Spectral statistics of permutation matrices

Idan Oren\textsuperscript{1,2} and Uzy Smilansky\textsuperscript{1,2}

\textsuperscript{1}Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot 76100, Israel
\textsuperscript{2}School of Mathematics, Cardiff University, Cardiff CF24 4AG, UK

We compute the mean two-point spectral form factor and the spectral number variance for permutation matrices of large order. The two-point correlation function is expressed in terms of generalized divisor functions, which are frequently discussed in number theory. Using classical results from number theory and casting them in a convenient form, we derive expressions which include the leading and next to leading terms in the asymptotic expansion, thus providing a new point of view on the subject, and improving some known results.

1. Introduction

The spectral statistics of the eigenvalues of permutation matrices of large dimension have been studied during the past decade from various points of view (e.g. \cite{1–5}). In this article, we approach this subject by computing the spectral form factor \cite{6} defined in the following way. Consider an $N \times N$ unitary matrix $U$ drawn from a random ensemble with a given probability distribution over its elements. The form factor is

$$K(t; N) = \mathbb{E}\left[\frac{1}{N} |\text{tr} U(t)|^2\right], \quad t \in \mathbb{N}. \quad (1.1)$$

Here and throughout, $\mathbb{E}[\cdot]$ stands for the expectation with respect to the random ensemble. As the spectra of the matrices are confined to the unit circle, $K(t; N)$ is the Fourier transform of the spectral two-point correlation function

$$K(t; N) = \frac{1}{2\pi} \int_0^{2\pi} ds e^{its} \mathbb{E}\left[\frac{1}{N} \sum_{l,m=1}^N \delta(s - (\theta_l - \theta_m))\right], \quad (1.2)$$

where the eigenvalues of $U$ are $\{\exp(i\theta_l)\}_1^N$. When the eigenphases $\theta_l$ are uniformly distributed on the unit circle, it is convenient to express their differences in units...
of the mean spacing $2\pi/N$, and for $N \to \infty$ the form factor limits to a function of the single parameter $\tau = t/N$. Explicit expressions for the form factor are known for various matrix ensembles, the most prominent ones being Dyson’s circular ensembles. For the Poissonian ensemble where the eigenphases are independent random variables with uniform distribution, $K(\tau) = 1$. The study of the form factor is motivated by its successful application in ‘quantum chaos’ [7] and spectral graph theory [8,9].

A closely related quantity of interest is the number variance $\langle \text{Var}(N(L)) \rangle$, which measures the fluctuations in the number of eigenphases $N(L; \theta)$ in an arc $[\theta, \theta + 2\pi(L/N)]$ which covers $L$ eigenphases on average,

$$\langle \text{Var}(N(L)) \rangle = \int_0^{2\pi} \frac{d\theta}{2\pi} E[(N(L; \theta) - L)^2].$$

This can be written in terms of the form factor as

$$\langle \text{Var}(N(L)) \rangle = \frac{2}{N\pi^2} \sum_{t=1}^{\infty} \frac{\sin^2 \pi L(t/N)}{(t/N)^2} K(t; N).$$

The ensemble to be discussed in the sequel is the ensemble of permutation matrices of $N$ objects in the large $N$ limit. A permutation $\rho$ of $N$ objects is represented by an $N \times N$ unitary matrix $P_\rho$ with elements from $\{0, 1\}$. The spectrum of $P_\rho$ is determined by the cycle structure of $\rho$: let $\rho$ consist of cycles of lengths $k$ with multiplicities $a_k(\rho)$. Clearly,

$$N = \sum_{k=1}^{N} ka_k(\rho).$$

Each cycle in $\rho$ of length $k$ contributes to the spectrum $k$ roots of unity: $\{e^{2\pi i(m/k)}\}_{m=1}^{k}$, with multiplicity $a_k(\rho)$. Thus,

$$\text{tr} P_\rho^t = \sum_{k=1}^{N} a_k(\rho) \sum_{m=1}^{k} e^{2\pi i(m/k)t} = \sum_{k|t}^{\min(t,N)} ka_k(\rho).$$

By assigning a uniform probability distribution to the permutations, the form factor reads

$$K(t; N) = \frac{1}{N} E[|\text{tr} P_\rho^t|^2] = \frac{1}{N} E \left[ \sum_{k|t} k^2 a_k(\rho)^2 \right],$$

where the expectation is with respect to the ensemble of permutation matrices as defined above.

