Asymptotic prime partitions of integers

Johann Bartel,[1,*] R. K. Bhaduri,[2,†] Matthias Brack,[3,‡] and M. V. N. Murthy[4,§]

[1]*Institut Pluridisciplinaire Hubert Curien, Physique Théorique, Université de Strasbourg, F-67037 Strasbourg, France*
[2]*Department of Physics and Astronomy, McMaster University, Hamilton, Canada L8S4M1*
[3]*Institute of Theoretical Physics, University of Regensburg, D-93040 Regensburg, Germany*
[4]*The Institute of Mathematical Sciences, Chennai 600 113, India*

In this paper, we discuss $\mathcal{P}(n)$, the number of ways a given integer n may be written as a sum of primes. In particular, an asymptotic form $\mathcal{P}_{as}(n)$ valid for $n \to \infty$ is obtained analytically using standard techniques of quantum statistical mechanics. First, the bosonic partition function of primes, or the generating function of unrestricted prime partitions in number theory, is constructed. Next, the density of states is obtained using the saddle-point method for Laplace inversion of the partition function in the limit of large n. This gives directly the asymptotic number of prime partitions $\mathcal{P}_{as}(n)$. The leading term in the asymptotic expression grows exponentially as $\sqrt{n/\ln(n)}$ and agrees with previous estimates. We calculate the next-to-leading-order term in the exponent, proportional to $\ln[\ln(n)]/\ln(n)$, and we show that an earlier result in the literature for its coefficient is incorrect. Furthermore, we also calculate the next higher-order correction, proportional to $1/\ln(n)$ and given in Eq. (43), which so far has not been available in the literature. Finally, we compare our analytical results with the exact numerical values of $\mathcal{P}(n)$ up to $n \sim 8 \times 10^6$. For the highest values, the remaining error between the exact $\mathcal{P}(n)$ and our $\mathcal{P}_{as}(n)$ is only about half of that obtained with the leading-order approximation. But we also show that, unlike for other types of partitions, the asymptotic limit for the prime partitions is still quite far from being reached even for $n \sim 10^7$.

I. INTRODUCTION

Consider N identical ideal bosons that occupy a quantum system with equispaced single-particle energy levels at integer-valued n, with the lowest level at $n = 0$. This is simply the one-dimensional harmonic-oscillator spectrum with the zero-point energy set to zero. In the ground state of this system, all the bosons occupy the lowest level at $n = 0$. When a large excitation energy is given to the system, there are many ways by which this energy can be distributed among the N bosons. In fact, this is precisely the number $\mathcal{P}(n)$ of ways in which an integer n can be partitioned into a sum of integers less than or equal to n. The asymptotic form of $\mathcal{P}(n)$ (corresponding to $N \to \infty$ particles) is precisely the Hardy-Ramanujan formula [1] for the number partitions. The generating function in number theory is intimately connected to the bosonic partition function of statistical mechanics. It is interesting to note that this was written down by Hardy and Ramanujan years before the Bose-Einstein distribution was discovered in physics. In an earlier publication by some of the present authors [2], the asymptotic quantum density of states $\rho(E)$ was shown to be the $\mathcal{P}(n = E)$ known from number theory. This was done by performing the inverse Laplace transformation of the partition function using the saddle-point method.

It is obvious that the same technique of statistical mechanics may be applied to obtain any partition of a positive integer n, thus in particular also for its partition into *primes* p, if we start with a system whose single-particle levels are simply the primes p. The total energy now is given as a sum of primes, and the corresponding density of states is given by the *number of prime partitions* $\mathcal{P}(n)$. For our calculations, we require to convert the sum over primes into a continuous integral. For this, we need the density of primes, which may be deduced from the well-known prime number theorem (see the Appendix). The leading-order (LO) analytical expression for $\mathcal{P}(n)$ in the asymptotic limit $n \to \infty$ is available in the literature [3,4]. Corrections to the LO asymptotic result have been derived by Vaughan [5] using the saddle-point method. While our LO result, multiplied by a preexponential factor, agrees with that given by Vaughan [5], our next-to-leading-order (NLO) term in the exponent has a different coefficient $(-\frac{1}{2})$ from that given by Vaughan $(+1)$, which we are convinced is incorrect. Furthermore, while only an error estimate was given in Ref. [5] for the remaining terms beyond the NLO correction, we give a precise analytical expression for the next higher-order correction in the exponent of $\mathcal{P}(n)$. Our asymptotic result, which is denoted by $\mathcal{P}_{as}(n)$ in Eq. (43), is compared numerically with the exactly computed $\mathcal{P}(n)$. Although our asymptotic form comes much closer to the true $\mathcal{P}(n)$ than that of Ref. [5] for large n, we find that all asymptotic expressions discussed here are still quite far from reaching the exact $\mathcal{P}(n)$, even for numbers as large as $n \sim 10^7$. The reason for this slow approach to the asymptotic form will be discussed after presenting the numerical results.

The plan of our paper is as follows. In Sec. II A, we present our tools of statistical mechanics for a system whose single-particle spectrum is given by distinct primes p and whose total energy E is distributed among N bosonic particles. In Sec. II B, an analytical asymptotic form of the canonical bosonic partition function $Z(\beta)$ is derived and checked by an exact numerical computation of $Z(\beta)$. In Sec. II C, we obtain the many-body density of states $\rho(E)$ by Laplace inversion

*johann.bartel@iphc.cnrs.fr
†bhaduri@physics.mcmaster.ca
‡matthias.brack@ur.de
§murthy@imsc.res.in

of $Z(\beta)$ using the saddle-point approximation. Sections II D and II E are devoted to deriving our main result $\mathcal{P}_{as}(n)$ for the number of prime partitions, as given in Eq. (43), whereby the continuous energy variable E is identified with the discrete number n. In Sec. III, our asymptotic result, as well as other expressions, are compared numerically with the exact function $\mathcal{P}(n)$ for the prime partitions, which we have computed up to $n \sim 8 \times 10^6$. We conclude our paper with a short summary in Sec. IV. Some details about the density of primes, relevant to our analysis, are presented in the Appendix.

