cos\, t_3 \mathrm{d}t_1 + \sin\, t_3 \sin\, t_1 \mathrm{d}t_2, \\ \omega_2 &= \sin\, t_3 \mathrm{d}t_1 - \cos\, t_3 \sin\, t_1 \mathrm{d}t_2, \\ \omega_3 &= \cos\, t_1 \mathrm{d}t_2 + \mathrm{d}t_3,\end{aligned} \tag{A.13}$$

and one checks explicitly

$$\mathrm{d}\omega_a = \frac{1}{2}\epsilon_{abc}\,\omega_b \wedge \omega_c, \tag{A.14}$$

as it should be according to (A.3).

A.2 Metric and spin connection

The metric on $SU(2) = \mathbb{S}^3$ is induced from the metric on $\mathbb{C}^2$

$$\mathrm{d}s^2 = r^2\left(\mathrm{d}|\alpha|^2 + |\alpha|^2 \mathrm{dArg}\alpha^2 + \mathrm{d}|\beta|^2 + |\beta|^2 \mathrm{dArg}\beta^2\right), \tag{A.15}$$

where r is the radius of the three-sphere. A simple calculation leads to

$$\mathrm{d}s^2 = \frac{r^2}{4}\left(\mathrm{d}t_1^2 + \mathrm{d}t_2^2 + \mathrm{d}t_3^2 + 2\cos\, t_1\, \mathrm{d}t_2 \mathrm{d}t_3\right), \tag{A.16}$$

with inverse metric

$$G^{-1} = \frac{4}{r^2}\begin{pmatrix}1 & 0 & 0 \\ 0 & \csc^2\, t_1 & -\cot\, t_1 \,\csc\, t_1 \\ 0 & -\cot\, t_1 \,\csc\, t_1 & \csc^2\, t_1\end{pmatrix} \tag{A.17}$$

and volume element

$$(\det G)^{1/2} = \frac{r^3 \sin\, t_1}{8}. \tag{A.18}$$

The volume of $\mathbb{S}^3$ is then

$$\int_{SU(2)} (\det G)^{1/2} \mathrm{d}t_1\, \mathrm{d}t_2\, \mathrm{d}t_3 = 2\pi^2\, r^3 \tag{A.19}$$

which is the standard result. The only nonzero Christoffel symbols of this metric are

$$\Gamma^1_{23} = \frac{1}{2}\sin t_1, \qquad \Gamma^2_{13} = \Gamma^3_{12} = -\frac{1}{2\sin t_1}, \qquad \Gamma^3_{13} = \Gamma^2_{12} = \frac{1}{2}\cot t_1. \tag{A.20}$$

We can use the Maurer–Cartan forms to analyze the differential geometry of $\mathbb{S}^3$. The dreibein of $\mathbb{S}^3$ is proportional to ω_a, and we have

$$e^a_\mu = \frac{r}{2}\left(\omega_a\right)_\mu. \tag{A.21}$$

In terms of forms, we have

$$e^a = e^a_\mu \mathrm{d}x^\mu = \frac{r}{2}\omega_a. \tag{A.22}$$

Indeed, one can explicitly check that

$$e^a_\mu e^b_\nu \eta_{ab} = G_{\mu\nu}. \tag{A.23}$$

The inverse vierbein is defined by

$$E^\mu_a = \eta_{ab} G^{\mu\nu} e^b_\mu, \tag{A.24}$$

which can be used to define left-invariant vector fields

$$\ell_a = E^\mu_a \frac{\partial}{\partial x^\mu}. \tag{A.25}$$

Let us give their explicit expression in components:

$$\begin{aligned}\ell_1 &= \frac{2}{r}\left(\cos t_3 \frac{\partial}{\partial t_1} + \frac{\sin t_3}{\sin t_1}\frac{\partial}{\partial t_2} - \sin t_3 \cot t_1 \frac{\partial}{\partial t_3}\right), \\ \ell_2 &= \frac{2}{r}\left(\sin t_3 \frac{\partial}{\partial t_1} - \frac{\cos t_3}{\sin t_1}\frac{\partial}{\partial t_2} + \cos t_3 \cot t_1 \frac{\partial}{\partial t_3}\right), \\ \ell_3 &= \frac{2}{r}\frac{\partial}{\partial t_3}.\end{aligned} \tag{A.26}$$

Of course, they obey

$$e^a(\ell_b) = \delta^a_b, \tag{A.27}$$

as well as the following commutation relations

$$[\ell_a, \ell_b] = -\frac{2}{r}\epsilon_{abc}\ell_c. \tag{A.28}$$

This can be checked by direct computation. If we now introduce the operators L_a through

$$\ell_a = \frac{2\mathrm{i}}{r} L_a. \tag{A.29}$$

we see that they satisfy the standard commutation relations of the $SU(2)$ angular momentum operators:

$$[L_a, L_b] = \mathrm{i}\epsilon_{abc}L_c. \tag{A.30}$$

The spin connection $\omega^a{}_b$ is characterized by

$$\mathrm{d}e^a + \omega^a{}_b \wedge e^b = 0. \tag{A.31}$$

Imposing no torsion one finds the explicit expression,

$$\omega^a{}_{b\mu} = -E^\nu_b \left(\partial_\mu e^a_\nu - \Gamma^\lambda_{\mu\nu} e^a_\lambda\right), \tag{A.32}$$

or, equivalently,

$$\partial_\mu e^a_\nu = \Gamma^\lambda_{\mu\nu} e^a_\lambda - e^b_\nu \omega^a_{b\mu}. \tag{A.33}$$

In our case we find

$$\omega^a{}_b = \frac{1}{r}\epsilon^a{}_{bc} e^c. \tag{A.34}$$

A.3 Laplace–Beltrami operator and scalar spherical harmonics

The scalar Laplacian on $\mathbb{S}^3$ can be calculated in coordinates from the general formula

$$-\Delta^0 \phi = \frac{1}{\sqrt{\det G}} \sum_{m,n} \frac{\partial}{\partial x^m}\left(\sqrt{\det G}\, G^{mn} \frac{\partial \phi}{\partial x^n}\right), \tag{A.35}$$

or equivalently

$$-\Delta^0 = G^{\mu\nu}\partial_\mu\partial_\nu - G^{\mu\nu}\Gamma^\rho_{\mu\nu}\partial_\rho. \tag{A.36}$$

In this case it reads

$$-\Delta^0 = \frac{4}{r^2}\left(\frac{\partial^2}{\partial t_1^2} + \cot\, t_1 \frac{\partial}{\partial t_1} + \csc^2\, t_1 \frac{\partial^2}{\partial t_2^2} + \csc^2\, t_1 \frac{\partial^2}{\partial t_3^2} - 2\csc\, t_1 \cot\, t_1 \frac{\partial^2}{\partial t_2 \partial t_3}\right). \tag{A.37}$$

It is easy to check that it can be written, in terms of left-invariant vector fields, as

$$-\Delta^0 = \sum_a \ell_a^2. \tag{A.38}$$

To see this, we write

$$\sum_a \ell_a^2 = \sum_a E^\mu_a E^\nu_a \partial_\mu \partial_\nu + \sum_a E^\mu_a \frac{\partial E^\nu_a}{\partial x^\mu}\frac{\partial}{\partial x^\nu}. \tag{A.39}$$

The first term reproduces the first term in (A.36). We now use the identity

$$\partial_\mu E^\nu_b = E^\nu_c \omega^c_{b\mu} - \Gamma^\nu_{\mu\lambda} E^\lambda_b. \tag{A.40}$$

After contraction with E^μ_a and use of the explicit form of the spin connection, we see that only the second term survives, which is indeed the second term in (A.36).

The Peter–Weyl theorem says that any square-integrable function on $\mathbb{S}^3 \simeq SU(2)$ can be written as a linear combination of

$$S_j^{mn}, \qquad m, n = 1, \cdots, d_j \tag{A.41}$$

where

$$S_j : SU(2) \to M_{d_j \times d_j} \tag{A.42}$$

is the representation of spin j and dimension d_j, and $M_{d_j \times d_j}$ are the inversible square matrices of rank d_j. The function S_j^{mn} is just the (m, n)-th entry of the matrix. The eigenvalues of the Laplacian might be calculated immediately by noticing that, in terms of the $SU(2)$ angular momentum operators, it reads

$$\Delta^0 = \frac{4}{r^2}\mathbf{L}^2, \tag{A.43}$$

and since the possible eigenvalues of $\mathbf{L}^2$ are

$$j(j+1), \qquad j = 0, \frac{1}{2}, \cdots , \tag{A.44}$$

we conclude that the eigenvalues of the Laplacian are of the form

$$\lambda_j = \frac{4}{r^2} j(j+1), \qquad j = 0, \ \frac{1}{2}, \cdots \tag{A.45}$$

Notice that the dependence on r is the expected one from dimensional analysis. The degeneracy of these eigenvalues is

$$d_j^2 = (2j+1)^2 \tag{A.46}$$

which is the dimension of the matrix $M_{d_j \times d_j}$.

A.4 Vector spherical harmonics

The space of one-forms on $\mathbb{S}^3$ can be decomposed in two different sets. One set is spanned by gradients of S_j^{mn}, and it is proportional to

$$S_j^{mq}(T_a)_j^{qn}\omega_a. \tag{A.47}$$

The other set is spanned by the so-called *vector spherical harmonics*,

$$V_{j\pm}^{mn}, \qquad \epsilon = \pm 1, \quad m = 1, \cdots , d_{j\pm\frac{1}{2}}, \quad n = 1, \cdots , d_{j\mp\frac{1}{2}}, \tag{A.48}$$

see Appendix B of [205] for a useful summary of their properties. The $\epsilon = \pm 1$ corresponds to two linear combinations of the ω_a which are independent from the one appearing in (A.47). The vector spherical harmonics are in the representation

$$\left(j \pm \frac{1}{2}, j \mp \frac{1}{2} \right) \tag{A.49}$$

of $SU(2) \times SU(2)$. We will write them, as in [205], as V^α, where

$$\alpha = (j, m, m', \epsilon), \tag{A.50}$$

and we will regard them as one-forms. They satisfy the properties

$$\mathrm{d}^\dagger V^\alpha = 0, \qquad *\mathrm{d}V^\alpha = -\epsilon(2j+1)V^\alpha. \tag{A.51}$$

It follows that

$$* \, \mathrm{d} * \mathrm{d}V^\alpha = -\Delta^1 V^\alpha = (2j+1)^2 V^\alpha. \tag{A.52}$$

