CrossMark

Instanton Effects and Quantum Spectral Curves

Johan Källén and Marcos Mariño

Abstract. We study a spectral problem associated to the quantization of a spectral curve arising in local mirror symmetry. The perturbative WKB quantization condition is determined by the quantum periods, or equivalently by the refined topological string in the Nekrasov–Shatashvili (NS) limit. We show that the information encoded in the quantum periods is radically insufficient to determine the spectrum: there is an infinite series of instanton corrections, which are non-perturbative in $\hbar$, and lead to an exact WKB quantization condition. Moreover, we conjecture the precise form of the instanton corrections: they are determined by the standard or unrefined topological string free energy, and we test our conjecture successfully against numerical calculations of the spectrum. This suggests that the non-perturbative sector of the NS refined topological string contains information about the standard topological string. As an application of the WKB quantization condition, we explain some recent observations relating membrane instanton corrections in ABJM theory to the refined topological string.

1. Introduction

This paper is motivated by two different, but related problems. The first problem is the non-perturbative structure of topological string theory. Topological strings, like many other string models, are only defined perturbatively, and it is natural to ask whether one can define them non-perturbatively or find new non-perturbative sectors. In the last years, there have been many different proposals addressing this problem, but none of them seems to be conclusive. A *bona fide* non-perturbative definition must be based on a manifestly well-defined quantity, at least for a certain range of the relevant parameters of the model. This quantity should have an asymptotic expansion, for small values of the string coupling constant, which reproduces the original perturbative expansion. The asymptotic expansion and the perturbative expansion can only differ in quantities which are non-analytic at the origin (like for example instanton

effects). Of course, one might find different non-perturbative definitions of the same quantity, all of them differing in non-analytic terms. This is for example what happens in two-dimensional gravity [1]. In some cases, a reasonable physical criterium might single out one non-perturbative definition.

In this paper, we will analyze the non-perturbative structure of refined topological strings in the Nekrasov–Shatashvili (NS) limit [2]. This theory depends on the Calabi–Yau (CY) moduli and on a string coupling constant which is usually denoted by $\hbar$. In the original proposal of [2], this refined string was related, for some special geometries, to quantum integrable systems. It was later pointed out in [3–5] that the NS limit of the topological string free energy can be computed by quantizing a spectral curve, given by the mirror Calabi–Yau or a limit thereof. Using the quantum spectral curve, one can construct quantum periods, depending on $\hbar$, which define the free energy through an $\hbar$-deformed version of special geometry. The quantum periods have a nice interpretation in terms of one-dimensional Quantum Mechanics: the quantized spectral curve defines a spectral problem, and the quantum periods are quantum-corrected WKB periods, which lead to a quantization condition, at all orders in $\hbar$.

In this paper, we study a spectral problem related to the quantization of the curve describing the mirror of local $\mathbb{P}^1 \times \mathbb{P}^1$. We show that the quantum periods of [5] (which lead to a perturbative series in $\hbar$) are insufficient to solve the spectral problem: there is an infinite series of non-perturbative corrections in $\hbar$, of the instanton type. This is a well-known phenomenon in ordinary Quantum Mechanics. For example, in the double-well potential, the standard WKB quantization condition is insufficient to determine the spectrum, even after including all perturbative corrections in $\hbar$: one should also take into account instantons tunneling between the two vacua, and including these leads to a non-perturbative, exact quantization condition [6–9]. In the problem at hand, a similar phenomena occurs, but it is even more dramatic: for some values of $\hbar$, the quantization condition based on the quantum periods leads to an unphysical divergent expression. Instanton corrections are needed to cure the divergence.

A first-principle calculation of these instanton corrections is difficult, but we conjecture their precise form: they involve the *standard* (i.e., unrefined) topological string free energy. In particular, we write down an *exact* WKB quantization condition involving both the perturbative quantum periods and the instanton corrections. We perform a very precise test of this conjecture for some value of $\hbar$, by comparing the exact quantization condition to the numerical calculation of the spectrum. The agreement is excellent.

Our conjecture gives a novel realization of the Gopakumar–Vafa invariants of local $\mathbb{P}^1 \times \mathbb{P}^1$ in terms of a spectral problem. More generally, it suggests that the spectral problems associated to quantum spectral curves in local mirror symmetry involve the standard topological string, non-perturbatively. It also suggests that we should *define* non-perturbative topological strings on local CYs through a well-defined spectral problem associated to the quantization of the mirror curve. If the structure we find in our particular example does

actually generalize to other cases, this definition will lead in a single strike to the refined NS string (as the perturbative sector) and the conventional topological string (as the non-perturbative sector). Notice that this non-perturbative definition, in the example of local $\mathbb{P}^1 \times \mathbb{P}^1$, satisfies the above two criteria: it is based on a well-defined quantity (the spectrum), and it reproduces perturbatively the quantum B-period (which is the derivative of the free energy). Of course, the most interesting aspect of this definition is the appearance of the unrefined topological string free energy in the non-perturbative sector.

The second motivation for our work comes from ABJM theory [10]. In [11], the partition function of this theory on a three sphere was obtained by using localization techniques, and written as a matrix integral. Its full 't Hooft $1/N$ expansion was obtained in [12] by using large N techniques. In the paper [13], this matrix integral was written as the thermal partition function of an ideal, one-dimensional Fermi gas. One advantage of this approach, as compared to the standard large N techniques, is that one can also compute non-perturbative effects due to membrane instantons, which go beyond the 't Hooft expansion. In [14], building on previous work [15–18], it was conjectured that these non-perturbative effects are encoded in the quantum periods of local $\mathbb{P}^1 \times \mathbb{P}^1$. In this paper, we prove this conjecture to a large extent. The basic idea is simple: the spectral problem associated to the one-particle Hamiltonian of the Fermi gas is nothing but the spectral problem studied in this paper, i.e., it is a specialization of the spectral problem appearing in the quantization of the spectral curve of local $\mathbb{P}^1 \times \mathbb{P}^1$. The WKB analysis of the ABJM spectral problem leads immediately to the connection with these quantum periods. The grand potential of ABJM theory can be computed once we know the spectrum from the WKB quantization condition, and this makes it possible to derive many aspects of this grand potential which were conjectured in [14].

Many of our results on the WKB approach to the spectral problem are dual to the results on the grand potential of ABJM theory. For example, the fact that quantum periods are divergent for some values of $\hbar$ and should get non-perturbative corrections which cancel these divergences is a dual version of the HMO cancelation mechanism of [16]. Our conjecture on the instanton corrections to the spectral problem was motivated to a large extent by the known worldsheet instanton corrections to the grand potential of ABJM theory.

The organization of this paper is as follows: in Sect. 2, we present the spectral problem we will focus on. In Sect. 3, we do a WKB analysis of this problem. We first show that quantum periods are insufficient, we conjecture the form of the instanton corrections, and we perform a detailed test against the numerical calculation of the spectrum. In Sect. 4, we derive from the results in Sect. 3 the structure of the grand potential of ABJM theory, proving in this way some of the conjectures in [14]. Finally, in Sect. 5, we state our conclusions and directions for further research. Appendix A contains some results on Mellin transforms which are used in Sect. 4 whereas in Appendix B we find details of a calculation of the first quantum correction to the perturbative WKB quantization condition.

2. The Spectral Problem

We will consider a spectral problem arising in the quantization of the spectral curve describing the local CY known as local $\mathbb{P}^1 \times \mathbb{P}^1$. This is a spectral problem for a difference equation, with an appropriate and natural choice of analyticity and boundary conditions. The resulting problem can be equivalently formulated in terms of an integral equation which plays a crucial role in the Fermi gas approach to ABJM theory [13]. In this section, we will first consider the integral equation formulation, and then the formulation in terms of a difference equation.

Let us consider the following integral kernel

$$\rho(x_1, x_2) = \frac{1}{2\pi k} \frac{1}{\left(2\cosh\frac{x_1}{2}\right)^{1/2}} \frac{1}{\left(2\cosh\frac{x_2}{2}\right)^{1/2}} \frac{1}{2\cosh\left(\frac{x_1 - x_2}{2k}\right)}. \tag{2.1}$$

Here, k is a real parameter. This is a particular case of the family of kernels studied in [19,20]. The kernel (2.1) defines a non-negative, Hermitian, Hilbert–Schmidt operator $\hat{\rho}$ by

$$\langle x|\hat{\rho}|x'\rangle = \rho(x, x'). \tag{2.2}$$

Therefore, the spectral problem

$$\int_{-\infty}^{\infty} \rho(x_1, x_2)\phi(x_2)\mathrm{d}x_2 = \mathrm{e}^{-E}\phi(x_1) \tag{2.3}$$

leads to a discrete spectrum E_n, $n = 0, 1, 2, \ldots$, if we require the corresponding eigenfunctions $\phi_n(x)$ to be square-integrable. We will order the eigenvalues as

$$E_0 \leq E_1 \leq E_2 \leq \cdots. \tag{2.4}$$

In addition, by evaluating $\mathrm{Tr}\hat{\rho}^2$, we obtain the upper bound

$$\mathrm{e}^{-E_0} \leq \left(\int \rho^2(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2\right)^{1/2} \leq \frac{1}{4k}, \tag{2.5}$$

therefore $E_0 \geq 0$ for $k \geq 1/4$.

It is easy to reformulate (2.3) as a spectral problem for a difference equation. To do this, we write $\hat{\rho}$ as [13]

$$\hat{\rho} = \mathrm{e}^{-\frac{1}{2}U(\hat{x})}\mathrm{e}^{-T(\hat{p})}\mathrm{e}^{-\frac{1}{2}U(\hat{x})}. \tag{2.6}$$

In this equation, $\hat{x}, \hat{p}$ are canonically conjugate operators,

$$[\hat{x}, \hat{p}] = \mathrm{i}\hbar, \tag{2.7}$$

where

$$\hbar = 2\pi k, \tag{2.8}$$

and

$$U(x) = \log\left(2\cosh\frac{x}{2}\right), \qquad T(p) = \log\left(2\cosh\frac{p}{2}\right). \tag{2.9}$$

In this paper, we will use $\hbar$ and k interchangeably. The spectral problem (2.3) can now be written as

$$\hat{\rho}|\phi\rangle = \mathrm{e}^{-E}|\phi\rangle. \tag{2.10}$$

Let us now define

$$|\psi\rangle = e^{\frac{1}{2}U(\hat{x})}|\phi\rangle. \tag{2.11}$$

It follows that

$$e^{U(\hat{x})}e^{T(\hat{p})}|\psi\rangle = e^E|\psi\rangle \tag{2.12}$$

or, equivalently, in the coordinate representation,

$$\psi\left(x + i\pi k\right) + \psi\left(x - i\pi k\right) = \frac{e^E}{2\cosh\left(\frac{x}{2}\right)}\psi(x). \tag{2.13}$$

This difference equation is only equivalent to the original problem (2.3) provided some analyticity and boundary conditions are imposed on the function $\psi(x)$. Following [20], let us denote by $\mathcal{S}_a$ the strip in the complex x-plane defined by

$$|\text{Im}(x)| < a. \tag{2.14}$$

Let us also denote by $A(\mathcal{S}_a)$ those functions g which are bounded and analytic in the strip, continuous on its closure, and for which $g(x+iy) \to 0$ as $x \to \pm\infty$ through real values, when $y \in \mathbb{R}$ is fixed and satisfies $|y| < a$. It can be seen, by using the results in [20], that the equivalence of (2.13) and (2.3) requires that $\psi(x)$ belongs to the space $A(\mathcal{S}_{\pi k})$.

The operator $\hat{\rho}$ can be used to define a quantum Hamiltonian in the usual way [13],

$$\hat{\rho} = e^{-\hat{H}}, \tag{2.15}$$

whose classical limit is simply

$$H_{\text{cl}}(x,p) = T(p) + U(x). \tag{2.16}$$

In [13], it was noticed that the curve

$$\exp\left(T(p) + U(x)\right) = e^E, \tag{2.17}$$

defining the classical limit of the spectral problem, is a specialization of the curve describing the mirror of the Calabi–Yau known as local $\mathbb{P}^1 \times \mathbb{P}^1$. Let us write this curve as in [5]

$$e^u + z_1e^{-u} + e^v + z_2e^{-v} = 1, \tag{2.18}$$

where u, v are complex coordinates. Then, the curve (2.17) can be seen to be equal to the curve (2.18) after the change of variables

$$u = \frac{x + p}{2} - E, \qquad v = \frac{x - p}{2} - E \tag{2.19}$$

and the specialization

$$z_1 = z_2 = z, \tag{2.20}$$

where we have denoted, for convenience,

$$z = e^{-2E}. \tag{2.21}$$

Notice that the above change of variables is essentially a canonical transformation, since it preserves the symplectic form, up to an overall constant,

$$\mathrm{d}u \wedge \mathrm{d}v = -\frac{1}{2}\mathrm{d}x \wedge \mathrm{d}p. \tag{2.22}$$

The curve (2.18) can be quantized, leading to a *quantum spectral curve*. One simply promotes u, v to quantum operators $\hat{u}$, $\hat{v}$ satisfying canonical commutation relations. The equation satisfied by wave functions is

$$\left(\mathrm{e}^{\hat{u}} + z_1\mathrm{e}^{-\hat{u}} + \mathrm{e}^{\hat{v}} + z_2\mathrm{e}^{-\hat{v}} - 1\right)|\psi\rangle = 0. \tag{2.23}$$

We can now regard (2.13) as a particular case of (2.23) by promoting the classical change of variables (2.19) to a quantum one,

$$\hat{u} = \frac{\hat{x} + \hat{p}}{2} + \frac{\mathrm{i}\pi k}{4} - E, \qquad \hat{v} = \frac{\hat{x} - \hat{p}}{2} - \frac{\mathrm{i}\pi k}{4} - E, \tag{2.24}$$

while the specialization (2.20) has now the quantum correction[1]

$$z_1 = q^{1/2}z, \qquad z_2 = q^{-1/2}z, \tag{2.25}$$

where

$$q = \mathrm{e}^{\frac{\mathrm{i}\hbar}{2}} = \mathrm{e}^{\pi \mathrm{i} k}. \tag{2.26}$$

In terms of $\hbar$ as defined in (2.8), we have

$$[\hat{v}, \hat{u}] = \frac{\mathrm{i}\hbar}{2}. \tag{2.27}$$

We conclude that the results for the quantum spectral curve (2.23) obtained for example in [5,14] can be specialized to study the spectral problem (2.3) and (2.13). We emphasize that our original spectral problem is defined by the integral kernel (2.1). The results of [5] will be used to obtain the all-orders perturbative WKB series for the quantum volume associated to the spectral problem (2.3).

