

## Random matrix theory and the Riemann zeros II: $n$ -point correlations

E.B. Bogomolny<sup>1</sup>, J.P. Keating  
Basic Research Institute in the  
Mathematical Sciences  
HP Laboratories Bristol  
HPL-BRIMS-96-13  
April, 1996

random matrix theory;  
Riemann zeta function;  
spectral statistics;  
quantum chaos

Montgomery (1973) has conjectured that the non-trivial zeros of the Riemann zeta-function are pairwise distributed like eigenvalues of matrices in the Gaussian Unitary Ensemble (GUE) of random matrix theory (RMT). In this respect, they provide an important model for the statistical properties of the energy levels of quantum systems whose classical limits are strongly chaotic. We generalise this connection by showing that for all  $n \geq 2$  the  $n$ -point correlation function of the zeros is equivalent to the corresponding GUE result in the appropriate asymptotic limit. Our approach is based on previous demonstrations for the particular cases  $n=2,3,4$  (Keating 1993, Bogomolny and Keating 1995). It relies on several new combinatorial techniques, first for evaluating the multiple prime sums involved using a Hardy-Littlewood prime-correlation conjecture, and second for expanding the GUE correlation-function determinant. This constitutes the first complete demonstration of RMT behaviour for all orders of correlation in a simple deterministic model.

Internal Accession Date Only

---

<sup>1</sup> Division de Physique Théorique, Institut de Physique Nucléaire, 91406 Orsay Cedex, France.

## Random matrix theory and the Riemann zeros II: $n$ -point correlations

E.B. Bogomolny<sup>†1</sup> & J.P. Keating<sup>‡</sup>

† Division de Physique Théorique  
Institut de Physique Nucléaire  
91406 Orsay Cedex  
France

‡ School of Mathematics  
University of Bristol  
University Walk  
Bristol BS8 1TW  
UK

and

BRIMS  
Hewlett-Packard Laboratories  
Filton Road, Stoke Gifford  
Bristol BS12 6QZ  
UK

### Abstract

Montgomery (1973) has conjectured that the non-trivial zeros of the Riemann zeta-function are pairwise distributed like the eigenvalues of matrices in the Gaussian Unitary Ensemble (GUE) of random matrix theory (RMT). In this respect, they provide an important model for the statistical properties of the energy levels of quantum systems whose classical limits are strongly chaotic. We generalise this connection by showing that for all  $n \geq 2$  the  $n$ -point correlation function of the zeros is equivalent to the corresponding GUE result in the appropriate asymptotic limit. Our approach is based on previous demonstrations for the particular cases  $n=2, 3, 4$  (Keating 1993, Bogomolny and Keating 1995). It relies on several new combinatorial techniques, first for evaluating the multiple prime sums involved using a Hardy-Littlewood prime-correlation conjecture, and second for expanding the GUE correlation-function determinant. This constitutes the first complete demonstration of RMT behaviour for all orders of correlation in a simple, deterministic model.

---

<sup>1</sup>On leave from the L.D Landau Institute for Theoretical Physics 142432 Chernogolovka, Russia

## 1. Introduction

The Riemann zeta-function may be defined by the Euler product over primes  $p$

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1} \quad (1)$$

when  $\text{Re } s > 1$ , and then by analytic continuation to the rest of the complex plane. It has infinitely many complex (*non-trivial*) zeros; that is, zeros at points  $s_n$  such that  $\text{Im } s_n \neq 0$ . According to the Riemann hypothesis, these all satisfy  $\text{Re } s_n = 1/2$  (Edwards 1974, Titchmarsh 1986). We shall assume this hypothesis to be correct and consider the statistical distribution of the imaginary parts  $E_n = \text{Im } s_n$ .

A convenient function with which to describe the distribution of zeros is the density (derivative of the counting function)

$$d(E) = \sum_n \delta(E - E_n). \quad (2)$$

This may be split into two parts:

$$d(E) = \bar{d}(E) + d_f(E), \quad (3)$$

where

$$\bar{d}(E) = \frac{1}{2\pi} \left\{ \text{Re} \frac{\Gamma'(\frac{1}{4} - i\frac{E}{2})}{\Gamma(\frac{1}{4} - i\frac{E}{2})} - \ln \pi \right\} \quad (4)$$

is the mean of  $d$ , and  $d_f$ , which represents the fluctuations about  $\bar{d}$ , is given by

$$d_f(E) = -\frac{1}{\pi} \sum_p \sum_{k=1}^{\infty} \frac{\ln p}{p^{k/2}} \cos(Ek \ln p) \quad (5)$$

(Berry 1986, Keating 1993). A statistical measure of the zero-distribution is then provided by the autocorrelation properties of  $d_f$ .

One of the main reasons for interest in the fluctuation statistics is that extensive numerical computations (Odlyzko 1989) support the conjecture of Montgomery (1973) that the irreducible 2-point correlation function

$$\tilde{R}_2(x_1, x_2; E) = \left\langle d_f(E + x_1) d_f(E + x_2) \right\rangle_E, \quad (6)$$

where

$$\langle f(E) \rangle_E = \frac{1}{E} \int_0^E f(E') dE', \quad (7)$$

is given asymptotically as  $E \rightarrow \infty$  by

$$\tilde{R}_2(x_1, x_2; E) \sim -\frac{\sin^2[(x_1 - x_2)\pi\bar{d}(E)]}{\pi^2(x_1 - x_2)^2}; \quad (8)$$

or, equivalently, that

$$\lim_{E \rightarrow \infty} \frac{\tilde{R}_2\left(\frac{y_1}{\bar{d}(E)}, \frac{y_2}{\bar{d}(E)}; E\right)}{\bar{d}^2(E)} = -\frac{\sin^2[(y_1 - y_2)\pi]}{\pi^2(y_1 - y_2)^2}. \quad (9)$$

The importance of this follows from the fact that the analogous 2-point correlation function defined for the eigenvalues of matrices in the Gaussian Unitary Ensemble (GUE) of random matrix theory (RMT) is given by the same formula, (9) (Mehta 1991). Hence Montgomery's conjecture suggests that the pairwise distribution of the Riemann zeros is identical to that of the eigenvalues of large random Hermitian matrices. This in turn suggests a fundamental link with quantum chaology, because there is now considerable numerical evidence (see, for example, Bohigas 1991) and a semiclassical theory (Hannay & Ozorio de Almeida 1984, Berry 1985) to support the conjecture (Bohigas *et al.* 1984a, Berry 1987) that generically the energy levels of any quantum system whose classical limit is strongly chaotic are also pairwise correlated in exactly the same way as the eigenvalues in RMT. Specifically, in systems where the classical trajectories do not possess time-reversal symmetry the 2-point energy-level correlation function, defined as in (6), should be the same as that of the GUE; that is, it should also be given by (9). The Riemann zeros thus provide an extremely valuable mathematical model for understanding eigenvalue distributions (Berry 1985, 1986, 1988; Keating 1991a, 1993; Argaman *et al.* 1993; Bogomolny & Keating 1995).

Assuming the Riemann hypothesis, Montgomery (1973) proved his conjecture for sufficiently large (but fixed) values of the unfolded correlation distance  $y=(y_1-y_2)$ . A heuristic demonstration for all values of  $y$  based on the Hardy-Littlewood twin-prime conjecture (Hardy & Littlewood 1923) was given in Keating (1991a, 1993). This heuristic approach was subsequently used to establish similar results for other  $L$ -functions by Bogomolny & Leboeuf (1994). Our intention here is to develop the method further and use it to calculate the irreducible  $n$ -point correlation function

$$\tilde{R}_n(x_1, x_2, \dots, x_n; E) = \left\langle \prod_{i=1}^n d_f(E + x_i) \right\rangle_E. \quad (10)$$

In particular, we shall show that the natural generalization of the Montgomery conjecture holds for all  $n \geq 2$ ; that is, in the asymptotic limit  $E \rightarrow \infty$ ,  $\tilde{R}_n$  coincides with the corresponding GUE result:

$$\tilde{R}_n(x_1, \dots, x_n; E) \sim \text{dets} \quad (11)$$

where the elements  $s_{ij}$  of the matrix  $s$  are given by

$$s_{ij} = s(x_i, x_j) = \frac{\sin[(x_i - x_j)\pi\bar{d}(E)]}{\pi(x_i - x_j)}(1 - \delta_{ij}), \quad (12)$$

$\delta_{ij}$  being the Kronecker symbol, or, equivalently, that  $\tilde{R}_n\left(\frac{y_1}{\bar{d}(E)}, \dots, \frac{y_n}{\bar{d}(E)}; E\right) / [\bar{d}(E)]^n$  tends to a well-defined limit when  $E \rightarrow \infty$ , as in (9).

As before, our analysis is based on the representation (5) of  $d_f$ . Substituting this into (10) gives an  $n$ -fold sum involving the primes, which we evaluate using conjecture D of Hardy & Littlewood (1923). The specific method of calculation is the same as that employed in Bogomolny & Keating 1995 - hereinafter referred to as BK - where  $R_3$  and  $R_4$  were computed explicitly and shown to coincide with the corresponding RMT formulae. The main components introduced there were the inclusion/exclusion expansion (section 2 of BK), and the evaluation of what were called type-II off-diagonal contributions to the prime sums (section 4.2(iv) of BK). Here these will be coupled with a series of new combinatorial devices to extend the approach to all higher orders of correlation. Specifically, we calculate the leading order asymptotics as  $E \rightarrow \infty$  of the general term in the expression for  $R_n$  (section 2) for all  $n \geq 2$ . In section 3, we then show that the result is exactly equivalent to the corresponding term in the expansion of the GUE determinant (11). This constitutes the first demonstration of RMT behaviour for all orders of correlation in a simple deterministic mathematical model (i.e. without any assumption of randomness, or system averaging).

The skewness  $\gamma_1$  and excess  $\gamma_2$ , statistics which contain information about correlations of order up to  $n=4$ , have been computed numerically for the Riemann zeros by Odlyzko (1987). His data show close agreement with the GUE forms, and so might appear to provide some support for our conclusions when  $n \leq 4$ . Unfortunately, in both cases the lower order correlations dominate, and so it is not clear to what extent his results can actually be used to obtain information about the irreducible correlations for  $n > 2$ . Statistics that include contributions from  $R_n$  with  $n > 4$  have, to the best of our knowledge, not previously been computed.

The fact that our method of calculation is based on the prime sum (5), that is, on the direct analogue of Gutzwiller's semiclassical trace formula (Gutzwiller 1971) for the quantum spectral density of classically chaotic systems, offers the hope that it may shed new light on the way in which RMT behaviour arises in the analogous higher-order energy-level statistics, in the same way that understanding the mechanism by which prime correlations determine the 2-point zero-statistics (Goldston & Montgomery 1984, Keating 1991a, 1993) has led to further insights into the corresponding semiclassical theory (Argaman *et al.* 1993). In this sense, the fact that we are able to demonstrate a direct link to RMT is crucial, because in all other exactly solvable quantum chaological models the periodic orbits display number-theoretical degeneracies which cause the 2-point statistics to be close to Poissonian in form (Keating 1991b, Bogomolny *et al.* 1994, Luo & Sarnak 1994).