Their degeneracy is

$$2 d_{j+\frac{1}{2}} d_{j-\frac{1}{2}} = 4j(2j+2). \tag{A.53}$$

A.5 Spinors

Using the dreibein, we define the “locally inertial" gamma matrices as

$$\gamma_a = E_a^\mu \gamma_\mu, \tag{A.54}$$

which satisfy the relations

$$\{\gamma_a, \gamma_b\} = 2\delta_{ab}, \qquad [\gamma_a, \gamma_b] = 2\mathrm{i}\epsilon_{abc}\gamma_c. \tag{A.55}$$

The standard definition of a covariant derivative acting on a spinor is

$$\nabla_\mu = \partial_\mu + \frac{1}{4}\omega_\mu^{ab}\gamma_a\gamma_b = \partial_\mu + \frac{1}{8}\omega_\mu^{ab}[\gamma_a,\gamma_b]. \tag{A.56}$$

Using the commutation relations of the gamma matrices γ_a and the explicit expression for the spin connection (A.34) we find

$$\begin{aligned}\nabla_\mu =&\partial_\mu + \frac{\mathrm{i}}{4r}\epsilon_{abc}\epsilon_{abd}e^c_\mu\gamma_d = \partial_\mu + \frac{\mathrm{i}}{2}e^c_\mu\gamma_c\\ =&\partial_\mu + \frac{\mathrm{i}}{2r}\gamma_\mu.\end{aligned} \tag{A.57}$$

It follows that the Dirac operator is

$$-\mathrm{i}\not{D} = -\mathrm{i}\gamma^\mu\partial_\mu + \frac{3}{2r} = -\mathrm{i}\gamma^a E^\mu_a\partial_\mu + \frac{3}{2r} = -\mathrm{i}\gamma^a\ell_a + \frac{3}{2r}. \tag{A.58}$$

Let us now introduce the spin operators

$$S_a = \frac{1}{2}\gamma_a, \tag{A.59}$$

which satisfy the $SU(2)$ algebra

$$[S_a, S_b] = \mathrm{i}\epsilon_{abc}S_c. \tag{A.60}$$

In terms of the S_a and the $SU(2)$ operators L_a, the Dirac operator reads

$$-\mathrm{i}\not{D} = \frac{1}{r}\left(4\mathbf{L}\cdot\mathbf{S} + \frac{3}{2}\right). \tag{A.61}$$

The calculation of the spectrum of this operator is as in standard Quantum Mechanics: we introduce the total angular momentum

$$\mathbf{J} = \mathbf{L} + \mathbf{S}, \tag{A.62}$$

so that

$$4\mathbf{L}\cdot\mathbf{S} = 2\left(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2\right). \tag{A.63}$$

Since $\mathbf{S}$ corresponds to spin $s = 1/2$, and $\mathbf{L}$ to j, the possible eigenvalues of $\mathbf{J}$ are $j \pm 1/2$, and we conclude that the eigenvalues of (A.61) are (we set $r = 1$)

$$2\left(\left(j\pm\frac{1}{2}\right)\left(j\pm\frac{1}{2}+1\right) - j(j+1)\right) = \begin{cases}2j+\frac{3}{2} & \text{for } +\\ -2j-\frac{1}{2} & \text{for } -,\end{cases}, \tag{A.64}$$

with degeneracies

$$d_{j\pm\frac{1}{2}} = \left(2\left(j\pm\frac{1}{2}\right)\right)(2j+1) = \begin{cases}2(j+1)(2j+1) & \text{for } +\\ 2j(2j+1) & \text{for } -.\end{cases} \tag{A.65}$$

These can be written in a more compact form as

$$\lambda_n^\pm = \pm\left(n+\frac{1}{2}\right), \quad d_n^\pm = n(n+1), \qquad n = 1,2,\cdots \tag{A.66}$$

B Quantum corrections in ABJM theory at order $\mathcal{O}(\hbar^4)$

In this Appendix, we give some details on the computation of the order $\mathcal{O}(\hbar^4)$ corrections to the grand canonical potential of ABJM theory, which confirm the general arguments of section 3.4.

The Baker–Campbell–Hausdorff formula applied to (3.2.39) gives

$$\begin{aligned}H_{\mathrm{W}}(q,p) = {}& T + U + \frac{1}{12}[T,[T,U]_\star]_\star + \frac{1}{24}[U,[T,U]_\star]_\star + \frac{1}{360}[[[[T,U]_\star,U]_\star,U]_\star,T]_\star\\ &- \frac{1}{480}[[[[U,T]_\star,U]_\star,T]_\star,U]_\star + \frac{1}{360}[[[[U,T]_\star,T]_\star,T]_\star,U]_\star + \frac{1}{120}[[[[T,U]_\star,T]_\star,U]_\star,T]_\star\\ &+ \frac{7}{5760}[[[[T,U]_\star,U]_\star,U]_\star,U]_\star - \frac{1}{720}[[[[U,T]_\star,T]_\star,T]_\star,T]_\star + \cdots\end{aligned} \tag{B.1}$$

This leads to the next correction to the Wigner transform of the Hamiltonian

$$\begin{aligned} H_{\mathrm{W}}^{(2)} = & \frac{1}{144}T'(p)T'''(p)U^{(4)}(q) - \frac{1}{288}U'(q)U'''(q)T^{(4)}(q) \\ & - \frac{1}{240}\left(U'(q)\right)^2 U''(q)\left(T''(p)\right)^2 + \frac{1}{60}\left(T'(p)\right)^2 T''(p)\left(U''(q)\right)^2 \\ & - \frac{1}{80}\left(U'(q)\right)^2 U''(q)T'(p)T'''(p) + \frac{1}{120}\left(T'(p)\right)^2 T''(p)U'(q)U'''(q) \\ & + \frac{7}{5760}\left(U'(q)\right)^4 T^{(4)}(p) - \frac{1}{720}\left(T'(p)\right)^4 U^{(4)}(q). \end{aligned} \tag{B.2}$$

The computation of $J_2(\mu)$ also involves the $\mathcal{G}_r$ defined in (3.4.3) up to order $\mathcal{O}(\hbar^4)$. Due to (3.4.5) only the terms with $r \leq 6$ are needed. A long but straightforward calculation leads finally to

$$J_2(\mu) = -\frac{\pi^2}{4320} - \frac{\pi^2}{2880}\left(104 + 5\pi^2 - 134\mu + 30\mu^2\right)\mathrm{e}^{-2\mu} + \mathcal{O}\left(\mu^2\mathrm{e}^{-4\mu}\right). \tag{B.3}$$

Notice that no polynomial in μ is generated, as expected from the analysis in section 3.4.

C Normalization of the ABJM matrix model

Here we shall fix the overall normalization of the matrix model. As explained in the beginning of Section 2.2, to fix the normalization we must fix the coefficient of the cosh in the denominator. This term appears as a consequence of integrating out the matter hypermultiplets at one-loop. For general supersymmetric Chern–Simons-matter theories, the contribution of a hypermultiplet in representation R is given by [2]

$$\log Z[a] = \log \prod_{\rho}\prod_{n=1}^{\infty}\left(\frac{n + 1/2 + \mathrm{i}\rho(a)}{n - 1/2 - \mathrm{i}\rho(a)}\right)^n \tag{C.1}$$

where ρ are the weights of the representation, and a is the element in the Cartan algebra given by

$$a = \frac{1}{2\pi}\mathrm{diag}\left(\mu_1, \cdots, \mu_{N_1}, \nu_1, \cdots, \nu_{N_2}\right). \tag{C.2}$$

In [2] the one-loop determinant is evaluated up to a multiplicative constant,

$$Z[a] = \prod_{\rho}\left(C\cosh\left(\pi\rho(a)\right)\right)^{-1/2}. \tag{C.3}$$

The constant C can be determined by setting $a = 0$ in (C.1)

$$-\frac{1}{2}\log C = \log\prod_{n=1}^{\infty}\left(\frac{n+1/2}{n-1/2}\right)^n. \tag{C.4}$$

This is a divergent constant, but as usual when considering determinants on compact manifolds, we can compute it by using ζ-function regularization. Let us define

$$\zeta_Z(s) = \sum_{n=1}^{\infty}\left(\frac{n}{\left(n+\frac{1}{2}\right)^s} - \frac{n}{\left(n-\frac{1}{2}\right)^s}\right). \tag{C.5}$$

The regularization of the quantity appearing in (C.4) is then $-\zeta_Z'(0)$. An elementary calculation shows that

$$\zeta_Z(s) = -\left(2^s - 1\right)\zeta(s) \tag{C.6}$$

where $\zeta(s)$ is the standard Riemann zeta function. Therefore,

$$-\zeta_Z'(0) = -\frac{\log 2}{2} \tag{C.7}$$

and $C = 2$.

D Giant Wilson loops in Chern–Simons theory

Chern–Simons theory on $\mathbb{S}^3$ is a particular case of the lens space matrix model when $b = 1$ and the second cut collapses to zero size, *i.e.*, $t_1 = t$, $t_2 = 0$. It gives the leading behavior of the Wilson loop in ABJM theory when $\lambda_2 \ll \lambda_1$, as discussed in Section 2.6.

Here we consider the behavior of the giant Wilson loops, those in high dimensional symmetric or antisymmetric representations presented in Section 2.7.2, in this limit. In this case it is easy to calculate explicitly the action (2.7.35), since the integral

$$g(Y) = -\int_0^Y \frac{\mathrm{d}Y'}{Y'} \log(h(Y')) , \qquad h(Y) = \frac{1}{2}\left[1 + Y + \sqrt{(1+Y)^2 - 4\mathrm{e}^t Y}\right] \tag{D.1}$$

can be obtained in closed form

$$\begin{aligned} g(Y) = \frac{\pi^2}{6} - \frac{1}{2}\log^2(h(Y)) + \log(h(Y))\Big(\log\left(1 - \mathrm{e}^{-t}h(Y)\right) - \log(1 - h(Y))\Big) \\ - \mathrm{Li}_2(h(Y)) + \mathrm{Li}_2\left(\mathrm{e}^{-t}h(Y)\right) - \mathrm{Li}_2(\mathrm{e}^{-t}). \end{aligned} \tag{D.2}$$

Here we used the dilogarithm identity

$$\mathrm{Li}_2(1-x) = \frac{\pi^2}{6} - \mathrm{Li}_2(x) - \log(x)\log(1-x). \tag{D.3}$$

The solution of the saddle point equation (2.7.25) is obtained by setting in (2.7.33)

$$\kappa = -4\mathrm{i}\sinh\frac{t}{2}, \qquad B = \frac{t}{2\pi\mathrm{i}} + \frac{1}{2} \tag{D.4}$$

and we find

$$Y_* = -\frac{1 - \mathrm{e}^{-2\pi\mathrm{i}\nu}}{1 - \mathrm{e}^{2\pi\mathrm{i}\nu + t}}. \tag{D.5}$$

The action (2.7.35) is

$$\begin{aligned} \eta\, A_\eta &= -2\pi\mathrm{i}\nu\log(\eta Y_*) + g(Y_*) \\ &= -2\pi\mathrm{i}\nu\log\eta - 2\pi^2\nu^2 + 2\pi\mathrm{i}\nu t + \frac{\pi^2}{6} + \mathrm{Li}_2\left(\mathrm{e}^{2\pi\mathrm{i}\nu - t}\right) - \mathrm{Li}_2\left(\mathrm{e}^{2\pi\mathrm{i}\nu}\right) - \mathrm{Li}_2\left(\mathrm{e}^{-t}\right). \end{aligned} \tag{D.6}$$

Notice that this expression is exact in t.