The difference Eq. (2.13) has the structure of Baxter's TQ equation, which determines the spectrum of a quantum integrable system and can be regarded as a quantization of the spectral curve of the classical system. This similarity is not surprising: as it is well known, the curve (2.18), describing the mirror of local $\mathbb{P}^1 \times \mathbb{P}^1$, can be regarded as a relativistic deformation of the spectral curve of the periodic Toda chain [21]. The Baxter equations for Toda and relativistic Toda have been studied in [22] and [23], respectively. Similar difference equations also appear in the study of $\mathcal{N} = 2$ supersymmetric gauge theories in the NS limit [24,25]. (2.13) is also a close cousin of the difference equation studied in [26], which has its origin in the integral equation of the 't Hooft model.

Unfortunately, the spectral problem (2.3) does not seem to be exactly solvable, and one has to use numerical or approximate methods.

---

[1] This change of variables was previously observed by Kazumi Okuyama.

3. The Exact WKB Quantization Condition

3.1. Perturbative WKB Quantization

We will now analyze the spectral problem (2.13) by using the WKB method. At leading order in $\hbar$, the WKB method for bound states is the Bohr–Sommerfeld quantization condition. In this approximation, one calculates the classical volume of phase space as a function of the energy, $\mathrm{vol}_0(E)$ (here, the subscript 0 means that we are working at zero order in the $\hbar$ expansion). The quantization condition says that this volume should be a half-integer multiple[2] of the volume of an elementary cell in phase space, $2\pi\hbar$, and one obtains

$$\mathrm{vol}_0(E) = 2\pi\hbar\left(n + \frac{1}{2}\right), \qquad n = 0, 1, 2, \ldots. \tag{3.1}$$

The classical volume of phase space was already determined in [13]. It is given by a period integral on the curve (2.17). Since this is an elliptic curve, it has two periods, the A and the B periods. The spectral problem we are looking at involves the B period, and we find

$$\mathrm{vol}_0(E) = \oint_B \lambda, \qquad \lambda = p(x)\mathrm{d}x, \tag{3.2}$$

where $p(x)$ is obtained by solving (2.17). The calculation in [13] expresses this period in terms of a Meijer G-function

$$\mathrm{vol}_0(E) = \frac{\mathrm{e}^E}{\pi} G_{3,3}^{2,3}\left(\frac{\mathrm{e}^{2E}}{16} \left|\begin{array}{c} \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \\ 0, 0, -\frac{1}{2} \end{array}\right.\right) - 4\pi^2 = 8E^2 - \frac{4\pi^2}{3} + \mathcal{O}\left(E\,\mathrm{e}^{-2E}\right). \tag{3.3}$$

In the WKB method, we are interested in large energies as compared to $\hbar$, i.e., in large quantum numbers. It is then useful to have a basis of classical periods of the curve which is appropriate for the $E \gg 1$ regime. Since the curve (2.17) is a specialization of the mirror of local $\mathbb{P}^1 \times \mathbb{P}^1$, the relevant periods are nothing but the large radius periods of this CY. Let us now review some basic facts about these periods.

In the full CY (2.18), i.e., for generic values of z_1, z_2, it is useful to consider two different A periods and two different B periods. The A periods are given by

$$\Pi_{A_I}(z) = \log z_I + \widetilde{\Pi}_A(z_1, z_2), \qquad I = 1, 2, \tag{3.4}$$

where

$$\widetilde{\Pi}_A(z_1, z_2) = 2 \sum_{\substack{k, l \geq 0, \\ (k,l) \neq (0,0)}} \frac{\Gamma(2k + 2l)}{\Gamma(1 + k)^2 \Gamma(1 + l)^2} z_1^k z_2^l$$

$$= 2z_1 + 2z_2 + 3z_1^2 + 12z_1 z_2 + 3z_2^2 + \cdots \tag{3.5}$$

---

[2] The fact that this is a half-integer, and not an integer, can be shown by using a next-to-leading WKB analysis, as in [22]: the presence of an inverse square-root factor in the WKB wave function leads to an extra phase in going around a cut.

There are two independent B-periods, $\Pi_{B_I}(z_1, z_2)$, $I = 1, 2$, which are related by the exchange of z_1 and z_2,

$$\Pi_{B_2}(z_1, z_2) = \Pi_{B_1}(z_2, z_1). \tag{3.6}$$

The B_1 period is given by

$$\Pi_{B_1}(z_1, z_2) = -\frac{1}{8} \left(\log^2 z_1 - 2 \log z_1 \log z_2 - \log^2 z_2 \right)$$
$$+ \frac{1}{2} \log z_2 \, \widetilde{\Pi}_A(z_1, z_2) + \frac{1}{4} \widetilde{\Pi}_B(z_1, z_2), \tag{3.7}$$

where

$$\widetilde{\Pi}_B(z_1, z_2) = 8 \sum_{\substack{k,l \geq 0, \\ (k,l) \neq (0,0)}} \frac{\Gamma(2k+2l)}{\Gamma(1+k)^2 \Gamma(1+l)^2} \left(\psi(2k+2l) - \psi(1+l) \right) z_1^k z_2^l$$
$$= 8z_1 + 22z_1^2 + 40z_1 z_2 + 4z_2^2 + \cdots \tag{3.8}$$

In our spectral problem we have $z_1 = z_2$ classically [see (2.20)]. In this limit, one has

$$\widetilde{\Pi}_A(z) \equiv \widetilde{\Pi}_A(z,z) = \sum_{\ell \geq 1} \widehat{a}_\ell^{(0)} z^\ell, \qquad \widetilde{\Pi}_B(z) \equiv \widetilde{\Pi}_B(z,z) = \sum_{\ell \geq 1} \widehat{b}_\ell^{(0)} z^\ell, \tag{3.9}$$

where

$$\widehat{a}_\ell^{(0)} = \frac{1}{\ell} \left(\frac{\Gamma\left(\ell + \frac{1}{2}\right)}{\Gamma(\frac{1}{2})\ell!} \right)^2 16^\ell,$$

$$\widehat{b}_\ell^{(0)} = \frac{4}{\ell} \left(\frac{\Gamma\left(\ell + \frac{1}{2}\right)}{\Gamma(\frac{1}{2})\ell!} \right)^2 16^\ell \left[\psi\left(\ell + \frac{1}{2}\right) - \psi(\ell+1) + 2\log 2 - \frac{1}{2\ell} \right]. \tag{3.10}$$

The classical B-period becomes

$$\Pi_B(z) \equiv \Pi_{B_{1,2}}(z,z) = \frac{1}{4} \left(\log z \right)^2 + \frac{1}{2} \log z \, \widetilde{\Pi}_A(z) + \frac{1}{4} \widetilde{\Pi}_B(z), \tag{3.11}$$

and one finds that the volume of phase space can be written in terms of this period as

$$\mathrm{vol}_0(E) = 8\Pi_B \left(\mathrm{e}^{-2E} \right) - \frac{4\pi^2}{3}$$
$$= 8E^2 - \frac{4\pi^2}{3} - 8E \sum_{\ell \geq 1} \widehat{a}_\ell^{(0)} \mathrm{e}^{-2\ell E} + 2 \sum_{\ell \geq 1} \widehat{b}_\ell^{(0)} \mathrm{e}^{-2\ell E}. \tag{3.12}$$

It is well known that the Bohr–Sommerfeld quantization condition has perturbative corrections in $\hbar$. These can be obtained in a straightforward way by solving the Eq. (2.13) with a WKB ansatz,

$$\psi(x, \hbar) = \exp\left(\frac{1}{\hbar} S(x, \hbar) \right), \tag{3.13}$$

where

$$S(x, \hbar) = \sum_{n \geq 0} S_n(x) \hbar^n, \tag{3.14}$$

and interpreting $\partial_x S(x, \hbar)\mathrm{d}x$ as a "quantum" differential. The leading order approximation gives

$$S_0'(x) = p(x) \tag{3.15}$$

and reproduces the Bohr–Sommerfeld quantization condition. Using this quantum differential, we can define the perturbative, "quantum" volume of phase space as

$$\mathrm{vol_p}(E; \hbar) = \oint_B \partial_x S(x, \hbar)\mathrm{d}x. \tag{3.16}$$

We could calculate these corrections directly in the Eq. (2.13). However, it is more illuminating to obtain them as particular cases of the quantum corrections for the spectral curve (2.23). Indeed, as explained in [3–5] and reviewed in [14], these corrections promote the classical periods $\Pi_{A_I}(z_1, z_2)$, $\Pi_{B_I}(z_1, z_2)$ to quantum A-periods

$$\Pi_{A_I}(z_1, z_2; \hbar) = \log z_I + \widetilde{\Pi}_A(z_1, z_2; \hbar), \quad I = 1, 2, \tag{3.17}$$

and quantum B-periods $\Pi_{B_I}(z_1, z_2; \hbar)$, $I = 1, 2$. As in the classical case, there are two of them, but they are related by the exchange of the moduli,

$$\Pi_{B_2}(z_1, z_2; \hbar) = \Pi_{B_1}(z_2, z_1; \hbar). \tag{3.18}$$

The quantum counterpart of (3.7) is

$$\Pi_{B_1}(z_1, z_2; \hbar) = -\frac{1}{8}\left(\log^2 z_1 - 2\log z_1 \log z_2 - \log^2 z_2\right) + \frac{1}{2}\log z_2\, \widetilde{\Pi}_A(z_1, z_2; \hbar)$$
$$+ \frac{1}{4}\widetilde{\Pi}_B(z_1, z_2; \hbar). \tag{3.19}$$

These quantum periods can be computed systematically in a power series in $z_{1,2}$ [5,14]. One finds, to the very first orders,

$$\widetilde{\Pi}_A(z_1, z_2; \hbar) = 2(z_1 + z_2) + 3\left(z_1^2 + z_2^2\right) + 2\left(4 + q + q^{-1}\right)z_1 z_2 + \frac{20}{3}\left(z_1^3 + z_2^3\right)$$
$$+ 2\left(16 + 6q + 6q^{-1} + q^2 + q^{-2}\right)z_1 z_2(z_1 + z_2) + \mathcal{O}\left(z_i^4\right),$$
$$\widetilde{\Pi}_B(z_1, z_2; \hbar) = 8\left[\frac{q+1}{2(q-1)}\log q\right]z_1 + 4\left[1 + \frac{5q^2 + 8q + 5}{2(q^2 - 1)}\log q\right]z_1^2$$
$$+ 8\left[1 + \frac{(1+q)^3}{2q(q-1)}\log q\right]z_1 z_2 + 4z_2^2 + \mathcal{O}\left(z_i^3\right), \tag{3.20}$$

where q is given in (2.26).

Let us now come back to the problem of calculating (3.16). This is a quantum period for the spectral curve defined by (2.13), but this curve is just a specialization of (2.23) with the dictionary (2.25) and after a canonical transformation. Therefore, (3.16) should be a combination of the quantum

periods of local $\mathbb{P}^1 \times \mathbb{P}^1$, specialized to the "slice" (2.25). Let us denote

$$\widetilde{\Pi}_A(z;\hbar) \equiv \widetilde{\Pi}_A\left(q^{1/2}z, q^{-1/2}z;\hbar\right) = \sum_{\ell\geq 1} \widehat{a}_\ell(\hbar)z^\ell,$$

$$\widetilde{\Pi}_B(z;\hbar) \equiv \frac{1}{2}\left(\widetilde{\Pi}_B\left(q^{1/2}z, q^{-1/2}z;\hbar\right) + \widetilde{\Pi}_B\left(q^{-1/2}z, q^{1/2}z;\hbar\right)\right) = \sum_{\ell\geq 1} \widehat{b}_\ell(\hbar)z^\ell,$$
(3.21)

where $\widehat{a}_\ell(\hbar)$, $\widehat{b}_\ell(\hbar)$ have the $\hbar$-expansion,

$$\widehat{a}_\ell(\hbar) = \sum_{n=0}^\infty \widehat{a}_\ell^{(n)}\hbar^{2n}, \qquad \widehat{b}_\ell(\hbar) = \sum_{n=0}^\infty \widehat{b}_\ell^{(n)}\hbar^{2n}. \tag{3.22}$$

Requiring the combination of quantum periods to have the correct classical limit, and that only even powers of $\hbar$ appear, we find,

$$\mathrm{vol_p}(E;\hbar) = 4\Pi_{B_1}\left(q^{1/2}z, q^{-1/2}z;\hbar\right) + 4\Pi_{B_2}\left(q^{1/2}z, q^{-1/2}z;\hbar\right) - \frac{4\pi^2}{3} - \frac{\hbar^2}{12}$$

$$= 8E^2 - \frac{4\pi^2}{3} + \frac{\hbar^2}{24} - 8E\sum_{\ell\geq 1}\widehat{a}_\ell(\hbar)\mathrm{e}^{-2\ell E} + 2\sum_{\ell\geq 1}\widehat{b}_\ell(\hbar)\mathrm{e}^{-2\ell E}. \tag{3.23}$$

The third term in the last line of (3.23) is an E-independent correction to the quantum period which was already computed in [13]. Notice that the perturbative quantum volume has an $\hbar$-expansion of the form,

$$\mathrm{vol_p}(E) = \sum_{n\geq 0}\mathrm{vol}_n(E)\hbar^{2n}. \tag{3.24}$$

The first term in this expansion is the function of E given in (3.3). By the WKB expansion of the wave function (3.13), (3.14), it is possible to find an exact expression for the first quantum correction, which reads

$$\mathrm{vol}_1(E) = \frac{\mathrm{e}^{-E}\left(\left(32\,\mathrm{e}^{-2E} - 1\right)E(k) - K(k)\right)}{6\left(16\,\mathrm{e}^{-2E} - 1\right)}, \tag{3.25}$$

where $K(k)$, $E(k)$ are elliptic integrals of the first and second kind, respectively, and their modulus is given by

$$k^2 = 1 - \frac{\mathrm{e}^{2E}}{16}. \tag{3.26}$$

The derivation of this result is sketched in Appendix B.