The only statistics that contain information about correlations of order  $n > 2$  to have been studied numerically for real systems are, again, the skewness and excess (Bohigas *et al.* 1984b, Bohigas *et al.* 1985). Here too, numerical agreement with the RMT conjecture was found, but, for the reasons outlined above, it is not clear to what extent the data determine the irreducible behaviour beyond  $n=2$ .

As discussed in BK, the results that follow from this work concern the universal range of correlations, such as, for example, when the limit in (9) is taken with  $y_1$  and  $y_2$  fixed. We thus ignore the interesting nonuniversal features (Berry 1985, 1988) that emerge beyond the Montgomery conjecture if some or all of the unfolded correlation lengths  $y_j$  are taken to be asymptotically increasing functions of the mean density  $\bar{d}(E)$ , which itself increases like

$$\bar{d}(E) \sim \frac{1}{2\pi} \ln E \quad (13)$$

when  $E \rightarrow \infty$ , as may be seen by substituting Stirling's formula for the  $\Gamma$ -function in (4). Also, it should again be emphasised that the calculations presented here are essentially heuristic, in that they are based on the unproved conjecture D of Hardy & Littlewood (1923) and, in addition, involve several uncontrolled approximations. Our results are thus complementary to those of Montgomery (1973), Hejhal (1994), and Rudnick & Sarnak (1994), which are rigorous but restricted only to the range of correlation lengths over which the diagonal contributions (BK, and section 2.3 of the present paper) dominate the various multiple-prime-sums; that is, when the prime correlations, which form the main focus of our attention, can be neglected.

## 2. $\tilde{R}_n$ for the Riemann zeros

We seek to calculate the irreducible  $n$ -point correlation function  $\tilde{R}_n$ , defined by (10), for the Riemann zeros. From (5) we have that

$$d_f(E + x_i) = \frac{-1}{2\pi} \sum_p \sum_{k=1}^{\infty} \frac{\ln p}{\sqrt{p^k}} \left\{ \exp[i(E + x_i)k \ln p] + \exp[-i(E + x_i)k \ln p] \right\}. \quad (14)$$

As  $E \rightarrow \infty$ , the leading order asymptotics of the correlations is governed by the  $k=1$  term in the second sum (Keating 1991a, Keating 1993, BK), and so in this limit

$$\begin{aligned} \tilde{R}_n(x_1, \dots, x_n; E) \approx \frac{(-1)^n}{(2\pi)^n} \sum_{p_1, \dots, p_n} \left\langle \prod_{i=1}^n \frac{\ln p_i}{\sqrt{p_i}} \left\{ \exp[iE \ln p_i + ix_i \ln p_i] \right. \right. \\ \left. \left. + \exp[-iE \ln p_i - ix_i \ln p_i] \right\} \right\rangle_E. \end{aligned} \quad (15)$$

Expanding the product then gives

$$\begin{aligned} \tilde{R}_n(x_1, \dots, x_n; E) \approx & \frac{(-1)^n}{(2\pi)^n} \sum_{a+b=n} \sum_{\substack{p_1, \dots, p_a \\ q_1, \dots, q_b}} \prod_{i=1}^a \frac{\ln p_i}{\sqrt{p_i}} \exp[i\tilde{x}_i \ln p_i] \prod_{j=1}^b \frac{\ln q_j}{\sqrt{q_j}} \exp[-i\tilde{x}_j \ln q_j] \\ & \times \left\langle \exp \left[ iE \ln \left( \frac{p_1 p_2 \dots p_a}{q_1 q_2 \dots q_b} \right) \right] \right\rangle_E, \end{aligned} \quad (16)$$

where the sum includes all partitions of the primes  $p_1, \dots, p_n$  into two groups,  $p_1, \dots, p_a$  and  $q_1, \dots, q_b$  ( $a+b=n$ ), and the associated correlation variables  $x_1, \dots, x_n$  have, for temporary notational convenience, been split into two corresponding groups  $\tilde{x}_1, \dots, \tilde{x}_a$  and  $\tilde{x}_1, \dots, \tilde{x}_b$ .

Because of the average (denoted by the angular brackets), the main contribution to (16) comes from prime combinations that satisfy

$$p_1 p_2 \dots p_a \approx q_1 q_2 \dots q_b. \quad (17)$$

Consider first the case when none of the primes  $p_i$  on the left side of (17) is equal to a prime  $q_j$  on the right; i.e.  $p_i \neq q_j$   $1 \leq i \leq a$ ,  $1 \leq j \leq b$ . (We shall return to the case when cancellations do occur in section 2.3.) Then the most general way for the approximation (17) to hold is if

$$\begin{aligned} (p)_1 t_1 &= (q)_1 s_1 + k_1 \\ (p)_2 t_2 &= (q)_2 s_2 + k_2 \\ &\vdots \quad \quad \quad \vdots \\ (p)_r t_r &= (q)_r s_r + k_r \end{aligned} \quad (18)$$

where the two groups of primes  $p_1, \dots, p_a$  and  $q_1, \dots, q_b$  have each been split into  $r$  blocks (products of terms), the  $i$ th blocks,  $(p)_i$  and  $(q)_i$ , containing  $a_i$  and  $b_i$  primes respectively (hence  $\sum_{i=1}^r a_i = a$  and  $\sum_{i=1}^r b_i = b$ ); the integers  $s_i$ ,  $t_i$  and  $k_i$  satisfy the coprimality conditions  $(t_i, s_i) = (t_i, k_i) = (s_i, k_i) = ((p)_i, k_i) = ((q)_i, k_i) = 1$  for each  $i = 1, \dots, r$ ; and  $\prod_{i=1}^r t_i = \prod_{i=1}^r s_i$ .

We now employ the approach developed in BK for the special cases  $n=3$  and  $n=4$ . In particular, we generalize the method of evaluation of the type-I and type-II off-diagonal contributions given there in sections 4.2(ii) and 4.2(iv). To avoid unnecessary repetition, we refer readers to that paper for a full discussion and justification of the assumptions and approximations involved. Briefly, in the  $i$ th equation in (18)  $k_i$  is treated as small compared to the other terms. This allows us to approximate the corresponding contribution to (16) using

$$\exp \left[ iE \ln \left\{ \frac{(p)_i}{(q)_i} \right\} \right] \approx \exp \left[ iE \ln \left( \frac{s_i}{t_i} \right) + iE \frac{k_i}{(q)_i s_i} \right], \quad (19)$$

and so to replace a sum over one of the primes from either of the products  $(p)_i$  or  $(q)_i$  by a sum over  $k_i$ , with the appropriate probability weighting  $P$  defined by equations (7) and (11)

of BK. This sum over  $k_i$  may then be approximated using the inclusion/exclusion expansion developed in section 2 of BK to incorporate the coprimality conditions stated after (18). The final result follows after summing over all possible ways of splitting the sets of primes  $p_1, \dots, p_a$  and  $q_1, \dots, q_b$  into  $r$  subgroups as above, summing over  $r$ , and evaluating the sum over the  $(a, b)$ -partitions in (16).

In order to disentangle the various combinatorial divisions involved, we shall present the calculations of this section in three parts: the first will deal with the special case when  $r=1$  (called the type-I off-diagonal contributions in BK); the second will then extend the results obtained to include the additional permutations that arise when  $r>1$  (i.e. to calculate what were called the type-II contributions); and finally, in the third section, we shall evaluate the contributions from terms in (16) for which the condition  $p_i \neq q_j$  in (17) is not satisfied (these correspond to the diagonal and, more generally, the semidiagonal contributions discussed in BK).

## 2.1 Type-I off-diagonal contributions

As explained above, we begin by evaluating the contributions from combinations corresponding to  $r=1$  in (18). For these it follows that  $t_1=s_1=1$ , and that

$$p_1 p_2 \dots p_a = q_1 q_2 \dots q_b + k, \quad (20)$$

with  $(k, p_n) = 1 \forall n$  ( $1 \leq n \leq a$ ) and  $(k, q_m) = 1 \forall m$  ( $1 \leq m \leq b$ ). We treat  $k$  as small with respect to the other terms in (20) (in line with (17)) and so, using the approximation (19), a given  $(a, b)$ -partition in (16) satisfies

$$\begin{aligned} \tilde{R}_n^{(a,b)}(x_1, \dots, x_n; E) &\approx \frac{(-1)^n}{(2\pi)^n} \sum_{\substack{p_1, \dots, p_a \\ q_1, \dots, q_b}} \Delta(p_1 p_2 \dots p_a, q_1 q_2 \dots q_b) \times \\ &\prod_{i=1}^a \frac{\ln p_i}{\sqrt{p_i}} \exp[i\tilde{x}_i \ln p_i] \prod_{j=1}^b \frac{\ln q_j}{\sqrt{q_j}} \exp[-i\bar{x}_j \ln q_j] \times \\ &\sum_{k \neq 0} \exp\left(\frac{iEk}{q_1 q_2 \dots q_b}\right) P(p_1 p_2 \dots p_a, q_1 q_2 \dots q_b, k), \end{aligned} \quad (21)$$

where  $P$  is the probability density associated with the solutions of (20) (defined by equations (7) and (11) of BK, as noted above), and  $\Delta$  is a characteristic function which fixes the condition (17). The full type-I contribution to  $\tilde{R}_n$  is obtained by summing over all interchanges of the primes  $p_1, \dots, p_a$  and  $q_1, \dots, q_b$  between the two groups, and hence of the associated correlation variables  $\tilde{x}_1, \dots, \tilde{x}_a$  and  $\bar{x}_1, \dots, \bar{x}_b$ , and then over all positive integers  $a$  and  $b$  subject to  $a+b=n$ .