II. PARTITIONS INTO PRIMES

A. *N*-body system with a single-particle spectrum of primes

We set out to formulate our method for a fictitious bound system whose discrete, nondegenerate single-particle energies are given by the primes $p = 2, 3, 5, \dots$. We do not know of any physical system having this property (see also the last paragraph of our conclusions). But the use of quantum-statistical methods together with semiclassical "trace formulas" [6,7] for purely mathematical spectra can be very enlightening. A famous example is the spectrum of the nontrivial zeros of the Riemann ζ function. The quest for a Hamiltonian with this spectrum (see Ref. [8] for a recent attempt) has motivated the research of many physicists and mathematicians, and the hope is that it may even lead to a proof of the Riemann hypothesis. (For two nice reviews about this topic, see Refs. [9,10].) A trace formula for the prime spectrum is given in the Appendix.

Consider now a large number N of bosonic particles occupying these levels described by the prime spectrum. The total energy E of the system is given by

$$E = \sum_p n_p \, p. \tag{1}$$

Here and in the following, the sums $\sum_p$ run over all primes, and n_p are the occupancies of the levels, which may be zero or positive integers such that

$$\sum_p n_p = N. \tag{2}$$

In other words, the total energy E in (1) is given by any of its partitions into primes, restricted by the particle number conservation (2). The possible number of such partitions shall be denoted by $\mathcal{P}_N(E)$, where the subscript N keeps track of the number of particles. Although E is therefore necessarily integer, we treat it as a continuous variable as in statistical mechanics. It is important to realize that $\mathcal{P}_N(E)$ is identical to the many-body density of states $\rho_N(E)$ that is related to the canonical N-body partition function $Z_N(\beta)$ by

$$Z_N(\beta) = \int_0^\infty dE \, \rho_N(E) \exp(-\beta E), \tag{3}$$

where $\beta = 1/kT$ is the inverse temperature. Note that this expression, with $\rho_N(E) = \mathcal{P}_N(E)$, has the familiar form of the generating function of partitions used in number theory [1]. Since Eq. (3) formally is a Laplace transform, the density of states $\rho_N(E)$ can be obtained from the partition function by its

inverse Laplace transform

$$\rho_N(E) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} d\beta \, Z_N(\beta) \exp(\beta E). \tag{4}$$

We shall perform this Laplace inversion in the saddle-point approximation.

In terms of the single-particle spectrum, the canonical partition function $Z_N(\beta)$ may be written, after taking the limit $N \to \infty$, as

$$Z_\infty(\beta) = \prod_p \frac{1}{1 - e^{-\beta p}}, \tag{5}$$

where the product runs over all primes p. For simplicity, we shall henceforth omit the subscript "∞" from $Z(\beta)$ as well as from the functions $\rho(E)$ and $\mathcal{P}(E)$. Having taken the limit $N \to \infty$ implies that the partitions of the total energy now are unrestricted, admitting any number of summands allowed by the value of the energy (1). Taking the Laplace inverse of $Z(\beta)$ according to (4) thus leads to $\rho(E) = \mathcal{P}(n = E)$ in the limit $N \to \infty$.

In doing the transform Eq. (4), we define the function

$$S(\beta) = \beta E + \ln Z(\beta), \tag{6}$$

which formally defines the canonical entropy. We now evaluate the inverse Laplace transform in Eq. (4) using the method of steepest descent, or the saddle-point method. Hereby one is looking for a stationary point β_0 of the function $S(\beta)$ appearing in the exponent of the inverse Laplace integral, which corresponds to a saddle point in the complex β plane. This leads to the *saddle-point equation* (or saddle-point condition)

$$\frac{\partial S(\beta)}{\partial \beta} \bigg|_{\beta_0} = E + \frac{Z'(\beta_0)}{Z(\beta_0)} = 0. \tag{7}$$

If this equation has a solution β_0, which will be a function $\beta_0(E)$ of the energy, one evaluates the derivatives of $S(\beta)$ at β_0:

$$S^{(n)}(\beta_0) = \frac{\partial^n S(\beta)}{\partial \beta^n} \bigg|_{\beta_0}. \tag{8}$$

The result of the inverse Laplace transform is then given by

$$\rho(E) = \frac{e^{S(\beta_0)}}{\sqrt{2\pi S^{(2)}(\beta_0)}} [1 + \cdots], \tag{9}$$

where the dots indicate so-called cumulants involving higher derivatives of the entropy, which become more important for large β (see, e.g., Ref. [11]). Since we are interested here in the limit $\beta \to 0$ relevant for the asymptotics of large E, we shall neglect them.

B. Asymptotic partition function

Taking the logarithm of the partition function (5) gives a sum over all primes p, which we may also write as an integral,

$$\ln Z(\beta) = -\sum_p \ln(1 - e^{-\beta p}) = -\int_{x_0}^\infty dx \, g(x) \ln(1 - e^{-\beta x}), \tag{10}$$

where $g(x)$ is the exact density of primes given by the sum of δ-function distributions,

$$g(x) = \sum_p \delta(x - p), \qquad (11)$$

and x_0 is any real number smaller than the lowest prime: $x_0 < 2$. The sum of distributions (11) may be decomposed, as in many exact trace formulas for spectral distributions [7], into an average part $g_{av}(x)$, which is a continuous function of x, and an oscillating part representing itself as a sum of harmonics whose superposition results in the discreteness of the spectrum described by Eq. (11). (See the Appendix for the case of the prime spectrum.) The main object of the present study is the asymptotic behavior of $\mathcal{P}(n)$ for large n, for which the use of $g_{av}(x)$ in (10) is sufficient. For large x, the discreteness of $g(x)$ may be ignored, and the resulting $\mathcal{P}_{as}(n)$ is a smooth function of n as a continuous variable.

Therefore, we now replace the exact $g(x)$ in (10) by the average prime density $g_{av}(x)$. In doing so, we define the logarithm of the average partition function

$$\ln Z_{av}(\beta) = -\int_a^\infty dx\, g_{av}(x) \ln(1 - e^{-\beta x}), \qquad (12)$$

where the constant a must be chosen carefully as discussed in the following. As a specific choice of $g_{av}(x)$, we use the asymptotic prime density that is well-known from number theory (see the Appendix):

$$g_{av}(x) = 1/\ln(x). \qquad (13)$$

Since we are only interested in asymptotic results, it will be sufficient to look at the limit $\beta \to 0$, i.e., the high-temperature limit of the partition function.