The Eq. (3.16) gives the full series of perturbative $\hbar$ corrections to the classical phase-space volume. The perturbatively exact quantization condition involves the quantum B-periods of the spectral curve, and it reads

$$\mathrm{vol_p}(E;\hbar) = 2\pi\hbar\left(n + \frac{1}{2}\right), \qquad n = 0, 1, 2, \ldots \tag{3.27}$$

We can now use (3.27) to compute the quantum-corrected spectrum. For example, we can use the explicit expressions (3.3) and (3.25) to compute the energies E_n perturbatively, as a power series expansion around $\hbar = 0$, i.e.,

$$E_n^{\mathrm{p}} = \sum_{\ell=0}^{\infty} E_{n,\ell} \hbar^\ell. \tag{3.28}$$

The leading order term corresponds to the zero of $\mathrm{vol}_0(E)$, and as found in [13] this is,

$$E_{n,0} = 2 \log 2. \tag{3.29}$$

Plugging now the series (3.28) in (3.27), we find

$$E_{n,1} = \frac{2n+1}{8}, \quad E_{n,2} = -\frac{2n^2+2n+1}{128}. \tag{3.30}$$

These values agree with a calculation starting directly from the density operator (2.6) [28].

3.2. Non-Perturbative WKB Quantization

As we have seen, the perturbative WKB quantization condition (3.27) makes it possible to compute the energies as a power series in $\hbar$. However, when we consider finite values of $\hbar$, we find a key problem: the coefficients $\widehat{b}_\ell(\hbar)$ appearing in the expansion (3.23) *diverge* for any rational k, and lead to a nonsensical WKB expansion when $\hbar$ approaches 2π times a rational number. At the same time, there is no physical source for this divergence in the spectral problem itself: the eigenvalues for E appearing in (2.13) and in the associated integral Eq. (2.3) are perfectly well defined for any real value of $\hbar$, and in particular for rational k. As a matter of fact they can be computed numerically, as we will see in the next subsection. We stress that the divergence problem in (3.27) is not an artifact of the large E expansion used to obtain (3.23). This leads to an asymptotic expansion for the energy levels valid for large quantum numbers, which should be well defined. It can be shown, by using the BPS structure of the refined topological string free energy [29,30], that the quantum B-periods of *any* local CY manifold are divergent for an infinite number of values of $\hbar$. In particular, the WKB quantization condition for B-periods written down in [5] has an infinite number of poles in the complex $\hbar$-plane.

We conclude that the expression (3.23) is *incomplete*, and there *must* be an extra correction which makes the quantum volume of phase space finite and leads to a reasonable quantization condition. Since (3.27) already incorporates all the perturbative information available, this correction must be non-perturbative in $\hbar$.

The possibility of having instanton corrections to the quantum volume of phase space was already anticipated in [13]. As pointed out there, to understand these corrections we need the geometric approach to the WKB method developed in for example [31–35]. In this approach, the perturbative WKB quantization condition is associated to a classical periodic orbit of energy E and the quantum fluctuations around it. The classical action of this trajectory is the classical B-period, and the perturbative $\hbar$ corrections promote it to a quantum period. In our example, this perturbative analysis leads to the result

(3.23). In fact, the classical periodic orbits can be described in detail. The classical Hamiltonian (2.16) leads to the equations of motion

$$\dot{x} = \frac{1}{2}\tanh\frac{p}{2}, \qquad \dot{p} = -\frac{1}{2}\tanh\frac{x}{2}. \tag{3.31}$$

For an orbit of energy E, we find

$$\dot{x} = \frac{1}{2}\sqrt{1 - 16\,e^{-2E}\cosh^2\frac{x}{2}}, \tag{3.32}$$

which can be integrated in terms of Jacobi's elliptic sine,

$$\tanh\frac{x}{2} = k\,\mathrm{sn}\left(\frac{t}{4}, k\right), \tag{3.33}$$

where the modulus k is now given by

$$k^2 = 1 - 16\,e^{-2E}. \tag{3.34}$$

The function $\mathrm{sn}(u,k)$ is doubly periodic in u. It has a real period given by

$$\omega_1 = 4K(k), \tag{3.35}$$

where $K(k)$ is the complete elliptic integral of the first kind. The action around this periodic trajectory is

$$S_B(E) = \int_{-8K(k)}^{8K(k)} p(t)\dot{x}(t)\mathrm{d}t = \mathrm{vol}_0(E), \tag{3.36}$$

where we took into account a relative factor of 4 coming from $u = t/4$.

However, as pointed out in [31,32], to obtain a non-perturbative quantization condition one should take into account as well *complexified* trajectories. These trajectories are allowed because the curve (2.17) has genus one and therefore it has an imaginary period, on top of the real period associated to the real periodic orbit (an example of such a situation was discussed recently in [37]). In our case, this is just the imaginary period of the Jacobi sine function,

$$\omega_2 = 2\mathrm{i}K'(k), \tag{3.37}$$

where

$$K'(k) = K(k'), \qquad k'^2 = 1 - k^2. \tag{3.38}$$

In the complexified orbit, time is imaginary, as in [31,32]: $t = \mathrm{i}\tau$. If we now use the relation

$$\mathrm{sn}(\mathrm{i}u, k) = \mathrm{i}\,\mathrm{tn}(u, k'), \tag{3.39}$$

where $\mathrm{tn} = \mathrm{sn}/\mathrm{cn}$, we find that x is also imaginary: $x = \mathrm{i}\theta$. The equation for the complex trajectory becomes

$$\tan\frac{\theta}{2} = k\,\mathrm{tn}\left(\frac{\tau}{4}, k'\right). \tag{3.40}$$

The real and complexified trajectories, (3.33), (3.40), are represented in Fig. 1 for the value of the energy $E = 3$.

Geometrically, the trajectory (3.33) describes a closed orbit in phase space, along the hypersurface of constant energy $H_{\mathrm{cl}}(x,p) = E$. This is the
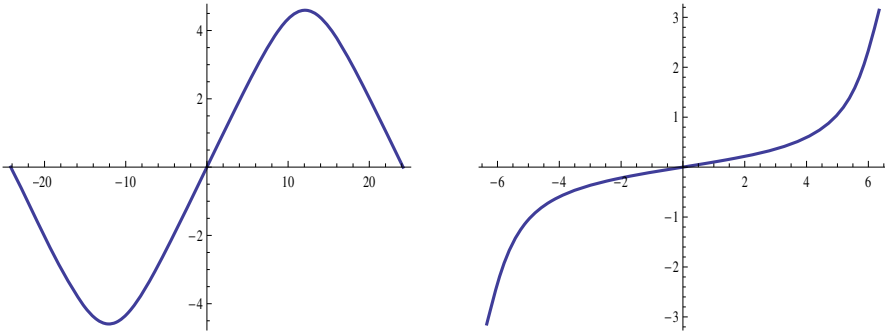
FIGURE 1. For a given energy E, we have a real periodic trajectory given by (3.33), which is shown in the figure on the *left* for $E = 3$. The *horizontal axis* represents the time t and it runs through a full period, from $-8K(k)$ to $8K(k)$. There is also an imaginary periodic trajectory for $\theta = \text{Im}(q)$, given by (3.40), which is shown in the figure on the *right*, also for $E = 3$. Here, the imaginary time τ runs from $-4K(k')$ to $4K(k')$

Fermi surface of the ideal Fermi gas introduced in [13]. After complexifying the exponentiated variable $e^{x/2}$, this closed orbit becomes a torus. The imaginary trajectory (3.40) is a closed orbit around the A-cycle of this torus, while (3.33) is now regarded as a closed orbit around the B-cycle. We depict both orbits in Fig. 2.

In the original Hamiltonian, θ has a periodicity of 4π, therefore the relevant action is

$$S_A(E) = 2i \int_{-4K'(k)}^{4K'(k)} p(\tau)\dot{\theta}(\tau)\mathrm{d}\tau = -4\pi i \Pi_A(z), \qquad (3.41)$$

where we have denoted

$$\Pi_A(z) = \log z + \widetilde{\Pi}_A(z, z). \qquad (3.42)$$

The contribution of such a complex trajectory to the quantization condition is of the form

$$\exp\left(\frac{i}{\hbar}S_A\right) = \exp\left(\frac{2}{k}\Pi_A(z)\right). \qquad (3.43)$$

Including the quantum corrections simply promotes the classical period to its quantum counterpart, as already noticed in [31,32]. As in [34,35], we will call the exponentiated quantum period associated to a cycle a *Voros multiplier*. The quantum A-period, specialized to the "slice" (2.25), is

$$\Pi_{A_I}\left(q^{1/2}z, q^{-1/2}z; \hbar\right) = \log z \pm \frac{i\pi k}{2} + \sum_{\ell \geq 1} \widehat{a}_\ell(\hbar)z^\ell, \qquad I = 1, 2, \quad (3.44)$$
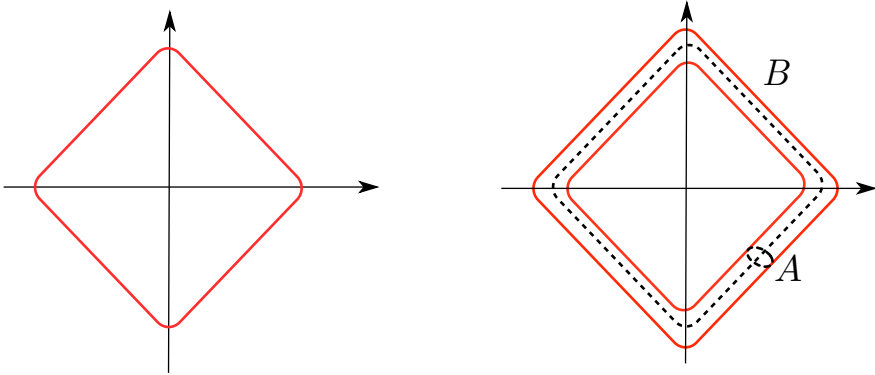
FIGURE 2. The trajectory (3.33) describes a closed orbit in the phase space (x, p), represented schematically in the figure on the *left*. After complexifying the exponentiated variable $e^{x/2}$, this closed orbit becomes a torus, as shown in the figure on the *right*. The imaginary trajectory (3.40) is a closed orbit around the A-cycle of the torus

where the $\pm$ sign corresponds to $I = 1, 2$, respectively. We conclude from (3.43) that the appropriate Voros multiplier for the A-period in this theory is

$$\exp\left[\frac{2}{k}\Pi_{A_I}\left(q^{1/2}z, q^{-1/2}z; \hbar\right)\right] = -e^{-4E_{\text{eff}}/k}, \tag{3.45}$$

where

$$E_{\text{eff}} = E - \frac{1}{2}\sum_{\ell=1}^{\infty}\widehat{a}_\ell(\hbar)e^{-2\ell E}. \tag{3.46}$$

This result was also obtained in [13].[3] In general, the non-perturbative correction to the quantum volume is a formal power series in the Voros multiplier for the quantum A-period. In our case, it takes the form,

$$\text{vol}_{\text{np}}(E; \hbar) = \sum_{m=1}^{\infty} s_m(k)(-1)^m e^{-4mE_{\text{eff}}/k}. \tag{3.47}$$

This is clearly non-perturbative in $\hbar$ (or equivalently, in k), and it is invisible in the standard perturbative correction to the WKB condition.

The calculation of $s_m(k)$ from first principles is difficult. In the case of the Schrödinger equation studied in [31–35], the instanton corrections are determined by the vanishing of the so-called Jost function, and their calculation requires a detailed analysis of the spectral problem and of the WKB wave function. We will now present a conjecture for the form of the instanton corrections for the spectral problem (2.13). To write down our formula, let

---

[3] In [13], what is here called the A-period was called the B-period, and viceversa. The notation we are using agrees with the standard conventions for periods in local Calabi–Yau manifolds.

us consider the worldsheet instanton corrections to the unrefined topological string free energy on local $\mathbb{P}^1 \times \mathbb{P}^1$, in the Gopakumar–Vafa form [27]:

$$
\begin{aligned}
F(T_1, T_2, g_s) \\
= \sum_{g \geq 0} \sum_{w \geq 1} \sum_{d_1, d_2} \frac{(-1)^{g-1}}{w} n_g^{d_1, d_2} \left(q_s^{w/2} - q_s^{-w/2} \right)^{2g-2} e^{-w(d_1 T_1 + d_2 T_2)}.
\end{aligned} \quad (3.48)
$$

Here,

$$
q_s = e^{g_s}, \quad (3.49)
$$

and g_s is the topological string coupling constant. In (3.48), T_1, T_2 are the complexified Kähler classes, corresponding to the two compact $\mathbb{P}^1$s in the geometry, and $n_g^{d_1, d_2}$ are the Gopakumar–Vafa invariants of local $\mathbb{P}^1 \times \mathbb{P}^1$ for genus g and degrees d_1, d_2. Along the "slice"

$$
T_1 = T_2 = T \quad (3.50)
$$

the natural invariant is the diagonal one,

$$
n_g^d = \sum_{d_1 + d_2 = d} n_g^{d_1, d_2}. \quad (3.51)
$$

We are now ready to state our conjecture about the form of the instanton corrections. We claim that the coefficients $s_m(k)$ in (3.47) are given by

$$
s_m(k) = -4\pi k \sin \left(\frac{4\pi m}{k} \right) d_m(k), \quad (3.52)
$$

where

$$
d_m(k) = \frac{1}{m} \sum_{g \geq 0} \sum_{d|m} d \, n_g^d \left(2 \sin \frac{2\pi m}{dk} \right)^{2g-2}. \quad (3.53)
$$

The coefficients $d_m(k)$ have a simple interpretation: they are the coefficients of e^{-mT} in the topological string free energy of local $\mathbb{P}^1 \times \mathbb{P}^1$, along the slice (3.50), and for

$$
g_s = \frac{4\pi \mathrm{i}}{k}. \quad (3.54)
$$

We then conjecture that the non-perturbative contribution to the quantum volume of phase space is

$$
\mathrm{vol}_{\mathrm{np}}(E; \hbar) = -4\pi k \sum_{m=1}^{\infty} \sin \left(\frac{4\pi m}{k} \right) (-1)^m d_m(k) e^{-4m E_{\mathrm{eff}}/k}, \quad (3.55)
$$

and the total quantum volume is

$$
\mathrm{vol}(E; \hbar) = \mathrm{vol}_{\mathrm{p}}(E; \hbar) + \mathrm{vol}_{\mathrm{np}}(E; \hbar). \quad (3.56)
$$

Notice that the non-perturbative volume (3.55) is also divergent for rational values of k. We will now show that the total volume (3.56) is well defined, i.e., the divergences in $\mathrm{vol}_{\mathrm{p}}(E; \hbar)$ cancel against the divergences in $\mathrm{vol}_{\mathrm{np}}(E; \hbar)$.