Figure 1 shows the numerical computation of the form factor for $N = 500$. The rapid fluctuations almost hide the mean smooth behaviour of the form factor. The fluctuations are not the result of any approximation or an insufficient sample of matrices taken in the simulations. Rather, it is inherent to the problem, and its origin is number theoretical. In the remainder of the article, we shall provide an asymptotic expression for $K(t; N)$ and show that, after appropriate smoothing, it is described by a simple formula which will be given explicitly.

The spectral statistics of the random permutation matrices have been studied by various authors. Focusing our attention on the uniform probability measure (Ewens parameter $\theta = 0$ [10]), and on subjects directly relevant to this work, the following asymptotic results in the limit $N \to \infty$ are known:

- On average, the eigenphases are uniformly distributed on the unit circle: $E[N(L; \theta)] = L$ [1].
- The leading term in the mean number variance is $\Sigma_2(L) \asymp \frac{1}{6} \log L$ [2].
Figure 1. The form factor $K(t, N)$ for $N = 500$ (1.7) is the wildly fluctuating graph (grey). Its smooth version is indicated by the thick dark line.

The leading expression for the spectral two-point correlation function is explicitly given in [4],

$$K(\tau) \approx 2 \int_0^{\infty} \cos 2\pi \tau x \left\{ \frac{1}{2} + \frac{1}{x^2} \sum_{n=1}^{\left\lfloor x \right\rfloor} n \right\} dx. \quad (1.8)$$

The Fourier transform of the above gives the leading asymptotic expressions for the form factor which coincides with the expression derived here in §3b.

This approach provides an alternative way of deriving the known results, and enables us to present the next terms in the $1/N$ asymptotic expansion, as will be shown in §3c.

2. The spectral form factor

The following theorem [2,11] enables us to proceed with the computation of $K(t, N)$.

Theorem 2.1 (Wieand). If a permutation $\rho$ is chosen at random, then, for large $N$, the numbers $a_2(\rho), a_3(\rho), \ldots, a_k(\rho)$ are asymptotically independent Poisson random variables with parameters $1/2, 1/3, \ldots, 1/k$. In particular,

$$\mathbb{E}[a_k] = \begin{cases} \frac{1}{k} & \text{if } k \leq N, \\ 0 & \text{otherwise.} \end{cases}$$

For $k \neq m$

$$\mathbb{E}[a_ka_m] = \begin{cases} \frac{1}{km} & \text{if } k + m \leq N, \\ 0 & \text{otherwise.} \end{cases}$$
where

\[ \sigma_k = \begin{cases} 
\frac{1}{k} & \text{if } 1 \leq k \leq \frac{N}{2}, \\
1 - \frac{1}{k^2} & \text{if } \frac{N}{2} < k \leq N, \\
0 & \text{otherwise}.
\]

A straightforward application of the theorem provides an explicit expression of the form factor. It is convenient to write it as a sum of two terms: the first is the leading term which remains finite in the limit of interest here \((t, N \to \infty, t/N \to \tau)\); the second term vanishes in this limit.

\[ K(t, N) = \frac{1}{N} \sigma_1(t; N) + \frac{1}{N} C(t, N) \quad (2.1) \]

and

\[ C(t, N) = \begin{cases} 
\sigma_0(t)^2 & \text{if } 0 < t \leq \frac{N}{2}, \\
\sigma_0(t)^2 - 1 & \text{if } \frac{N}{2} < t \leq N, \\
\sigma_0(t, N)^2 + \sigma_0 \left( t, \frac{N}{2} \right) & \text{if } N < t.
\]

The restricted divisor functions \(\sigma_a(t, N)\) [12] are defined by

\[ \sigma_a(t, N) = \sum_{k \mid t} k^a \quad \text{for } a \in \mathbb{R}, \quad (3.1) \]

where \(\sigma_a(t, N)\) generalize the well-known divisor functions \(\sigma_a(t)\) [13]. Indeed, for \(t \leq N\), \(\sigma_a(t, N) = \sigma_a(t)\).