The integrand (13) in (12) has a pole at $x = 1$, which becomes relevant when $a < 1$. We therefore define the following principal-value integral:

$$I(a,\beta) = -\lim_{\epsilon \to 0}\left[\int_a^{1-\epsilon} dx\, \frac{1}{\ln(x)} \ln(1 - e^{-\beta x}) \right.$$
$$\left. + \int_{1+\epsilon}^\infty dx\, \frac{1}{\ln(x)} \ln(1 - e^{-\beta x}) \right], \quad (a \neq 1), \quad (14)$$

which in the following is denoted by the symbol $\fint_a^\infty dx(\cdots)$, so that

$$\ln Z_{av}(a,\beta) = I(a,\beta) = -\fint_a^\infty dx\, \frac{1}{\ln(x)} \ln(1 - e^{-\beta x}). \quad (15)$$

This integral exists for any $a \neq 1$ and for finite β. We now make the change of variable $y = \beta x$ to obtain

$$I(a,\beta) = \frac{1}{\beta \ln(\beta)} \fint_{a\beta}^\infty dy\, \frac{1}{\left[1 - \frac{\ln(y)}{\ln(\beta)}\right]} \ln(1 - e^{-y}). \quad (16)$$

To make the next step more clear, we define

$$\tau = 1/\beta \qquad (17)$$

and rewrite (16) as

$$I(a,\tau) = -\frac{\tau}{\ln(\tau)} \fint_{a/\tau}^\infty dy\, \frac{1}{\left[1 + \frac{\ln(y)}{\ln(\tau)}\right]} \ln(1 - e^{-y}), \quad (18)$$

which we want to evaluate asymptotically in the high-temperature limit $\tau \to \infty$. We split it into two parts, writing

$$I(a,\tau) = -\frac{\tau}{\ln(\tau)}\left[\fint_{a/\tau}^\tau dy\, \frac{1}{\left[1 + \frac{\ln(y)}{\ln(\tau)}\right]} \ln(1 - e^{-y}) \right.$$
$$\left. + \int_\tau^\infty dy\, \frac{1}{\left[1 + \frac{\ln(y)}{\ln(\tau)}\right]} \ln(1 - e^{-y}) \right]. \quad (19)$$

If we fix a to an arbitrary value in the limits $1 < a < 2$ and take $\tau > 1$, we may approximate the first integral by the first term of the binomial expansion of its denominator and write

$$I(a,\tau) \simeq -\frac{\tau}{\ln(\tau)}\left[\int_{a/\tau}^\tau dy \left[1 - \frac{\ln(y)}{\ln(\tau)}\right] \ln(1 - e^{-y}) \right.$$
$$\left. + \int_\tau^\infty dy\, \frac{1}{\left[1 + \frac{\ln(y)}{\ln(\tau)}\right]} \ln(1 - e^{-y}) \right]. \quad (20)$$

In the limit $\tau \to \infty$, the second integral goes to zero and the first integral gives the asymptotic approximation

$$I_{as}(\tau) = -\frac{\tau}{\ln(\tau)} \int_0^\infty dy \left(1 - \frac{\ln(y)}{\ln(\tau)}\right) \ln(1 - e^{-y}), \quad (21)$$

which no longer depends on the value of a. Using (17) and (15), we obtain the following asymptotic form for the logarithm of the partition function, which we call $\ln Z_{as}(\beta)$ and which we can evaluate analytically:

$$\ln Z_{as}(\beta) = \frac{1}{\beta \ln(\beta)} \int_0^\infty dy \left(1 + \frac{\ln(y)}{\ln(\beta)}\right) \ln(1 - e^{-y})$$
$$= -\frac{f_1}{\beta \ln(\beta)} + \frac{f_2}{\beta \ln^2(\beta)}, \qquad (22)$$

where

$$f_1 = \frac{\pi^2}{6}, \quad f_2 = \frac{C\pi^2}{6} + \sum_k \frac{\ln(k)}{k^2} = 1.887\,03. \quad (23)$$

Here $C = 0.577\,216$ is the Euler constant, and the sum over k has been evaluated numerically (with $k_{max} \sim 10\,000$).

We want now to test the quality of the approximation (22), which should become accurate in the limit $\beta \to 0$. To that purpose we first integrate the principal-value integral $\ln Z_{av}(a,\beta)$ in (15). Here we choose $a = 0$ for definiteness; we emphasize that this choice is *a priori* independent of the fact that $1 < a < 2$ was used to derive the approximation (22). Then we compare it to the exact function (10) and to the approximation $\ln Z_{as}(\beta)$ in (22). The results are shown in Fig. 1.

We see that both approximations approach the exact values closely for small β, while $\ln Z_{av}(\beta)$ is better than $\ln Z_{as}(\beta)$ for the largest values of β. In Fig. 2, we see the same in a region of smaller values for β. The approximation $\ln Z_{as}(\beta)$ given in (22) crosses the exact curve near $\beta \sim 0.008$ and appears to stay below it for $\beta \to 0$. It reveals itself as an excellent asymptotic approximation to the exact $\ln Z(\beta)$ in the small-β limit.

Using the analytical form (22) of the partition function, the inverse Laplace transform (4) can now be done in the saddle-point approximation as outlined at the end of Sec. II A.
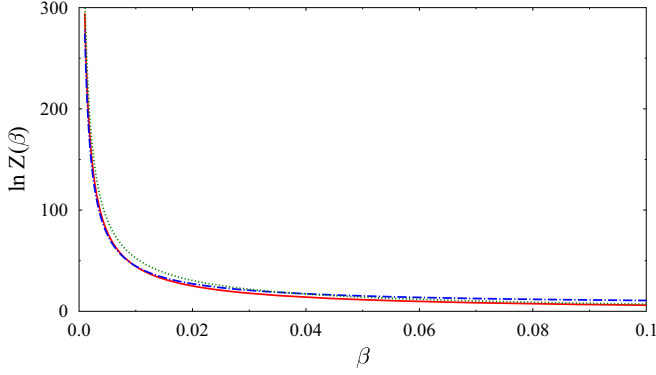
FIG. 1. Logarithm $\ln Z(\beta)$ of the partition function plotted vs β. Solid line (red): exact function (10). Dotted line (green): numerically integrated principal-value integral $\ln Z_{av}(a,\beta)$ in (15) with $a=0$. Dash-dotted (blue) line: asymptotic approximation $\ln Z_{as}(\beta)$ in (22).