The sum over  $k$  in (21) may be evaluated using the inclusion/exclusion expansion formula, equation (20) of BK, to incorporate the coprimality conditions stated after (20)

above. In this case the inclusion/exclusion series involves all combinations of the primes  $p_1, \dots, p_a, q_1, \dots, q_b$ . Thus, approximating the prime sums by integrals using

$$\sum_p g(p) \approx \int_1^\infty \frac{g(x)}{\ln x} dx = \int_0^\infty g(\exp\{y\}) \frac{\exp\{y\}}{y} dy \quad (22)$$

and replacing  $\Delta$  by a Dirac  $\delta$ -function, the result is

$$\begin{aligned} \tilde{R}_n^{(a,b)}(x_1, \dots, x_n; E) \approx & \frac{(-1)^n}{(2\pi)^n} \int_0^\infty \dots \int_0^\infty \prod_{i=1}^a dw_i \exp[i\tilde{x}_i w_i] \prod_{j=1}^b dz_j \exp[-i\tilde{x}_j z_j] \times \\ & \delta\left(\sum_{i=1}^a w_i - \sum_{j=1}^b z_j\right) \left[ (\ln E - \sum_{j=1}^b z_j) \Theta\left(\sum_{j=1}^b z_j - \ln E\right) \right. \\ & - \sum_{\alpha_1=1}^a (\ln E + w_{\alpha_1} - \sum_{j=1}^b z_j) \Theta\left(\sum_{j=1}^b z_j - w_{\alpha_1} - \ln E\right) \\ & \left. - \sum_{\beta_1=1}^b (\ln E + z_{\beta_1} - \sum_{j=1}^b z_j) \Theta\left(\sum_{j=1}^b z_j - z_{\beta_1} - \ln E\right) + \dots \right], \end{aligned} \quad (23)$$

where  $\Theta$  is the unit-step function, and successive pairs of terms in the square brackets involve adding increasing numbers of the variables  $w_i, z_j$  to  $\ln E$  (equation (49) of BK represents an explicit example of this expansion for the case when  $n=4$ ). The general term in (23), corresponding to  $p_{\alpha_1} p_{\alpha_2} \dots p_{\alpha_s} q_{\beta_1} q_{\beta_2} \dots q_{\beta_t} |k$  in the inclusion/exclusion expansion, has sign  $(-1)^{s+t}$  and takes the form

$$\begin{aligned} \tilde{r}_n^{(a,b,s,t)}(x_1, \dots, x_n; E) = & \frac{(-1)^{n+s+t}}{(2\pi)^n} \int_0^\infty \dots \int_0^\infty \prod_{i=1}^a dw_i \exp[i\tilde{x}_i w_i] \prod_{j=1}^b dz_j \exp[-i\tilde{x}_j z_j] \\ & \times \delta\left(\sum_{i=1}^a w_i - \sum_{j=1}^b z_j\right) \left( \ln E + \sum_{\mu=1}^s w_{\alpha_\mu} + \sum_{\nu=1}^t z_{\beta_\nu} - \sum_{j=1}^b z_j \right) \\ & \times \Theta\left(\sum_{j=1}^b z_j - \sum_{\mu=1}^s w_{\alpha_\mu} - \sum_{\nu=1}^t z_{\beta_\nu} - \ln E\right). \end{aligned} \quad (24)$$

In order to evaluate the  $n$ -fold integral in (24) it is convenient to first rewrite it in the form

$$\begin{aligned} \tilde{r}_n^{(a,b,s,t)}(x_1, \dots, x_n; E) = & \frac{(-1)^{n+s+t}}{(2\pi)^n} \int_0^\infty \dots \int_0^\infty \prod_{i=1}^a dw_i \exp[i\tilde{x}_i w_i] \prod_{j=1}^b dz_j \exp[-i\tilde{x}_j z_j] \\ & \times \int_{-\infty}^\infty du (\ln E - u) \Theta(u - \ln E) \delta(\tilde{W} - Z - u) \delta(\tilde{Z} - W - u), \end{aligned} \quad (25)$$

where

$$W = \sum_{\mu=1}^s w_{\alpha_\mu}, \quad (26)$$

$$Z = \sum_{\nu=1}^t z_{\beta_\nu}, \quad (27)$$

$$\tilde{W} = \sum_{i=1}^a w_i - W, \quad (28)$$

and

$$\tilde{Z} = \sum_{j=1}^b z_j - Z. \quad (29)$$

We now re-label the terms in the  $\tilde{W}$ -sum (i.e., those not in the  $W$ -sum) as follows

$$\tilde{W} = \sum_{\mu'=1}^{a-s} w'_{\alpha_{\mu'}}; \quad (30)$$

and similarly for the  $\tilde{Z}$ -sum,

$$\tilde{Z} = \sum_{\nu'=1}^{b-t} z'_{\beta_{\nu'}}. \quad (31)$$

We also temporarily re-label the associated correlation variables by  $\tilde{x}_{\alpha_\mu}$  (multiplying  $w_{\alpha_\mu}$  in the product in (25)),  $\bar{x}_{\beta_\nu}$  (multiplying  $z_{\beta_\nu}$ ),  $\tilde{x}'_{\alpha_{\mu'}}$  (multiplying  $w'_{\alpha_{\mu'}}$ ), and  $\bar{x}'_{\beta_{\nu'}}$  (multiplying  $z'_{\beta_{\nu'}}$ ). The integrals in (25) are evaluated in terms of these variables in Appendix A. The result is

$$\begin{aligned} \tilde{r}_n^{(a,b,s,t)}(x_1, \dots, x_n; E) &= -\frac{(-1)^n}{(2\pi i)^n} \sum_{\mu'=1}^{a-s} \sum_{\nu'=1}^{b-t} \frac{\exp\left[i(\tilde{x}'_{\alpha_{\mu'}} - \bar{x}'_{\beta_{\nu'}}) \ln E\right]}{\left(\tilde{x}'_{\alpha_{\mu'}} - \bar{x}'_{\beta_{\nu'}}\right)^2} \\ &\quad \times \prod_{\substack{\mu''=1 \\ \mu'' \neq \mu'}}^{a-s} \frac{1}{\left(\tilde{x}'_{\alpha_{\mu''}} - \tilde{x}'_{\alpha_{\mu'}}\right)} \prod_{\substack{\nu''=1 \\ \nu'' \neq \nu'}}^{b-t} \frac{1}{\left(\bar{x}'_{\beta_{\nu''}} - \bar{x}'_{\beta_{\nu'}}\right)} \\ &\quad \times \prod_{\mu=1}^s \frac{1}{\left(\tilde{x}_{\alpha_\mu} - \bar{x}'_{\beta_{\nu'}}\right)} \prod_{\nu=1}^t \frac{1}{\left(\tilde{x}'_{\alpha_{\mu'}} - \bar{x}_{\beta_\nu}\right)}. \end{aligned} \quad (32)$$

What the somewhat cumbersome notation introduced above now makes clear is that  $\tilde{r}_n^{(a,b,s,t)}$  is a sum of terms, each with a pair of correlation variables in its phase. One of the pair is associated with a  $p$ -prime (and so is denoted by  $\tilde{x}$ ) that is not an element of the corresponding inclusion/exclusion set (and hence it is also primed), and the other is

associated with a  $q$ -prime (and so is denoted by  $\bar{x}$ ) that is likewise not in the inclusion/exclusion set (hence  $\bar{x}'$ ). The amplitude is then given by a product over the remaining  $n-2$  variables.

It follows that the total contribution of type-I terms to the correlation function is itself a sum of exponentials involving the differences of pairs of  $x$ -variables, the net amplitude associated with each such term being obtained from that in (32) by first summing the inclusion/exclusion series (i.e., summing over all values of  $s$  and  $t$  in the appropriate ranges:  $0 \leq s \leq a$ ,  $0 \leq t \leq b$ ), and then second, by summing over all of the  $(a, b)$ -divisions in (16). We now calculate this amplitude in the general case in which, retaining for the moment the notation introduced above, the phase is  $(\tilde{x}'_{\alpha_\mu} - \bar{x}'_{\beta_\nu}) \ln E$ . Denoting it by  $A_{\alpha_\mu, \beta_\nu}$ , we have that

$$A_{\alpha_\mu, \beta_\nu} = -\frac{(-1)^n}{(2\pi i)^n (\tilde{x}'_{\alpha_\mu} - \bar{x}'_{\beta_\nu})^2} \sum_{(a,b)} \sum_{s=0}^a \sum_{t=0}^b \prod_{\substack{\mu'=1 \\ \mu' \neq \mu}}^{a-s} \frac{1}{(\tilde{x}'_{\alpha_{\mu'}} - \bar{x}'_{\alpha_{\mu'}})} \prod_{\substack{v'=1 \\ v' \neq \nu}}^{b-t} \frac{1}{(\bar{x}'_{\beta_{v'}} - \bar{x}'_{\beta_{v'}})} \times \prod_{\mu=1}^s \frac{1}{(\tilde{x}_{\alpha_\mu} - \bar{x}'_{\beta_\nu})} \prod_{v=1}^t \frac{1}{(\tilde{x}'_{\alpha_\mu} - \bar{x}_{\beta_\nu})}. \quad (33)$$

Rewriting (33) in the form

$$A_{\alpha_\mu, \beta_\nu} = -\frac{(-1)^n}{(2\pi i)^n (\tilde{x}'_{\alpha_\mu} - \bar{x}'_{\beta_\nu})^2} \sum_{(a,b)} \sum_{s=0}^a \sum_{t=0}^b \prod_{\substack{i=1 \\ i \neq \alpha_\mu}}^a \frac{1}{(\tilde{x}'_{\alpha_\mu} - \tilde{x}_i)} \prod_{\substack{j=1 \\ j \neq \beta_\nu}}^b \frac{1}{(\bar{x}_j - \bar{x}'_{\beta_\nu})} \times \prod_{\mu=1}^s \frac{(\tilde{x}'_{\alpha_\mu} - \tilde{x}_{\alpha_\mu})}{(\tilde{x}_{\alpha_\mu} - \bar{x}'_{\beta_\nu})} \prod_{v=1}^t \frac{(\bar{x}_{\beta_\nu} - \bar{x}'_{\beta_\nu})}{(\tilde{x}'_{\alpha_\mu} - \bar{x}_{\beta_\nu})}, \quad (34)$$

the sums over  $s$  and  $t$  can be evaluated directly. In doing so we may now drop the distinction between primed and unprimed variables. The result is that

$$A_{\alpha, \beta} = -\frac{(-1)^n}{(2\pi i)^n (\tilde{x}_\alpha - \bar{x}_\beta)^2} \sum_{(a,b)} \prod_{\substack{i=1 \\ i \neq \alpha}}^a \frac{1}{(\tilde{x}_\alpha - \tilde{x}_i)} \prod_{\substack{j=1 \\ j \neq \beta}}^b \frac{1}{(\bar{x}_j - \bar{x}_\beta)} \times \prod_{\gamma=1}^a \left( 1 + \frac{\tilde{x}_\alpha - \tilde{x}_\gamma}{\tilde{x}_\gamma - \bar{x}_\beta} \right) \prod_{\delta=1}^b \left( 1 + \frac{\bar{x}_\delta - \bar{x}_\beta}{\tilde{x}_\alpha - \bar{x}_\delta} \right), \quad (35)$$

or, combining the products,

$$A_{\alpha, \beta} = -\frac{(-1)^n (\tilde{x}_\alpha - \bar{x}_\beta)^{n-2}}{(2\pi i)^n} \sum_{(a,b)} \prod_{\substack{p=1 \\ x_p \neq \tilde{x}_\alpha}}^n \frac{1}{(\tilde{x}_\alpha - x_p)} \prod_{\substack{q=1 \\ x_q \neq \bar{x}_\beta}}^n \frac{1}{(x_q - \bar{x}_\beta)}. \quad (36)$$

We are now in a position to drop the distinction between  $\tilde{x}$ -variables and  $\bar{x}$ -variables. Hence (36) may be written as

$$A_{\alpha,\beta} = -\frac{(-1)^n (x_\alpha - x_\beta)^{n-4}}{(2\pi i)^n} \sum_{(a,b)} \prod_{\substack{p=1 \\ p \neq \alpha \\ p \neq \beta}}^n \frac{1}{(x_\alpha - x_p)(x_p - x_\beta)}. \quad (37)$$

The summand in (37) is clearly the same for all  $(a, b)$ -divisions, the number of which is  $2^{n-2}$ , that is, the number of ways of splitting  $n-2$  variables  $(x_j, j \neq \alpha, \beta)$  into two groups ( $\tilde{x}_j$  or  $\bar{x}_j$ ). Consequently,

$$A_{\alpha,\beta} = -\frac{2^{n-2} (-1)^n (x_\alpha - x_\beta)^{n-4}}{(2\pi i)^n} \prod_{\substack{p=1 \\ p \neq \alpha \\ p \neq \beta}}^n \frac{1}{(x_\alpha - x_p)(x_p - x_\beta)}. \quad (38)$$

Alternatively, we may write this in the form

$$A_{\alpha,\beta} = -\frac{2^{n-2} (-1)^n}{(2\pi i)^n (x_\alpha - x_\beta)^2} \prod_{\substack{p=1 \\ p \neq \alpha \\ p \neq \beta}}^n \left( \frac{1}{x_\alpha - x_p} + \frac{1}{x_p - x_\beta} \right), \quad (39)$$

or, equivalently, by expanding out the product,

$$A_{\alpha,\beta} = -\frac{2^{n-2} (-1)^n}{(2\pi i)^n (x_\alpha - x_\beta)^2} \sum_{\{\gamma\} + \{\delta\} = n-2} \prod_{\{\gamma\}} \frac{1}{x_\alpha - x_\gamma} \prod_{\{\delta\}} \frac{1}{x_\delta - x_\beta}, \quad (40)$$

where the sum involves all partitions of the  $n-2$  non-phase variables into two groups,  $\{\gamma\}$  and  $\{\delta\}$ , with elements  $x_\gamma$  and  $x_\delta$  respectively.

