Two-point spectral correlations for the square billiard

Two-point spectral correlations for the square billiard

R D Connors† and J P Keating†‡

† School of Mathematics, University of Bristol, Bristol BS8 1TW, UK
‡ BRIMS, Hewlett-Packard Laboratories, Filton Road, Stoke Gifford, Bristol BS12 6QZ, UK

Abstract. We investigate the two-point correlations in the quantum spectrum of the square billiard. This system is unusual in that the degeneracy of the energy levels increases in the semiclassical limit in such a way that the average level separation is not given by the inverse of the mean density of states. Hence, for example, the standard level spacings distribution does not tend to the Poissonian limit expected for integrable systems. In this paper we calculate the leading-order asymptotic form of a degeneracy-weighted two-point correlation function using a combination of probabilistic techniques and classical number theory. The result exhibits number-theoretical fluctuations about a mean which is a sum of two terms: one having the usual (constant) Poissonian form and the second representing a small correction which decays as the inverse of the correlation distance. This is confirmed by numerical computations.

1. Introduction

The energy levels of generic, integrable systems are conjectured to be Poisson distributed in the semiclassical limit [4]. This belief is supported both by theories based on semiclassical asymptotics [4, 2, 19, 16, 5, 17] and by extensive numerical computations [3, 6].

The square billiard, although completely integrable, is non-generic in this respect. Its levels, when suitably scaled, are given by

$$E_{m,n} = m^2 + n^2 \qquad (m, n) \in \mathbb{N}^2, (m, n) \neq (0, 0) \tag{1}$$

and so the density of states may be written in the form

$$d(E) = \sum_{n=1}^{\infty} r_2(n) \, \delta(E - n) \tag{2}$$

where $r_2(n)$ is the number of ways that n can be written as a sum of two squares. Unlike the case of typical systems, these levels become increasingly degenerate as $E_{m,n} \to \infty$. One consequence of this is that the mean density $\overline{d} = \overline{r_2(n)} = \frac{\pi}{4}$ (obtained by counting the number of lattice points in a quarter of a circle) does not equal the inverse of the average non-zero level spacing, and so the standard spectral unfolding does not lead to a limiting statistical distribution. For example, the level spacings distribution tends to a δ-function at the origin.

These degeneracies have their origin in the number-theoretical nature of the levels. It is a classical result of Landau [15] that if

$$B_2(n) = \begin{cases} 1 & \text{if } r_2(n) > 0 \\ 0 & \text{if } r_2(n) = 0 \end{cases} \tag{3}$$

then the local average of $B_2(n)$ as $n \to \infty$ is given by

$$\overline{B_2}(n) \simeq \frac{C_2}{\sqrt{\ln n}} \tag{4}$$

where

$$C_2 = \sqrt{\frac{1}{2} \prod_r \left(\frac{1}{1 - \frac{1}{r^2}} \right)} \tag{5}$$

with r denoting the subset of primes congruent to 3 modulo 4. The product in (5) converges to give $C_2 \simeq 0.764$. For this to be consistent with $\overline{r_2}(n) = \frac{\pi}{4}$, the logarithmically increasing separation between the levels implied by (4) must be compensated by a corresponding logarithmic increase in the mean degeneracy.

An important consequence of this is that the two-point correlation function $\overline{B_2(n)B_2(n+h)}$, which is proportional to the probability that both n and $n+h$ are energy levels, tends, in the limit $n \to \infty$, to zero for any fixed h. Our main aim here is to show that this problem does not occur if the levels are weighted by their degeneracies; that is, for $\overline{r_2(n)r_2(n+h)}$ the limit is finite. The limiting form we find in this case shows two levels of structure. The function itself depends strongly on the prime factorization of h and so exhibits large fluctuations. However, a local average with respect to h reveals an underlying trend that corresponds to a sum of two terms: one, a constant, being the usual Poisson correlation function, and the other representing a small deviation from pairwise randomness which decays as the inverse of the correlation distance. This behaviour is strikingly reminiscent of that implied by the Hardy–Littlewood conjectures [10] for the primes [14].

The methods we use to calculate these correlations involve a combination of probabilistic ideas and number-theoretic results. This provides the most direct approach to obtaining the leading-order behaviour in the semiclassical limit. The price to be paid is that the analysis is essentially heuristic, in that we cannot give rigorous estimates for the size of the error terms. For this reason, the results are also tested against numerical computations.