We can test (D.6) in all details against a direct calculation of correlators. Indeed, the VEVs $\langle \mathrm{Tr}_R\, U\rangle$ for the Chern–Simons matrix model on $\mathbb{S}^3$ are proportional to quantum dimensions (see for example [90]):

$$\langle \mathrm{Tr}_R\, U\rangle = q^{\kappa_R/2 + \ell(R)N/2}\mathrm{dim}_q(R). \tag{D.7}$$

In this equation,

$$q = \mathrm{e}^{g_s}, \tag{D.8}$$

$\ell(R)$ is the number of boxes in R, and κ_R is the framing factor, given by

$$\kappa_R = \sum_i l_i(l_i - 2i + 1), \tag{D.9}$$

where l_i are the lenghts of the rows in the diagrams. The quantum dimensions of the symmetric and antisymmetric representations are given by

$$\mathrm{dim}_q(\mathcal{R}_n^\eta) = \frac{q^{\eta n(n-1)/4}\mathrm{e}^{nt/2}}{[n]!}\prod_{i=1}^n (1 - \mathrm{e}^{-t}q^{-\eta(i-1)}), \tag{D.10}$$

where

$$[n]! = \prod_{i=1}^n (q^{i/2} - q^{-i/2}) = q^{\frac{1}{4}n(n+1)}\prod_{i=1}^n (1 - q^{-i}). \tag{D.11}$$

At large n we rescale

$$\xi = \frac{i}{n}, \qquad q^{-i} = \exp(-g_s i) \to \mathrm{e}^{-2\pi\mathrm{i}\eta\nu\xi} \tag{D.12}$$

so that

$$\log([n]!) \approx \frac{1}{g_s}\left(-\pi^2\nu^2 + 2\pi \mathrm{i}\eta\nu \int_0^1 \mathrm{d}\xi \log(1-\mathrm{e}^{-2\pi\mathrm{i}\eta\nu\xi})\right). \tag{D.13}$$

This gives the following contribution to the action

$$\pi^2\nu^2 + \frac{\pi^2}{6} - \mathrm{Li}_2\big(\mathrm{e}^{-2\pi\mathrm{i}\eta\nu}\big) = \eta\left(\pi^2\nu^2 - 2\pi\mathrm{i}\nu\log\eta + \frac{\pi^2}{6} - \mathrm{Li}_2\big(\mathrm{e}^{-2\pi\mathrm{i}\nu}\big)\right). \tag{D.14}$$

To derive the expression on the right hand side we used, for $\eta=-1$ the dilogarithm identity

$$\mathrm{Li}_2(\mathrm{e}^x) = -\mathrm{Li}_2(\mathrm{e}^{-x}) + \frac{\pi^2}{3} - \frac{x^2}{2} \pm \pi\mathrm{i}x. \tag{D.15}$$

The product in the numerator of both the symmetric and antisymmetric representations can be written in a unified form as

$$2\pi\mathrm{i}\eta\nu\int_0^1 \mathrm{d}\xi\log(1-\mathrm{e}^{-t}\mathrm{e}^{-2\pi\mathrm{i}\nu\xi}) = \eta\left(\mathrm{Li}_2(\mathrm{e}^{-t-2\pi\mathrm{i}\nu}) - \mathrm{Li}_2(\mathrm{e}^{-t})\right). \tag{D.16}$$

The prefactors in (D.7) and (D.10) contribute

$$\eta(-3\pi^2\nu^2 + 2\pi\mathrm{i}\nu t). \tag{D.17}$$

Together with (D.14) and (D.16) this exactly reproduces (D.6).

In the antisymmetric representation the result can also be written as

$$-2\pi\mathrm{i}\nu(t+2\pi\mathrm{i}\nu) + \frac{\pi^2}{6} + \mathrm{Li}_2(\mathrm{e}^{-t}) - \mathrm{Li}_2(\mathrm{e}^{-t-2\pi\mathrm{i}\nu}) - \mathrm{Li}_2(\mathrm{e}^{2\pi\mathrm{i}\nu}). \tag{D.18}$$

This expression agrees at leading order with the D6-brane calculation (2.7.21) and should be the full answer in the limit of $\lambda_2 = 0$. In this expression we see the expected symmetry [159]

$$n \leftrightarrow N - n \tag{D.19}$$

which is

$$2\pi\mathrm{i}\nu \leftrightarrow -t - 2\pi\mathrm{i}\nu. \tag{D.20}$$

E Strongly coupled density of eigenvalues and tropical geometry

In this Appendix we rederive some of the results for the $\mathcal{N}=3$ theory by using the approach of [38], and we compare it in detail to our tropical methods.

The starting point of [38] is an analysis of the ABJM matrix model (2.2.3) in the ABJM slice, at large N but fixed k, which corresponds to the strongly coupled limit of the theory. Let us see how this is done, following closely the steps in [38]. The behaviour at large N of the equilibrium eigenvalues of the matrix model is

$$\mu_k = N^{1/2}x_k + \mathrm{i}\ell_k, \qquad \nu_k = N^{1/2}x_k - \mathrm{i}\ell_k, \qquad k = 1,\cdots,N, \tag{E.1}$$

where x_k, ℓ_k are of order one at large N. At large N the eigenvalues x_k, ℓ_k become dense, so that

$$\frac{k}{N} \to \xi \in [0,1] \tag{E.2}$$

and they are described by the functions

$$\rho(x) = \frac{\mathrm{d}\xi}{\mathrm{d}x}, \qquad \ell(x). \tag{E.3}$$

It is shown in [38] that, when N is large, the free energy of the matrix model can be written as

$$-F = N^{3/2}\left[\frac{k}{\pi}\int \mathrm{d}x\, x\rho(x)\ell(x) + \int \mathrm{d}x\,\rho^2(x) f\left(2\ell(x)\right) - \frac{m}{2\pi}\left(\int \mathrm{d}x\,\rho(x) - 1\right)\right]. \tag{E.4}$$

Here, $f(t)$ is a periodic function of t, with period 2π, and given by

$$f(t) = \pi^2 - t^2, \qquad t \in [-\pi, \pi]. \tag{E.5}$$

The last term in (E.4) involves, as usual, a Lagrange multiplier m imposing the normalization of $\rho(x)$. Notice that our sign convention is opposite to the one chosen in [38]. Varying this functional w.r.t. $\rho(x)$ and $\ell(x)$ one obtains the two equations

$$\begin{aligned} 2\pi\rho(x)f'\left(2\ell(x)\right) &= -kx, \\ 4\pi\rho(x)f\left(2\ell(x)\right) &= m - 2kx\ell(x), \end{aligned} \tag{E.6}$$

which are solved by

$$\rho(x) = \frac{m}{4\pi^3}, \qquad \ell(x) = \frac{\pi^2 kx}{2m}. \tag{E.7}$$

The support of $\rho(x)$, $\ell(x)$ is the interval $[-x_*, x_*]$. One fixes x_* and m from the normalization of ρ and by minimizing $-F$. This gives

$$x_* = \pi\sqrt{\frac{2}{k}}, \qquad m = \frac{2\pi^3}{x_*}. \tag{E.8}$$

Evaluating the free energy for the functions (E.7) and the values (E.8) of x_*, m, one reproduces the result of section 2.5 for the free energy.

The above results can be easily compared with our tropical analysis. The value of x_* gives (up to a factor $N^{1/2}$) the position of the endpoint A, and it is in accord with the value of (2.9.11), since

$$A = N^{1/2} x_*. \tag{E.9}$$

The fact that the density $\rho(x)$ is constant follows from our result for the tropical limit of the curve. Indeed, the density of eigenvalues (normalized as to have an integral along the cut equal to one) is given by the well-known formula

$$\rho(z) = \frac{1}{8\pi^2\lambda} \operatorname{disc} y(z) \tag{E.10}$$

where $\operatorname{disc} y(z)$ is the discontinuity of the curve through the cut $[-A, A]$. In our case this is just the constant $2A$, and it is given by the horizontal separation between the two diagonals in Fig. 2.9. Changing variables from $z = N^{1/2}x$ to x we find indeed,

$$\rho(x) = \frac{\sqrt{2k}}{4\pi}, \tag{E.11}$$

in precise agreement with the result of [38].

The inclusion of fundamental matter in the approach of [38] is straightforward. $-F$ includes now the extra term

$$\frac{N^{3/2}N_f}{2} \int \mathrm{d}x\, \rho(x)|x| \tag{E.12}$$

which is the large N limit of the operator in the exponential of (2.10.1) (as in (2.10.27)). The new saddle point equations are

$$\begin{aligned} 2\pi\rho(x)f'\left(2\ell(x)\right) &= -kx, \\ 4\pi\rho(x)f\left(2\ell(x)\right) &= m - 2kx\ell(x) - \pi N_f|x|, \end{aligned} \tag{E.13}$$

with solution

$$\rho(x) = \frac{m - \pi N_f|x|}{4\pi^3}, \qquad \ell(x) = \frac{k\pi^2 x}{2\left(m - \pi N_f|x|\right)}. \tag{E.14}$$

Normalization of the density and minimization of $-F$ lead to

$$x_* = \frac{2\pi}{\sqrt{2k + N_f}}, \qquad m = 2\pi^2 \frac{k + N_f}{\sqrt{2k + N_f}}. \tag{E.15}$$

A straightforward calculation of $-F$ reproduces (2.8.9).