C. Saddle-point approximation

To find the saddle point β_0, we isolate the most singular terms in $S(\beta)$ in the high-temperature limit. We first write the entropy, using (22) above, in the form

$$S(\beta) = \beta E - \frac{f_1}{\beta \ln(\beta)} + \frac{f_2}{\beta \ln^2(\beta)}. \tag{24}$$

Since the entropy above is given up to order $1/\ln^2(\beta)$, all further calculations will be done up to this order. To begin with, we need the following derivatives of the entropy:

$$S^{(1)}(\beta) = E + \frac{f_1}{\beta^2 \ln(\beta)} + \frac{f_1}{\beta^2 \ln^2(\beta)} - \frac{f_2}{\beta^2 \ln^2(\beta)} + \cdots, \tag{25}$$

$$S^{(2)}(\beta) = -\frac{2f_1}{\beta^3 \ln(\beta)} - \frac{3f_1}{\beta^3 \ln^2(\beta)} + \frac{2f_2}{\beta^3 \ln^2(\beta)} + \cdots. \tag{26}$$

The saddle-point equation (7) can therefore be written in the following form:

$$\beta E = -\frac{f_1}{\beta \ln(\beta)} + \frac{f_2}{\beta \ln^2(\beta)} - \frac{f_1}{\beta \ln^2(\beta)} + \cdots. \tag{27}$$

This is a transcendental equation whose solution β_0 can only be obtained iteratively as outlined in Sec. III D below. However, we may use the above equation directly to write $S(\beta)$ in terms
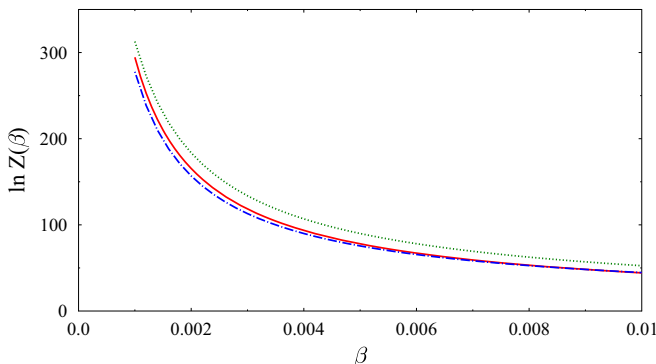

FIG. 2. Same as Fig. 1, but shown in the limit of small β.

of the as yet undetermined β_0 as

$$S(\beta_0) = 2\beta_0 E + \frac{f_1}{\beta_0 \ln^2(\beta_0)} + \cdots, \tag{28}$$

and similarly,

$$S^{(2)}(\beta_0) = \frac{1}{\beta_0^2}\left[2\beta_0 E - \frac{f_1}{\beta_0 \ln^2(\beta_0)} + \cdots\right]. \tag{29}$$

With Eq. (9) we obtain the asymptotic density of states as

$$\rho(E) = \frac{\exp\left(2\beta_0 E + \frac{f_1}{\beta_0 \ln^2(\beta_0)} + \cdots\right)}{\sqrt{(2\pi/\beta_0^2)\left[2\beta_0 E - \frac{f_1}{\beta_0 \ln^2(\beta_0)} + \cdots\right]}}. \tag{30}$$

This is the same as the asymptotic expression for the number of prime partitions $\mathcal{P}(n) = \rho(E = n)$. Any further analysis requires the explicit solution $\beta_0(E)$, which we derive next.

D. Saddle-point solution to leading order

To leading order $O[1/\ln(\beta_0)]$, the saddle-point equation (27) reads

$$\beta E = -\frac{f_1}{\beta \ln \beta}. \tag{31}$$

We now solve this equation iteratively. Let $\tau = 1/\beta$,

$$\frac{f_1}{E} = \frac{\ln(\tau)}{\tau^2}. \tag{32}$$

We start by assuming the solution to be of the form

$$\tau = a_1 E^{a_2}[\ln(E)]^{a_3}, \tag{33}$$

where a_1, a_2, and a_3 are constants to be determined using Eq. (31). Upon substitution, assuming large E, we get

$$\frac{f_1}{E} = \frac{1}{a_1^2 E^{2a_2}(\ln E)^{2a_3}}[\ln a_1 + a_2 \ln(E) + a_3 \ln \ln(E)]$$

$$\approx \frac{a_2}{a_1^2 E^{2a_2}(\ln E)^{2a_3-1}}. \tag{34}$$

First we determine the leading term, comparing powers, to find the solutions

$$a_3 = \frac{1}{2}, \quad a_2 = \frac{1}{2}, \quad a_1^2 = \frac{a_2}{f_1} = \frac{3}{\pi^2}. \tag{35}$$

Thus we have the leading solution given by

$$\tau = \frac{1}{\beta_0} = \sqrt{\frac{3}{\pi^2} E \ln(E)}. \tag{36}$$

To leading order, therefore, we have the following result for the density, or equivalently for unrestricted prime partitions:

$$\rho(E) = \frac{e^{S(\beta_0)}}{\sqrt{2\pi S''(\beta_0)}} = \frac{e^{2\pi\sqrt{E/[3\ln(E)]}}}{\sqrt{4E^{3/2}[3\ln(E)]^{1/2}}}. \tag{37}$$

Apart from the prefactor, it is well known [3,4] that $\ln[\rho(E)] \approx 2\pi\sqrt{E/(3\ln E)}$. In the paper by Vaughan [5], the prefactor has also been given by calculating $\sqrt{2\pi S^{(2)}(\beta)}$, which agrees with the calculation given here.

Next we consider corrections to the asymptotic result given in Eq. (37).

E. Higher-order corrections

The results of the previous subsection may be further improved by including additional terms that were neglected so far. This is done by assuming the saddle-point solution to be of the form

$$\beta_0 = \pi \sqrt{\frac{1}{3E \ln(E)}} \left[1 + a\, \frac{\ln[\ln(E)]}{\ln(E)} + b\, \frac{1}{\ln(E)} \cdots \right], \quad (38)$$

where a, b are arbitrary coefficients to be determined using the equation above up to order $1/\ln(E)$. The form of the solution is suggested by the transcendental equation (34) itself. Since the left-hand side of (34) is a monomial in E, the only way this can be satisfied is to have additional corrections to cancel the nonleading terms. The saddle-point condition to go beyond the leading order is given in Eq. (27), which is rewritten more conveniently as

$$E = -\frac{f_1}{\beta_0^2 \ln \beta_0} \left[1 + \frac{1 - f_2/f_1}{\ln \beta_0} + O(1/\ln^2 \beta_0) \right]. \quad (39)$$