Equation (40) is the main result of this section. It represents the amplitude of the type-I contribution to  $\tilde{R}_n$  that has as its phase factor  $\exp[i(x_\alpha x_\beta) \ln E]$ . The complete type-I contribution is then given by the sum over all values of  $\alpha$  and  $\beta$  in the range  $1 \leq \alpha \leq n, 1 \leq \beta \leq n$  such that  $\alpha \neq \beta$ . We note that when  $n=3$  the expression reduces correctly to equation (33) of BK, as may most easily be seen from the form (38) above. Similarly, when  $n=4$  it is identical to the corresponding formula obtained by adding equations (43) and (51) of BK.

## 2.2 Type-II off-diagonal contributions

We now evaluate the type-II off-diagonal contributions to (16), which come from subdivisions of the form (18) with  $r > 1$ . Consider first the  $i$ th equation in (18), together with the associated coprimality conditions stated thereafter. The sum (16) over the  $a_i$  primes  $p_j^{(i)}$

and the  $b_i$  primes  $q_l^{(i)}$  contained, respectively, in the products  $(p)_i$  and  $(q)_i$  may be written, using (19), in the form

$$\rho_{a_i, b_i, s_i, t_i}^{(i)} = \frac{(-1)^{n_i}}{(2\pi)^{n_i}} \sum_{p_j^{(i)} \in (p)_i} \sum_{q_l^{(i)} \in (q)_i} \Delta((p)_i, t_i, (q)_i, s_i) \prod_{j=1}^{a_i} \frac{\ln p_j^{(i)}}{\sqrt{p_j^{(i)}}} \exp[i\tilde{x}_j^{(i)} \ln p_j^{(i)}] \times \prod_{l=1}^{b_i} \frac{\ln q_l^{(i)}}{\sqrt{q_l^{(i)}}} \exp[-i\bar{x}_l^{(i)} \ln q_l^{(i)}] \sum_{k_i \neq 0} \exp\left(\frac{iEk_i}{(q)_i, s_i}\right) P((p)_i, t_i, (q)_i, s_i, k_i), \quad (41)$$

so that the the total type-II contribution to  $\tilde{R}_n$  is

$$\tilde{R}_n(x_1, \dots, x_n; E) \approx \sum_r \sum_{\text{perms}} \prod_{i=1}^r \rho_{a_i, b_i, s_i, t_i}^{(i)}, \quad (42)$$

the first sum running over the different possible values of  $r$ , and the second over all permutations of the terms involved in a given  $r$ -fold subdivision of the form (18), including all permissible integers  $s_i$  and  $t_i$ , and all values of  $a_i$  and  $b_i$  subject to  $a_i + b_i = n_i$ . Note that in (41) we have, for temporary notational convenience, again denoted the two sets of variables corresponding to  $p$ -primes and  $q$ -primes by  $\tilde{x}$  and  $\bar{x}$ .

As in the calculation of the type-I contributions, the sum over  $k_i$  in (41) may be evaluated in terms of an inclusion/exclusion expansion (now with respect to all of the primes  $p_j^{(i)}$ ,  $q_l^{(i)}$  and all of the prime factors of the integers  $t_i$  and  $s_i$ ), the prime-sums may be approximated by integrals using (22), and the characteristic function  $\Delta$  may be replaced by a Dirac  $\delta$ -function. The result, obtained by applying equation (20) of BK, is

$$\rho_{a_i, b_i, s_i, t_i}^{(i)} = \frac{(-1)^{n_i}}{(2\pi)^{n_i}} \frac{\exp\{iE \ln(s_i/t_i)\}}{s_i} \int_0^\infty \dots \int_0^\infty \prod_{j=1}^{a_i} dw_j \exp[i\tilde{x}_j^{(i)} w_j] \times \prod_{l=1}^{b_i} dz_l \exp[-i\bar{x}_l^{(i)} z_l] \delta\left(\ln t_i + \sum_{j=1}^{a_i} w_j - \ln s_i - \sum_{l=1}^{b_i} z_l\right) \times \left[ \left(\ln E - \ln s_i - \sum_{l=1}^{b_i} z_l\right) \Theta\left(\sum_{l=1}^{b_i} z_l + \ln s_i - \ln E\right) + \dots \right], \quad (43)$$

where the square brackets contain the inclusion/exclusion expansion, the first term of which is shown explicitly.

The integral (43) is similar in form to (23) and may be evaluated directly using the results of Appendix A. Specifically, it may be calculated by replacing  $\ln E$  by  $\ln E - \ln s_i$  in the  $z$ -integrals, and by  $\ln E - \ln t_i$  in the  $w$ -integrals (this is most easily seen by writing the integrand of (43) in the same form as in (25)). The result, after summing the inclusion/exclusion expansion with respect to the primes  $p_j^{(i)}$  and  $q_l^{(i)}$  (but not yet with respect to the prime factors of  $t_i$  and  $s_i$ ) as in the type-I off-diagonal contributions of the

previous section, is thus a series of terms of the form  $A_{\alpha_i, \beta_i}^{(i)} \exp\{i(x_{\alpha_i} - x_{\beta_i}) \ln E\}$  ( $1 \leq \alpha_i \leq a_i$ ,  $1 \leq \beta_i \leq b_i$ ,  $\alpha_i \neq \beta_i$ ), with  $A_{\alpha_i, \beta_i}^{(i)}$  proportional to  $A_{\alpha_i, \beta_i}$  given by (40) with  $n$  replaced by  $n_i$ :

$$A_{\alpha_i, \beta_i}^{(i)} = \frac{A_{\alpha_i, \beta_i}}{s_i} \left[ \exp\{iE \ln(s_i/t_i) + ix_{\beta_i} \ln s_i - ix_{\alpha_i} \ln t_i\} + \dots \right], \quad (44)$$

where the square brackets now contain the inclusion/exclusion expansion with respect to the prime factors of  $s_i$  and  $t_i$ , which we denote  $\{\sigma_m^{(i)}\}$  and  $\{\tau_m^{(i)}\}$  respectively. The terms in this expansion involve multiplying  $E$  in (41) by all possible combinations of  $\sigma_m^{(i)}$  and  $\tau_m^{(i)}$ , since these appear in the coprimality conditions on  $k_i$  (just as the corresponding expansion in the last section involved multiplying  $E$  by all combinations of the primes in the sums in (21)). The result is a sum, each term of which is associated with a particular combination. The general term corresponding to when both  $\sigma = \sigma_{m_1}^{(i)} \sigma_{m_2}^{(i)} \dots \sigma_{m_r}^{(i)}$  and  $\tau = \tau_{m_1}^{(i)} \tau_{m_2}^{(i)} \dots \tau_{m_r}^{(i)}$  divide  $k_i$  is obtained by replacing  $\ln E$  by  $\ln E + \ln \sigma + \ln \tau - \ln s_i$  in the  $z$ -integrals in (23), by  $\ln E + \ln \sigma + \ln \tau - \ln t_i$  in the  $w$ -integrals, and by including the appropriate sign, namely  $(-1)^{m_i + m_i}$ . Summing the series then gives

$$A_{\alpha_i, \beta_i}^{(i)} = \frac{A_{\alpha_i, \beta_i}}{s_i} \exp\{iE \ln(s_i/t_i) + ix_{\beta_i} \ln s_i - ix_{\alpha_i} \ln t_i\} \\ \times \prod_{\sigma_m^{(i)}} \left[ 1 - \exp\{i(x_{\alpha_i} - x_{\beta_i}) \ln \sigma_m^{(i)}\} \right] \prod_{\tau_m^{(i)}} \left[ 1 - \exp\{i(x_{\alpha_i} - x_{\beta_i}) \ln \tau_m^{(i)}\} \right]. \quad (45)$$

It thus follows that the  $r$ -sum in (42) involves a series of terms of the form

$$A_{\alpha_1, \beta_1, \dots, \alpha_r, \beta_r}^{(r)} \exp\{i(x_{\alpha_1} - x_{\beta_1} + x_{\alpha_2} - x_{\beta_2} + \dots + x_{\alpha_r} - x_{\beta_r}) \ln E\}, \quad (46)$$

where, using (40),

$$A_{\alpha_1, \beta_1, \dots, \alpha_r, \beta_r}^{(r)} = \frac{2^{n-2r} (-1)^{n+r}}{(2\pi i)^n} \prod_{i=1}^r \frac{1}{(x_{\alpha_i} - x_{\beta_i})^2} \\ \times \sum_{\text{perms}} f(x_1, \dots, x_n) g(x_{\alpha_1}, x_{\beta_1}, \dots, x_{\alpha_r}, x_{\beta_r}), \quad (47)$$

with

$$f(x_1, \dots, x_n) = \sum_{\substack{\{\gamma\}_1 + \{\delta\}_1 = n_1 - 2 \\ \{\gamma\}_2 + \{\delta\}_2 = n_2 - 2 \\ \vdots \\ \{\gamma\}_r + \{\delta\}_r = n_r - 2}} \prod_{i=1}^r \prod_{\{\gamma\}_i} \frac{1}{(x_{\alpha_i} - x_{\gamma})} \prod_{\{\delta\}_i} \frac{1}{(x_{\delta}^{(i)} - x_{\beta_i})}, \quad (48)$$

in which the sum has the same meaning as in (40) for the correlation variables  $x^{(i)}$  associated with the  $i$ th equation in (18) (i.e., it involves all partitions of the  $n_i-2$  variables  $x^{(i)} \neq x_{\alpha_i} \neq x_{\beta_i}$  into two groups  $\{\gamma\}_i$  and  $\{\delta\}_i$ , for each  $1 \leq i \leq r$ );

$$g(x_{\alpha_1}, x_{\beta_1}, \dots, x_{\alpha_r}, x_{\beta_r}) = \prod_{i=1}^r \frac{1}{s_i} \prod_{\sigma_m^{(i)}} \left[ \exp\{ix_{\beta_i} \ln \sigma_m^{(i)}\} - \exp\{ix_{\alpha_i} \ln \sigma_m^{(i)}\} \right] \times \prod_{\tau_m^{(i)}} \left[ \exp\{-ix_{\alpha_i} \ln \tau_m^{(i)}\} - \exp\{-ix_{\beta_i} \ln \tau_m^{(i)}\} \right]; \quad (49)$$

and where we have used the condition  $\prod_{i=1}^r s_i = \prod_{i=1}^r t_i$ .