Let us now compare this with the tropical approach. First of all, we have again the equality (E.9) between the endpoints of the cut in both approaches, involving now the value of x_* obtained in (E.15) and the value of A obtained in (2.11.37). The density of eigenvalues in the theory with matter can be obtained from the planar resolvent (2.11.26) as

$$\rho(x) = \frac{N^{1/2}}{8\pi^2\lambda}\text{disc}\left(y_{\text{p}}(z) + \mu y_{\text{m}}(z)\right), \tag{E.16}$$

where μ is defined in (2.8.8). From Fig. 2.11 and (2.11.33) we read off immediately

$$\frac{1}{2}\text{disc}\, y_{\text{p}}(z) = A - K, \qquad \frac{1}{2}\text{disc}\, y_{\text{m}}(z) = 2A - K - |z|, \tag{E.17}$$

and we deduce

$$\rho(x) = \frac{N^{1/2}}{4\pi^2\lambda}\left[(1+\mu)A - \mu N^{1/2}|x|\right]. \tag{E.18}$$

Plugging in the value of A (2.11.37), we recover precisely the form of $\rho(x)$ given in (E.14). This shows explicitly that the piece-wise linear densities obtained with the method in [38] correspond to the tropical curves obtained in this paper.

One advantage of the method of [38] is that it gives the large N limit of the free energy, at strong coupling, without directly using the resolvent of the model. Therefore, this method is useful when the resolvent is difficult to write down. If on the contrary one is interested in calculating the resolvent of the model (to study for example weak coupling expansions), then the tropical approach developed in this paper provides a powerful method to extract the strong coupling limit from the resolvent. In addition, the method of [38] assumes as an *ansatz* the scaling (E.1) of the eigenvalues, while the approach based on studying the resolvent and its tropical limit provides a *bona fide* solution to the large N theory without any further assumptions.

F Metric

In this appendix we follow the notations in [103] (with the replacement $\chi \to 2\chi$). The metric on $\text{AdS}_4 \times \mathbb{CP}^3$ is

$$\mathrm{d}s^2 = \frac{L^3}{4k}\left(\mathrm{d}s^2_{\text{AdS}_4} + 4\mathrm{d}s^2_{\mathbb{CP}^3}\right). \tag{F.1}$$

For the AdS_4 part we may use the global Lorentzian metric

$$\mathrm{d}s^2_{\text{AdS}_4} = -\cosh^2\rho\,\mathrm{d}t^2 + \mathrm{d}\rho^2 + \sinh^2\rho\big(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\psi^2\big). \tag{F.2}$$

The metric on $\mathbb{CP}^3$ can be written in terms of four complex projective coordinates z_i as

$$\mathrm{d}s^2_{\mathbb{CP}^3} = \frac{1}{\rho^2}\sum_{i=1}^{4}\mathrm{d}z_i\,\mathrm{d}\bar{z}_i - \frac{1}{\rho^4}\left|\sum_{i=1}^{4} z_i\,\mathrm{d}\bar{z}_i\right|^2, \qquad \rho^2 = \sum_{i=1}^{4}|z_i|^2. \tag{F.3}$$

In the following we choose a specific representations in terms of angular coordinates (used also in [197, 198]). We start by parametrizing $\mathbb{S}^7 \subset \mathbb{C}^4$ as

$$\begin{aligned} z_1 &= \cos\frac{\alpha}{2}\cos\frac{\vartheta_1}{2}\,\mathrm{e}^{\mathrm{i}(2\varphi_1+2\chi+\zeta)/4}, \qquad & z_3 &= \sin\frac{\alpha}{2}\cos\frac{\vartheta_2}{2}\,\mathrm{e}^{\mathrm{i}(2\varphi_2-2\chi+\zeta)/4}, \\ z_2 &= \cos\frac{\alpha}{2}\sin\frac{\vartheta_1}{2}\,\mathrm{e}^{\mathrm{i}(-2\varphi_1+2\chi+\zeta)/4}, \qquad & z_4 &= \sin\frac{\alpha}{2}\sin\frac{\vartheta_2}{2}\,\mathrm{e}^{\mathrm{i}(-2\varphi_2-2\chi+\zeta)/4}. \end{aligned} \tag{F.4}$$

The metric on $\mathbb{S}^7$ is then given by

$$\begin{aligned} \mathrm{d}s^2_{\mathbb{S}^7} =& \frac{1}{4}\bigg[\mathrm{d}\alpha^2 + \cos^2\frac{\alpha}{2}(\mathrm{d}\vartheta_1^2 + \sin^2\vartheta_1\,\mathrm{d}\varphi_1^2) + \sin^2\frac{\alpha}{2}(\mathrm{d}\vartheta_2^2 + \sin^2\vartheta_2\,\mathrm{d}\varphi_2^2) \\ &+ \sin^2\frac{\alpha}{2}\cos^2\frac{\alpha}{2}(2\mathrm{d}\chi + \cos\vartheta_1\,\mathrm{d}\varphi_1 - \cos\vartheta_2\,\mathrm{d}\varphi_2)^2 + \frac{1}{4}(\mathrm{d}\zeta + 2A)^2\bigg], \end{aligned} \tag{F.5}$$

$$A = \cos\alpha\,\mathrm{d}\chi + \cos^2\frac{\alpha}{2}\cos\vartheta_1\,\mathrm{d}\varphi_1 + \sin^2\frac{\alpha}{2}\cos\vartheta_2\,\mathrm{d}\varphi_2. \tag{F.6}$$

The angle ζ appears only in the last term and if we drop it we end up with the metric on $\mathbb{CP}^3$

$$\begin{aligned}\mathrm{d}s^2_{\mathbb{CP}^3} = \frac{1}{4}\Big[&\mathrm{d}\alpha^2 + \cos^2\frac{\alpha}{2}(\mathrm{d}\vartheta_1^2 + \sin^2\vartheta_1\,\mathrm{d}\varphi_1^2) + \sin^2\frac{\alpha}{2}(\mathrm{d}\vartheta_2^2 + \sin^2\vartheta_2\,\mathrm{d}\varphi_2^2)\\ &+ \sin^2\frac{\alpha}{2}\cos^2\frac{\alpha}{2}(2\mathrm{d}\chi + \cos\vartheta_1\,\mathrm{d}\varphi_1 - \cos\vartheta_2\,\mathrm{d}\varphi_2)^2\Big].\end{aligned} \tag{F.7}$$

The ranges of the angles are $0 \le \alpha, \vartheta_1, \vartheta_2 \le \pi$ and $0 \le \varphi_1, \varphi_2, \chi \le 2\pi$.

In addition to the metric, the supergravity background has the dilaton, and the 2-form and 4-form field strengths from the Ramond-Ramond (RR) sector

$$e^{2\Phi} = \frac{L^3}{k^3}\,, \qquad F_4 = \frac{3}{8}L^3\,\mathrm{d}\Omega_{\mathrm{AdS}_4}\,, \qquad F_2 = \frac{k}{2}\,\mathrm{d}A\,. \tag{F.8}$$

Here $\mathrm{d}\Omega_{\mathrm{AdS}_4}$ is the volume form on AdS_4 and F_2 is proportional to the Kähler form on $\mathbb{CP}^3$.

To write down the general D-brane action in this background one also needs the potentials for these forms. The one-form potential is, up to gauge transformations

$$C_1 = \frac{k}{2}A\,, \tag{F.9}$$

with A defined in (F.6). It is easy to write down C_3, the three-form potential for F_4 and C_5, its magnetic dual, but they are not required for our calculation in Section 2.14.1.

The relation between the parameters of the string background and of the field theory are (for $\alpha' = 1$ and in the supergravity and tree-level limit)

$$\frac{L^3}{4k} = \pi\sqrt{\frac{2N}{k}} = \pi\sqrt{2\lambda}\,. \tag{F.10}$$

G Killing spinors

To write down the Killing spinors it is useful to start in 11-dimensions with the AdS_4 metric in (F.2) and the $\mathbb{S}^7$ metric in (F.5).

We take the elfbeine (ignoring the factor of L^3/k)

$$\begin{aligned} &e^0 = \frac{1}{2}\cosh\rho\,\mathrm{d}t\,, \quad e^1 = \frac{1}{2}\mathrm{d}\rho\,, \quad e^2 = \frac{1}{2}\sinh\rho\,\mathrm{d}\theta\,, \quad e^3 = \frac{1}{2}\sinh\rho\sin\theta\,\mathrm{d}\psi\,,\\ &e^4 = \frac{1}{2}\mathrm{d}\alpha, \qquad e^5 = \frac{1}{2}\cos\frac{\alpha}{2}\,\mathrm{d}\vartheta_1, \qquad e^6 = \frac{1}{2}\sin\frac{\alpha}{2}\,\mathrm{d}\vartheta_2,\\ &e^7 = \frac{1}{2}\cos\frac{\alpha}{2}\sin\frac{\alpha}{2}\Big(\cos\vartheta_1\,\mathrm{d}\varphi_1 - \cos\vartheta_2\,\mathrm{d}\varphi_2 + 2\mathrm{d}\chi\Big),\\ &e^8 = \frac{1}{2}\cos\frac{\alpha}{2}\sin\vartheta_1\,\mathrm{d}\varphi_1, \quad e^9 = \frac{1}{2}\sin\frac{\alpha}{2}\sin\vartheta_2\,\mathrm{d}\varphi_2\,,\\ &e^\natural = -\frac{1}{4}\Big(\mathrm{d}\zeta + 2\cos^2\frac{\alpha}{2}\cos\vartheta_1\,\mathrm{d}\varphi_1 + 2\sin^2\frac{\alpha}{2}\cos\vartheta_2\,\mathrm{d}\varphi_2 + 2\cos\alpha\,\mathrm{d}\chi\Big)\,.\end{aligned} \tag{G.1}$$

Killing spinor equation for this background comes from the supersymmetry transformation of the gravitino

$$\delta\Psi_\mu = D_\mu\epsilon - \frac{1}{288}\Big(\Gamma_\mu^{\ \nu\lambda\rho\sigma} - 8\delta_\mu^\nu\Gamma^{\lambda\rho\sigma}\Big)F_{\nu\lambda\rho\sigma}\epsilon\,, \qquad D_\mu\epsilon = \partial_\mu\epsilon + \frac{1}{4}\omega_\mu^{ab}\gamma_{ab}\epsilon\,. \tag{G.2}$$

The 4-form corresponding to the $\mathrm{AdS}_4 \times \mathbb{S}^7$ solution is $F_{\nu\lambda\rho\sigma} = 6\,\varepsilon_{\nu\lambda\rho\sigma}$, where the epsilon symbol is the volume form on AdS_4 (so the indices take the values $0, 1, 2, 3$). Plugging this into the variation above one finds the Killing spinor equation

$$\begin{aligned} D_\mu\epsilon &= \hat\gamma\Gamma_\mu\epsilon\,, \qquad && \mu = 0, 1, 2, 3\\ D_\mu\epsilon &= \frac{1}{2}\hat\gamma\Gamma_\mu\epsilon\,, && \mu = 4, 5, \cdots, 9, 10\end{aligned} \tag{G.3}$$

where μ runs over all 11 coordinates, and $\hat\gamma = \gamma^{0123}$. Note that small γ have tangent-space indices while capital Γ carry curved-space indices.