Before showing this, let us give some indications on the origin of this conjecture. As we will see in the next section, the spectral problem (2.13) appears in the Fermi gas approach to ABJM theory. It turns out that the grand potential of ABJM theory is closely related to the volume of phase space. The perturbative part (3.23) leads to the non-perturbative membrane corrections to the grand potential, while the non-perturbative part (3.55) leads to the worldsheet instanton corrections to the grand potential. The conjecture (3.55) is inspired by the known form of these corrections [12,16,36]. The requirement that divergences should cancel in the total volume (3.56) is a dual manifestation of the HMO cancelation mechanism discovered in [16]. According to this mechanism, the divergences in the worldsheet instanton part of the grand potential of ABJM theory should cancel against the divergences in the membrane instanton part, since the total grand potential is well defined and finite for any k. We have here a similar mechanism, which is based this time on the fact that the spectral problem is well defined for any value of k. The cancelation mechanism in the quantum volume is simpler however than the HMO mechanism, since it only involves simple poles, while the HMO mechanism involves double poles.

Let us now verify that $\mathrm{vol}(E;\hbar)$ is well defined for any value of k. Since the non-perturbative contribution is defined in terms of E_{eff}, instead of E, let us reexpress the perturbative part $\mathrm{vol}_{\mathrm{p}}(E;\hbar)$ in terms of this variable. This defines a new set of coefficients $\widetilde{b}_\ell(k)$ as follows,

$$\mathrm{vol}_{\mathrm{p}}(E;\hbar) = 8E_{\mathrm{eff}}^2 - \frac{4\pi^2}{3} + \frac{\hbar^2}{24} + 4\pi^2 k \sum_{\ell=1}^{\infty} \widetilde{b}_\ell(k) \mathrm{e}^{-2\ell E_{\mathrm{eff}}}. \tag{3.57}$$

These coefficients were first introduced in [18], in the context of ABJM theory, and their geometric meaning was uncovered in [14]: they can be expressed in terms of the refined BPS invariants $N_{j_L,j_R}^{d_1,d_2}$ of local $\mathbb{P}^1 \times \mathbb{P}^1$ as

$$\widetilde{b}_\ell(k) = -\frac{\ell}{2\pi} \sum_{j_L,j_R} \sum_{\ell=dw} \sum_{d_1+d_2=d} N_{j_L,j_R}^{d_1,d_2} q^{\frac{w}{2}(d_1-d_2)}$$
$$\times \frac{\sin\frac{\pi k w}{2}(2j_L+1)\sin\frac{\pi k w}{2}(2j_R+1)}{w^2 \sin^3\frac{\pi k w}{2}}. \tag{3.58}$$

This formula is based on the fact that the combination of B-periods appearing in (3.23) is a derivative of the refined topological string free energy, which in turn can be written in terms of refined BPS invariants (see for example [30] for a summary of these facts, and a list of values of the refined invariants for low degrees). On the other hand, the coefficient (3.53) can be expressed in terms of the same invariants by the formula,

$$d_m(k) = \sum_{j_L,j_R} \sum_{m=dn} \sum_{d_1+d_2=d} N_{j_L,j_R}^{d_1,d_2} \frac{2j_R+1}{\left(2\sin\frac{2\pi n}{k}\right)^2} \frac{\sin\left(\frac{4\pi n}{k}(2j_L+1)\right)}{\sin\frac{4\pi n}{k}} \frac{1}{n},$$
$$\tag{3.59}$$

see [14] for a derivation. We can now use a simplified version of the argument appearing in [14] to check that the singularities in (3.56) cancel. First of all, notice that the singularities appear when k takes the form

$$k = \frac{2n}{w} = \frac{2m}{\ell}. \tag{3.60}$$

The singularities are simple poles. The poles appearing in $\mathrm{vol}_{\mathrm{np}}(E; \hbar)$ are of the form

$$(-1)^m \frac{8m}{w^3 \left(k - \frac{2n}{w}\right)} (1 + 2j_L)(1 + 2j_R) N_{j_L, j_R}^{d_1, d_2} \mathrm{e}^{-\frac{2mw}{n} E_{\mathrm{eff}}}. \tag{3.61}$$

The corresponding poles appearing in (3.57) are of the form

$$-\mathrm{e}^{\pi\mathrm{i} k w (d_1 - d_2)/2} \frac{8m}{w^3 \left(k - \frac{2n}{w}\right)} (-1)^{n(2j_L + 2j_R - 1)} (1 + 2j_L)(1 + 2j_R) N_{j_L, j_R}^{d_1, d_2} \mathrm{e}^{-2\ell E_{\mathrm{eff}}}. \tag{3.62}$$

By using (3.60), one notices that

$$\mathrm{e}^{\pi\mathrm{i} k w (d_1 - d_2)/2} = (-1)^m, \tag{3.63}$$

and it is easy to see that all poles in (3.61) cancel against the poles in (3.62), for any value of E_{eff}, provided that

$$(-1)^{n(2j_L + 2j_R - 1)} = 1. \tag{3.64}$$

This can be seen to be the case by a geometric argument explained in [14]. We conclude that $\mathrm{vol}(E; \hbar)$ is well defined and finite for any real value of $\hbar$, as a series in e^{-2E}.

The exact WKB quantization condition reads now

$$\mathrm{vol}(E; \hbar) = 2\pi\hbar \left(n + \frac{1}{2}\right), \qquad n = 0, 1, 2, \ldots, \tag{3.65}$$

where $\mathrm{vol}(E; \hbar)$ is a sum of the perturbative part (3.23) and the non-perturbative part (3.55). This condition determines the energy levels E_n as functions of n and $\hbar$. As in similar examples of exact WKB quantization conditions, the total $\mathrm{vol}(E)$ is a trans-series involving various small parameters, on top of $\hbar$ itself. On one hand, we have of course $\hbar$, but we also have exponentially small quantities in $\hbar$, and non-analytic functions of $\hbar$ at $\hbar = 0$, like the trigonometric functions of $1/k$ appearing in (3.55). We can solve this quantization at small $\hbar$, as one does for example in the case of the double-well potential in Quantum Mechanics [8,9]. To give a flavor of the type of expressions one finds, let us calculate the first non-perturbative correction to the perturbative series in (3.28). To do this calculation, we need the leading term of E_{eff} in an expansion around $\hbar = 0$. We find, after using (3.10),

$$E_{\mathrm{eff}}(E) = E_{\mathrm{eff}}^{(0)}(E) + \mathcal{O}(\hbar^2), \tag{3.66}$$

where

$$E_{\text{eff}}^{(0)}(E) = E - 2e^{-2E} {}_4F_3\left(1,1,\frac{3}{2},\frac{3}{2};2,2,2;16e^{-2E}\right). \qquad (3.67)$$

After expanding it around $E = 2\log 2$, we find,

$$E_{\text{eff}}^{(0)}(E) = \frac{4\text{K}}{\pi} + \frac{1}{\pi}\left(E - 2\log 2\right)\left(1 - \log\left(\frac{E - 2\log 2}{8}\right)\right)$$
$$+ \mathcal{O}\left((E - 2\log 2)^2\right), \qquad (3.68)$$

where K is the Catalan number. A simple calculation shows that

$$E_n = E_n^{\text{p}} + E_n^{\text{np}}, \qquad (3.69)$$

where, as $\hbar \to 0$,

$$E_n^{\text{np}} \approx -\frac{1}{4\pi}\left(\frac{2n+1}{64}\right)^{2n+1}\sin\left(\frac{4\pi}{k}\right) d_1(k)e^{-2n-1}\hbar^{2n+2}e^{-A/k}, \qquad (3.70)$$

and

$$A = \frac{16\text{K}}{\pi}. \qquad (3.71)$$

It seems that E_n^{np} is a trans-series involving the small parameter $e^{-A/k}$, $\log\hbar$ and trigonometric functions of $1/k$.

Before closing this subsection, let us summarize the two most important consequences of our proposal for the exact quantization condition:

1. In solving the spectral problem (2.13) associated to the quantum curve of local $\mathbb{P}^1 \times \mathbb{P}^1$, the information encoded in the quantum periods (or, equivalently, in the NS limit of the refined topological string), evaluated with the perturbative WKB method as in [3–5], is *not* enough. The perturbative quantization condition is insufficient and even ill defined for some values of $\hbar$, and has to be corrected by an infinite series of instantons. In the terminology of [34,35], we can say that the WKB method of [3–5] gives the Voros multipliers of the problem, but does not contain the information about the Jost function solving the spectral problem.

2. The information about the instanton corrections involves the *standard* topological string free energy, and in particular the standard Gopakumar–Vafa invariants. This is a statement dual to the conjecture in [14], where evidence was given that the non-perturbative corrections to the standard topological string free energy involves the refined topological string. The fact that the unrefined topological string and the NS refined topological string are intimately related in a non-perturbative treatment is reminiscent of the observations made in [38].

Although our conjecture is well motivated by the relationship to ABJM theory, it can be tested in detail. After all, our conjecture gives the exact WKB quantization condition of the spectral problem, and we can (and should) compare it to the actual values of the energy levels.

3.3. Testing the Non-Perturbative WKB Quantization Condition

The goal of this subsection is to compute the energy levels using (3.65), but this time at large quantum number n and fixed k. In this computation, the non-perturbative part of $\mathrm{vol}(E; \hbar)$ plays a crucial role. We will then compare the energy levels obtained in this way with numerical values obtained directly from the integral Eq. (2.3), and find an impressive agreement.

Since the numerical spectrum is easier to compute with high accuracy when $k = 1, 2$, we will study (3.65) for these values of k (corresponding to $\hbar = 2\pi$ and $\hbar = 4\pi$, respectively). In these cases, we can see from (3.23) and (3.55) that $\mathrm{vol}(E; \hbar)$ is given as an expansion in powers of e^{-2E}. In fact, for $k = 1$, the coefficients of the odd powers

$$\mathrm{e}^{-2(2m+1)E}, \quad m = 0, 1, 2, \ldots, \tag{3.72}$$

vanish. Therefore, for $k = 1$, we will have an expansion in powers of e^{-4E}. Let us write the total quantum volume, for $k = 1$ and $k = 2$, as

$$\frac{\mathrm{vol}(E; \hbar)}{2\pi\hbar} = \widehat{C}(\hbar)E^2 + \widehat{n}_0(\hbar) + \sum_{\ell=1}^{\infty} \left(A_\ell(\hbar)E + B_\ell(\hbar)\right)\mathrm{e}^{-\frac{4\ell E}{k}}, \tag{3.73}$$

where

$$\widehat{C}(\hbar) = \frac{4}{\pi\hbar}, \qquad \widehat{n}_0(\hbar) = -\frac{2\pi}{3\hbar} + \frac{\hbar}{48\pi}. \tag{3.74}$$

The coefficients $B_\ell(\hbar)$ will in general have contributions from both the perturbative and the non-perturbative part.

Since the volume is given as a large E expansion, we will calculate the energy levels in an asymptotic expansion for large quantum numbers. Let us assume an ansatz for the solution of the exact WKB quantization condition of the form

$$E_n = E_n^{(0)} + \sum_{\ell=1}^{\infty} E_n^{(\ell)}\mathrm{e}^{-\frac{4\ell E_n^{(0)}}{k}}. \tag{3.75}$$

Plugging it into (3.65) we find,

$$E_n^{(0)} = \sqrt{\frac{n + 1/2 - \widehat{n}_0(\hbar)}{\widehat{C}(\hbar)}},$$

$$E_n^{(1)} = -\frac{1}{2\widehat{C}(\hbar)E_n^{(0)}}\left(B_1(\hbar) + A_1(\hbar)E_n^{(0)}\right),$$

$$E_n^{(2)} = \frac{1}{2\widehat{C}(\hbar)E_n^{(0)}}\left(-B_2(\hbar) - A_2(\hbar)E_n^{(0)} + \frac{4}{k}A_1(\hbar)E_n^{(1)}E_n^{(0)} - \widehat{C}(\hbar)(E_n^{(1)})^2\right.$$

$$\left. - A_1(\hbar)E_n^{(1)} + \frac{4}{k}B_1(\hbar)E_n^{(1)}\right), \tag{3.76}$$

as well as

$$E_n^{(3)} = \frac{1}{2\widehat{C}(\hbar)E_n^{(0)}} \left(-B_3(\hbar) - A_3(\hbar)E_n^{(0)} - A_2(\hbar)\left(E_n^{(1)} - \frac{8}{k}E_n^{(0)}E_n^{(1)} \right) \right.$$

$$+ \frac{8}{k}B_2(\hbar)E_n^{(1)} - 2\widehat{C}(\hbar)E_n^{(1)}E_n^{(2)} + B_1(\hbar)\left(\frac{4}{k}E_n^{(2)} - \frac{8}{k^2}(E_n^{(1)})^2 \right)$$

$$+ A_1(\hbar)\left(\frac{4}{k}(E_n^{(1)})^2 - \frac{8}{k^2}E_n^{(0)}(E_n^{(1)})^2 - E_n^{(2)} + \frac{4}{k}E_n^{(0)}E_n^{(2)} \right) \right). \quad (3.77)$$

To lowest order, plugging in the values we find

$$E_n^{(0)} = \frac{\sqrt{k}\pi}{\sqrt{2}} \left(n + \frac{1}{2} + \frac{1}{3k} - \frac{k}{24} \right)^{1/2}. \quad (3.78)$$

This expression is valid for any $k > 0$, provided n is large enough. The leading growth of the eigenvalues derived from the lowest order approximation,

$$E_n^2 - E_0^2 \approx \frac{k\pi^2}{2}n, \qquad n \gg 1, \quad (3.79)$$

has been verified numerically in [15] for $k = 1$, by computing the spectrum for $n = 0, \ldots, 6$.