We expand the unknowns on the right-hand side to the desired order $1/\ln(E)$ in the limit of large E,

$$\beta_0^2 = \frac{\pi^2}{3E \ln(E)} \left[1 + 2a\, \frac{\ln[\ln(E)]}{\ln(E)} \right.$$
$$\left. + 2b\, \frac{1}{\ln(E)} + O\{1/\ln^2(E)\} \right],$$

$$\ln(\beta_0) = -\frac{1}{2} \ln(E) \left[1 + \frac{\ln[\ln(E)]}{\ln(E)} - \ln\left(\frac{\pi^2}{3}\right) \frac{1}{\ln(E)} \right.$$
$$\left. - 2a\, \frac{\ln[\ln(E)]}{\ln^2(E)} - 2b\, \frac{1}{\ln^2(E)} \right],$$

and

$$\beta_0^2 \ln(\beta_0) = -\frac{\pi^2}{6E} \left[1 + (2a+1) \frac{\ln[\ln(E)]}{\ln(E)} \right.$$
$$\left. + \frac{2b - \ln(\pi^2/3)}{\ln(E)} + O\{1/\ln^2(E)\} \right].$$

Substituting these in Eq. (39), we determine the coefficients a, b as

$$a = -\frac{1}{2}, \quad b = \ln(\pi/\sqrt{3}) + \left(\frac{f_2}{f_1} - 1\right),$$

and therefore

$$\beta_0 = \pi \sqrt{\frac{1}{3E \ln(E)}} \left[1 - \frac{1}{2} \frac{\ln[\ln(E)]}{\ln(E)} \right.$$
$$\left. + \frac{\ln(\pi/\sqrt{3}) + (f_2/f_1 - 1)}{\ln(E)} \cdots \right]. \quad (40)$$

The density of prime partitions is then obtained by substituting the above solution into

$$\rho(E) = \frac{\exp\left[2\beta_0 E \left(1 + \frac{(\beta_0 E)^2}{2 f_1 E}\right) + \cdots \right]}{\sqrt{2\pi \left(2\beta_0 E/\beta_0^2\right)\left[1 - \frac{(\beta_0 E)^2}{2 f_1 E} + \cdots\right]}}, \quad (41)$$

where we have kept the NLO term in the density consistent with the order to which the solution has been obtained. Substituting $\beta_0 E$ from Eq. (40), we finally obtain

$$\rho(E)$$
$$= \frac{\exp\left\{ 2\pi \sqrt{\frac{E}{3\ln(E)}} \left[1 - \frac{1}{2} \frac{\ln[\ln(E)]}{\ln(E)} + \frac{[f_2/f_1 + \ln(\pi/\sqrt{3})]}{\ln(E)} + \cdots \right] \right\}}{\sqrt{\{4[3\ln(E)]^{1/2} E^{3/2} + \cdots\}}}. \quad (42)$$

Identifying $\rho(E)$ with $\mathcal{P}(n = E)$, the above equation gives the asymptotic prime partitions of an integer n. The first correction to the exponent given above, proportional to $\ln[\ln(E)]/\ln(E)$, is similar to that given by Vaughan [5] except that its coefficient here is $-\frac{1}{2}$ instead of $+1$. In the following section, we shall test the approximation obtained by ignoring all higher-order terms indicated by the dots above, thus defining

$$\mathcal{P}_{\mathrm{as}}(n) = \frac{1}{2[3\ln(n)]^{\frac{1}{4}} n^{\frac{3}{4}}} \exp\left\{ 2\pi \sqrt{\frac{n}{3\ln(n)}} \left[1 - \frac{1}{2} \frac{\ln[\ln(n)]}{\ln(n)} \right.\right.$$
$$\left.\left. + \frac{[f_2/f_1 + \ln(\pi/\sqrt{3})]}{\ln(n)} \right] \right\}. \quad (43)$$

This is the main result of our paper. A few comments might be in order here before we compare our result with exact numerical values.

(i) The leading term in the exponent, namely $2\pi\sqrt{\frac{n}{3\ln(n)}}$, agrees with the previously known results [3–5].

(ii) The prefactor given by $\{2[3\ln(n)]^{\frac{1}{4}} n^{\frac{3}{4}}\}^{-1}$ agrees also with that given by Vaughan [5].

(iii) The first correction term to the exponential, given by $-\frac{1}{2} \frac{\ln[\ln(n)]}{\ln(n)}$, has also been calculated by Vaughan [5] but with a coefficient $+1$ instead of our coefficient $-\frac{1}{2}$, which we believe is its correct value.

(iv) While Vaughan [5] has only given an estimate of the remaining error beyond the first correction in the exponential, we have been able to determine the exact coefficient of the successive term $\propto 1/\ln(n)$ in the exponential. As far as we know, this term has not been given in the literature so far.

III. NUMERICAL STUDIES OF THE ASYMPTOTIC FUNCTION $\mathcal{P}_{\mathrm{as}}(n)$

A. Evaluation of the exact database for $\mathcal{P}(n)$

We evaluate the prime partition $\mathcal{P}(n)$ using a standard method. Given an integer n, find the distinct primes that divide n. The sum of distinct prime factors that decompose n is denoted by $\mathscr{S}(n)$ [12]. For example, $\mathscr{S}(4) = 2$ since $4 = 2 \times 2$ has only one distinct prime that divides it, $\mathscr{S}(6) = 5$ since $6 = 2 \times 3$, or $\mathscr{S}(52) = 15$ since $52 = 2 \times 2 \times 13$. (Note: if a prime factor occurs several times, it should only be counted once.) Once the sum of prime factors $\mathscr{S}(n)$ is generated in a table, the following recursion relation [13] is used to compute the prime partitions (without any restriction):

$$\mathcal{P}(n) = \frac{1}{n} \left[\mathscr{S}(n) + \sum_{k=1}^{n-1} \mathscr{S}(k)\mathcal{P}(n-k) \right], \quad (44)$$

which involves all prime partitions of integers less than n. This procedure is very time consuming for large n. We have been able to compute $\mathcal{P}(n)$ for n up to 8 654 775. But, as we shall see, even this large number is not sufficient to reach the asymptotics of $\mathcal{P}(n)$.