The divisor functions \(\sigma_a(t)\) display considerable fluctuations (see, for example, the drawings of \(\sigma_a(t)\) in [14] for \(a = 0, 1, \ldots\)). The fluctuation of \(\sigma_1(t)\) which provides the dominant contribution to \(K(t, N)\) obtain their lowest value \((t + 1)\), for prime \(t\). They can also reach values of order \(t \log t\). A further smoothing is required to deduce the smooth behaviour of \(K(t, N)\). The smoothing procedure used here is explained in the next section.

3. The smooth version of the form factor

The computation of the restricted divisor functions \(\sigma_a(t; N)\) is simplified by recalling that when \(k\) is a divisor of \(t\), then \(t/k\) is also a divisor. Hence to every divisor, \(k \leq \lfloor \sqrt{t} \rfloor\) corresponds to the divisor \(t/k\) in the interval \([\lfloor \sqrt{t} \rfloor, t]\). Therefore,

\[ \sigma_a(t; N) = \sum_{k \mid t} k^a + \sum_{k \mid t: k \geq \lfloor t/N \rfloor} \left( \frac{t}{k} \right)^a. \quad (3.1) \]

When \(t < N\), \(\sigma_a(t; N) = \sigma_a(t)\), which takes the simpler form

\[ \sigma_a(t) = \sum_{k \mid t} \left( k^a + \left( \frac{t}{k} \right)^a \right). \quad (3.2) \]

In this case and if \(t\) is itself a square of an integer, the last term in the sum above is replaced by \(t^a/2\).

It is convenient to define the indicator \(\chi_t(k)\), which takes the value 1 if \(k \mid t\) and 0 otherwise. Thus,

\[ \sigma_a(t; N) = \sum_{k=1}^{\lfloor \min(\sqrt{t}, N) \rfloor} \chi_t(k) k^a + \sum_{k=\lfloor t/N \rfloor}^{\lfloor \sqrt{t} \rfloor} \chi_t(k) \left( \frac{t}{k} \right)^a. \quad (3.3) \]

The smoothing of the form factor will be achieved using the following procedure.
Figure 2. The domains in the $k, r$ plane enclosing the integer lattice points used in the smoothing procedure (3.5).

(a) Smoothing of divisor functions

Let $F(t) : \mathbb{N} \to \mathbb{R}$ be an arbitrary function. We define its smooth form $\langle F(t) \rangle$ at $t$ in the following way. Choose a smoothing interval of size $\Delta$ so that $t \gg \Delta \gg 1$. Then,

$$
\langle F(t) \rangle = \frac{1}{2\Delta/2 + 1} \sum_{|r-t| \leq \Delta/2} F(r). \tag{3.4}
$$

The choice of $\Delta$ is dictated by the properties of $F(t)$ and the required degree of smoothing. The functions to be smoothed here are all of the form $\sigma_a(t; N) = \sum_{k \leq N} k^a$, where $a \in \mathbb{R}$, and it will be shown that the optimal choice of $\Delta$ is $\sqrt{t}$.

We shall first consider the smoothing in the case $t < N$. Then, $\sigma_a(t; N) = \sigma_a(t)$ and

$$
\langle \sigma_a(t) \rangle = \frac{1}{2\Delta/2 + 1} \sum_{|r-t| \leq \Delta/2} \sum_{k=1}^{\lfloor \sqrt{t} \rfloor} \chi_r(k) \left( k^a + \left( \frac{r}{k} \right)^a \right). \tag{3.5}
$$