The general solution to these equations is

$$e^{\frac{\alpha}{4}(\hat{\gamma}\gamma_4-\gamma_{7\natural})}e^{\frac{\vartheta_1}{4}(\hat{\gamma}\gamma_5-\gamma_{8\natural})}e^{\frac{\vartheta_2}{4}(\gamma_{79}+\gamma_{46})}e^{-\frac{\xi_1}{2}\hat{\gamma}\gamma_\natural}e^{-\frac{\xi_2}{2}\gamma_{58}}e^{-\frac{\xi_3}{2}\gamma_{47}}e^{-\frac{\xi_4}{2}\gamma_{69}}e^{\frac{\rho}{2}\hat{\gamma}\gamma_1}e^{\frac{t}{2}\hat{\gamma}\gamma_0}e^{\frac{\theta}{2}\gamma_{12}}e^{\frac{\psi}{2}\gamma_{23}}\epsilon_0 = \mathcal{M}\epsilon_0\,, \tag{G.4}$$

where the ξ_i are given by

$$\xi_1 = \frac{2\varphi_1+2\chi+\zeta}{4}\,, \qquad \xi_2 = \frac{-2\varphi_1+2\chi+\zeta}{4}\,, \qquad \xi_3 = \frac{2\varphi_2-2\chi+\zeta}{4}\,, \qquad \xi_4 = \frac{-2\varphi_2-2\chi+\zeta}{4}\,. \tag{G.5}$$

In (G.4) ϵ_0 is a constant 32-component spinor and the Dirac matrices were chosen such that $\gamma_{0123456789\natural} = 1$. A similar calculation in a different coordinate system was done in [199].

To see which Killing spinors survive the orbifolding from M-theory to type IIA, we write the spinor ϵ_0 in a basis which diagonalizes

$$i\hat{\gamma}\gamma_\natural\epsilon_0 = s_1\epsilon_0\,, \qquad i\gamma_{58}\epsilon_0 = s_2\epsilon_0\,, \qquad i\gamma_{47}\epsilon_0 = s_3\epsilon_0\,, \qquad i\gamma_{69}\epsilon_0 = s_4\epsilon_0\,. \tag{G.6}$$

All the s_i take values ± 1 and by our conventions on the product of all the Dirac matrices, the number of negative eigenvalues is even. Now consider a shift along the ζ circle, which changes all the angles by $\xi_i \to \xi_i + \delta/4$, the Killing spinors transform as

$$\mathcal{M}\epsilon_0 \to \mathcal{M}e^{i\frac{\delta}{8}(s_1+s_2+s_3+s_4)}\epsilon_0\,. \tag{G.7}$$

This transformation is a symmetry of the Killing spinor when two of the s_i eigenvalues are positive and two negative and not when they all have the same sign (unless δ is an integer multiple of 4π). Note that on $\mathbb{S}^7$ the radius of the ζ circle is 8π, so the $\mathbb{Z}_k$ orbifold of $\mathbb{S}^7$ is given by taking $\delta = 8\pi/k$. The allowed values of the s_i are therefore

$$(s_1,s_2,s_3,s_4) \in \begin{Bmatrix} (+,+,-,-),\ (+,-,+,-),\ (+,-,-,+), \\ (-,+,+,-),\ (-,+,-,+),\ (-,-,+,+) \end{Bmatrix} \tag{G.8}$$

Each configuration represents four supercharges, so the orbifolding breaks 1/4 of the supercharges (except for $k = 1, 2$) and leaves 24 unbroken supersymmetries.

Bibliography

[1] O. Aharony, O. Bergman, D. L. Jafferis and J. Maldacena, "N=6 superconformal Chern-Simons-matter theories, M2-branes and their gravity duals," JHEP **0810**, 091 (2008) [arXiv:0806.1218 [hep-th]].

[2] A. Kapustin, B. Willett and I. Yaakov, "Exact Results for Wilson Loops in Superconformal Chern-Simons Theories with Matter," JHEP **1003**, 089 (2010) [arXiv:0909.4559 [hep-th]].

[3] O. Aharony, O. Bergman, D. L. Jafferis, "Fractional M2-branes," JHEP **0811**, 043 (2008). [arXiv:0807.4924 [hep-th]].

[4] I. R. Klebanov and A. A. Tseytlin, "Entropy of near-extremal black p-branes," Nucl. Phys. B **475** (1996) 164 [hep-th/9604089].

[5] M. Mariño and P. Putrov, "Exact Results in ABJM Theory from Topological Strings," JHEP **1006**, 011 (2010) [arXiv:0912.3074 [hep-th]].

[6] N. Drukker, M. Mariño, P. Putrov, "From weak to strong coupling in ABJM theory," Commun. Math. Phys. **306**, 511-563 (2011). [arXiv:1007.3837 [hep-th]].

[7] R. Couso Santamaria, M. Mariño, P. Putrov, "Unquenched flavor and tropical geometry in strongly coupled Chern-Simons-matter theories," [arXiv:1011.6281 [hep-th]].

[8] N. Drukker, M. Mariño, P. Putrov,"Nonperturbative aspects of ABJM theory," JHEP **1111**, 141 (2011). [arXiv:1103.4844 [hep-th]].

[9] M. Marino and P. Putrov, "ABJM theory as a Fermi gas," J. Stat. Mech. **1203** (2012) P03001 [arXiv:1110.4066 [hep-th]].

[10] N. Seiberg and E. Witten, "Electric-magnetic duality, monopole condensation, and confinement in $\mathcal{N} = 2$ supersymmetric Yang-Mills theory," Nucl. Phys. B **426**, 19 (1994) [Erratum-ibid. B **430**, 485 (1994)] [hep-th/9407087].

[11] N. A. Nekrasov, "Seiberg-Witten prepotential from instanton counting," Adv. Theor. Math. Phys. **7**, 831 (2004) [hep-th/0206161].

[12] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity," Adv. Theor. Math. Phys. **2**, 231-252 (1998). [hep-th/9711200].

[13] V. Pestun, "Localization of gauge theory on a four-sphere and supersymmetric Wilson loops," arXiv:0712.2824 [hep-th].

[14] A. Gustavsson, "Algebraic structures on parallel M2-branes," Nucl. Phys. B **811** (2009) 66 [arXiv:0709.1260].

[15] J. Bagger and N. Lambert, "Gauge symmetry and supersymmetry of multiple M2-branes," Phys. Rev. D **77** (2008) 065008 [arXiv:0711.0955].

[16] M. Mariño, "Lectures on localization and matrix models in supersymmetric Chern-Simons-matter theories," J. Phys. A **44**, 463001 (2011) [arXiv:1104.0783 [hep-th]].

[17] D. Gaiotto and D. L. Jafferis, "Notes on adding D6 branes wrapping RP3 in AdS4 x CP3," arXiv:0903.2175 [hep-th].

[18] K. M. Lee and H. U. Yee, “New $AdS_4 \times X_7$ Geometries with $\mathcal{N} = 6$ in M Theory,” JHEP **0703**, 012 (2007) [arXiv:hep-th/0605214].

[19] S. H. Shenker, “The Strength of nonperturbative effects in string theory,” in O. Álvarez, E. Marinari and P. Windey (eds.), *Random surfaces and quantum gravity,* Plenum, New York, 1991, pp. 191-200.

[20] J. Polchinski, “Combinatorics of boundaries in string theory,” Phys. Rev. **D50**, 6041-6045 (1994). [hep-th/9407031].

[21] E. J. Martinec, “The Annular report on noncritical string theory,” [hep-th/0305148].

[22] S. Y. Alexandrov, V. A. Kazakov, D. Kutasov, “Nonperturbative effects in matrix models and D-branes,” JHEP **0309**, 057 (2003). [hep-th/0306177].

[23] H. Fuji, S. Hirano, S. Moriyama, “Summing Up All Genus Free Energy of ABJM Matrix Model,” JHEP **1108**, 001 (2011). [arXiv:1106.4631 [hep-th]].

[24] B. Eynard and N. Orantin, “Invariants of algebraic curves and topological expansion," math-ph/0702045v4.

[25] M. C. N. Cheng, R. Dijkgraaf, C. Vafa, “Non-Perturbative Topological Strings And Conformal Blocks,” [arXiv:1010.4573].

[26] D. R. Gulotta, J. P. Ang and C. P. Herzog, “Matrix Models for Supersymmetric Chern-Simons Theories with an ADE Classification,” JHEP **1201**, 132 (2012) [arXiv:1111.1744 [hep-th]].

[27] M. Aganagic, R. Dijkgraaf, A. Klemm, M. Mariño, C. Vafa, “Topological strings and integrable hierarchies,” Commun. Math. Phys. **261**, 451-516 (2006). [hep-th/0312085].

[28] B. Eynard, “Large N expansion of convergent matrix integrals, holomorphic anomalies, and background independence,” JHEP **0903**, 003 (2009). [arXiv:0802.1788 [math-ph]].

[29] B. Eynard, M. Mariño, “A Holomorphic and background independent partition function for matrix models and topological strings,” J. Geom. Phys. **61**, 1181-1202 (2011). [arXiv:0810.4273 [hep-th]].

[30] M. Mariño, S. Pasquetti, P. Putrov, “Large N duality beyond the genus expansion,” JHEP **1007**, 074 (2010). [arXiv:0911.4692 [hep-th]].

[31] I. K. Kostov, “Solvable statistical models on a random lattice,” Nucl. Phys. Proc. Suppl. **45A**, 13-28 (1996). [hep-th/9509124].

[32] C. A. Tracy, H. Widom, “Fredholm determinants and the mKdV/sinh-Gordon hierarchies," Commun. Math. Phys. **179**, 1-10 (1996).

[33] A. B. Zamolodchikov, “Painleve III and 2-d polymers,” Nucl. Phys. **B432**, 427-456 (1994). [hep-th/9409108].

[34] C. A. Tracy, H. Widom, “Proofs of two conjectures related to the thermodynamic Bethe ansatz,” Commun. Math. Phys. **179**, 667-680 (1996). [solv-int/9509003].

[35] D. Berenstein, “A Toy model for the AdS / CFT correspondence,” JHEP **0407**, 018 (2004) [hep-th/0403110].

[36] H. Lin, O. Lunin and J. M. Maldacena, “Bubbling AdS space and 1/2 BPS geometries,” JHEP **0410**, 025 (2004) [hep-th/0409174].

[37] P. Horava, C. A. Keeler, “Noncritical M-theory in 2+1 dimensions as a nonrelativistic Fermi liquid,” JHEP **0707**, 059 (2007). [hep-th/0508024].