The permutations to be included in (47) involve, first, all interchanges of the variables  $x_{\gamma}^{(i)}$  and  $x_{\delta}^{(i)}$  between the different ( $i \neq j$ ) sums in (48), giving rise to an amplitude factor associated with  $f$  of the form

$$\sum_{\{\gamma\}_1 + \{\delta\}_1 + \dots + \{\gamma\}_r + \{\delta\}_r = n-2r} \prod_{i=1}^r \prod_{\{\gamma\}_i} \prod_{\{\delta\}_i} \frac{1}{(x_{\alpha_i} - x_{\gamma}^{(i)})(x_{\delta}^{(i)} - x_{\beta_i})}, \quad (50)$$

where the sum is over all partitions of the  $n-2r$  correlation variables not appearing in the phase in (46) into  $2r$  groups  $\{\gamma\}_i$  and  $\{\delta\}_i$ . In addition, they also involve summing over all values of the  $2r$  integers  $s_i$  and  $t_i$  appearing in (18), subject to  $\prod_{i=1}^r s_i = \prod_{i=1}^r t_i$  and  $(s_i, t_i) = 1$  (for an explicit example, see equation (59) of BK, which represents the  $n=4$  case). Now it follows from the first of these conditions that, for any  $i$ , each prime factor of  $s_i$  must also be a factor of some  $t_j$ . The second condition then implies that  $i \neq j$  in this link. Hence the  $(s_i, t_i)$ -sum is equivalent to summing over all common factors for each such pairwise connection.

The combinatorial contribution from a given connection linking  $s_i$  and  $t_j$  ( $i \neq j$ ) may be calculated as follows. If  $(s_i, t_j) = m$ , the corresponding term in (49) is

$$g_{i,j}(m) = \frac{1}{m} \prod_{p|m} \left[ \exp\{ix_{\beta_i} \ln p\} - \exp\{ix_{\alpha_i} \ln p\} \right] \times \left[ \exp\{-ix_{\alpha_j} \ln p\} - \exp\{-ix_{\beta_j} \ln p\} \right]. \quad (51)$$

Summing over all values of the common factor  $m$  gives

$$g_{i,j} = \sum_{m=1}^{\infty} g_{i,j}(m) = \prod_p \left( 1 + \frac{1}{p-1} \left[ \exp\{ix_{\beta_i} \ln p\} - \exp\{ix_{\alpha_i} \ln p\} \right] \times \left[ \exp\{-ix_{\alpha_j} \ln p\} - \exp\{-ix_{\beta_j} \ln p\} \right] \right), \quad (52)$$

where the product includes all primes  $p$ . Hence  $\ln g_{i,j}$  is a sum over primes and so may be estimated using (22) (for further details see the appendix of BK). The result is that

$$g_{i,j} = \frac{(x_{\alpha_i} - x_{\alpha_j})(x_{\beta_i} - x_{\beta_j})}{(x_{\alpha_i} - x_{\beta_j})(x_{\beta_i} - x_{\alpha_j})}. \quad (53)$$

It remains to evaluate the net contribution from all different pairwise connections. Clearly from (49),  $g$  itself is a product of these, and so summing over all values of  $s_i$  and  $t_j$  obviously gives  $\prod_{i \neq j} g_{i,j} = \prod_{i > j} g_{i,j}^2$ . Finally, it might appear on first sight that interchanges of the phase variables, such as  $x_{\alpha_i} \leftrightarrow x_{\alpha_j}$  etc., should also be included in the permutation sum in (47), however these are in fact already taken care of by the  $(s_i, t_i)$ -sum just evaluated, as shown at the end of section 4.2(iv) in BK.

The main result of this section now follows from combining the expressions derived above. It is that the type-II off-diagonal terms lead to contributions of the form (46) for all  $1 \leq r \leq [n/2]$ , where [...] denotes the integer part and

$$A_{\alpha_1, \beta_1, \dots, \alpha_r, \beta_r}^{(r)} = \frac{2^{n-2r} (-1)^{n+r}}{(2\pi i)^n} \left[ \frac{\prod_{1 \leq j < i \leq r} (x_{\alpha_i} - x_{\alpha_j})(x_{\beta_i} - x_{\beta_j})}{\prod_{1 \leq i, j \leq r} (x_{\alpha_i} - x_{\beta_j})} \right]^2 \times \quad (54)$$

$$\sum_{\{\gamma\}_1 + \{\delta\}_1 + \dots + \{\gamma\}_r + \{\delta\}_r = n-2r} \prod_{i=1}^r \prod_{\{\gamma\}_i} \prod_{\{\delta\}_i} \frac{1}{(x_{\alpha_i} - x_{\gamma}^{(i)})(x_{\delta}^{(i)} - x_{\beta_i})}.$$

When  $r=1$  this expression reduces to (40), and in the case when  $n=4$  and  $r=2$  it coincides with the corresponding result, equation (61), of BK.

### 2.3 Diagonal and semidiagonal contributions

As already noted, the calculations of the previous two sections concern the particular case in which none of the primes  $p_i$  on the left side of (17) is equal to a prime  $q_j$  on the right. We now turn to the situations in which this condition is relaxed. If, for example,  $p_i = q_j$ , the corresponding contribution to (16) from this pair is

$$R_{i,j}^{(\text{diag})} = \frac{1}{(2\pi i)^2} \sum_p \frac{\ln^2 p}{p} \exp[i(x_i - x_j) \ln p], \quad (55)$$

where we have dropped the tilde/bar from the variables, because in this case the distinction is not important. The prime sum may be approximated using (22) to give

$$R_{i,j}^{(\text{diag})} \approx \frac{-1}{(2\pi)^2} \int_0^\infty y \exp[i(x_i - x_j)y] dy. \quad (56)$$

Evaluating the integral and adding on the  $i \leftrightarrow j$  permutation we thus have that

$$R_{i,j}^{(\text{diag})} + R_{j,i}^{(\text{diag})} \approx \frac{1}{2\pi^2(x_i - x_j)^2}. \quad (57)$$

In the general term (46) of our expression for  $\tilde{R}_n$ , the factor (57) multiplies the result (54) of the previous section with the variables  $x_i$  and  $x_j$  excluded from the partition sum and  $n$  replaced by  $n-2$ . Clearly this procedure may be repeated for any number of possible pairings of the variables not appearing in the phase. The final result for the amplitude is then given by the sum over all such combinations. These contributions were called semidiagonal in BK, or, if  $a=b$  in (17) and all of the  $n$  terms are paired, diagonal.

### 3. $\tilde{R}_n$ for the GUE

Our aim next is to show that the correlation function  $\tilde{R}_n$  calculated for the Riemann zeros in the last section coincides with the corresponding result for the GUE of random matrix theory, as given by (11) and (12), in the limit  $E \rightarrow \infty$ . To do so we derive a formula for the general term in the expansion of the determinant (11). This will turn out to be identical to the leading order asymptotic approximation, (46) and (54), which was obtained from the explicit (trace-type) formula (5) involving the primes. In addition, we shall also identify the terms in the expansion that correspond to the diagonal and off-diagonal contributions described in section 2.3.

The starting point for this calculation is the determinant expansion formula

$$\tilde{R}_n(x_1, \dots, x_n) = \sum_{\substack{\text{cycle} \\ \text{combinations}}} (-1)^{n-m} \prod_{j=1}^m \left[ s(x_1^{(j)}, x_2^{(j)}) s(x_2^{(j)}, x_3^{(j)}) \right. \\ \left. \times s(x_3^{(j)}, x_4^{(j)}) \dots s(x_{n(j)}^{(j)}, x_1^{(j)}) \right], \quad (58)$$

where the sum includes all cycle combinations (or multicycles),  $j$  labels the cycles in each combination,  $m$  is the number of cycles in each combination, and  $n(j)$  is the number of elements of the  $j$ th cycle (hence  $\sum_{j=1}^m n(j) = n$ ). The total number of terms in the sum is  $n!$ . We shall begin by first evaluating the contribution from one given cycle.

#### 3.1 Contribution from a single cycle

Consider a given  $N$ -cycle

$$C_N = s(x_1, x_2) s(x_2, x_3) s(x_3, x_4) \dots s(x_N, x_1) \quad (59)$$

with  $s$  as in (12). Clearly  $C_1=0$ , since  $s(x_1, x_1)=0$ . For  $N>1$ , using



elements (and these other cycles will arise in the sum over cycle combinations in (58)), and from multicycle combinations. We now show how these various contributions may be summed to give the general term in the expansion of  $\tilde{R}_n$ .

### 3.2 Permutations within a cycle

Consider first the  $\mu_i$  elements in the block  $\alpha_i \leq m < \beta_i$ . These give as a net contribution to the product in (63)

$$a_{\alpha_i, \beta_i} = \sum_{\text{perms}} \frac{1}{(x_{\alpha_i} - x_{\alpha_i+1})(x_{\alpha_i+1} - x_{\alpha_i+2}) \cdots (x_{\alpha_i+\mu_i} - x_{\beta_i})}, \quad (64)$$

where the sum includes all permutations of the  $\mu_i$  associated correlation variables  $x_{\alpha_i+1}, \dots, x_{\alpha_i+\mu_i}$ . Using the result (C8), derived in appendix C, this may be evaluated to give

$$a_{\alpha_i, \beta_i} = \frac{1}{x_{\alpha_i} - x_{\beta_i}} \sum_{\{\gamma\}_i, \{\delta\}_i = \mu_i} \prod_{\{\gamma\}_i} \frac{1}{x_{\alpha_i} - x_{\gamma}^{(i)}} \prod_{\{\delta\}_i} \frac{1}{x_{\delta}^{(i)} - x_{\beta_i}}, \quad (65)$$

where the sum now includes all partitions of these variables into two groups  $\{\gamma\}_i$  and  $\{\delta\}_i$ , with elements  $x_{\gamma}^{(i)}$  and  $x_{\delta}^{(i)}$ .