Before extracting the leading terms in an asymptotic expansion in powers of $\sqrt{t}$, one should note the following. The sum above extends over the integer lattice points in the $(r, k)$ plane inside a (slightly deformed) trapezoid with corners at $(t - \lfloor \Delta/2 \rfloor, 1)$, $(t + \lfloor \Delta/2 \rfloor, 1)$, $(t + \lfloor \Delta/2 \rfloor, \lfloor \sqrt{t} + \lfloor \Delta/2 \rfloor \rfloor)$ and $(t - \lfloor \Delta/2 \rfloor, \lfloor \sqrt{t} - \lfloor \Delta/2 \rfloor \rfloor)$ (figure 2). Most of the lattice points are covered by the rectangle $(t - \lfloor \Delta/2 \rfloor, 1)$, $(t + \lfloor \Delta/2 \rfloor, 1)$, $(t + \lfloor \Delta/2 \rfloor, \lfloor \sqrt{t} \rfloor)$ and $(t - \lfloor \Delta/2 \rfloor, \lfloor \sqrt{t} \rfloor)$. The two domains differ by two triangles (marked in grey in figure 2), one just below the line $r = \sqrt{t}$ (for $r < t$) and the other just above this line (for $r > t$). Note that $\sqrt{t} \pm \lfloor \Delta/2 \rfloor$ differs from $\sqrt{t}$ by $\approx \pm (1/2\sqrt{t}) \lfloor \Delta/2 \rfloor$, which is smaller than $1/4$ in absolute value. Thus, the number of lattice points inside the triangles is bounded by $\Delta$. The sum (3.5) over the trapezoid therefore equals the sum over the rectangle plus the contribution of each of the triangles, which are added with opposite signs. As $t$ is varying, the total contribution from the two triangles fluctuates about zero, thus contributing a fluctuating correction which vanishes upon further smoothing. Another important observation applies to the second term in (3.5) only. It is proportional to $r^a$ and replacing it by $t^a$ incurs an error bounded by $(a(a - 1)/2) t^{a-2} \Delta^2$, because the summation over symmetric points with respect to the middle line $r = t$ removes the linear terms in $\Delta$. The above should be borne in mind when we proceed now with some approximations, whose accuracy will be estimated in light of these observations.
The leading terms in the asymptotic expansion of (3.5) are obtained by restricting the sum to the points inside the rectangle outlined above, and by replacing \( r^d \) by \( t^d \) in the second term. In the resulting sum, the order of summations can be interchanged, so that

\[
\langle \sigma_a(t) \rangle = \sum_{k=1}^{\lfloor \sqrt{t} \rfloor} \frac{1}{k} \left( k^a + \left( \frac{t}{k} \right)^a \right) + \mathcal{R}_a(t),
\]

(3.6)

where \( \mathcal{R}_a(t) \) is a reminder to be discussed later. The expression in the curly brackets equals the number of times the integer \( k \) divides the integers \( r \) in the interval \([t − \lfloor t/2 \rfloor, t + \lfloor t/2 \rfloor]\), normalized by the number of integers in the interval. As \( 1 \leq k \leq \lfloor \sqrt{t} \rfloor \), for any \( \Delta \geq \lfloor \sqrt{t} \rfloor \), the curly bracket can be approximated by \( 1/k \) for all the \( k \) values appearing in the \( k \)-sum. This introduces an error term of order \( 1/\Delta \) whose sign fluctuates depending on \( k \) and \( t \), and which is included in the reminder term \( \mathcal{R}_a(t) \).

The minimal smoothing interval is \( \Delta = \lfloor \sqrt{t} \rfloor \) and is chosen because it provides the best possible resolution of details which survive smoothing. With this choice,

\[
\langle \sigma_a(t) \rangle = \sum_{k=1}^{\lfloor \sqrt{t} \rfloor} \frac{1}{k} \left( k^a + \left( \frac{t}{k} \right)^a \right) + \mathcal{R}_a(t).
\]

(3.7)

In particular, for \( a = 0 \),

\[
\langle \sigma_0(t) \rangle = 2 \sum_{k=1}^{\lfloor \sqrt{t} \rfloor} \frac{1}{k} = 2H_{\lfloor \sqrt{t} \rfloor} + \mathcal{R}_0(t),
\]