B. Numerical study of $\mathcal{P}_{\mathrm{as}}(n)$

Using the above-derived database for the exact $\mathcal{P}(n)$, we now test various approximations for their asymptotic behavior. Rather than calculating the exponentially growing full function $\mathcal{P}(n)$, we look at its logarithm. We compare numerically the logarithm of the exact $\mathcal{P}(n)$ with that of the following approximations:

(i) To lowest order (LO), we set the prefactor of the exponent in (43) to unity, ignoring its denominator, and just keep the leading exponential term

$$\mathcal{P}_0(n) = \exp\left\{2\pi\sqrt{\frac{n}{3\ln(n)}}\right\}, \tag{45}$$

an asymptotic result that has been known for a long time [3,4].

(ii) The next expression is that of Vaughan [5]:

$$\mathcal{P}_V(n) = \frac{1}{2[3\ln(n)]^{1/4}n^{3/4}}$$
$$\times \exp\left\{2\pi\sqrt{\frac{n}{3\ln(n)}}\left[1 + \frac{\ln[\ln(n)]}{\ln(n)}\right]\right\}. \tag{46}$$

We repeat the fact that the NLO correction term in the exponent here has a different coefficient $(+1)$ from the coefficient $(-\frac{1}{2})$ in our result (43).

(iii) The third approximation we investigate is our asymptotic result (43) derived in the previous section.

The numerical comparison of the above three expressions with the exact prime partitions is now discussed in several steps.

We first plot $\ln\mathcal{P}(n)$ versus n for the various approximations in Fig. 3. The solid (black) curve gives the exact values
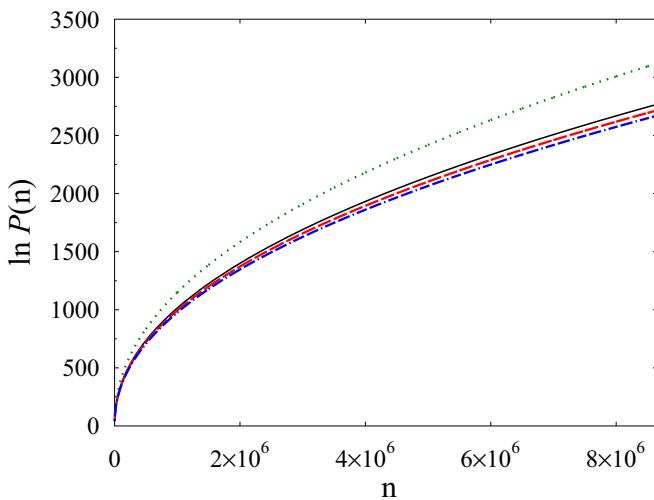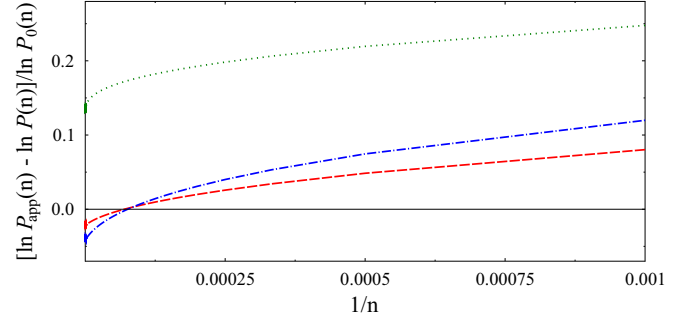

FIG. 4. Relative differences $[\ln\mathcal{P}_{\mathrm{app}}(n) - \ln\mathcal{P}(n)]/\ln\mathcal{P}_0(n)$ plotted vs $1/n$. Dashed (red): present (43), dash-dotted (blue): LO term (45), dotted (green): Vaughan (46).

$\ln\mathcal{P}(n)$. Our present approximation (43), shown by the dashed (red) line, comes closest to it, improving noticeably over the LO approximation $\ln\mathcal{P}_0(n)$ (45), shown by the dash-dotted (blue) line, in that the remaining error is reduced by about a factor of 2 for $n \gtrsim 10^6$. The expression (46) of Vaughan, shown by the dotted (green) curve, overshoots the exact values substantially and is actually much worse than the LO approximation—which does not appear to have been noticed so far.

From this figure, however, we cannot assess how the various approximations approach the correct asymptotics, since all curves increase monotonously. To this purpose, we next show in Fig. 4 the relative differences of the approximated logarithms, $[\ln\mathcal{P}_{\mathrm{app}}(n) - \ln\mathcal{P}(n)]/\ln\mathcal{P}_0(n)$, and plot them versus $1/n$ so that they should tend to zero for $n \to \infty$ (i.e., toward the left vertical axis in the figure). Shown are, with the same symbols (and colors) as above, our present approximation (43), the leading term (45), and that of Vaughan (46), all in the region $0 \leqslant 1/n \leqslant 0.001$ (i.e., $n > 1000$). Here we see that sign changes occur in the two lowest curves: at $n \sim 5800$ for (43), and at $n \sim 13\,000$ for (45). They hence approach zero from below, while the curve of Vaughan (46) stays far up on the positive side.

To see to which extent the two lower curves approach the asymptotic result 0, we now focus on the largest values of n available in our computation and further reduce the scale to $1/n \leqslant 10^{-5}$, shown in Fig. 5. Vaughan's curve cannot be seen

FIG. 3. Logarithms $\ln\mathcal{P}(n)$ in various approximations. Solid line (black): exact numerical values, dashed (red): $\ln\mathcal{P}_{\mathrm{as}}(n)$ (43), dash-dotted (blue): LO $\ln\mathcal{P}_0(n)$ (45), dotted (green): Vaughan (46).
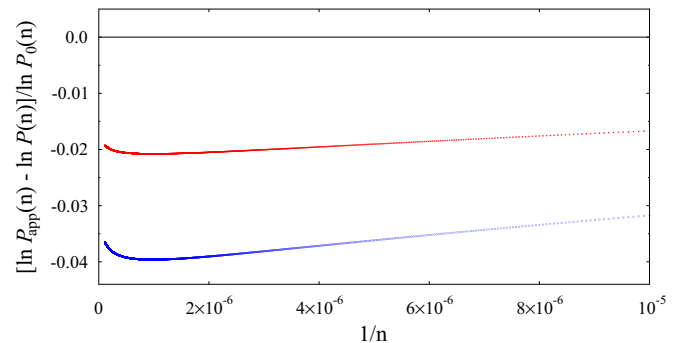

FIG. 5. Same as Fig. 4 in the lowest region $1/n \leqslant 10^{-5}$. Upper curve (red): our result (43); lower curve (blue): LO result (45). The Vaughan curve is not seen at this scale.

at this scale. Clearly, our result reduces the remaining error of the LO approximation by nearly a factor 2 in this region. We also notice that around $n \sim 10^6$, the curves have reached a minimum and then bend upward, so that they do appear to go asymptotically to zero. However, the differences are still finite even for our largest value $n_{max} = 8\,654\,775$. We must therefore question how far one has to go to reach the correct asymptotics $\mathcal{P}(n)$. Although the two curves in Fig. 5 clearly bend up toward zero for $1/n \to 0$, the slopes at n_{max} are such that there may still be a long way to go—too long perhaps to be covered by any numerical computation of the exact $\mathcal{P}(n)$.