[38] C. P. Herzog, I. R. Klebanov, S. S. Pufu, T. Tesileanu, “Multi-Matrix Models and Tri-Sasaki Einstein Spaces,” Phys. Rev. **D83**, 046001 (2011). [arXiv:1011.5487 [hep-th]].

[39] D. Martelli, J. Sparks, “The large N limit of quiver matrix models and Sasaki-Einstein manifolds,” Phys. Rev. **D84**, 046008 (2011). [arXiv:1102.5289 [hep-th]].

[40] S. Cheon, H. Kim, N. Kim, "Calculating the partition function of N=2 Gauge theories on S^3 and AdS/CFT correspondence," JHEP **1105**, 134 (2011). [arXiv:1102.5565 [hep-th]].

[41] M. Gabella, D. Martelli, A. Passias, J. Sparks, "The free energy of N=2 supersymmetric AdS_4 solutions of M-theory," JHEP **1110**, 039 (2011) [arXiv:1107.5035 [hep-th]].

[42] D. L. Jafferis, I. R. Klebanov, S. S. Pufu, B. R. Safdi, "Towards the F-Theorem: N=2 Field Theories on the Three-Sphere," JHEP **1106**, 102 (2011). [arXiv:1103.1181 [hep-th]].

[43] N. Hama, K. Hosomichi, S. Lee, N. Hama, K. Hosomichi, S. Lee, "Notes on SUSY Gauge Theories on Three-Sphere," JHEP **1103**, 127 (2011). [arXiv:1012.3512 [hep-th]].

[44] J. Wess and J. Bagger, *Supersymmetry and supergravity*, Princeton University Press, 1992.

[45] S. Deser, R. Jackiw, S. Templeton, "Three-Dimensional Massive Gauge Theories," Phys. Rev. Lett. **48**, 975-978 (1982).

[46] E. Witten, "Quantum field theory and the Jones polynomial," Commun. Math. Phys. **121**, 351 (1989).

[47] L. C. Jeffrey, "Chern-Simons-Witten invariants of lens spaces and torus bundles, and the semiclassical approximation," Commun. Math. Phys. **147** (1992) 563.

[48] L. Rozansky, "A large k asymptotics of Witten's invariant of Seifert manifolds," hep-th/9303099, Commun. Math. Phys. **171** (1995) 279.

[49] L. Rozansky, "Witten's invariant of three-dimensional manifolds: Loop expansion and surgery calculus," in Kauffman, L.H. (ed.), *Knots and applications*, p. 271-299. [hep-th/9401060].

[50] D. S. Freed and R. E. Gompf, "Computer calculation of Witten's three manifold invariant," Commun. Math. Phys. **141** (1991) 79.

[51] D. H. Adams, "A note on the Faddeev-Popov determinant and Chern-Simons perturbation theory," Lett. Math. Phys. **42**, 205 (1997) [arXiv:hep-th/9704159]. "Perturbative expansion in gauge theories on compact manifolds," arXiv:hep-th/9602078.

[52] D. H. Adams, "The semiclassical approximation for the Chern-Simons partition function," Phys. Lett. B **417**, 53 (1998) [arXiv:hep-th/9709147].

[53] S. W. Hawking, "Zeta Function Regularization of Path Integrals in Curved Space-Time," Commun. Math. Phys. **55**, 133 (1977).

[54] D. B. Ray and I. M. Singer, "R-torsion and the Laplacian on Riemannian manifolds," Advances in Math. **7** (1971) 145-210. "Analytic torsion," in *Partial differential equations*, Proc. Sympos. Pure Math., XXIII, American Mathematica Society, 1973, pp. 167-181.

[55] A. S. Schwarz, "The Partition Function of Degenerate Quadratic Functional and Ray-Singer Invariants," Lett. Math. Phys. **2**, 247-252 (1978).

[56] T. Friedmann, E. Witten, "Unification scale, proton decay, and manifolds of G(2) holonomy," Adv. Theor. Math. Phys. **7**, 577-617 (2003). [hep-th/0211269].

[57] M. Atiyah, *The geometry and physics of knots*, Cambridge University Press, 1990.

[58] D. H. Adams and S. Sen, "Phase and scaling properties of determinants arising in topological field theories," Phys. Lett. B **353**, 495 (1995) [arXiv:hep-th/9506079].

[59] M. Asorey, F. Falceto, J. L. López, G. Luzón, "Universality and ultraviolet regularizations of Chern-Simons theory," Nucl. Phys. **B429**, 344-374 (1994). [hep-th/9403117].

[60] M. Mariño, "Chern-Simons theory and topological strings," Rev. Mod. Phys. **77**, 675-720 (2005). [hep-th/0406005].

[61] J. R. Quine and J. Choi, "Zeta regularized products and functional determinants on spheres," Rocky Mountain J. Math. **26** (1996) 719-729.

[62] B. Willett, I. Yaakov, "N=2 Dualities and Z Extremization in Three Dimensions," [arXiv:1104.0487 [hep-th]].

[63] C. Nash, D. J. O'Connor, "Determinants of Laplacians, the Ray-Singer torsion on lens spaces and the Riemann zeta function," J. Math. Phys. **36**, 1462-1505 (1995). [hep-th/9212022].

[64] L. Weng and Y. You, "Analytic torsions of spheres," Internat. J. Math. **7** (1996) 109-125.

[65] J. S. Dowker, "Functional determinants on spheres and sectors," J. Math. Phys. **35**, 4989-4999 (1994). [hep-th/9312080].

[66] E. Witten, "Anti-de Sitter space and holography," Adv. Theor. Math. Phys. **2**, 253 (1998) [arXiv:hep-th/9802150].

[67] G. W. Gibbons, S. W. Hawking, "Action Integrals and Partition Functions in Quantum Gravity," Phys. Rev. **D15**, 2752-2756 (1977).

[68] E. Poisson, *A relativist's toolkit*, Cambridge University Press, 2004.

[69] M. Henningson, K. Skenderis, "The Holographic Weyl anomaly," JHEP **9807**, 023 (1998). [hep-th/9806087].

[70] V. Balasubramanian and P. Kraus, "A stress tensor for anti-de Sitter gravity," Commun. Math. Phys. **208**, 413 (1999) [hep-th/9902121].

[71] R. Emparan, C. V. Johnson and R. C. Myers, "Surface terms as counterterms in the AdS/CFT correspondence," Phys. Rev. D **60**, 104001 (1999) [arXiv:hep-th/9903238].

[72] S. de Haro, S. N. Solodukhin, K. Skenderis, "Holographic reconstruction of space-time and renormalization in the AdS / CFT correspondence," Commun. Math. Phys. **217**, 595-622 (2001). [hep-th/0002230].

[73] K. Skenderis, "Lecture notes on holographic renormalization," Class. Quant. Grav. **19**, 5849 (2002) [arXiv:hep-th/0209067].

[74] E. Kiritsis, *String theory in a nutshell*, Princeton University Press, 2007.

[75] I. R. Klebanov, G. Torri, "M2-branes and AdS/CFT," Int. J. Mod. Phys. **A25**, 332-350 (2010). [arXiv:0909.1580 [hep-th]].

[76] M. J. Duff, B. E. W. Nilsson, C. N. Pope, "Kaluza-Klein Supergravity," Phys. Rept. **130**, 1-142 (1986).

[77] O. Bergman and S. Hirano, "Anomalous radius shift in AdS_4/CFT_3," JHEP **0907**, 016 (2009) [arXiv:0902.1743].

[78] O. Aharony, A. Hashimoto, S. Hirano, P. Ouyang, "D-brane Charges in Gravitational Duals of 2+1 Dimensional Gauge Theories and Duality Cascades," JHEP **1001**, 072 (2010). [arXiv:0906.2390 [hep-th]].

[79] M. Blau, G. Thompson, "Localization and diagonalization: A review of functional integral techniques for low dimensional gauge theories and topological field theories," J. Math. Phys. **36**, 2192-2236 (1995). [hep-th/9501075].

[80] G. W. Moore, E. Witten, "Integration over the u plane in Donaldson theory," Adv. Theor. Math. Phys. **1**, 298-387 (1998). [hep-th/9709193].

[81] J. M. F. Labastida, C. Lozano, "Lectures in topological quantum field theory," [hep-th/9709192].

[82] D. L. Jafferis, "The Exact Superconformal R-Symmetry Extremizes Z," [arXiv:1012.3210 [hep-th]].

[83] S. Cordes, G. W. Moore, S. Ramgoolam, "Lectures on 2-d Yang-Mills theory, equivariant cohomology and topological field theories," Nucl. Phys. Proc. Suppl. **41**, 184-244 (1995). [hep-th/9411210].

[84] H. -C. Kao, K. -M. Lee, T. Lee, "The Chern-Simons coefficient in supersymmetric Yang-Mills Chern-Simons theories," Phys. Lett. **B373**, 94-99 (1996). [hep-th/9506170].

[85] M. Mariño, “Chern-Simons theory, matrix integrals, and perturbative three manifold invariants,” Commun. Math. Phys. **253**, 25-49 (2004). [hep-th/0207096].

[86] C. Beasley, E. Witten, “Non-Abelian localization for Chern-Simons theory,” J. Diff. Geom. **70**, 183-323 (2005). [hep-th/0503126].

[87] M. Blau, G. Thompson, “Chern-Simons theory on S1-bundles: Abelianisation and q-deformed Yang-Mills theory,” JHEP **0605**, 003 (2006). [hep-th/0601068].

[88] M. Aganagic, A. Klemm, M. Mariño and C. Vafa, “Matrix model as a mirror of Chern-Simons theory,” JHEP **0402**, 010 (2004) [arXiv:hep-th/0211098].

[89] P. Di Francesco, P. H. Ginsparg, J. Zinn-Justin, “2-D Gravity and random matrices,” Phys. Rept. **254**, 1-133 (1995). [hep-th/9306153].

[90] M. Mariño, “Les Houches lectures on matrix models and topological strings,” [hep-th/0410165].

[91] N.I. Muskhelishvili, *Singular integral equations*, Dover, 2008.

[92] A. Migdal, “Loop equations and $1/N$ expansion,” Phys. Rept. **102**, 199 (1983).

[93] L. Álvarez-Gaumé and J. L. Mañes, “Supermatrix models,” Mod. Phys. Lett. A **6**, 2039 (1991).

[94] S. A. Yost, “Supermatrix models,” Int. J. Mod. Phys. A **7**, 6105 (1992) [arXiv:hep-th/9111033].

[95] R. Dijkgraaf and C. Vafa, “N = 1 supersymmetry, deconstruction, and bosonic gauge theories,” arXiv:hep-th/0302011.