(3.8)

where \( H_n \) is the \( n \)th harmonic number. For general \( a \in \mathbb{N}^+ \)

\[
\langle \sigma_a(t) \rangle = \frac{B_a(\lfloor \sqrt{t} \rfloor + 1) - B_a}{a} + t^a(\zeta(a + 1) - \zeta(a + 1, \lfloor \sqrt{t} \rfloor + 1)) + \mathcal{R}_a(t),
\]

(3.9)

where \( B_a(x) \) are the Bernoulli polynomials, \( B_a \) are the Bernoulli numbers and \( \zeta(a,x) \) is the generalized (Hurwitz) \( \zeta \) function: \( \zeta(a, s) = \sum_{n=0}^{\infty} (n + s)^{-a} \). Explicit expressions can be obtained by using the asymptotic expression \( \zeta(k,s) = s^{-k}(s/(k - 1) + 1/2) + O(s^{-(k+1)}) \) and similar expressions for \( B_a(x) \). In this work, only \( \langle \sigma_0(t) \rangle \) and \( \langle \sigma_1(t) \rangle \) are required,

\[
\langle \sigma_0(t) \rangle = \sum_{k=1}^{\lfloor \sqrt{t} \rfloor} \frac{1}{k} (1 + 1) = \log t + 2 \gamma + \mathcal{R}_0(t)
\]

(3.10)

and

\[
\langle \sigma_1(t) \rangle = t \zeta(2) - \frac{1}{2} + \frac{1}{4 \sqrt{t}} + \mathcal{R}_1(t).
\]

(3.11)

These results are consistent with the known asymptotic expressions [15–18]: for \( \langle \sigma_0(t) \rangle \),

\[
\sum_{s=1}^{t} \sigma_0(s) = t \log t + t(2 \gamma - 1) + O(t^d),
\]

(3.12)

where \( \gamma \) is the Euler constant and the lowest known upper bound for \( \theta \) is 131/416 [19].

For \( \langle \sigma_a(t) \rangle \), with \( a \geq 1 \),

\[
\sum_{s=1}^{t} \sigma_a(s) = \frac{\zeta(a + 1)}{a + 1} t^{a+1} + O(t^d).
\]

(3.13)

The magnitude of the fluctuations of \( \sigma_a(t) \) about their smooth form for \( a \geq 2 \) and large \( t \) can be bounded using an identity due to Ramanujan [16]

\[
\frac{1}{\zeta(a + 1)} < \frac{\sigma_a(t)}{t^a \zeta(a + 1)} = \frac{\sigma_a(t)}{\langle \sigma_a(t) \rangle} < \frac{\zeta(a)}{\zeta(a + 1)}. \tag{3.14}
\]

Both bounds converge rapidly to 1 as \( a \) increases. The divisor functions with \( a = 0, 1 \) fluctuate much more. A measure of their fluctuations is given by comparing their RMS with their mean. Using another result of Ramanujan [16], one can show that, for \( a = 0 \), the mean square amplitude exceeds the mean by a factor which increases as \( \sqrt{\log t} \). MN Huxley (2012, private
communication) showed that \( \langle \sigma_0^2(t) \rangle = ((\pi^2/6)t^2) \zeta(3)/\zeta(4) \), from which it follows that the ratio between the variance of the fluctuation and the mean (smooth) term approaches a constant. These results imply that the reminder terms \( R_o(t) \) are fluctuating functions of \( t \) with mean which approaches zero as \( t \) grows. For the relevant cases to this study, \( R_1(t) = O(t^{1/2}) \).

The computation of the restricted divisor functions follows the same lines and to avoid repetitions only the results will be quoted,

\[
\langle \sigma_0(t; N) \rangle = \begin{cases} 
\log t + 2\gamma - \sum_{r=1}^{\lfloor t/N \rfloor} \frac{1}{r} & \text{for } t \leq N^2, \\
\log N + \gamma & \text{for } t > N^2.
\end{cases}
\]  

(3.15)

Similarly,

\[
\langle \sigma_1(t; N) \rangle = \begin{cases} 
\frac{\pi^2}{6} - \sum_{r=1}^{\lfloor t/N \rfloor} \frac{1}{r^2} & \text{for } t \leq N^2, \\
N & \text{for } t > N^2.
\end{cases}
\]  

(3.16)

(b) The smooth form factor: leading-order term

The evaluation of the leading term of the smoothed form factor (2.1) makes use of the asymptotic expression for \( \langle \sigma_1(t; N) \rangle \).