The result for the corresponding sum over the permutations between the cycle elements in the block  $\beta_i \leq m < \alpha_{i+1}$  clearly takes a form identical to (65). In this case the partition sum involves the  $\nu_i$  variables  $x_{\beta_i+1}, \dots, x_{\beta_i+\nu_i}$ , and the factor  $(-1)^{\nu_i}$  in (63) may be incorporated by reversing the signs of the terms in the products. Thus summing over all permutations of the elements of a single  $N$ -cycle within their respective blocks gives rise to contributions to  $\tilde{R}_n$  with an exponent (phase) as in (63), and an amplitude

$$\begin{aligned} \tilde{a}_N^{(r)} = & \frac{(-1)^r}{(2\pi i)^N} \frac{1}{(x_{\alpha_1} - x_{\beta_1})(x_{\beta_1} - x_{\alpha_2}) \cdots (x_{\alpha_r} - x_{\beta_r})(x_{\beta_r} - x_{\alpha_1})} \\ & \times \sum_{\substack{\{\gamma\}_1 + \{\delta\}_1 = \mu_1 \\ \{\gamma\}'_1 + \{\delta\}'_1 = \nu_1 \\ \vdots \\ \{\gamma\}_r + \{\delta\}_r = \mu_r \\ \{\gamma\}'_r + \{\delta\}'_r = \nu_r}} \prod_{i=1}^r \prod_{\{\gamma\}_i} \frac{1}{x_{\alpha_i} - x_{\gamma}^{(i)}} \prod_{\{\delta\}_i} \frac{1}{x_{\delta}^{(i)} - x_{\beta_i}} \prod_{\{\gamma\}'_i} \frac{1}{x_{\gamma'}^{(i)} - x_{\beta_i}} \prod_{\{\delta\}'_i} \frac{1}{x_{\alpha_{i+1}} - x_{\delta'}^{(i)}}, \quad (66) \end{aligned}$$

where the sum includes all partitions of the non-phase variables associated with each block into two groups.

Finally, we must now include permutations of the cycle elements between the different blocks. This obviously involves allowing for permutations of the non-phase variables between the  $2r$  different groups in (67). The result is that the amplitude associated with the exponent in (63) arising from all permutations of a single  $N$ -cycle is

$$a_N^{(r)} = \frac{(-1)^r 2^{N-2r}}{(2\pi i)^N} \frac{1}{(x_{\alpha_1} - x_{\beta_1})(x_{\beta_1} - x_{\alpha_2}) \dots (x_{\alpha_r} - x_{\beta_r})(x_{\beta_r} - x_{\alpha_1})} \quad (67)$$

$$\times \sum_{\{\gamma\}_1 + \{\delta\}_1 + \dots + \{\gamma\}_r + \{\delta\}_r = N-2r} \prod_{i=1}^r \prod_{\{\gamma\}_i} \frac{1}{x_{\alpha_i} - x_{\gamma}^{(i)}} \prod_{\{\delta\}_i} \frac{1}{x_{\delta}^{(i)} - x_{\beta_i}}$$

where now the sum includes all partitions of the  $N-2r$  non-phase variables into  $2r$  groups, and the factor of  $2^{N-2r}$  appears because each variable can contribute via two of the groups (primed and unprimed) in (66). This expression represents the main result of the present section.

When  $r=0$  the amplitude takes a particularly simple form, not immediately apparent from (67). Clearly in this case the only contributions come from cycles with no blocks - that is cycles in which all  $N$  signs  $\sigma_i$  in (61) are the same. Consider, for example, the case when  $\sigma_i = +1 \forall i$  (the  $\sigma_i = -1$  case differs only by a factor of  $(-1)^N$ ). Then, from (63),

$$a_N^{(0)} = \frac{1}{(2\pi i)^N} \sum_{\text{perms}} \frac{1}{(x_1 - x_2)(x_2 - x_3) \dots (x_N - x_1)} \quad (68)$$

where the sum includes all permutations of the  $N$  variables  $x_1, \dots, x_N$ . This may be written in the form

$$a_N^{(0)} = \frac{1}{(2\pi i)^N} \sum_{i=1}^N S_{N-1}(x_i, x_i), \quad (69)$$

where  $S_{N-1}$  is the corresponding permutation sum involving the  $N-1$  variables  $x_j, j \neq i$ . Using the identity (C2), proved in appendix C, we then have that

$$a_N^{(0)} = \frac{(-1)^{N-1}}{(2\pi i)^N} \sum_{i=1}^N (x_i - x_i)^{N-2} \prod_{\substack{j=1 \\ j \neq i}}^N \frac{1}{(x_j - x_k)^2}. \quad (70)$$

From this one sees immediately that

$$a_N^{(0)} = 0 \quad \text{if } N > 2 \quad (71)$$

and

$$a_2^{(0)} = \frac{1}{2\pi^2 (x_1 - x_2)^2}. \quad (72)$$

The second of these results, (72), may be verified directly, since it represents the non-oscillatory term in the sum of  $s(x_1, x_2)s(x_2, x_1)$  and its permutation ( $1 \leftrightarrow 2$ ), where  $s$  is given by (12).

### 3.3 Multicycle contributions

It now remains to extend the single- $N$ -cycle amplitude calculated in the last section to cover the multicycle combinations that appear in (58). We again focus on a general term in the expansion of the GUE correlation function  $\tilde{R}_n$  with a phase as in (63). Since the mean density  $\bar{d}(E)$  is given asymptotically by (13), this phase is, to leading order as  $E \rightarrow \infty$ , the same as that in (46). Our aim, therefore, is to evaluate the sum in (58) using the block-decomposition structure described in section 3.2, and to compare the net amplitude with the result (54) for the Riemann zeros.

The combinations involved in (58) take the form of a product of cycles whose combined length is  $n$ , and so the exponent in (63) may in general arise if different cycles each contribute subsets of the variables  $x_{\alpha_i}, x_{\beta_i}$ . The amplitude of each combination is clearly a product of the corresponding single-cycle terms, given by (67), (71) and (72). The problem is to sum over all possible divisions of the  $2r$  phase variables among different numbers of cycles.

The first point to note is that, for  $N \neq 2$ , each  $N$ -cycle must contribute a non-zero number of variables to the overall phase of any combination of which it is a part. This follows immediately from the vanishing of the amplitude when  $r=0$  (recall that  $r$  labels the number of variables in the phase associated with the cycle in (63)), as determined by (72) when  $N > 2$ , and by the fact that 1-cycles automatically give zero since  $s(x, x)=0$  (this is a consequence of the fact that we are working with the irreducible correlation function  $\tilde{R}_n$  - in the full correlation function  $R_n$  the factor  $(1-\delta_{ij})$  does not appear in (12), and so in that case the 1-cycles contribute a constant amplitude equal to the mean density  $\bar{d}$ ). Hence only 2-cycles can contribute to the amplitude but not the phase of a multicycle.

We begin by ignoring this special 2-cycle behaviour; that is, we first treat the case in which each cycle in a combination contributes a non-zero number of terms to the overall phase. We then return to the 2-cycles in the next-but-one subsection, 3.3(ii).

#### 3.3(i) General $N$ -cycle contributions

Neglecting the special status of the 2-cycles allows us to consider the general case in which each cycle, labelled  $j$ , in a combination in (58) contributes  $r_j \neq 0$  pairs of variables to the overall phase. We then seek the corresponding net amplitude when these phase variables are as in (46) and (63); that is, when there are  $r$  pairs in total, and so for a combination of  $m$  cycles  $\sum_{j=1}^m r_j = r$ . The amplitude for a particular combination, including all permutations among the  $n(j)$  elements of each cycle, is clearly  $\prod_{j=1}^m a_{n(j)}^{(r_j)}$ , where  $a$  is given by (67). The net result then follows after summing over all divisions of the  $2r$  phase variables amongst the various possible cycle combinations consistent with  $\sum_{j=1}^m n(j) = n$ .

Consider first the second factor in (67), the sum over partitions of the variables not appearing in the phase. Taking the product over all cycles and including permutations of these variables between different cycles clearly gives

$$\sum_{\{\gamma\}_1 + \{\delta\}_1 + \dots + \{\gamma\}_r + \{\delta\}_r = n-2r} \prod_{i=1}^r \prod_{\{\gamma\}_i} \frac{1}{x_{\alpha_i} - x_{\gamma}^{(i)}} \prod_{\{\delta\}_i} \frac{1}{x_{\delta}^{(i)} - x_{\beta_i}}, \quad (73)$$

where the sum includes all partitions of the  $n-2r$  variables into  $2r$  groups.

Now consider the part of the amplitude coming from the first term in (67). For a given combination, this will be

$$\prod_{j=1}^m \frac{1}{\left(x_{\alpha_1}^{(j)} - x_{\beta_1}^{(j)}\right)\left(x_{\beta_1}^{(j)} - x_{\alpha_2}^{(j)}\right) \dots \left(x_{\alpha_r}^{(j)} - x_{\beta_r}^{(j)}\right)\left(x_{\beta_r}^{(j)} - x_{\alpha_1}^{(j)}\right)}, \quad (74)$$

where  $x^{(j)}$  denotes a variable associated with the  $j$ th cycle. Clearly (74) is itself a product of cycles amongst the  $2r$  phase variables. Summing over all permutations amongst these cycles thus gives

$$F(x_{\alpha_1}, x_{\beta_1}, \dots, x_{\alpha_r}, x_{\beta_r}) = \sum_{\substack{\text{cycle} \\ \text{combinations}}} (-1)^{2r-m} \times \prod_{j=1}^m \frac{1}{\left(x_{\alpha_1}^{(j)} - x_{\beta_1}^{(j)}\right)\left(x_{\beta_1}^{(j)} - x_{\alpha_2}^{(j)}\right) \dots \left(x_{\alpha_r}^{(j)} - x_{\beta_r}^{(j)}\right)\left(x_{\beta_r}^{(j)} - x_{\alpha_1}^{(j)}\right)}, \quad (75)$$

the factor  $(-1)^m$  coming from (58).

The crucial point is that  $F$  may now be recognised as the determinant of the  $2r \times 2r$  matrix with elements

$$f_{2i-1, 2j} = -f_{2i, 2j-1} = \frac{1}{\left(x_{\alpha_i} - x_{\beta_j}\right)}, \quad (76)$$

$$f_{2i-1, 2j-1} = f_{2i, 2j} = 0, \quad (77)$$

where  $1 \leq i, j \leq r$ . This determinant can be evaluated by first shifting all of the even columns to form the first  $r$  columns of an intermediate matrix, and then moving all of the odd rows of the intermediate to form the first  $r$  rows of a new matrix  $\tilde{f}$ . The reason is that  $\tilde{f}$  is block-diagonal. The two  $r \times r$  nonzero blocks on the diagonal are each in the form of a Cauchy matrix (Cauchy 1841, Knuth 1973), the elements of the first block being given by

$$\tilde{f}_{i,j}^{(+)} = \frac{1}{x_{2i-1} - x_{2j}}, \quad (78)$$

$1 \leq i, j \leq r$ , and those of the second block by

$$\tilde{f}_{i,j}^{(-)} = \frac{1}{x_{2i} - x_{2j-1}}. \quad (79)$$

Both blocks have the same determinant, which may be calculated from Cauchy's formula. The result is then that

$$F(x_{\alpha_1}, x_{\beta_1}, \dots, x_{\alpha_r}, x_{\beta_r}) = \left[ \frac{\prod_{1 \leq j < i \leq r} (x_{\alpha_i} - x_{\alpha_j})(x_{\beta_i} - x_{\beta_j})}{\prod_{1 \leq i, j \leq r} (x_{\alpha_i} - x_{\beta_j})} \right]^2. \quad (80)$$

It follows that summing the amplitude product over all cycle combinations except for the  $r=0$  contributions from 2-cycles leads to a general term in the expansion of  $\tilde{R}_n$  of the form

$$A_{\alpha_1, \beta_1, \dots, \alpha_r, \beta_r}^{(r)} \exp\left\{2\pi i(x_{\alpha_1} - x_{\beta_1} + x_{\alpha_2} - x_{\beta_2} + \dots + x_{\alpha_r} - x_{\beta_r})\bar{d}(E)\right\}, \quad (81)$$

where, from (67), (73) and (80),  $A^{(r)}$  is given by exactly the same expression (54) as for the Riemann zeros, ignoring the diagonal contributions in that case. The fact that the phase is also asymptotically the same as that in (46) follows from (13), as noted above. Hence the full equality of the  $n$ -point correlation functions is demonstrated once it has been established that there is a one-to-one link between the 2-cycles and the diagonal terms.