[96] R. Dijkgraaf, S. Gukov, V. A. Kazakov and C. Vafa, “Perturbative analysis of gauged matrix models,” Phys. Rev. D **68**, 045007 (2003) [arXiv:hep-th/0210238].

[97] D. Gaiotto and E. Witten, “Janus Configurations, Chern-Simons Couplings, And The Theta-Angle in N=4 Super Yang-Mills Theory,” arXiv:0804.2907 [hep-th].

[98] K. Hosomichi, K. M. Lee, S. Lee, S. Lee and J. Park, “N=5,6 Superconformal Chern-Simons Theories and M2-branes on Orbifolds,” JHEP **0809**, 002 (2008) [arXiv:0806.4977 [hep-th]].

[99] N. Drukker and D. Trancanelli, “A supermatrix model for N=6 super Chern-Simons-matter theory,” JHEP **1002**, 058 (2010) [arXiv:0912.3006 [hep-th]].

[100] N. Halmagyi and V. Yasnov, “The spectral curve of the lens space matrix model,” JHEP **0911**, 104 (2009) [arXiv:hep-th/0311117].

[101] N. Halmagyi, T. Okuda and V. Yasnov, “Large N duality, lens spaces and the Chern-Simons matrix model,” JHEP **0404**, 014 (2004) [arXiv:hep-th/0312145].

[102] A. Brini and A. Tanzini, “Exact results for topological strings on resolved Y(p,q) singularities,” Commun. Math. Phys. **289**, 205 (2009) [arXiv:0804.2598 [hep-th]].

[103] N. Drukker, J. Plefka and D. Young, “Wilson loops in 3-dimensional N=6 supersymmetric Chern-Simons Theory and their string theory duals,” JHEP **0811**, 019 (2008) [arXiv:0809.2787 [hep-th]].

[104] B. Chen and J. B. Wu, “Supersymmetric Wilson Loops in N=6 Super Chern-Simons-matter theory,” Nucl. Phys. B **825**, 38 (2010) [arXiv:0809.2863 [hep-th]].

[105] S. J. Rey, T. Suyama and S. Yamaguchi, “Wilson Loops in Superconformal Chern-Simons Theory and Fundamental Strings in Anti-de Sitter Supergravity Dual,” JHEP **0903**, 127 (2009) [arXiv:0809.3786 [hep-th]].

[106] J. K. Erickson, G. W. Semenoff and K. Zarembo, “Wilson loops in N = 4 supersymmetric Yang-Mills theory,” Nucl. Phys. B **582**, 155 (2000) [arXiv:hep-th/0003055].

[107] K. M. Lee and S. Lee, “1/2-BPS Wilson Loops and Vortices in ABJM Model,” arXiv:1006.5589.

[108] R. Gopakumar and C. Vafa, “On the gauge theory/geometry correspondence,” Adv. Theor. Math. Phys. **3**, 1415 (1999) [arXiv:hep-th/9811131].

[109] B. Haghighat, A. Klemm and M. Rauch, "Integrability of the holomorphic anomaly equations," JHEP **0810**, 097 (2008) [arXiv:0809.1674 [hep-th]].

[110] S. Hosono, A. Klemm and S. Theisen, "Lectures on mirror symmetry," hep-th/9403096.

[111] T. M. Chiang, A. Klemm, S. T. Yau and E. Zaslow, "Local mirror symmetry: Calculations and interpretations," Adv. Theor. Math. Phys. **3**, 495 (1999) [hep-th/9903053].

[112] J. A. Minahan, O. O. Sax and C. Sieg, "A limit on the ABJ model," arXiv:1005.1786.

[113] M. Aganagic, V. Bouchard and A. Klemm, "Topological Strings and (Almost) Modular Forms," Commun. Math. Phys. **277**, 771 (2008) [arXiv:hep-th/0607100].

[114] N. Drukker, J. Gomis and D. Young, "Vortex Loop Operators, M2-branes and Holography," JHEP **0903** (2009) 004 [arXiv:0810.4344].

[115] D. R. Gulotta, C. P. Herzog, S. S. Pufu, "From Necklace Quivers to the F-theorem, Operator Counting, and T(U(N))," [arXiv:1105.2817 [hep-th]].

[116] D. R. Gulotta, C. P. Herzog, S. S. Pufu, "Operator Counting and Eigenvalue Distributions for 3D Supersymmetric Gauge Theories," JHEP **1111**, 149 (2011) [arXiv:1106.5484 [hep-th]].

[117] I. R. Klebanov, "String theory in two-dimensions," [hep-th/9108019].

[118] D. Boulatov, V. Kazakov, "One-dimensional string theory with vortices as the upside down matrix oscillator," Int. J. Mod. Phys. **A8**, 809-852 (1993). [hep-th/0012228].

[119] V. A. Kazakov, I. K. Kostov, N. A. Nekrasov, "D particles, matrix integrals and KP hierarchy," Nucl. Phys. **B557**, 413-442 (1999). [hep-th/9810035].

[120] K. Becker, M. Becker and A. Strominger, "Five-Branes, Membranes And Nonperturbative String Theory," Nucl. Phys. B **456**, 130 (1995) [arXiv:hep-th/9507158].

[121] E. Witten, "Quantum background independence in string theory," [hep-th/9306122].

[122] R. P. Feynman, *Statistical Mechanics*, Westview Press, Boulder, 1998.

[123] G. Parisi, *Statistical Field Theory*, Westview Press, Boulder, 1998.

[124] J. B. Kogut, "An Introduction to Lattice Gauge Theory and Spin Systems," Rev. Mod. Phys. **51**, 659 (1979).

[125] M. Hillery, R. F. O'Connell, M. O. Scully, E. P. Wigner, "Distribution functions in physics: Fundamentals," Phys. Rept. **106**, 121-167 (1984).

[126] Y. Imamura, K. Kimura, "On the moduli space of elliptic Maxwell-Chern-Simons theories," Prog. Theor. Phys. **120**, 509-523 (2008). [arXiv:0806.3727 [hep-th]].

[127] D. L. Jafferis, A. Tomasiello, "A Simple class of N=3 gauge/gravity duals," JHEP **0810**, 101 (2008). [arXiv:0808.0864 [hep-th]].

[128] S. Kharchev, A. Marshakov, A. Mironov, A. Morozov, S. Pakuliak, "Conformal matrix models as an alternative to conventional multimatrix models," Nucl. Phys. **B404**, 717-750 (1993). [hep-th/9208044].

[129] S. Hohenegger and I. Kirsch, "A note on the holography of Chern-Simons matter theories with flavour," JHEP **0904**, 129 (2009) [arXiv:0903.1730 [hep-th]].

[130] Y. Hikida, W. Li and T. Takayanagi, "ABJM with Flavors and FQHE," JHEP **0907**, 065 (2009) [arXiv:0903.2194 [hep-th]].

[131] H. Widom, "Asymptotic behavior of the eigenvalues of certain integral equations," Trans. Amer. Math. Soc. **109**, 278-295 (1963).

[132] E. P. Wigner, "On the quantum correction for thermodynamic equilibrium," Phys. Rev. **40**, 749-760 (1932).

[133] J. G. Kirkwood, "Quantum Statistics of Almost Classical Assemblies," Phys. Rev. **44**, 31-37 (1933).

[134] A. Voros, "Asymptotic $\hbar$-expansions of stationary quantum states," Ann. Inst. H. Poincaré Sect. A (N.S.) **26** (1977), no. 4, 343-403.

[135] B. Grammaticos, A. Voros, "Semiclassical Approximations For Nuclear Hamiltonians. 1. Spin Independent Potentials," Annals Phys. **123**, 359 (1979).

[136] M. Centelles, X. Viñas, M. Durand, P. Schuck, D. Von-Eiff, "Variational Wigner–Kirkwood $\hbar$ Expansion," Annals Phys. **266**, 207 (1998).

[137] A. Voros, "The return of the quartic oscillator. The complex WKB method," Ann. Inst. H. Poincaré A **39**, 211 (1983).

[138] E. Delabaere, H. Dillinger and F. Pham, "Exact semiclassical expansions for one-dimensional quantum oscillators," J. Math. Phys. **38**, 6126 (1997).

[139] O. Aharony, D. Jafferis, A. Tomasiello, A. Zaffaroni, "Massive type IIA string theory cannot be strongly coupled," JHEP **1011**, 047 (2010). [arXiv:1007.2451 [hep-th]].

[140] T. Suyama, "On Large N Solution of Gaiotto-Tomasiello Theory," JHEP **1010**, 101 (2010). [arXiv:1008.3950 [hep-th]].

[141] T. Suyama, "Eigenvalue Distributions in Matrix Models for Chern-Simons-matter Theories," Nucl. Phys. B **856**, 497 (2012) [arXiv:1106.3147 [hep-th]].

[142] M. x. Huang and A. Klemm, "Holomorphic anomaly in gauge theories and matrix models," JHEP **0709**, 054 (2007) [hep-th/0605195].

[143] T. W. Grimm, A. Klemm, M. Mariño and M. Weiss,"Direct integration of the topological string," JHEP **0708**, 058 (2007) [hep-th/0702187].

[144] M. Bershadsky, S. Cecotti, H. Ooguri and C. Vafa, "Kodaira-Spencer theory of gravity and exact results for quantum string amplitudes," Commun. Math. Phys. **165**, 311 (1994) [hep-th/9309140].

[145] J. Gomis and F. Passerini, "Holographic Wilson loops," JHEP **0608**, 074 (2006) [hep-th/0604007].

[146] J. Gomis and F. Passerini, "Wilson loops as D3-branes," JHEP **0701**, 097 (2007) [hep-th/0612022].

[147] E. Guadagnini, M. Martellini and M. Mintchev, "Wilson lines in Chern–Simons theory and link invariants," Nucl. Phys. B **330** (1990) 575.

[148] M. Alvarez and J. M. F. Labastida, "Analysis of observables in Chern–Simons perturbation theory," Nucl. Phys. B **395** (1993) 198 [hep-th/9110069].

[149] N. Drukker, D. J. Gross and A. A. Tseytlin, "Green-Schwarz string in $\mathrm{AdS}_5 \times \mathbb{S}^5$: Semiclassical partition function," JHEP **0004**, 021 (2000) [hep-th/0001204].

[150] M. Hanada, M. Honda, Y. Honma, J. Nishimura, S. Shiba and Y. Yoshida, "Numerical studies of the ABJM theory for arbitrary N at arbitrary coupling constant," arXiv:1202.5300 [hep-th].

[151] P. Candelas, X. C. De La Ossa, P. S. Green and L. Parkes, "A Pair of Calabi-Yau manifolds as an exactly soluble superconformal theory," Nucl. Phys. B **359**, 21 (1991).