To obtain subleading corrections, we have to calculate the coefficients $A_\ell(\hbar)$ and $B_\ell(\hbar)$, which are determined by the coefficients $\widehat{a}_\ell(\hbar), \widehat{b}_\ell(\hbar)$ in (3.23) and $s_m(k)$ in (3.55). For the first few values of ℓ, m, they can be read off from (3.20) and (3.21), and from the Gopakumar–Vafa invariants of local $\mathbb{P}^1 \times \mathbb{P}^1$, as listed in for example [16]. For the first five orders, we find

$$\frac{\mathrm{vol}(E; \hbar = 2\pi)}{4\pi^2} = \frac{2}{\pi^2}E^2 - \frac{7}{24} + \frac{1}{\pi^2}(8E + 1)\,\mathrm{e}^{-4E} - \frac{1}{\pi^2}\left(52E - \frac{1}{4} \right)\mathrm{e}^{-8E}$$

$$+ \frac{1}{\pi^2}\left(\frac{1472}{3}E - \frac{152}{9} \right)\mathrm{e}^{-12E} - \frac{1}{\pi^2}\left(5402E - \frac{13949}{48} \right)\mathrm{e}^{-16E}$$

$$+ \frac{1}{\pi^2}\left(\frac{323648}{5}E - \frac{317122}{75} \right)\mathrm{e}^{-20E} + \mathcal{O}\left(E\mathrm{e}^{-24E} \right),$$

$$\frac{\mathrm{vol}(E; \hbar = 4\pi)}{8\pi^2} = \frac{1}{\pi^2}E^2 - \frac{1}{12} + \frac{1}{\pi^2}(8E + 2)\,\mathrm{e}^{-2E} - \frac{1}{\pi^2}\left(52E + \frac{1}{2} \right)\mathrm{e}^{-4E}$$

$$+ \frac{1}{\pi^2}\left(\frac{1472}{3}E - \frac{304}{9} \right)\mathrm{e}^{-6E} - \frac{1}{\pi^2}\left(5402E - \frac{13949}{24} \right)\mathrm{e}^{-8E}$$

$$+ \frac{1}{\pi^2}\left(\frac{323648}{5}E - \frac{634244}{75} \right)\mathrm{e}^{-10E} + \mathcal{O}\left(E\mathrm{e}^{-12E} \right), \quad (3.80)$$

and the poles cancel, as expected.

Explicit calculations of the coefficients (3.21) and (3.53) in [14] indicate that (3.73) is a convergent series when $|q| = 1$ and for sufficiently large E. Notice that, for small $\hbar$, i.e., $q \to 1$, the coefficients in (3.21) are just the coefficients of the classical B-period, which are known to lead to a series with a finite radius of convergence. It seems that these convergence properties are preserved as long as q stays in the unit circle of the complex plane. It follows that the series (3.75) should also be convergent, for n large enough. Our

TABLE 1. The lowest and next-to-lowest energy eigenvalues for $k = 1$ were calculated analytically, including higher and higher orders of exponentially small corrections in (3.75)

Order	Energy levels for $k = 1$	
	E_0	E_1
0	1.97654203314	2.973469456
1	1.97575850097	2.973455218
2	1.97575795136	2.973455217
3	1.97575795105	2.973455217
Numerical value	1.97575795105	2.973455217

In the last line, numerical values are given. At each order of the approximation, we underline the digits which agree with the numerical result

TABLE 2. The lowest and next-to-lowest energy eigenvalues for $k = 2$ were calculated analytically, including higher and higher orders of exponentially small corrections in (3.75)

Order	Energy levels for $k = 2$	
	E_0	E_1
0	2.399431022965	3.953084066277
1	2.363040773485	3.951517001949
2	2.362388178770	3.951515902713
3	2.362377640277	3.951515902099
Numerical value	2.362377493014	3.951515902099

In the last line, numerical values are given. At each order of the approximation, we underline the digits which agree with the numerical result

numerical calculations indicate that, in fact, this series converges very rapidly already for $n = 0, 1$. The results for E_n, for $n = 0, 1$, as computed with the WKB quantization condition, and up to third order for $k = 1$ and $k = 2$ are listed in Tables 1 and 2, respectively.

We can now compare these results with numerical values obtained starting from the integral Eq. (2.3). To get numerical values with high accuracy we will use that, as shown in [15], this equation can be rewritten in the form of an eigenvalue equation for an infinite-dimensional matrix M. For the derivation we refer to [15], here we only quote the results that we need to numerically test the energy eigenvalues we have obtained. The matrix elements M_{nm} depend only on $m + n$. Such a matrix is called a Hankel matrix. Moreover, in this case, they are zero if $m + n$ is odd, that is, M has the following form

$$M = \begin{pmatrix} m_0 & 0 & m_1 & 0 & m_2 & 0 & \ldots \\ 0 & m_1 & 0 & m_2 & 0 & m_3 & \\ m_1 & 0 & m_2 & 0 & m_3 & 0 & \\ 0 & m_2 & 0 & m_3 & 0 & m_4 & \\ m_2 & 0 & m_3 & 0 & m_4 & 0 & \\ \vdots & & & & & & \ddots \end{pmatrix}. \tag{3.81}$$

Such a matrix can be decomposed into two blocks of Hankel matrices, M_+ and M_-:

$$M_+ = \begin{pmatrix} m_0 & m_1 & m_2 & \cdots \\ m_1 & m_2 & m_3 & \\ m_2 & m_3 & m_4 & \\ \vdots & & & \ddots \end{pmatrix}, \qquad M_- = \begin{pmatrix} m_1 & m_2 & m_3 & \cdots \\ m_2 & m_3 & m_4 & \\ m_3 & m_4 & m_5 & \\ \vdots & & & \ddots \end{pmatrix}. \quad (3.82)$$

The eigenspaces of M decompose into a direct product of the eigenspaces of $M_\pm$. Let the eigenvalues of M be denoted by λ_n, ordered such that

$$\lambda_0 > \lambda_1 > \lambda_2 > \cdots, \qquad (3.83)$$

and let the eigenvalues of $M_\pm$ be denoted by $\lambda_{\pm,n}$, ordered in the same way. We then have

$$\lambda_{+,n} = \lambda_{2n}, \quad \lambda_{-,n} = \lambda_{2n+1}. \qquad (3.84)$$

The relation between the eigenvalues of M and the energy eigenvalues is

$$E_n = -\log \lambda_n. \qquad (3.85)$$

Different values of k give different M_{nm}. For $k = 1$ we have (for $m + n$ even, otherwise we get zero)

$$M_{nm}^{k=1} = \frac{C_{\frac{m+n}{2}}}{2^{n+m+3}} \qquad (3.86)$$

where C_n is the Catalan number

$$C_n = \frac{(2n)!}{(n+1)!n!}. \qquad (3.87)$$

For $k = 2$, we have (again for $m + n$ even, otherwise we get zero)[4]

$$M_{nm}^{k=2} = \frac{1}{4\pi} \left[-\frac{2}{n+m+1} + \psi\left(\frac{n+m+3}{4}\right) - \psi\left(\frac{n+m+1}{4}\right) \right] \quad (3.88)$$

where $\psi(x)$ is the digamma function.

We have calculated the lowest and next-to-lowest energy eigenvalues obtained from these two matrices numerically. In principle, we have to diagonalize an infinite-dimensional matrix, but in practice we have to truncate the Hankel matrices to an $L \times L$ matrix. The eigenvalues of this truncated matrix $E_n(L)$ give numerical approximations to the exact E_n, and they converge to it as $L \to \infty$. To incorporate finite-size effects, let us assume that the eigenvalues depend on L as

$$E_n(L) = E_n + \sum_{j \geq 1} \frac{E_n^j}{L^j}. \qquad (3.89)$$

We can accelerate the convergence of the sequence $E_{n,L}$ to E_n by using for example Richardson extrapolation [39]. The results obtained with this method for $k = 1$ and $k = 2$ are given in Tables 1 and 2, respectively, with the displayed numerical accuracy. In all cases, the series of instanton corrections add

---

[4] Note that there is a factor of 2 missing in equation 2.41 in the first version of [15].

up to values closer and closer to the numerical value. For $k = 1$, and for the first excited state with $k = 2$, the exponential corrections of fourth order and higher are not visible in the numerical approximation to the eigenvalue, due to the limited accuracy of our calculation. In the case of the ground state energy E_0 for $k = 2$, the corrections are moderately large: the approximation obtained by including up to the third exponentially small correction is only correct to order 10^{-7}. However, we have checked that our quantization condition correctly reproduces the numerical eigenvalue E_0, which we have obtained with an accuracy of 18 significant digits, by including further exponentially small corrections.

It seems clear from this numerical analysis that our ansatz (3.55) for the non-perturbative corrections to the quantum volume is not only divergence-free: it is also in excellent agreement with the true spectrum of eigenvalues.

4. The Grand Potential of ABJM Theory

As we explained in the introduction, the second motivation for looking at the spectral problem (2.3) is the study of non-perturbative effects in the partition function of ABJM theory [10]. This partition function $Z(N, k)$ depends on two parameters: the rank of the gauge group N and the Chern–Simons level k. Building on [11,40], it was shown in [13] that $Z(N, k)$ can be written as

$$Z(N,k) = \frac{1}{N!} \sum_{\sigma \in S_N} (-1)^{\epsilon(\sigma)} \int \mathrm{d}^N x \prod_i \rho(x_i, x_{\sigma(i)}), \tag{4.1}$$

where $\rho(x_1, x_2)$ is the kernel defined in (2.1). This is nothing but the partition function of an ideal Fermi gas with one-particle energies E_n, $n \geq 0$, determined by the spectral problem (2.3). The grand potential of the Fermi gas is given by

$$J(\mu, k) = \sum_{n \geq 0} \log \left(1 + \mathrm{e}^{\mu - E_n} \right), \tag{4.2}$$

where the energy levels E_n are determined by the WKB quantization condition (3.65), which defines in fact an implicit function $E(n)$ for arbitrary values of n. To perform the sum over discrete energy levels, we will use the Euler–Maclaurin formula, which reads

$$\sum_{n \geq 0} f(n) = \int_0^\infty f(n)\mathrm{d}n + \frac{1}{2} \left(f(0) + f(\infty) \right)$$
$$+ \sum_{r \geq 1} \frac{B_{2r}}{(2r)!} \left(f^{(2r-1)}(\infty) - f^{(2r-1)}(0) \right). \tag{4.3}$$

Notice that, in general, this formula gives an asymptotic expansion for the sum. However, as noticed in [41], this asymptotic expansion is Borel summable under mild assumptions for the function $f(n)$, and this can be used to write an exact formula. In this paper, we will not explore this possibility. In our case,

the function $f(n)$ is given by

$$f(n) = \log\left(1 + e^{\mu - E(n)}\right). \tag{4.4}$$

Since $E(n) \to \infty$ as $n \to \infty$, we have $f(\infty) = 0$, $f^{(2r-1)}(\infty) = 0$ for all $r \geq 1$. The first terms of (4.3) give,

$$\int_{E_0}^{\infty} \frac{\mathrm{d}n(E)}{\mathrm{d}E} \log\left(1 + e^{\mu - E}\right)\mathrm{d}E + \frac{1}{2}f(0) = \frac{1}{2\pi\hbar}\int_{E_0}^{\infty} \frac{\mathrm{vol}(E)}{1 + e^{E - \mu}}\mathrm{d}E. \tag{4.5}$$

In deriving this equation, we first changed variables from n to E, used

$$n(E) = \frac{\mathrm{vol}(E)}{2\pi\hbar} - \frac{1}{2}, \tag{4.6}$$

we integrated by parts, and we took into account that

$$\frac{\mathrm{vol}(E_0)}{2\pi\hbar} = \frac{1}{2}, \tag{4.7}$$

as well as the asymptotic behavior

$$\mathrm{vol}(E) \approx E^2, \qquad E \to \infty. \tag{4.8}$$

We conclude that[5]

$$J(\mu, k) = \frac{1}{2\pi\hbar}\int_{E_0}^{\infty} \frac{\mathrm{vol}(E)\mathrm{d}E}{e^{E - \mu} + 1} - \sum_{r \geq 1} \frac{B_{2r}}{(2r)!}f^{(2r-1)}(0). \tag{4.9}$$

The first term in this formula is nothing but an integral transform of the quantum volume of phase space studied in the previous section. To analyze it, we just have to compute the integrals

$$R_\ell^{(j)}(E_0, \mu) = \int_{E_0}^{\infty} \frac{E^j e^{-2\ell E}}{e^{E - \mu} + 1}\mathrm{d}E, \qquad j = 0, 1, \qquad \ell \in \mathbb{N}, \tag{4.10}$$

which appear when considering the perturbative quantum volume (3.23), as well as the integrals

$$R_\sigma(E_0, \mu) = \int_{E_0}^{\infty} \frac{e^{-\sigma E}}{e^{E - \mu} + 1}\mathrm{d}E, \tag{4.11}$$

where $\sigma \notin \mathbb{N}$. These appear when we consider the non-perturbative quantum volume (3.55) for k arbitrary.

The calculation of these integrals is not elementary, and we will use a variant of the Mellin transform used in [8,9]. It is defined as

$$\widehat{g}(s) = \int_0^1 g(u)u^{-s-1}\mathrm{d}u. \tag{4.12}$$

This will make it possible to calculate analytically the integral over the density of states.

---

[5] In the first version of this paper, we did not include the corrections involving the derivatives of $f(n)$ at $n = 0$, and some of the resulting formulae were incorrect, as pointed out to us by Yasuyuki Hatsuda. We would like to thank him for his precious observations, which prompted us to find the correct formula for the grand potential.