Substituting in (2.1) and taking the limit \( t, N \rightarrow \infty, t/N \rightarrow \tau \), one finds that the leading term of \( K(t, N) \) depends on its arguments through \( \tau = t/N \) exclusively, and

\[
\langle K(\tau) \rangle = \begin{cases} 
\frac{\pi^2}{6}, & 0 < \tau \leq 1, \\
\tau \left( \frac{\pi^2}{6} - \sum_{r=1}^{\lfloor \tau \rfloor} \frac{1}{r^2} \right), & 1 < \tau.
\end{cases}
\]  

(3.17)

Writing \( \tau = M + \xi \) with \( M = \lfloor \tau \rfloor, \xi \in (0,1) \), we have,

\[
K(\tau) = \langle M + \xi \rangle \left( \zeta(2) - \sum_{r=1}^{M} \frac{1}{r^2} \right) = 1 + a_M + b_M \xi,
\]  

(3.18)

with \( a_0 = -1, b_0 = \pi^2/6 \). For large \( M \), and to the leading order, \( a_M = -1/(2(M+1)), b_M = 1/(M+1) \). Hence, for large \( \tau \),

\[
K(\tau) \approx 1 + \frac{(\tau - \lfloor \tau \rfloor) - 1/2}{\lfloor \tau \rfloor} \rightarrow 1.
\]  

(3.19)

Figure 3 summarizes graphically the results above, and illustrates the accuracy by which the leading-order term reproduces the numerically smoothed \( K(\tau) \).

(c) The smooth form factor: next to leading-order term

We shall examine in detail the next to leading-order corrections for the smooth form factor in the interval \( 0 < \tau \leq 1 \). They arise from two sources: the lower order terms in \( \langle \sigma_1(t; N) \rangle \) and \( C(t, N) \) in (2.1). The former are all of order \( 1/N \) or less, and the main contribution is the constant \(-1/2N\).

The largest term in \( C(t, N) \) (2.1) is because of the smooth value of \( \langle \sigma_0(t) \rangle^2 \). It can be computed by making use of the following asymptotic result due to Ramanujan [16] and Wilson [17]:

\[
\sum_{n=1}^{t} \sigma_0^2(n) = t \left[ A(\log t)^3 + B(\log t)^2 + C(\log t) + D \right] + O(t^{1/2+\epsilon}),
\]  

(3.20)

and the explicit expressions for \( A \) and \( B \) are given by \( A = 1/\pi^2 \) and \( B = (12\gamma - 3)/\pi^2 - (36/\pi^4)\zeta'(2) \). Hence, to leading order, \( \langle \sigma_0^2(t) \rangle = (1/\pi^2)(\log t)^3 \), and the corresponding correction
Figure 3. The numerically smoothed form factor $K(\tau)$ for $N = 500$ (oscillatory grey line) is compared with the theoretical expression (3.17) (dark line).

to the smooth form factor is obtained by division by $N$. The computation of $\langle \sigma_0^2(t, N) \rangle$ for $t > N$ requires much more delicate considerations which are beyond the power of the smoothing method proposed above.

4. The number variance

Previous studies [2] have shown that, in the limit of large $N$, the expected number of eigenvalues in an arc $[\alpha, \beta]$ on the unit circle with $\alpha, \beta$ being irrational fractions of $2\pi$ and independent over $\mathbb{Q}$ is $\mathbb{E}[N(\beta - \alpha)] = (\beta - \alpha/2\pi)N + o(\log N)$. In other words, the mean distribution of the eigenvalues on the unit circle is uniform. The variance of the $N(\beta - \alpha)$ was also computed in [2] and with the same restrictions on $\alpha, \beta$,

$$\text{Var}(N(\beta - \alpha)) = \frac{1}{6} \log L + o(\log L), \quad L = \frac{\beta - \alpha}{2\pi} N. \quad (4.1)$$