[152] N. Drukker, "1/4 BPS circular loops, unstable world-sheet instantons and the matrix model," JHEP **0609**, 004 (2006) [hep-th/0605151].

[153] K. Zarembo, "Supersymmetric Wilson loops," Nucl. Phys. B **643**, 157 (2002) [hep-th/0205160].

[154] S. Giombi, V. Pestun and R. Ricci, "Notes on supersymmetric Wilson loops on a two-sphere," arXiv:0905.0665.

[155] X. Arsiwalla, R. Boels, M. Mariño and A. Sinkovics, "Phase transitions in q-deformed 2d Yang-Mills theory and topological strings," Phys. Rev. D **73**, 026005 (2006) [hep-th/0509002].

[156] M. x. Huang and A. Klemm, "Holomorphicity and modularity in Seiberg-Witten theories with matter," arXiv:0902.1325.

[157] B. Eynard, M. Mariño and N. Orantin, "Holomorphic anomaly and matrix models," JHEP **0706**, 058 (2007) [hep-th/0702110].

[158] A. Klemm, M. Mariño and S. Theisen, "Gravitational corrections in supersymmetric gauge theory and matrix models," JHEP **0303**, 051 (2003) [hep-th/0211216].

[159] S. Yamaguchi, "Wilson loops of anti-symmetric representation and D5-branes," JHEP **0605** (2006) 037 [hep-th/0603208].

[160] O. Lunin, "1/2-BPS states in M theory and defects in the dual CFTs," JHEP **0710**, 014 (2007) [arXiv:0704.3442].

[161] S. A. Hartnoll and S. P. Kumar, "Higher rank Wilson loops from a matrix model," JHEP **0608**, 026 (2006) [hep-th/0605027].

[162] T. Suyama, "On Large N Solution of Gaiotto-Tomasiello Theory," JHEP **1010**, 101 (2010) [arXiv:1008.3950 [hep-th]].

[163] G. Veneziano, "Some Aspects Of A Unified Approach To Gauge, Dual And Gribov Theories," Nucl. Phys. B **117**, 519 (1976).

[164] A. Karch and E. Katz, "Adding flavor to AdS/CFT," JHEP **0206**, 043 (2002) [arXiv:hep-th/0205236]..

[165] G. Mikhalkin, "Tropical geometry and its applications," International Congress of Mathematicians. Vol. II, 827–852, Eur. Math. Soc., Zürich, 2006 [arXiv:math/0601041].

[166] S. Iwao, "Integration over tropical plane curves and ultradiscretization," Int. Math. Res. Not. IMRN **2010**, no. 1, 112–148 [arXiv:0812.1873].

[167] M. Aganagic, A. Klemm and C. Vafa, "Disk instantons, mirror symmetry and the duality web," Z. Naturforsch. A **57**, 1 (2002) [arXiv:hep-th/0105045].

[168] J. Erdmenger, N. Evans, I. Kirsch and E. Threlfall, "Mesons in Gauge/Gravity Duals - A Review," Eur. Phys. J. A **35**, 81 (2008) [arXiv:0711.4467 [hep-th]].

[169] C. Núñez, A. Paredes and A. V. Ramallo, "Unquenched flavor in the gauge/gravity correspondence," Adv. High Energy Phys. **2010**, 196714 (2010) [arXiv:1002.1088 [hep-th]].

[170] J. Ambjorn, L. Chekhov, C. F. Kristjansen and Yu. Makeenko, "Matrix model calculations beyond the spherical limit," Nucl. Phys. B **404**, 127 (1993) [Erratum-ibid. B **449**, 681 (1995)] [arXiv:hep-th/9302014].

[171] B. Eynard, "Topological expansion for the 1-hermitian matrix model correlation functions," JHEP **0411**, 031 (2004) [arXiv:hep-th/0407261].

[172] M. Mariño, "Open string amplitudes and large order behavior in topological string theory," JHEP **0803**, 060 (2008) [arXiv:hep-th/0612127].

[173] G. Akemann, "Higher genus correlators for the Hermitian matrix model with multiple cuts," Nucl. Phys. B **482**, 403 (1996) [arXiv:hep-th/9606004].

[174] G. Mikhalkin and I. Zharkov, "Tropical curves, their Jacobians and Theta functions," in *Curves and abelian varieties*, Contemp. Math. **465** 203–230, Amer. Math. Soc., Providence, 2008 [arXiv:math/0612267].

[175] B. Eynard and N. Orantin, "Algebraic methods in random matrices and enumerative geometry," arXiv:0811.3531.

[176] D. Gaiotto and A. Tomasiello, "The gauge dual of Romans mass," JHEP **1001**, 015 (2010) [arXiv:0901.0969 [hep-th]].

[177] E. Brézin, J. -C. Le Guillou, J. Zinn-Justin, "Perturbation Theory at Large Order. 2. Role of the Vacuum Instability," Phys. Rev. **D15**, 1558-1564 (1977).

[178] M. Mariño, R. Schiappa and M. Weiss, "Non-perturbative effects and the large-order behavior of matrix models and Topological Strings," arXiv:0711.1954.

[179] M. Mariño, R. Schiappa, M. Weiss, "Multi-instantons and multi-cuts," J. Math. Phys. **50**, 052301 (2009). [arXiv:0809.2619].

[180] A. Klemm, M. Mariño and M. Rauch, "Direct integration and non-perturbative effects in matrix models," JHEP **1010**, 004 (2010) [arXiv:1002.3846].

[181] J. T. Liu and R. Minasian, "Computing $1/N^2$ corrections in AdS/CFT," arXiv:1010.6074.

[182] W. Taylor, "M(atrix) theory: Matrix quantum mechanics as a fundamental theory," Rev. Mod. Phys. **73**, 419-462 (2001). [hep-th/0101126].

[183] F. David, "Phases of the large N matrix model and non-perturbative effects in 2-D gravity," Nucl. Phys. B **348**, 507 (1991); "Non-perturbative effects in matrix models and vacua of two-dimensional gravity," Phys. Lett. B **302**, 403 (1993) [hep-th/9212106].

[184] N. Seiberg and D. Shih, "Branes, rings and matrix models in minimal (super)string theory," JHEP **0402**, 021 (2004) [hep-th/0312170].

[185] V. A. Kazakov and I. K. Kostov, "Instantons in non-critical strings from the two-matrix model," hep-th/0403152.

[186] J.C. Le Guillou and J. Zinn-Justin, *Large order behavior of perturbation theory*, North Holland, 1990.

[187] S. Pasquetti and R. Schiappa, "Borel and Stokes nonperturbative phenomena in topological string theory and $c = 1$ matrix models," Annales Henri Poincare **11**, 351 (2010) [arXiv:0907.4082].

[188] S. Garoufalidis, A. Its, A. Kapaev and M. Mariño, "Asymptotics of the instantons of Painlevé I," arXiv:1002.3634 [math.CA].

[189] S. Hikami, E. Brézin, "Large order behavior of the $1/n$ expansion in zero-dimensions and one-dimensions," J. Phys. A **A12**, 759-770 (1979).

[190] B. Eynard, J. Zinn-Justin, "Large order behavior of 2-D gravity coupled to $d < 1$ matter," Phys. Lett. **B302**, 396-402 (1993). [hep-th/9301004].

[191] M. Mariño, "Nonperturbative effects and nonperturbative definitions in matrix models and topological strings," JHEP **0812**, 114 (2008). [arXiv:0805.3033].

[192] D. Ghoshal, C. Vafa, "$C = 1$ string as the topological theory of the conifold," Nucl. Phys. **B453**, 121-128 (1995). [hep-th/9506122].

[193] A. M. Polyakov, "De Sitter space and eternity," Nucl. Phys. **B797**, 199-217 (2008). [arXiv:0709.2899]; "Global warming of De Sitter space," talk at Strings 2008 (CERN).

[194] Y. Hikida, W. Li, T. Takayanagi, "ABJM with Flavors and FQHE," JHEP **0907**, 065 (2009). [arXiv:0903.2194].

[195] P. Koerber, L. Martucci, "From ten to four and back again: How to generalize the geometry," JHEP **0708**, 059 (2007). [arXiv:0707.1038].

[196] P. Koerber, "Lectures on Generalized Complex Geometry for Physicists," Fortsch. Phys. **59**, 169-242 (2011). [arXiv:1006.1536].

[197] M. Cvetic, H. Lu and C. N. Pope, "Consistent warped-space Kaluza-Klein reductions, half-maximal gauged supergravities and $\mathbb{CP}^n$ constructions," Nucl. Phys. B **597** (2001) 172 [hep-th/0007109].

[198] T. Nishioka and T. Takayanagi, "On Type IIA Penrose Limit and $\mathcal{N} = 6$ Chern-Simons Theories," JHEP **0808** (2008) 001 [arXiv:0806.3391].

[199] T. Nishioka, T. Takayanagi, "Fuzzy Ring from M2-brane Giant Torus," JHEP **0810**, 082 (2008). [arXiv:0808.2691].

[200] M. Kruczenski and A. Tirziu, "Matching the circular Wilson loop with dual open string solution at 1-loop in strong coupling," JHEP **0805** (2008) 064 [arXiv:0803.0315].

[201] V. Bouchard, A. Klemm, M. Mariño and S. Pasquetti, "Remodeling the B-model," Commun. Math. Phys. **287**, 117 (2009) [arXiv:0709.1453 [hep-th]].

[202] N. Drukker and D. J. Gross, "An exact prediction of N = 4 SUSYM theory for string theory," J. Math. Phys. **42**, 2896 (2001) [arXiv:hep-th/0010274].

[203] N. Drukker and B. Fiol, "All-genus calculation of Wilson loops using D-branes," JHEP **0502**, 010 (2005) [arXiv:hep-th/0501109].

[204] A. Kapustin, B. Willett, I. Yaakov, "Nonperturbative Tests of Three-Dimensional Dualities," JHEP **1010**, 013 (2010). [arXiv:1003.5694 [hep-th]].

[205] O. Aharony, J. Marsano, S. Minwalla, K. Papadodimas and M. Van Raamsdonk, "A First order deconfinement transition in large N Yang-Mills theory on a small $\mathbb{S}^3$," Phys. Rev. **D71**, 125018 (2005). [hep-th/0502149].

[206] A. Cagnazzo, D. Sorokin, L. Wulff, "String instanton in $AdS_4 \times \mathbb{CP}^3$," JHEP **1005**, 009 (2010). [arXiv:0911.5228 [hep-th]].

[207] N. Ja. Vilenkin and A.U. Klimyk, *Representations of Lie groups and special functions*, volume I, Kluwer, 1991.

[208] I. Bars, "Supergroups And Their Representations," Lectures Appl. Math. **21**, 17 (1983).