We will first consider the contribution to the integral due to the perturbative quantum volume. To have a more compact notation for the answer, it is useful to define the functions

$$
\begin{aligned}
\widehat{a}_\ell(s;\hbar) &= \widehat{a}_\ell(\hbar)\mathrm{e}^{-(2\ell-s)E_0}\left(1+(2\ell-s)E_0\right), \\
\widehat{b}_\ell(s;\hbar) &= \widehat{b}_\ell(\hbar)\mathrm{e}^{-(2\ell-s)E_0},
\end{aligned}
\tag{4.13}
$$

as well as the elementary integrals

$$
\mathcal{I}_j(n) = \int_{E_0}^\infty E^j \mathrm{e}^{-nE}\mathrm{d}E = \left(-\frac{\partial}{\partial n}\right)^j \left(\frac{1}{n}\mathrm{e}^{-nE_0}\right), \qquad j \geq 0. \tag{4.14}
$$

Using the results collected in Appendix A, one finds

$$
\begin{aligned}
\frac{1}{2\pi\hbar}\int_{E_0}^\infty \frac{\mathrm{vol_p}(E)\mathrm{d}E}{\mathrm{e}^{E-\mu}+1} &= \frac{2\mu^3}{3k\pi^2} + \left(\frac{1}{3k}+\frac{k}{24}\right)\mu + \widehat{A}(\hbar) \\
&\quad + \sum_{\ell\geq 1}\left(-\frac{\widehat{a}_\ell(\hbar)}{\pi^2 k}\mu^2 + \frac{\widehat{b}_\ell(\hbar)}{2\pi^2 k}\mu + \frac{\widehat{c}_\ell(\hbar)}{2\pi^2 k}\right)\mathrm{e}^{-2\ell\mu} - \sum_{\ell\geq 0}\frac{\widehat{d}_\ell(\hbar)}{2\pi^2 k}\mathrm{e}^{-(2\ell+1)\mu}. \quad (4.15)
\end{aligned}
$$

In this expression, the μ-independent function $\widehat{A}(\hbar)$ is given by

$$
\widehat{A}(\hbar) = -\frac{1}{2\pi^2 k}\left[\frac{4}{3}E_0^3 - \left(\frac{2\pi^2}{3}-\frac{\hbar^2}{48}\right)E_0 + \sum_{\ell\geq 1}\frac{\widehat{a}_\ell(0;\hbar)}{\ell^2} - \sum_{\ell\geq 1}\frac{\widehat{b}_\ell(0;\hbar)}{2\ell}\right],
\tag{4.16}
$$

and the coefficients $\widehat{c}_\ell$ and $\widehat{d}_\ell$ are given by

$$
\begin{aligned}
\widehat{c}_\ell(\hbar) &= -\frac{2\pi^2}{3}\widehat{a}_\ell(\hbar) + 2\widehat{a}_\ell(\hbar)E_0^2 - \widehat{b}_\ell(\hbar)E_0 - \sum_{m\neq\ell}\frac{\widehat{a}_m(2\ell;\hbar)}{(m-\ell)^2} + \sum_{m\neq\ell}\frac{\widehat{b}_m(2\ell;\hbar)}{2(m-\ell)} \\
&\quad + 4\mathcal{I}_2(-2\ell) - \left(\frac{2\pi^2}{3}-\frac{\hbar^2}{48}\right)\mathcal{I}_0(-2\ell), \\
\widehat{d}_\ell(\hbar) &= -4\sum_{m\geq 1}\frac{\widehat{a}_m(2\ell+1;\hbar)}{(2m-(2\ell+1))^2} + \sum_{m\geq 1}\frac{\widehat{b}_m(2\ell+1;\hbar)}{2m-(2\ell+1)} + 4\mathcal{I}_2(-(2\ell+1)) \\
&\quad - \left(\frac{2\pi^2}{3}-\frac{\hbar^2}{48}\right)\mathcal{I}_0(-(2\ell+1)). \tag{4.17}
\end{aligned}
$$

Although these expressions look complicated, the derivatives of these quantities w.r.t. E_0 have a simple expression in terms of the perturbative quantum volume. This is because

$$
\frac{\partial}{\partial E_0}\left(\frac{1}{2\pi\hbar}\int_{E_0}^\infty \frac{\mathrm{vol_p}(E)\mathrm{d}E}{\mathrm{e}^{E-\mu}+1}\right) = -\frac{1}{2\pi\hbar}\frac{\mathrm{vol_p}(E_0)}{\mathrm{e}^{E_0-\mu}+1},
\tag{4.18}
$$

and we deduce

$$\frac{\partial \widehat{A}(\hbar)}{\partial E_0} = -\frac{1}{2\pi\hbar}\mathrm{vol_p}(E_0),$$

$$\frac{\partial \widehat{c}_\ell(\hbar)}{\partial E_0} = -\frac{\mathrm{e}^{2\ell E_0}}{2\pi\hbar}\mathrm{vol_p}(E_0), \tag{4.19}$$

$$\frac{\partial \widehat{d}_\ell(\hbar)}{\partial E_0} = -\frac{\mathrm{e}^{(2\ell+1)E_0}}{2\pi\hbar}\mathrm{vol_p}(E_0),$$

where E_0 is regarded as an independent variable (in particular, independent of $\hbar$). We can use these expressions to obtain explicit formulae for the $\hbar$ expansion of $\widehat{A}(\hbar)$, $\widehat{c}_\ell(\hbar)$, $\widehat{d}_\ell(\hbar)$ in terms of integrals of $\mathrm{vol}_n(E)$. This makes it possible to evaluate analytically the infinite sums appearing in (4.16) and (4.17), order by order in $\hbar$.

Let us now consider the contribution to the grand potential coming from the non-perturbative part of the quantum volume. We expand the exponent of (3.47) and write

$$\mathrm{vol_{np}}(E) = \sum_{m=1}^{\infty}\sum_{\ell=0}^{\infty} s_{\ell,m}(k)(-1)^m \mathrm{e}^{-\left(\frac{4m}{k}+2\ell\right)E}, \tag{4.20}$$

where the coefficients $s_{\ell,m}(k)$ can be obtained from $s_n(k)$ and $\widehat{a}_r(\hbar)$, with $n \leq m$, $r \leq \ell$. By using now (A.13), with

$$\sigma = \frac{4m}{k} + 2\ell, \qquad m = 1, 2, \ldots, \quad \ell = 0, 1, \ldots, \tag{4.21}$$

we find

$$\frac{1}{2\pi\hbar}\int_{E_0}^{\infty}\frac{\mathrm{vol_{np}}(E)\mathrm{d}E}{\mathrm{e}^{E-\mu}+1} = -\frac{1}{2\hbar}\sum_{m=1}^{\infty}\csc\left(\frac{4\pi m}{k}\right)\sum_{\ell=0}^{\infty}s_{\ell,m}(k)(-1)^m\mathrm{e}^{-\left(\frac{4m}{k}+2\ell\right)\mu}$$

$$+\frac{1}{4\pi^2}\sum_{n\geq 0}\left\{\sum_{m=1}^{\infty}\sum_{\ell=0}^{\infty}\frac{(-1)^m s_{\ell,m}(k)}{4m+k(2\ell-n)}\mathrm{e}^{-\left(\frac{4m}{k}+2\ell\right)E_0}\right\}$$

$$\times (-1)^n \mathrm{e}^{-n(\mu-E_0)}. \tag{4.22}$$

If our conjecture (3.52) is true, we can write the first line of (4.22) as

$$\sum_{m=1}^{\infty} d_m(k)(-1)^m \mathrm{e}^{-4m\mu_{\mathrm{eff}}/k}, \tag{4.23}$$

where μ_{eff} is defined by the same equality as (3.46), i.e.,

$$\mu_{\mathrm{eff}} = \mu + \frac{\pi^2 k}{2}\sum_{\ell=1}^{\infty} a_\ell(k)\mathrm{e}^{-2\ell\mu}. \tag{4.24}$$

To obtain the grand potential, we have to add to the integral (4.15) the infinite series of corrections in (4.9). Since

$$f'(n) = -\frac{E'(n)}{\mathrm{e}^{E(n)-\mu}+1}, \tag{4.25}$$

it is easy to see that, when expanded around $\mu \to \infty$, these corrections can only lead to a constant term, plus a series in $e^{-\mu}$. In addition, if we evaluate these corrections perturbatively in $\hbar$, we find that $f^{(2r-1)}(0)$ is of order $\hbar^{2r-1}$.

We can now compare this calculation to the existing results on the perturbative grand potential of ABJM theory [13–18]

$$
\begin{aligned}
J(\mu, k) = & \frac{2\mu^3}{3k\pi^2} + \left(\frac{1}{3k} + \frac{k}{24} \right) \mu + A(k) \\
& + \sum_{\ell \geq 1} \left(a_\ell(k)\mu^2 + b_\ell(k)\mu + c_\ell(k) \right) e^{-2\ell\mu} \\
& + \sum_{m=1}^{\infty} d_m(k)(-1)^m e^{-4m\mu_{\mathrm{eff}}/k}.
\end{aligned}
\tag{4.26}
$$

As we have seen, the correction terms in (4.9) only affect the constant term and the series in $e^{-\mu}$. In particular, the terms in $\mu^2 e^{-2\ell\mu}$ and $\mu e^{-2\ell\mu}$ in $J(\mu, k)$ are completely captured by the integral (4.15). We conclude that:

$$
a_\ell(k) = -\frac{\widehat{a}_\ell(\hbar)}{\pi^2 k}, \qquad b_\ell(k) = \frac{\widehat{b}_\ell(\hbar)}{2\pi^2 k}.
\tag{4.27}
$$

This is the main conjecture in [14], which was tested there by direct computation. Here, we have derived this conjecture from the WKB solution of the spectral problem. Furthermore, (4.23) is exactly the non-perturbative contribution (in k) to the grand potential. It contains the contribution of worldsheet instantons, which is determined by the relation to standard topological string theory [12,13,16,36], as well as the contribution of bound states conjectured in [18]. In fact, our conjecture (3.52) was tailored to reproduce the worldsheet instanton contribution. Although we can not derive (4.23) from first principles (since we have not proved (3.52)), the appearance of the "effective" chemical potential μ_{eff} is a simple consequence of the fact that instanton corrections in the WKB method involve the exponentiated quantum A-period (3.45), as it was already pointed out in [13].

The conclusion of this analysis is that the conjecture of [14] relating the a, b coefficients of the grand potential to the refined topological string can be derived from the perturbative WKB analysis of the spectral problem, and the conjecture on the contribution of bound states in [18] can be also partially justified. Conversely, the known results about the worldsheet instanton contribution to the grand potential lead to the conjecture (3.55) for the non-perturbative correction to the WKB quantization condition.

However, our analysis of (4.9) leads to further results for the grand potential, since we have not yet considered the constant term and the series in $e^{-\mu}$. Let us first look at the constant term. First of all, it is easy to see that the constant part (in μ) of the series of corrections in (4.9), as $\mu \to \infty$, is given by

$$
\sum_{r \geq 1} \frac{B_{2r}}{(2r)!} E^{(2r-1)}(0),
\tag{4.28}
$$

and involve the derivatives of the function $E(n)$ at $n = 0$. We conclude that

$$A(k) = \widehat{A}(\hbar) + \frac{1}{4\pi^2} \sum_{m=1}^{\infty} \sum_{\ell=0}^{\infty} \frac{(-1)^m s_{\ell,m}(k)}{4m + 2\ell k} e^{-\left(\frac{4m}{k} + 2\ell\right)E_0} + \sum_{r \geq 1} \frac{B_{2r}}{(2r)!} E^{(2r-1)}(0).$$
(4.29)

The second term in the r.h.s. comes from the contribution of $n = 0$ appearing in (4.22), and it is non-perturbative in $\hbar$. We can still analyze this equality by looking at the perturbative expansion around $k = 0$. The l.h.s. has a power series expansion of the form [13,42]

$$A(k) = \frac{2\zeta(3)}{\pi^2 k} - \frac{k}{12} + \cdots$$
(4.30)

To test the equality of these two expressions, we can expand the first and third term in the r.h.s. of (4.29) in powers of k and compare the resulting series order by order. Notice however that each order in (4.16) is given by an infinite sum. At leading order, the third term in (4.29) does not contribute, and in the expression (4.16) we can use the values of the coefficients (3.10) as well as the leading order term for $n = 0$ in (3.28). If we plug in these values in (4.16), we reproduce (numerically) the first term in (4.30), testing in this way the proposed equality at leading order. Alternatively, writing

$$\widehat{A}(\hbar) = \sum_{n \geq 0} \widehat{A}_n k^{2n-1}$$
(4.31)

we can use the first equation in (4.19) and (3.3) to obtain the expression

$$\widehat{A}_0 = E_0 + \frac{2\zeta(3)}{\pi^2} - \frac{e^{E_0}}{8\pi^3} G_{4,4}^{2,4} \left(\frac{e^{2E_0}}{16} \left| \begin{array}{c} \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \\ 0, 0, -\frac{1}{2}, -\frac{1}{2} \end{array} \right. \right),$$
(4.32)

where the constant of integration is fixed by matching the expansion (4.16). With this expression we can perform an analytic check.[6]

Similarly, we can now look at the coefficients in the expansion in powers of $e^{-\mu}$ in both sides of (4.9). The equality in (4.9) implies that the coefficients $\widehat{c}_\ell(\hbar)$ are related to the coefficients $c_\ell(k)$ appearing in (4.26) by the equation

$$c_\ell(k) = \frac{\widehat{c}_\ell(\hbar)}{2\pi^2 k} + \frac{e^{2\ell E_0}}{4\pi^2} \sum_{m=1}^{\infty} \sum_{r=0}^{\infty} \frac{(-1)^m s_{r,m}(k)}{4m + 2k(r-\ell)} e^{-\left(\frac{4m}{k} + 2r\right)E_0}$$

$$- \sum_{r \geq 1} \frac{B_{2r}}{(2r)!} \left[f^{(2r-1)}(0) \right]_{-2\ell}, \qquad \ell \geq 1,$$
(4.33)

---

[6] This analytic check has also been independently performed by Yasuyuki Hatsuda in an unpublished note.

where the bracket $[\cdot]_{-2\ell}$ means the term in $e^{-2\ell\mu}$ in the expansion of $f^{(2r-1)}(0)$ at large μ. Finally, we obtain the condition,

$$\frac{\widehat{d}_\ell(\hbar)}{2\pi^2 k} + \frac{e^{(2\ell+1)E_0}}{4\pi^2} \sum_{m=1}^{\infty} \sum_{r=0}^{\infty} \frac{(-1)^m s_{r,m}(k)}{4m + k(2r-2\ell-1)} e^{-\left(\frac{4m}{k}+2r\right)E_0}$$

$$+ \sum_{r\geq 1} \frac{B_{2r}}{(2r)!} \left[f^{(2r-1)}(0)\right]_{-2\ell-1} = 0, \qquad \ell \geq 0. \tag{4.34}$$

As before, these equations can be tested in perturbation theory around $k = 0$. In this case, the terms coming from the last sum in (4.22) do not contribute. In fact, the function $c_\ell(k)$ is known to be an analytic function of k at the origin, therefore it should be given by the perturbative series around $k = 0$ of the r.h.s., while the non-perturbative terms in k in the r.h.s. should cancel. At leading order in k, the corrections involving $f^{(2r-1)}(0)$ do not contribute either, and the resulting equalities can be checked for the very first values of ℓ, by performing numerically the sums or by using the Eq. (4.19).