The explicit expression for the smooth form factor derived above allows one to compute the smooth number variance and provide an exact formula in the limit $N \to \infty$ and $\beta - \alpha \to 0$ while $((\beta - \alpha)/2\pi)N \to L$. This is the limit of interest in most studies of spectral statistics of random matrix ensembles. The computation of the mean number variance in terms of the form factor makes use of the definition

$$\langle \text{Var} N(L) \rangle = \frac{1}{2\pi} \int_0^{2\pi} d\theta \left( \int_{L/2}^{L/2} d\xi \left( \rho \left( \theta + \frac{2\pi N}{2\pi} \xi \right) - \frac{N}{2\pi} \right) \right)^2, \quad (4.2)$$

where the spectral density on the unit circle is given by

$$\rho(\phi) = \sum_{j=1}^N \delta(\phi - \phi_j). \quad (4.3)$$
An elementary computation using the definition (1.7) gives
\[
\langle \text{Var}(N(L)) \rangle = \frac{2}{\pi^2} \frac{1}{N} \sum_{t=1}^{\infty} \sin^2 \frac{\pi L(t/N)}{(t/N)^2} K(t,N) \rightarrow \frac{2}{\pi^2} \int_0^\infty \frac{d\tau}{\tau^2} \frac{\sin^2 \pi L \tau}{\tau^2} K(\tau). \tag{4.4}
\]
Substituting (3.18) above provides an explicit expression for the mean number variance
\[
\langle \text{Var}(N(L)) \rangle = \frac{2}{\pi^2} \left( \frac{\pi^2}{6} \int_0^1 d\tau \frac{\sin^2 \pi L \tau}{\tau} + \int_1^\infty d\tau \frac{\sin^2 \pi L \tau}{\tau^2} \right) + \sum_{M=1}^{\infty} \int_0^1 d\xi (aM + bM\xi) \frac{\sin^2 \pi L(M + \xi)}{(M + \xi)^2}. \tag{4.5}
\]
The leading contribution comes from the upper line and it reads explicitly
\[
\frac{1}{6} (\log(2\pi L + \gamma) - \text{Ci}(2L\pi)) + \frac{1}{\pi^2} (\pi^2 L - 2\pi L \text{Si}(2L\pi) - \cos(2\pi L) + 1), \tag{4.6}
\]
where \(\text{Si}(x)\) and \(\text{Ci}(x)\) are the sine and cosine integrals, respectively. The large \(L\) asymptotic behaviour is dominated by
\[
\frac{1}{6} \log(2\pi L) + \left( \frac{1}{\pi^2} + \frac{\gamma}{6} \right) + O\left( \frac{1}{L} \right), \tag{4.7}
\]
which is consistent with (4.1) if one recalls the fact that the limits taken here and in [2] are not quite the same. The leading terms in the expansion near \(L = 0\) are
\[
L - \left( 2 - \frac{\pi^2}{6} \right) L^2 + \left( \frac{2\pi^2}{9} - \frac{\pi^4}{36} \right) L^4 + O(L^5). \tag{4.8}
\]
The infinite sum in (4.5) can be evaluated term by term. The series converges rapidly, and most of its contribution comes from the term with \(M = 1\). Its large \(L\) asymptotic is \((\pi^2/6 - 1) \log 4 - 1 + O(1/L)\) and its expansion near \(L = 0\) is \((\pi L/2)^2(\pi^2 - 10)\). Thus, the large and small \(L\) asymptotics of the number variance are given by the main terms as quoted in (4.7) and (4.8), respectively.

**Acknowledgements.** It is a pleasure to thank Prof. Martin Huxley, Prof. Ofer Zeitouni and Dr Igor Wigman for instructive discussions and comments.

**Funding statement.** This work was supported by the Minerva Center for Nonlinear Physics, the Einstein (Minerva) Center at the Weizmann Institute and the Wales Institute of Mathematical and Computational Sciences (WIMCS). Support from the Israel Science Foundation grant no. ISF-861/11 (F.I.R.S.T.) is acknowledged.

**References**