We conclude that our result $\mathcal{P}_{as}(n)$ in (43) does appear to have the correct asymptotic behavior, but that even the included corrections beyond the LO are not sufficient to reach the exact partitions in our numerically accessible region.

IV. SUMMARY AND CONCLUSIONS

In this paper we have discussed $\mathcal{P}(n)$, the number of ways a given integer may be written as a sum of primes—a central theme in number theory. We have adopted methods used in quantum statistical mechanics where the central problem is the number of ways in which energy is distributed among particles occupying single-particle states. The partition function in statistical mechanics plays the role of the generating function of partitions. We have applied this method to the problem of prime partitions of an integer. The dominant integral is evaluated using the saddle-point method.

The main results of the paper may be summarized as follows:

(i) While the leading-order (LO) asymptotic form Eq. (45) has been known for some time, we derive non-leading-order (NLO) corrections to the exponent. There has not been much discussion in the literature on the prefactor to the exponential form (45), nor of the NLO corrections. An exception is Vaughan [5], who derived the same prefactor and also a NLO correction to the exponent in (45). We obtain a NLO contribution to the exponent of the same form, but with a different coefficient $(-\frac{1}{2})$ than that of Vaughan [5] $(+1)$. Our coefficient brings a considerable improvement of the asymptotics compared to that of Vaughan.

(ii) We also obtain a higher-order correction beyond NLO that, to the best of our knowledge, is not available in the literature.

(iii) We use a well-known algorithm to compute the exact prime partitions in order to compare analytical expressions for asymptotic prime partitions numerically. We have been able to do this up to more than 8 million in n. To our knowledge, a numerical comparison of exact results with asymptotic expressions has not been done up to this range before now. This is presumably also the reason why it has not been noticed so far that Vaughan's asymptotic expression (46) actually performs far worse than the lowest-order (LO) approximation (45) for $n \gtrsim 10^6$ (see Figs. 3 and 4).

(iv) It has been known from earlier work (see, for example, Ref. [2]) that for partitions of integers into integers, the asymptotic expressions for $p(n)$ are reached very rapidly—for n of the order of 100 or more. This is so because, as shown by Radmacher [14], the exact expression for integer partitions may be written as a convergent series. The kth term in the

series is of order $\exp(\frac{\pi}{k}\sqrt{\frac{2n}{3}})$. The leading term with $k = 1$ gives the Hardy-Ramanujan result. The first correction to the exponent is at $k = 2$ and therefore

$$p(n) \approx C_1(n) \exp\left(\pi\sqrt{\frac{2n}{3}}\right)\left[1 + C_2(n) \exp\left(-\frac{\pi}{2}\sqrt{\frac{2n}{3}}\right)\right],$$
(47)

where C_1 and C_2 are n-dependent prefactors. The correction to the exponent falls off exponentially. However, as seen from Eq. (43), the correction in the case of primes falls off logarithmically, which explains why the asymptotic limit is reached much more slowly for prime partitions, as compared to that of $p(n)$.

(v) Although both the exact $\mathcal{P}(n)$ and the asymptotic form $\mathcal{P}_{as}(n)$ given in (43) are monotonously increasing, their difference is not monotonic. In fact, we show that $\mathcal{P}_{as}(n)$ crosses $\mathcal{P}(n)$ around $n \sim 5800$ and approaches it from below for $n \to \infty$ (within the limits of our data). The remaining error has a maximum absolute value around $n \sim 10^6$ (see Fig. 5), beyond which it clearly tends toward zero.

(vi) Our main conclusion is that our result $\mathcal{P}_{as}(n)$ given in (43) is a clear improvement over the LO expression $\mathcal{P}_0(n)$ in (45). It appears to have the correct asymptotic behavior for $n \to \infty$, but even the corrections beyond the LO are not sufficient to reach the exact $\mathcal{P}(n)$ in the numerically accessible region.

Concerning partitions of integers n into smaller integers or into squares of integers, there exist physical quantum Hamiltonians that lead to these partitions, namely the harmonic oscillator and the infinite square-well potential [2]. Regarding the spectrum of primes, there have been attempts to construct potentials whose eigenvalues are the primes. Unfortunately, these potentials keep changing upon inclusion of more primes, and they have a fractal-like character [15–17]. That the prime spectrum can be reproduced from the nontrivial zeros of the Riemann ζ function is shown in the trace formula given in Eq. (A7) of the Appendix and illustrated in Fig. 6. As
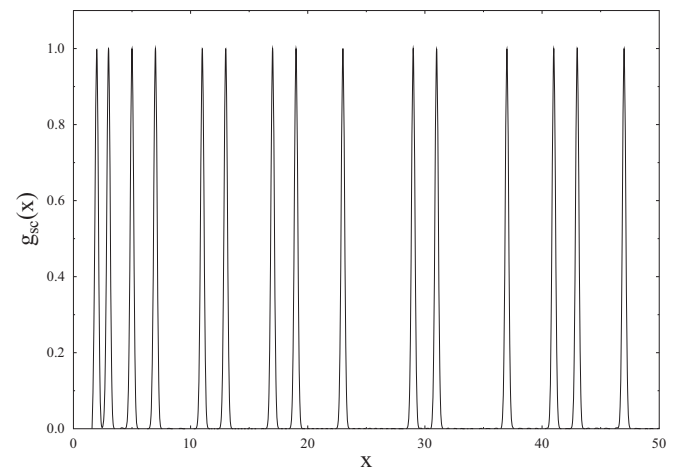

FIG. 6. Density of primes $g(x)$ obtained by the semiclassical expression $g_{sc}(x)$ in (A7), using the lowest 3000 Riemann zeros α and $m_{max} = 14$, coarse-grained with a Gaussian width $\gamma_{sh} = 0.1$.

we have mentioned already in the beginning of Sec. II A, semiclassical trace formulas can give insights into deep-lying mathematical connections. On the physical side, they provide the connection between a quantum spectrum and the periodic orbits of the corresponding classical Hamiltonian, and on the purely mathematical side, they connect spectral theory with symplectic geometry (in particular, with geodesics on Lagrangian manifolds) [6].