### 3.3 (ii) 2-cycle contributions

The arguments used above ignore the  $r=0$  contributions from the 2-cycles. If they are included, the general exponent in (81) may also arise when pairs of non-phase variables are removed from the partition sum (73) to form 2-cycle terms in the amplitude. The effect of this is to replace  $n$  by  $n-2x$  (the number of pairs removed) in the previous results, and then, using (72), to multiply by

$$a_2^{(0)} = \frac{1}{2\pi^2(x_i - x_j)^2} \quad (82)$$

for each such pair. This factor, and the resulting change, are identical to those associated with each diagonal term in the multiple prime sum for the Riemann zeros, as discussed in section 2.3. The 2-cycles and the diagonal terms are thus in one-to-one correspondence with respect to their contributions to the correlation function. Together with the results of the previous section, this completes our demonstration that the  $n$ -point statistics of the zeros are the same as those of the eigenvalues of matrices in the GUE for all orders  $n$ .

#### 4. Comments

We have demonstrated in this paper that the irreducible  $n$ -point correlation function  $\tilde{R}_n$  of the Riemann zeros and the corresponding correlation function of the eigenvalues of the GUE of random matrices are term-by-term asymptotically identical for all  $n$ . This represents, to the best of our knowledge, the first example in which all orders of correlation have been shown to exhibit random matrix behaviour in a single explicit example, that is, without any system averaging or randomness assumption. It should again be emphasised, however, that our method is essentially heuristic, in that it is based on the unproved Hardy-Littlewood conjecture, and because we have made no attempt to control the approximations involved in the analysis (it appears that the prime correlation effects, which form the main focus of our interest, are not amenable to the rigorous methods developed to deal with the diagonal terms by Montgomery (1973), Hejhal (1994), and Rudnick & Sarnak (1994)). For this reason, it may still be worth computing the irreducible statistics numerically for  $n > 2$  and to compare them with the GUE expression (11). Furthermore, we have concentrated solely on the universal statistical regime. The nonuniversal correlations could easily be investigated along the lines of Berry (1988), and these too may be worthy of numerical study when  $n > 2$ .

The fact that our approach was based directly on the analogue of the trace formula, the prime sum (5), suggests that it might be possible to generalise some of the combinatorial features of our method to study semiclassically the higher order correlations in the spectra of classically chaotic systems in the same way. Unfortunately, in that case one has no a priori knowledge of the periodic orbit statistics that correspond to the Hardy-Littlewood conjecture. However, it is clear that our results may be applied in the reverse direction: by starting from our representation for the general term in the expansion of the GUE determinant, as given by equations (54) and (81), it should be possible to take an inverse Fourier transform with respect to some suitable scaling parameter (e.g.  $1/\hbar$ ) and so obtain the general form of the  $n$ -point classical periodic orbit correlation function, as was done for the  $n=2$  case in Argaman *et al.* (1993).

Ultimately, it would of course be desirable to have for the zeta function a direct method of evaluating the multiple prime sum (16) that leads immediately to the determinant form (11) of the GUE result; that is, a method that does not require the determinant first to be expanded, and then equality demonstrated term-by-term. This would undoubtedly lead to a greater insight into the precise combinatorial links between (16) and (11). Unfortunately, we have so far failed to find such an approach.

#### Acknowledgements

JPK is grateful to the Institut de Physique Nucléaire in Orsay for generous hospitality during the period when this work was started, and to the Nuffield Foundation for supporting the collaboration in the UK.

## Appendix A

We are to evaluate  $\tilde{r}_n^{(a,b,s,t)}$ , defined by equations (25)-(29) in the text. Using the relabelled variables defined by (30), (31), and in the paragraph thereafter, we begin by calculating the integrals over  $w'_{\alpha_{\mu'}}$  ( $1 \leq \mu' \leq a-s$ ) and  $z_{\beta_v}$  ( $1 \leq v \leq t$ ):

$$I_1(u) = \int_0^\infty \dots \int_0^\infty \prod_{\mu'=1}^{a-s} dw'_{\alpha_{\mu'}} \exp[i\tilde{x}'_{\alpha_{\mu'}} w'_{\alpha_{\mu'}}] \prod_{v=1}^t dz_{\beta_v} \exp[-i\bar{x}_{\beta_v} z_{\beta_v}] \times \delta\left(\sum_{\mu'=1}^{a-s} w'_{\alpha_{\mu'}} - \sum_{v=1}^t z_{\beta_v} - u\right). \quad (\text{A1})$$

Expressing the  $\delta$ -function as a Fourier integral,

$$I_1(u) = \int_0^\infty \dots \int_0^\infty \prod_{\mu'=1}^{a-s} dw'_{\alpha_{\mu'}} \exp[i\tilde{x}'_{\alpha_{\mu'}} w'_{\alpha_{\mu'}}] \prod_{v=1}^t dz_{\beta_v} \exp[-i\bar{x}_{\beta_v} z_{\beta_v}] \times \int_{-\infty}^\infty \exp\left[-2\pi i \lambda \left(\sum_{\mu'=1}^{a-s} w'_{\alpha_{\mu'}} - \sum_{v=1}^t z_{\beta_v} - u\right)\right] d\lambda, \quad (\text{A2})$$

and shifting the correlation variables to ensure convergence

$$\tilde{x}'_{\alpha_{\mu'}} \rightarrow \tilde{x}'_{\alpha_{\mu'}} + i\varepsilon \quad (\text{A3})$$

$$\bar{x}_{\beta_v} \rightarrow \bar{x}_{\beta_v} - i\varepsilon \quad (\text{A4})$$

we have

$$I_1(u) = \frac{(-1)^t}{i^{a-s+t}} \int_{-\infty}^\infty \frac{\exp(2\pi i \lambda u)}{\prod_{\mu'=1}^{a-s} (2\pi \lambda - \tilde{x}'_{\alpha_{\mu'}} - i\varepsilon) \prod_{v=1}^t (2\pi \lambda - \bar{x}_{\beta_v} + i\varepsilon)} d\lambda. \quad (\text{A5})$$

Closing the  $\lambda$ -contour above the real axis and taking the limit  $\varepsilon \rightarrow 0$  then gives

$$I_1(u) = \frac{(-1)^t}{i^{a-s+t-1}} \sum_{\mu'=1}^{a-s} \frac{\exp(iu\tilde{x}'_{\alpha_{\mu'}})}{\prod_{\substack{\mu''=1 \\ \mu'' \neq \mu'}}^{a-s} (\tilde{x}'_{\alpha_{\mu''}} - \tilde{x}'_{\alpha_{\mu'}}) \prod_{v=1}^t (\tilde{x}'_{\alpha_{\mu'}} - \bar{x}_{\beta_v})}. \quad (\text{A6})$$

In a similar way we may also evaluate the integrals over  $w_{\alpha_\mu}$  ( $1 \leq \mu \leq s$ ) and  $z'_{\beta_{v'}}$  ( $1 \leq v' \leq b-t$ ) in (25):

$$I_2(u) = \int_0^\infty \dots \int_0^\infty \prod_{\mu=1}^s dw_{\alpha_\mu} \exp[i\tilde{x}_{\alpha_\mu} w_{\alpha_\mu}] \prod_{\nu'=1}^{b-t} dz'_{\beta_{\nu'}} \exp[-i\bar{x}'_{\beta_{\nu'}} z'_{\beta_{\nu'}}] \times \int_{-\infty}^\infty \exp\left[2\pi i \lambda \left(\sum_{\nu'=1}^{b-t} z'_{\beta_{\nu'}} - \sum_{\mu=1}^s w_{\alpha_\mu} - u\right)\right] d\lambda. \quad (\text{A7})$$

Again, shifting the variables

$$\tilde{x}_{\alpha_\mu} \rightarrow \tilde{x}_{\alpha_\mu} + i\varepsilon \quad (\text{A8})$$

$$\bar{x}'_{\beta_{\nu'}} \rightarrow \bar{x}'_{\beta_{\nu'}} - i\varepsilon \quad (\text{A9})$$

to ensure convergence, we find that

$$I_2(u) = \frac{(-1)^{b-t}}{i^{s+b-t}} \int_{-\infty}^\infty \frac{\exp(-2\pi i \lambda u)}{\prod_{\mu=1}^s (2\pi \lambda - \tilde{x}_{\alpha_\mu} - i\varepsilon) \prod_{\nu'=1}^{b-t} (2\pi \lambda - \bar{x}'_{\beta_{\nu'}} + i\varepsilon)} d\lambda. \quad (\text{A10})$$

In this case, closing the  $\lambda$ -contour below the real axis and taking the limit  $\varepsilon \rightarrow 0$  gives

$$I_1(u) = \frac{(-1)^{b-t}}{i^{a-s+t+1}} \sum_{\nu'=1}^{b-t} \frac{\exp(-iu\bar{x}'_{\beta_{\nu'}})}{\prod_{\mu=1}^s (\bar{x}'_{\beta_{\nu'}} - \tilde{x}_{\alpha_\mu}) \prod_{\substack{\nu''=1 \\ \nu'' \neq \nu'}}^{b-t} (\bar{x}'_{\beta_{\nu'}} - \bar{x}'_{\beta_{\nu''}})}. \quad (\text{A11})$$

Finally, multiplying  $I_1$  by  $I_2$  and then evaluating the integral over  $u$  in (25), we have that

$$\begin{aligned} \tilde{r}_n^{(a,b,s,t)}(x_1, \dots, x_n; E) &= -\frac{(-1)^n}{(2\pi i)^n} \sum_{\mu'=1}^{a-s} \sum_{\nu'=1}^{b-t} \frac{\exp[i(\tilde{x}'_{\alpha_{\mu'}} - \bar{x}'_{\beta_{\nu'}}) \ln E]}{(\tilde{x}'_{\alpha_{\mu'}} - \bar{x}'_{\beta_{\nu'}})^2} \\ &\quad \times \prod_{\substack{\mu''=1 \\ \mu'' \neq \mu'}}^{a-s} \frac{1}{(\tilde{x}'_{\alpha_{\mu'}} - \tilde{x}'_{\alpha_{\mu''}})} \prod_{\substack{\nu''=1 \\ \nu'' \neq \nu'}}^{b-t} \frac{1}{(\bar{x}'_{\beta_{\nu'}} - \bar{x}'_{\beta_{\nu''}})} \\ &\quad \times \prod_{\mu=1}^s \frac{1}{(\tilde{x}_{\alpha_\mu} - \bar{x}'_{\beta_{\nu'}})} \prod_{\nu=1}^t \frac{1}{(\tilde{x}'_{\alpha_{\mu'}} - \bar{x}_{\beta_\nu})}, \end{aligned} \quad (\text{A12})$$

where we have used the fact that  $a+b=n$ . This is the result quoted in the text.