It is instructive to perform a test of the various equalities we have written down, again in perturbation theory around $k = 0$, but at next-to-leading order in k. This can be done in a single strike by looking at (4.9) in perturbation theory. After taking into account the various expansions in $\hbar$, and after writing down[7]

$$J_{\mathrm{p}}(\mu, k) = \sum_{n\geq 0} J_n(\mu)\hbar^{2n-1}, \tag{4.35}$$

it is easy to see that a next-to-leading order test amounts to checking that

$$J_1(\mu) = \frac{1}{2\pi} \int_{E_0^{(0)}}^{\infty} \frac{\mathrm{vol}_1(E)\mathrm{d}E}{e^{E-\mu}+1} - \frac{1}{96(1+4\,e^{-\mu})}. \tag{4.36}$$

The second term in the r.h.s. has two contributions: one coming from the expansion of E_0^{p} at order $\hbar^2$, where we have taken into account (3.30), and another from the term $r = 1$ in the series of corrections in (4.9). The integral appearing in the r.h.s. can be evaluated explicitly, as a power series in e^μ, by using the explicit result (3.25). The result is,

$$\frac{1}{2\pi} \int_{E_0^{(0)}}^{\infty} \frac{\mathrm{vol}_1(E)\mathrm{d}E}{e^{E-\mu}+1} = -\sum_{n\geq 1} (r_n + s_n)(-1)^n e^{n\mu}, \tag{4.37}$$

where

$$r_n = \frac{1}{3} 2^{-2n-5}, \qquad s_n = -\frac{4^{-n-4} n(n^2-1)\,\Gamma\left(\frac{n}{2}\right)^2}{3\,\Gamma\left(\frac{n+3}{2}\right)^2}. \tag{4.38}$$

It is easy to see that

$$-\sum_{n\geq 1} r_n(-1)^n e^{n\mu} = \frac{1}{96(1+4\,e^{-\mu})}, \tag{4.39}$$

---

[7] $J_{\mathrm{p}}(\mu, k)$ denotes the perturbative, in k, part of $J(\mu, k)$.

and also that

$$s_n = \frac{Z_n^{(1)}}{n}, \tag{4.40}$$

where

$$J_1(\mu) = -\sum_{n \geq 1} \frac{Z_n^{(1)}}{n} \left(-\mathrm{e}^\mu\right)^n, \tag{4.41}$$

and the explicit expression for $Z_n^{(1)}$ was derived in [13] from the Wigner–Kirkwood expansion and written down in eq. (5.20) of that paper. This proves (4.36).

Although we have tested these equalities perturbatively, it would be also very interesting to test them non-perturbatively, for finite values of k, even numerically. To do this, one might need to perform a Borel resummation of the asymptotic series appearing in the Euler–Maclaurin expansion, following [41].

It should be stressed that the relationship between the quantum volume and the grand potential that we have analyzed in this section is not without subtleties. To understand this relationship better, we should have a clearer analytic understanding of the terms in the second line of (4.22). It is possible that such an improved understanding leads to corrections to our proposal (3.55), although such corrections do not seem to be present in the cases $k = 1$ and $k = 2$ that we have analyzed in detail.

We conclude that our calculation of the quantum volume for the spectral problem (2.13) proves, to a large extent, the conjectures made in [14] relating the membrane instanton part of the grand potential to the refined topological string in the NS limit. This is the content of our result (4.27). The incorporation of "bound states" by promoting μ to μ_{eff}, as proposed in [18], follows also from our WKB analysis. For a complete derivation of the known results, one should also prove the equalities (4.29), (4.33) and (4.34).

5. Conclusions and Open Problems

In this paper, we have analyzed in detail a spectral problem which appears in ABJM theory and in the theory of quantum spectral curves. This problem is a particular case of the quantization of the mirror curve of local $\mathbb{P}^1 \times \mathbb{P}^1$, for some particular values of the complex parameters z_1, z_2. This choice of parameters is very convenient since, after imposing appropriate analyticity conditions, it leads to a reformulation of the problem in terms of an integral Eq. (2.3) for which we have a lot of information —numerical and analytical—thanks to its rôle in ABJM theory.

Our main conclusion is that the WKB analysis of [3–5] (which reproduces the known results for the refined topological string in the NS limit) is insufficient for actually calculating the spectrum: there are instanton effects which should be added, just as in other quantum-mechanical situations [6–9]. Surprisingly, the relevant instanton series is essentially the free energy of the *standard* topological string. This shows that the NS limit of the refined string (i.e., the

choice $\epsilon_2 = 0$ in the Omega background) and the conventional topological string (i.e., the choice $\epsilon_1 = -\epsilon_2$) are intimately related in a non-perturbative treatment. This was already shown in the "dual" calculation of [14], by using the non-perturbative definition provided by the lens space matrix model.

Mathematically, the conjectural WKB quantization condition for the spectral problem (2.3) or (2.13) is very surprising. In particular, it contains information about the Gopakumar–Vafa invariants of local $\mathbb{P}^1 \times \mathbb{P}^1$. It would be interesting to derive this quantization condition analytically, perhaps along the lines of [26]. The spectral problem for integral equations like (2.3) is also related to a TBA equation [19,20], and it would be also interesting to understand the non-perturbative corrections from the point of view of the TBA.

Our exact quantization condition is also very interesting from the point of view of the WKB method. Usually, the perturbative WKB quantization leads to an *asymptotic* series in the energy which has to be resummed with some appropriate prescription. Instanton corrections are typically needed when the series is not Borel summable, and they cancel the non-perturbative ambiguities appearing in the resummation process, as in the double-well potential studied in [8,9]. In the case studied here, the WKB quantization condition leads to an infinite series which seems to be *convergent* in the energy plane. However, its coefficients develop simple poles at values of $\hbar$ of the form $2\pi n$, where n is a rational number. Instanton corrections are needed to cancel these non-physical poles, as in the closely related HMO mechanism of [18]. Therefore, the analytic mechanism for combining non-perturbative and perturbative corrections in our example is very different from standard quantum-mechanical situations: for generic, real values of $\hbar$, we have a WKB convergent series which requires however non-perturbative corrections.

Of course, the main question for future research is the following: is this story an accident of this example, or is it a general feature of local CY manifolds? One possible strategy to answer this question would be to look at the quantized curves of local CYs and take the spectral problem seriously. For example, the quantization of the mirror curve for local $\mathbb{P}^2$ leads to the operator [5]:

$$-1 + \mathrm{e}^{\hat{u}} + \mathrm{e}^{\hat{v}} + z\mathrm{e}^{-\hat{u}-\hat{v}}. \tag{5.1}$$

After a choice of polarization, we can require this operator to annihilate the wave function and obtain a difference equation. This equation can be analyzed as in [5], where it was shown that the perturbative WKB periods give the standard refined topological string on local $\mathbb{P}^2$, in the NS limit. However, this does not lead by itself to a quantized spectrum, since we need additional analyticity conditions on the wave function. One could impose a natural set of conditions similar to the ones we imposed on (2.13) (see [23] for a related discussion). As we mentioned before, the quantum B-period is singular for infinitely many values of $\hbar$. If the difference equation defined by (5.1) with the appropriate analyticity conditions leads to a well-defined spectral problem for these values of $\hbar$, then non-perturbative instanton corrections are needed.

It might be the case that these corrections involve the standard topological string, as in the example studied in this paper.

Our result suggests in fact a new point of view on the non-perturbative definition of topological string theory: start with a well-defined spectral problem arising in the quantization of the spectral curve, and find its exact WKB quantization condition. The perturbative part will be given by the refined topological string in the NS limit, and the non-perturbative corrections might lead to the Gopakumar–Vafa expansion of the usual topological string, just as in our example.

This non-perturbative definition depends ultimately on an appropriate definition of the spectral problem. There is however a family of CY geometries in which this problem should be well defined, thanks to the result of [2]: these are the A_{N-1} local geometries, which generalize local $\mathbb{P}^1 \times \mathbb{P}^1$. In this case, the spectral problem is the Baxter equation obtained by quantizing the spectral curve of the relativistic Toda lattice. However, this equation is an auxiliary tool to solve the original quantum integrable system, and the analyticity properties of the solution to the Baxter equation should follow from the original problem of determining the spectrum of the conserved Hamiltonians for the quantum, relativistic Toda lattice (this is well understood in the non-relativistic limit, see for example [22,43]). Interestingly, the A_{N-1} geometries also have a dual large N Chern–Simons description, as well as a matrix model description [44]. It would be very interesting to study non-perturbatively the quantum spectral curves of these geometries and understand their relation to the matrix models. In the case studied in this paper, the matrix model partition function is the partition function of an ideal gas of N fermions with the energy eigenvalues determined by the spectral problem.

Another possible extension of this paper is the analysis of the spectral problem appearing in the Fermi gas formulation of $\mathcal{N} = 3$ Chern–Simons–matter models. In [13], generalized kernels were defined for many of these theories; it would be interesting to study their spectrum with the techniques developed here. This would lead to new results for non-perturbative effects in the type IIA/M-theory duals to these theories.

We hope to report on these problems in the near future.

Note added: Since this paper was first posted, there have been various developments closely related to the problems raised here. It was noted in [45], by performing detailed numerical calculations of the spectrum for various values of k, that the proposal (3.55) for non-perturbative corrections is only valid in the so-called maximally supersymmetric cases $k = 1, 2$. A corrected quantization condition for the operator (2.1), which so far agrees with all existing data, has been proposed in [46,47].

Appendix A. Mellin Transform

We will use the Mellin transform as defined in (4.12). The inverse Mellin transform can be computed by inspection after using the following elementary integral:

$$\int_0^1 u^{n-s-1} (\log u)^m \, \mathrm{d}u = -\frac{\Gamma(m+1)}{(s-n)^{m+1}}. \tag{A.1}$$

Regular terms in s do not contribute to μ expansion of the inverse Mellin transform. Therefore, we first compute the Mellin transform, consider the polar part of the Laurent series, and perform the inverse Mellin transform by using (A.1).

We will do a Mellin transform of quantities in the grand canonical ensemble with respect to the variable

$$u = \mathrm{e}^{-\mu}. \tag{A.2}$$

The Mellin transform of the Fermi occupation number is

$$\int_0^1 \frac{u^{-s-1}}{\mathrm{e}^E u + 1} \mathrm{d}u = \mathrm{e}^{sE} I(s) + \sum_{k \geq 1} \frac{(-1)^k \mathrm{e}^{-kE}}{s+k}, \tag{A.3}$$

where

$$I(s) = \int_0^\infty \frac{x^s}{1+x} \mathrm{d}x = -\pi \csc(\pi s). \tag{A.4}$$

Notice that the function in the r.h.s. of (A.3) does not have poles at negative integer values of s. The function $I(s)$ has simple poles at the positive integers. The Laurent series around even non-negative integers is,

$$I(s) = -\frac{1}{s - 2\ell} - \frac{\pi^2}{6}(s - 2\ell) + \cdots \tag{A.5}$$

while for odd positive integers we find,

$$I(s) = \frac{1}{s - (2\ell+1)} + \frac{\pi^2}{6}(s - (2\ell+1)) + \cdots. \tag{A.6}$$

Using these results, it is easy to compute the Mellin transform of $J_\mathrm{p}(\mu, k)$:

$$\widehat{J}_\mathrm{p}(s, k) = \frac{1}{2\pi^2 k} \left\{ 4\mathcal{I}_2(-s) - \left(\frac{2\pi^2}{3} - \frac{\hbar^2}{48} \right) \mathcal{I}_0(-s) \right.$$

$$\left. -4 \sum_{\ell \geq 1} \widehat{a}_\ell(\hbar) \mathcal{I}_1(2\ell - s) + \sum_{\ell \geq 1} \widehat{b}_\ell(\hbar) \mathcal{I}_0(2\ell - s) \right\} I(s) + \cdots, \tag{A.7}$$

where $\mathcal{I}_j(n)$ are defined in (4.14), and the dots in (A.7) denote terms which are regular in s or come from the second term in (A.3). To implement the inverse Mellin transform, we notice that this function has a pole of order fourth at $s = 0$, with Laurent expansion

$$\frac{4}{\pi^2 k} \frac{1}{s^4} + \left(\frac{1}{3k} + \frac{k}{24} \right) \frac{1}{s^2} - \frac{\widehat{A}(\hbar)}{s} + \cdots \tag{A.8}$$

where $\widehat{A}(\hbar)$ is given in (4.16). At even positive integers, we have triple poles with the Laurent expansion

$$\frac{2}{\pi^2 k} \frac{\widehat{a}_\ell(\hbar)}{(s - 2\ell)^3} + \frac{1}{2\pi^2 k} \frac{\widehat{b}_\ell(\hbar)}{(s - 2\ell)^2} - \frac{1}{2\pi^2 k} \frac{\widehat{c}_\ell(\hbar)}{s - 2\ell} + \cdots \tag{A.9}$$

where $\widehat{c}_\ell$ is given in the first line of (4.17). At odd positive integers, we have simple poles with the structure

$$\frac{1}{2\pi^2 k}\frac{\widehat{d}_\ell(\hbar)}{s-(2\ell+1)}, \tag{A.10}$$

where $\widehat{d}_\ell$ is given in the second line of (4.17). Putting all the results together, we derive the result in (A.7).