APPENDIX: SOME DETAILS ABOUT THE DENSITY OF PRIMES

In this appendix, we discuss two approximations to the density of primes $g(x)$ defined in (11), which is related to the function $\pi(x)$ that counts the number of primes $p \leqslant x$ by a differentiation:

$$g(x) = \frac{d\pi(x)}{dx}. \qquad (A1)$$

Both $\pi(x)$ and $g(x)$ have been the object of a lot of research in number theory. $\pi(x)$ is a staircase function whose average part is given by the asymptotic form

$$\pi(x) \sim \frac{x}{\ln(x)}, \qquad (A2)$$

which is a consequence of the prime number theorem. A more refined asymptotic form is (see, e.g., [18])

$$\pi(x) \sim \frac{x}{\ln(x)} + \frac{x}{[\ln(x)]^2} + \cdots + (n-1)! \frac{x}{[\ln(x)]^n}. \quad (A3)$$

Differentiating it yields the asymptotic expression for the density of primes,

$$g(x) \sim 1/\ln(x), \qquad (A4)$$

whereby all higher-order terms coming from (A3) have canceled successively. In Sec. III we have used the above asymptotic form for the average prime density $g_{av}(x)$.

An alternative expression for $\pi(x)$ may be derived from a function studied by Riemann in 1859, called $J(x)$, and further discussed by Edwards [18],

$$J(x) = \sum_{n=1}^{\infty} \frac{1}{n} \sum_p \Theta(x - p^n) \qquad (x > 0). \qquad (A5)$$

Here p runs over all primes and n over all integers, and $\Theta(x)$ is the standard step function: $\Theta(x) = 1$ for $x \geqslant 0$ and $\Theta(x) = 0$ for $x < 0$. In a seminal paper [19], Riemann gave an expression for $J(x)$ in terms of zeros of the ζ function. We use his expression and employ the Mobius inversion formula (cf. [18])

$$\pi(x) = \sum_{m=1}^{\infty} \frac{\mu(m)}{m} J(x^{1/m}), \qquad (A6)$$

where $\mu(m)$ is the Mobius function $[\mu(1) = 1]$. Taking the derivative according to (A1), we obtain the following expression for the density of primes (given also in [9]):

$$g_{sc}(x) = \frac{1}{x \ln x} \sum_{m=1}^{\infty} \frac{\mu(m)}{m} \left[x^{1/m} - \frac{1}{(x^{2/m} - 1)} - 2 x^{1/2m} \sum_{\alpha} \cos\left(\frac{\alpha}{m} \ln x\right) \right]. \qquad (A7)$$

Here $\alpha > 0$ are the nontrivial zeros of the Riemann ζ function along the positive half-line, and the validity of the Riemann hypothesis has been assumed. This expression, which does not appear to be widely known, has the form of a semiclassical "trace formula" [6,7] and we have therefore denoted it with the subscript "sc" for "semiclassical." Ideally, $g_{sc}(x)$ should yield the exact prime density $g(x)$ in (11) if the sum over α is not truncated, and if the Riemann hypothesis is true.

We have tested Eq. (A7) numerically in order to convince ourselves of its validity. For practical purposes, we have coarse-grained it, replacing the δ functions in (11) by normalized Gaussians with a width γ, and correspondingly coarse-grained Eq. (A7) as described in Sec. 5.5 of [7]. Figure 6 shows the results, obtained using the lowest 3000 Riemann zeros α. We see that the coarse-grained trace formula indeed reproduces the Gaussian-smoothed density of primes, replacing the δ functions in (11) by Gaussians centered exactly at the primes p. (Note that the sum over m can be truncated for any finite value of x; in the situation described here, $m_{max} = 14$ was sufficient.)

We note that the average part of $g_{sc}(x)$ in (A7) is not suitable for use in Eq. (12), because it has a pole structure that cannot be integrated easily. Numerically we found it to be very well approximated by the familiar asymptotic expression (A4).

[1] G. H. Hardy and S. Ramanujan, Proc. London Math. Soc. **2 (XVII)**, 75 (1918).

[2] M. N. Tran, M. V. N. Murthy, and R. K. Bhaduri, Ann. Phys. (N.Y.) **311**, 204 (2004).

[3] K. F. Roth and G. Szekeres, Q. J. Math Oxf. Ser. 2 **5**, 241 (1954).

[4] Y. Yang, Trans. Am. Math. Soc. **352**, 2581 (2000).

[5] R. C. Vaughan, Ramanujan J. **15**, 109 (2008).

[6] M. C. Gutzwiller, J. Math. Phys. **12**, 343 (1971); see also *Chaos in Classical and Quantum Mechanics* (Springer-Verlag, New York, 1990).

[7] M. Brack and R. K. Bhaduri, *Semiclassical Physics* (Addison-Wesley, Reading, MA, 1997).

[8] C. M. Bender, D. C. Brody, and M. P. Müller, Phys. Rev. Lett. **118**, 130201 (2017).

[9] M. Berry and J. P. Keating, SIAM Rev. **41**, 236 (1999).

[10] P. Leboeuf, A. G. Monastra, and O. Bohigas, Reg. Chaot. Dyn. **6**, 205 (2001).

[11] A. Jelovic, Phys. Rev. C **76**, 017301 (2007).

[12] See http://oeis.org/A008472.

[13] N. J. A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences* (Academic, San Diego, 1995), pp. 20 and 21.

[14] H. Radmacher, Ann. Math. **44**, 416 (1943); see also G. E. Andrews, *The Theory of Partitions* (Cambridge University Press, Cambridge, 1976), p. 69.

[15] G. Mussardo, arXiv:cond-mat/9712010.

[16] S. K. Sekatskii, arXiv:0709.0364 (math-phys).

[17] D. Schumayer, B. P. van Zyl, and D. A. W. Hutchinson, Phys. Rev. E **78**, 056215 (2008).

[18] H. M. Edwards, *Riemann's Zeta Function*, Vol. 58 Pure and Applied Mathematics, a series of monographs and textbooks (Academic Press INC, New York, 1974).

[19] B. Riemann, Monatsber. Akad. Berlin 671 (1859).