## Appendix B

Given  $x_1, x_2, \dots, x_n$ , we here show that the product

$$\prod_{\substack{k=1 \\ k \neq j}}^n \frac{1}{x_j - x_k}$$

satisfies the following identities:

$$\sum_{j=1}^n x_j^m \prod_{\substack{k=1 \\ k \neq j}}^n \frac{1}{x_j - x_k} = 0 \quad (\text{B1})$$

for integers  $0 \leq m \leq n-2$ ;

$$\sum_{j=1}^n x_j^{n-1} \prod_{\substack{k=1 \\ k \neq j}}^n \frac{1}{x_j - x_k} = 1; \quad (\text{B2})$$

$$\sum_{j=1}^n \frac{1}{y - x_j} \prod_{\substack{k=1 \\ k \neq j}}^n \frac{1}{x_j - x_k} = \prod_{k=1}^n \frac{1}{y - x_k}; \quad (\text{B3})$$

and

$$\sum_{j=1}^n (y - x_j)^m \prod_{\substack{k=1 \\ k \neq j}}^n \frac{1}{x_j - x_k} = \begin{cases} 0 & \text{if } m = 0, 1, 2, \dots, n-2 \\ (-1)^{n-1} & \text{if } m = n-1 \end{cases}. \quad (\text{B4})$$

These will play a central role in the evaluation of the permutation sums in Appendix C.

### Proof of (B1) and (B2)

We use the fact that if  $m$  is an integer then

$$\sum_{j=1}^n x_j^m \prod_{\substack{k=1 \\ k \neq j}}^n \frac{1}{x_j - x_k} = \frac{1}{2\pi i} \oint \frac{z^m}{\prod_{i=1}^n (z - x_i)} dz, \quad (\text{B5})$$

where the contour of the integral is a circle of radius  $R$ , centre the origin, such that  $R > |x_i|$

$\forall i$ . We now let  $R \rightarrow \infty$ . If  $m < n-1$ , the integrand tends to zero faster than  $z^{-1}$  and so the value of the integral is zero, which proves (B1). If, on the other hand,  $m = n-1$ , the integrand tends to zero like  $1/z$ , and so the value of the integral is 1, which proves (B2).

### Proof of (B3)

We have that

$$\sum_{j=1}^n \frac{1}{y - x_j} \prod_{\substack{k=1 \\ k \neq j}}^n \frac{1}{x_j - x_k} - \prod_{k=1}^n \frac{1}{y - x_k} = \frac{1}{2\pi i} \oint \frac{1}{(y - z) \prod_{k=1}^n (z - x_k)} dz, \quad (\text{B6})$$

where the contour is the same as that in (B5) and, in addition, satisfies  $R > |y|$ . Again we let  $R \rightarrow \infty$ . The integral obviously tends to zero for  $n > 0$ . This then proves (B3).

#### Proof of (B4)

In this case we use

$$\sum_{j=1}^n (y - x_j)^m \prod_{\substack{k=1 \\ k \neq j}}^n \frac{1}{x_j - x_k} = \frac{1}{2\pi i} \oint \frac{(y - z)^m}{\prod_{k=1}^n (z - x_k)} dz, \quad (\text{B7})$$

where the contour is the same as in (B5). Letting  $R \rightarrow \infty$ , it is clear that the integral vanishes for  $m = 0, 1, 2, \dots, n-2$ , proving the result in the first case. When  $m = n-1$ , the integrand tends to zero like  $(-1)^{n-1}/z$  and so the value of the integral is  $(-1)^{n-1}$ , which proves (B4) in the one remaining case.

#### Appendix C

Using the results of Appendix B, we here show that

$$S_n(y, z) = \sum_{\text{perms}} \frac{1}{(y - x_1)(x_1 - x_2)(x_2 - x_3) \dots (x_{n-1} - x_n)(x_n - z)}, \quad (\text{C1})$$

where the sum includes all permutations of the  $n$  terms  $x_1, x_2, \dots$ , and  $x_n$ , is given by

$$S_n(y, z) = (y - z)^{n-1} \prod_{k=1}^n \frac{1}{(y - x_k)(x_k - z)}. \quad (\text{C2})$$

The proof is by induction. We start by observing that the permutation sum may be written in the form

$$S_n(y, z) = \sum_{\substack{i, j=1 \\ i \neq j}}^n \frac{1}{(y - x_i)(x_j - z)} S_{n-2}(x_i, x_j), \quad (\text{C3})$$

where  $S_{n-2}$  includes all permutations of the  $n-2$  terms  $x_p$  such that  $p \neq i \neq j$ . Assuming the result (C2), we then have that

$$S_n(y, z) = \sum_{\substack{i, j=1 \\ i \neq j}}^n \frac{1}{(y - x_i)(x_j - z)} (x_i - x_j)^{n-3} \prod_{p \neq i, j} \frac{1}{(x_i - x_p)(x_p - x_j)} \quad (\text{C4})$$

or

$$S_n(y, z) = (-1)^n \sum_{i,j=1}^n \frac{1}{(y-x_i)} \prod_{\substack{\alpha=1 \\ \alpha \neq i}}^n \frac{1}{(x_i-x_\alpha)} (x_i-x_j)^{n-1} \frac{1}{(z-x_j)} \prod_{\substack{\beta=1 \\ \beta \neq j}}^n \frac{1}{(x_j-x_\beta)}. \quad (\text{C5})$$

Substituting into (C5) the expansion

$$(x_i-x_j)^{n-1} = \sum_{n_1+n_2+n_3=n-1} \frac{(n-1)!}{n_1!n_2!n_3!} (x_i-y)^{n_1} (y-z)^{n_2} (z-x_j)^{n_3}, \quad (\text{C6})$$

where the sum includes all triples of non-negative integers  $(n_1, n_2, n_3)$  such that  $n_1+n_2+n_3=n-1$ , the sums over  $i$  and  $j$  may be evaluated using the identities (B3) and (B4) of appendix B. The result is precisely (C2). The equality is thus established on the basis of its obvious correctness when  $n=1$ .

Another useful expression for  $S_n$  may be obtained by rewriting (C2) in the form

$$S_n(y, z) = \frac{1}{y-z} \prod_{k=1}^n \left( \frac{1}{y-x_k} + \frac{1}{x_k-z} \right), \quad (\text{C7})$$

since expanding out the product gives

$$S_n(y, z) = \frac{1}{y-z} \sum_{\{\gamma\}+\{\delta\}=n} \prod_{\{\gamma\}} \frac{1}{y-x_\gamma} \prod_{\{\delta\}} \frac{1}{x_\delta-z}, \quad (\text{C8})$$

where the sum includes all partitions of the  $n$  terms  $x_1, x_2, \dots, x_n$  into two groups  $\{\gamma\}$  and  $\{\delta\}$ , with elements denoted by  $x_\gamma$  and  $x_\delta$  respectively. This is the result quoted in the text.

## References

- Argaman, N., Dittes, F.M., Doron, E., Keating, J.P., Kitaev, Ya., Sieber, M., & Smilansky, U., 1993, *Phys. Rev. Lett.* **71**, 4326-4329.
- Berry, M.V., 1985, *Proc. R. Soc. Lond.* **A400**, 229-251.
- Berry, M.V., 1986, 'Riemann's zeta function: a model for quantum chaos?' in *Quantum Chaos and Statistical Nuclear Physics* ed T. H. Seligman & H. Nishioka (*Lecture Notes in Physics* **263**) (Berlin: Springer), 1-17.
- Berry, M.V., 1987, *Proc. R. Soc. Lond.* **A413**, 183-198.
- Berry, M.V., 1988, *Nonlinearity* **1**, 399-407.
- Bogomolny, E.B., & Keating, J.P., 1995, 'Random matrix theory and the Riemann zeros I: three- and four-point correlations', *Nonlinearity*, in the press.
- Bogomolny, E., & Leboeuf, P., 1994, *Nonlinearity* **7**, 1155-1167.
- Bogomolny, E., Leyvraz, F., & Schmit, C., 1994, 'Distribution of eigenvalues for the modular group', preprint.

- Bohigas, O., 1991, 'Random matrix theories and chaotic dynamics', in *Les Houches Lecture Series*, vol 52, eds. M.J. Giannoni, A. Voros & J. Zinn-Justin (North-Holland, Amsterdam), 89-199.
- Bohigas, O., Giannoni, M.-J., & Schmit, C., 1984a, *Phys. Rev. Lett.* **52**, 1-4.
- Bohigas, O., Giannoni, M.-J., & Schmit, C., 1984b, *J. Physique Lett.* **45**, L1015-22.
- Bohigas, O., Haq, R.U., & Pandey, A., 1985 *Phys. Rev. Lett.* **54**, 1645-48.
- Cauchy, A., 1841, *Exercices d'Analyse et de Phys. Math.*, volume 2 (Paris), pp. 151-159.
- Edwards, H.M., 1973, *Riemann's zeta function* (Academic Press: New York).
- Goldston, D., & Montgomery, H.L., 1984, *Proc. Conf. at Oklahoma State Univ. 1984*, ed. A.C. Adolphson *et al*, 183-203.
- Gutzwiller, M.C., 1971, *J. Math. Phys.* **12**, 343-358.
- Hannay, J.H., & Ozorio de Almeida, A.M., 1984, *J. Phys. A* **17**, 3429-3440.
- Hardy, G.H., & Littlewood, J.E., 1923 *Acta Math.* **44**, 1-70.
- Hejhal, D.A., 1994, 'On the triple correlation of zeros of the zeta function', preprint.
- Keating, J.P., 1991a, 'The semiclassical sum rule and Riemann's zeta-function' in *Quantum Chaos* ed. H.A. Cerdeira, R. Ramaswamy, M.C. Gutzwiller & G. Casati (Singapore: World Scientific), 280-294.
- Keating, J.P., 1991b, *Nonlinearity* **4**, 309-341.
- Keating, J.P., 1993, 'The Riemann zeta-function and quantum chaology' in *Quantum Chaos* ed. G. Casati, I. Guarneri & U. Smilansky (Amsterdam: North-Holland), 145-185.
- Knuth, D.E., 1973, *The art of computer programming, volume 1, fundamental algorithms*, second edition, (Addison-Wesley, Reading, Massachusetts), p. 36.
- Lord Cherwell (1946), *Quart. J. of Math.* **18**, 46-62.
- Luo, W., & Sarnak, P., 1994, *Comm. Math. Phys.* **161**, 419-432.
- Mehta, M.L., 1991, *Random Matrices and the Statistical Theory of Energy Levels*, 2nd edition (New York: Academic Press).
- Montgomery, H.L., 1973, *Proc. Symp. Pure Math.* **24**, 181-193.
- Rudnick, Z., & Sarnak, P., 1994, *Comptes Rendus, Serie I* **319**, 1027-1032.
- Odlyzko, A.M., 1987, *Math. Comp.* **48**, 273-308.
- Odlyzko, A.M., 1989, 'The  $10^{20}$ -th zero of the Riemann zeta function and 70 million of its neighbors', preprint, AT&T Bell Laboratories.
- Titchmarsh, E.C., 1986, *The theory of the Riemann zeta-function* (Oxford University Press).