Equivalently, one can use the Mellin transform to calculate the integrals (4.10). Their values are easily found to be,

$$R_\ell^{(0)}(E_0,\mu) = \sum_{\substack{n=1,\\n\neq 2\ell}} \frac{(-1)^n}{2\ell-n}\mathrm{e}^{-(2\ell-n)E_0}\mathrm{e}^{-n\mu} + (\mu-E_0)\,\mathrm{e}^{-2\ell\mu} + \frac{\mathrm{e}^{-2\ell E_0}}{2\ell},$$

$$R_\ell^{(1)}(E_0,\mu) = \sum_{\substack{n=1,\\n\neq 2\ell}} (-1)^n\left(\frac{1}{(2\ell-n)^2}+\frac{E_0}{2\ell-n}\right)\mathrm{e}^{-(2\ell-n)E_0}\mathrm{e}^{-n\mu}$$

$$+ \left(\frac{\pi^2}{6}+\frac{\mu^2-E_0^2}{2}\right)\mathrm{e}^{-2\ell\mu} + \frac{(1+2\ell E_0)}{(2\ell)^2}\mathrm{e}^{-2\ell E_0}. \tag{A.11}$$

These expressions are valid for real $\mu > E_0$. Finally, we compute the integral (4.11). The Mellin transform is

$$\widehat{R}_\sigma(E_0,s) = \mathcal{I}_0(\sigma-s)I(s) + \cdots. \tag{A.12}$$

If σ is not an integer, we have two types of simple poles in this expression: poles at $s=\sigma$ coming from $\mathcal{I}_0(\sigma-s)$, and poles coming from $I(s)$ at integer values $s=n$ (like before, there are no poles at negative integers). A simple calculation shows that the inverse Mellin transform is given by

$$R_\sigma(E_0,\mu) = -\pi\csc(\pi\sigma)\mathrm{e}^{-\sigma\mu} + \sum_{n\geq 0}\frac{(-1)^n}{\sigma-n}\mathrm{e}^{-(\sigma-n)E_0}\mathrm{e}^{-n\mu}, \tag{A.13}$$

which again is valid for real $\mu > E_0$. Notice that the first term is what we would obtain with the Sommerfeld expansion of the Fermi factor used in for example [48], and the second term is a non-perturbative correction. This result can be also written down in terms of the Hurwitz Lerch transcendent $\Phi(a,b,c)$,

$$R_\sigma(E_0,\mu) = -\pi\csc(\pi\sigma)\mathrm{e}^{-\sigma\mu} + \frac{1}{\sigma}\mathrm{e}^{-\sigma E_0} + \mathrm{e}^{E_0(1-\sigma)-\mu}\Phi\left(\mathrm{e}^{E_0-\mu},1,1-\sigma\right). \tag{A.14}$$

Appendix B. Higher Order Quantum Volume

In this appendix, we will sketch how to derive the result (3.25). It is obtained by solving Eq. (2.13) with the ansatz (3.13) and then performing the period integral (3.16). When solving (2.13), we will follow the approach given in [49, 50].

Let us write the difference Eq. (2.13) as

$$\left[\mathrm{e}^{\frac{\mathrm{i}\hbar}{2}\partial_x} + \mathrm{e}^{-\frac{\mathrm{i}\hbar}{2}\partial_x} \right] \psi(x) = 2r(x)\psi(x) \tag{B.1}$$

where

$$r(x) = \frac{\mathrm{e}^E}{4\cosh\left(\frac{x}{2}\right)}. \tag{B.2}$$

Plugging in the ansatz (3.13) in the above equation and expanding around $\hbar = 0$ keeping $S(x,\hbar)$ fixed we get the following equation:

$$\cosh\left(\frac{1}{2}S^{(1)}(x,\hbar) - \frac{\hbar^2}{2^3}\frac{S^{(3)}(x,\hbar)}{3!} + \cdots \right) \exp\left(-\frac{\mathrm{i}\hbar}{2^2}\frac{S^{(2)}(x,\hbar)}{2!} + \cdots \right) = r(x), \tag{B.3}$$

where

$$S^{(m)}(x,\hbar) = \partial_x^m S(x,\hbar). \tag{B.4}$$

We notice that a derivative of $S(x,\hbar)$ of order m always comes with a factor $\hbar^{m-1}$, $m > 0$. If we plug in the expansion (3.14), we can therefore first solve for $\partial_x S_0$ and then find $\partial_x S_n(x)$, $n > 0$, in terms of derivatives of $\partial_x S_m(x)$ with $m < n$. For the first three orders, we find

$$\partial_x S_0(x) = 2\cosh^{-1}\left[r(x) \right] = p(x),$$

$$\partial_x S_1(x) = \frac{\mathrm{i}r(x)S_0^{(2)}(x)}{4\sqrt{r(x)^2 - 1}}, \tag{B.5}$$

$$\partial_x S_2(x) = \frac{r(x)}{64\left(r(x)^2 - 1\right)^{3/2}}\left[\left(S_0^{(2)}(x)\right)^2 + 16\left(r(x)^2 - 1\right)S_1^{(2)}(x) \right] + \frac{S_0^{(3)}(x)}{48}.$$

The perturbative part of the quantum volume is given by (3.16), and it has an expansion for small $\hbar$ given in (3.24). To find $\mathrm{vol}_n(E)$ we use (B.5) and calculate the integral. To lowest order, we find the result (3.3). The order $\hbar$ contribution vanishes as it should and at order $\hbar^2$ we obtain the result (3.25).

Acknowledgements

References

[1] Di Francesco, P., Ginsparg, P.H., Zinn-Justin, J.: 2-D Gravity and random matrices. Phys. Rept. **254**, 1 (1995). arXiv:hep-th/9306153

[2] Nekrasov, N.A., Shatashvili, S.L.: Quantization of integrable systems and four dimensional gauge theories. arXiv:0908.4052 [hep-th]

[3] Mironov, A., Morozov, A.: Nekrasov functions and exact Bohr–Zommerfeld integrals. JHEP **1004**, 040 (2010). arXiv:0910.5670 [hep-th]

[4] Mironov, A., Morozov, A.: Nekrasov functions from exact BS periods: the case of SU(N). J. Phys. A **43**, 195401 (2010). arXiv:0911.2396 [hep-th]

[5] Aganagic, M., Cheng, M.C.N., Dijkgraaf, R., Krefl, D., Vafa, C.: Quantum geometry of refined topological strings. JHEP **1211**, 019 (2012). arXiv:1105.0630 [hep-th]

[6] Zinn-Justin, J.: Multi-Instanton Contributions In Quantum Mechanics, 1 and 2. Nucl. Phys. B **192**, 125 (1981)

[7] Zinn-Justin, J.: Multi-Instanton Contributions In Quantum Mechanics, 1 and 2. Nucl. Phys. B **218**, 333 (1983)

[8] Zinn-Justin, J., Jentschura, U.D.: Multi-instantons and exact results I: conjectures, WKB expansions, and instanton interactions. Ann. Phys. **313**, 197 (2004). arXiv:quant-ph/0501136

[9] Zinn-Justin, J., Jentschura, U.D.: Multi-instantons and exact results II: specific cases, higher-order effects, and numerical calculations. Ann. Phys. **313**, 269 (2004). arXiv:quant-ph/0501137

[10] Aharony, O., Bergman, O., Jafferis, D.L., Maldacena, J.: N=6 superconformal Chern–Simons-matter theories, M2-branes and their gravity duals. JHEP **0810**, 091 (2008). arXiv:0806.1218 [hep-th]

[11] Kapustin, A., Willett, B., Yaakov, I.: Exact results for Wilson Loops in superconformal Chern–Simons theories with matter. JHEP **1003**, 089 (2010). arXiv:0909.4559 [hep-th]

[12] Drukker, N., Mariño, M., Putrov, P.: From weak to strong coupling in ABJM theory. Commun. Math. Phys. **306**, 511 (2011). arXiv:1007.3837 [hep-th]

[13] Mariño, M., Putrov, P.: ABJM theory as a Fermi gas. J. Stat. Mech. **1203**, P03001 (2012). arXiv:1110.4066 [hep-th]

[14] Hatsuda, Y., Mariño, M., Moriyama, S., Okuyama, K.: Non-perturbative effects and the refined topological string. arXiv:1306.1734 [hep-th]

[15] Hatsuda, Y., Moriyama, S., Okuyama, K.: Exact results on the ABJM fermi gas. JHEP **1210**, 020 (2012). arXiv:1207.4283 [hep-th]

[16] Hatsuda, Y., Moriyama, S., Okuyama, K.: Instanton effects in ABJM theory from fermi gas approach. JHEP **1301**, 158 (2013). arXiv:1211.1251 [hep-th]

[17] Calvo, F., Mariño, M.: Membrane instantons from a semiclassical TBA. JHEP **1305**, 006 (2013). arXiv:1212.5118 [hep-th]

[18] Hatsuda, Y., Moriyama, S., Okuyama, K.: Instanton bound states in ABJM theory. JHEP **1305**, 054 (2013). arXiv:1301.5184 [hep-th]

[19] Zamolodchikov, A.B.: Painlevé III and 2-d polymers. Nucl. Phys. B **432**, 427 (1994). arXiv:hep-th/9409108

[20] Tracy, C.A., Widom, H.: Proofs of two conjectures related to the thermodynamic Bethe ansatz. Commun. Math. Phys. **179**, 667–680 (1996). arXiv:solv-int/9509003

[21] Nekrasov, N.: Five dimensional gauge theories and relativistic integrable systems. Nucl. Phys. B **531**, 323 (1998). arXiv:hep-th/9609219

[22] Gaudin, M., Pasquier, V.: The periodic Toda chain and a matrix generalization of the Bessel function's recursion relations. J. Phys. A **25**, 5243 (1992)

[23] Sergeev, S.M.: Quantization scheme for modular q-difference equations. Theor. Math. Phys. **213**, 422 (2005). arXiv:nlin/0402008

[24] Poghossian, R.: Deforming SW curve. JHEP **1104**, 033 (2011). arXiv:1006.4822 [hep-th]

[25] Fucito, F., Morales, J.F., Pacifici, D.R., Poghossian, R.: Gauge theories on Ω-backgrounds from non commutative Seiberg–Witten curves. JHEP **1105**, 098 (2011). arXiv:1103.4495 [hep-th]

[26] Fateev, V.A., Lukyanov, S.L., Zamolodchikov, A.B.: On mass spectrum in 't Hooft's 2D model of mesons. J. Phys. A **42**, 304012 (2009). arXiv:0905.2280 [hep-th]

[27] Gopakumar, R., Vafa, C.: M theory and topological strings. 2. arXiv:hep-th/9812127

[28] Hatsuda, Y., Moriyama, S., Okuyama, K.: Unpublished.

[29] Iqbal, A., Kozcaz, C., Vafa, C.: The Refined topological vertex. JHEP **0910**, 069 (2009). arXiv:hep-th/0701156

[30] Choi, J., Katz, S., Klemm, A.: The refined BPS index from stable pair invariants. Commun. Math. Phys. **328**, 903 (2014). arXiv:1210.4403 [hep-th]

[31] Balian, R., Parisi, G., Voros, A.: Discrepancies from asymptotic series and their relation to complex classical trajectories. Phys. Rev. Lett. **41**, 1141 (1978)

[32] Balian, R., Parisi, G., Voros, A.: Quartic oscillator. In: Feynman Path Integrals, Lecture Notes in Physics, vol. 106, pp. 337 (1979)

[33] Voros, A.: The return of the quartic oscillator. The complex WKB method. Ann. Inst. H. Poincaré A **39**, 211 (1983)

[34] Delabaere, E., Dillinger, H., Pham, F.: Exact semiclassical expansions for one-dimensional quantum oscillators. J. Math. Phys. **38**, 6126 (1997)

[35] Delabaere, E., Pham, F.: Resurgent methods in semi-classical asymptotics. Ann. Inst. Henri Poincaré **71**, 1 (1999)

[36] Mariño, M., Putrov, P.: Exact results in ABJM theory from topological strings. JHEP **1006**, 011 (2010). arXiv:0912.3074 [hep-th]

[37] Basar, G., Dunne, G.V., Unsal, M.: Resurgence theory, ghost-instantons, and analytic continuation of path integrals. JHEP **1310**, 041 (2013). arXiv:1308.1108 [hep-th]

[38] Lockhart, G., Vafa, C.: Superconformal partition functions and non-perturbative topological strings. arXiv:1210.5909 [hep-th]

[39] Bender, C.M., Orszag, S.A.: Advanced mathematical methods for scientists and engineers. Springer, New York (1999)

[40] Kapustin, A., Willett, B., Yaakov, I.: Nonperturbative tests of three-dimensional dualities. JHEP **1010**, 013 (2010). arXiv:1003.5694 [hep-th]

[41] Costin, O., Garoufalidis, S.: Resurgence of the Euler–MacLaurin summation formula. Ann. de l'Institut Fourier **58**, 893 (2008). arXiv:math/0703641

[42] Hanada, M., Honda, M., Honma, Y., Nishimura, J., Shiba, S., Yoshida, Y.: Numerical studies of the ABJM theory for arbitrary N at arbitrary coupling constant. JHEP **1205**, 121 (2012). arXiv:1202.5300 [hep-th]

[43] Kharchev, S., Lebedev, D.: Integral representation for the eigenfunctions of quantum periodic Toda chain. Lett. Math. Phys. **50**, 53 (1999). arXiv:hep-th/9910265

[44] Aganagic, M., Klemm, A., Mariño, M., Vafa, C.: Matrix model as a mirror of Chern–Simons theory. JHEP **0402**, 010 (2004). arXiv:hep-th/0211098

[45] Wang, X.F., Wang, X., Huang, M.X.: A note on instanton effects in ABJM theory. JHEP **1411**, 100 (2014). arXiv:1409.4967 [hep-th]

[46] Grassi, A., Hatsuda, Y., Mariño, M.: Topological strings from Quantum Mechanics. arXiv:1410.3382 [hep-th]

[47] Grassi, A., Hatsuda, Y., Mariño, M.: Quantization conditions and functional equations in ABJ(M) theories. arXiv:1410.7658 [hep-th]

[48] Klemm, A., Mariño, M., Schiereck, M., Soroush, M.: ABJM Wilson loops in the Fermi gas approach. arXiv:1207.0611 [hep-th]

[49] Dingle, R.B., Morgan, G.J.: WKB methods for difference equations. I. Appl. Sci. Res. **18**, 221–237 (1967)

[50] Dingle, R.B., Morgan, G.J.: WKB methods for difference equations. II. Appl. Sci. Res. **18**, 238–245 (1968)

Johan Källén and Marcos Mariño
Département de Physique Théorique et Section de Mathématiques
Université de Genève
1211 Geneva, Switzerland
e-mail: johan.kallen@unige.ch
        marcos.marino@unige.ch