2. Counting representations

In this section we review the basic ideas underlying probabilistic number theory and show how these can be used to calculate the averages of the counting functions B_2 and r_2. As an example, we shall demonstrate that the results obtained using these methods are consistent with the classical asymptotic formulae for $\overline{r_2}(n)$ and $\overline{B_2}(n)$ discussed in the introduction. This then provides the justification for our use of these same methods to evaluate the associated autocorrelation functions in subsequent sections.

We begin by introducing some notation. Henceforth, the primes will be denoted by q; p will denote primes congruent to 1 modulo 4; and r will denote primes congruent to 3 modulo 4. The prime decomposition of an integer n into prime powers may therefore be written uniquely as

$$n = 2^{m_2(n)} \prod_p p^{m_p(n)} \prod_r r^{m_r(n)}.$$

Thus in general, $m_q(n)$ is the power to which q is raised in the prime decomposition of n.

The reason for dividing primes into residue classes modulo 4 is that it is a classical result in number theory that [9, 11]

$$B_2(n) = \begin{cases} 1 & \text{if } 2|m_r(n) \ \forall \ r \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

and

$$r_2(n) = \begin{cases} \prod_p (m_p(n) + 1) & \text{if } 2|m_r(n) \ \forall \ r \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

That is, an integer, n, is representable as a sum of two squares if and only if all primes that are congruent to 3 modulo 4 occur to even (including zero) powers in its prime factorization. The number of such representations in turn depends on the exponents of the primes that are congruent to 1 modulo 4.

It is the dependence on prime decomposition that opens the way for us to use probabilistic techniques. These methods are reviewed in [18, 12]. They have previously been employed to re-derive (4) in [1], and applied to other quantum chaological calculations in [13, 14]. The basic idea is that in averaging number-theoretical functions such as $r_2(n)$ and $B_2(n)$ with respect to n, one can treat the prime factors of n as being statistically independent, with the probability $P(q^m||n)$ that q appears with exponent m in the prime factorization of n being

$$P(q^m||n) = \frac{1}{q^m} - \frac{1}{q^{m+1}} = \frac{1}{q^m}\left(1 - \frac{1}{q}\right). \tag{8}$$

Here a natural interpretation of the first equality follows from the fact that one in q^m integers are divisible by q^m, and one in $q^{(m+1)}$ are divisible by higher powers of q. Similarly in the second equality, the first factor may be viewed as representing the probability that q^m divides n, and the second as representing the probability that the result of the division is then itself not divisible by q. For further details and examples, see [18].

It follows that the probability that a given prime r occurs to an even power in the decomposition of n is

$$P(m_r(n) \text{ even}) = \sum_{m=0}^{M} \frac{1}{r^{2m}}\left(1 - \frac{1}{r}\right) \tag{9}$$

where M is some integer such that $n < r^{2M+2}$; higher powers clearly cannot contribute. However, the sum converges rapidly, and so as $n \to \infty$

$$P(m_r(n) \text{ even}) \simeq \sum_{m=0}^{\infty} \frac{1}{r^{2m}}\left(1 - \frac{1}{r}\right) \tag{10}$$

$$= \left(1 - \frac{1}{r}\right)\left(1 - \frac{1}{r^2}\right)^{-1}.$$

The local average of $B_2(n)$ with respect to n corresponds to the probability that n can be written as a sum of squares. It thus corresponds to the probability that all primes $r < n$ have even exponents and so is given by

$$\overline{B_2}(n) \simeq \prod_{r \leqslant n} \left(1 - \frac{1}{r}\right)\left(1 - \frac{1}{r^2}\right)^{-1}. \tag{11}$$

The product of factors $(1 - r^{-2})^{-1}$ converges to give a constant, while the product over the remaining terms may be estimated using the prime number theorem and the fact that, by Dirichlet's theorem [11], the density of primes that are congruent to 3 modulo 4 is $\frac{1}{2}$. The result is

$$\overline{B_2}(n) \simeq \text{constant} \times \frac{1}{\sqrt{\ln n}}. \tag{12}$$

This corresponds directly to Landau's asymptotic formula (4).   Unfortunately, the probabilistic method is not refined enough to provide the correct value of the constant. Essentially, this is because the prime products diverge, and only the leading-order asymptotic form of $\ln \overline{B_2}(n)$ can be recovered as $n \to \infty$. However, if we take this value as given, then the constants that appear in the asymptotics of all other related quantities may be obtained by normalizing with respect to $\overline{B_2}(n)$. As an example, we will now obtain using this approach the corresponding results for $\overline{r_2}(n)$ and $\overline{r_2^2}(n)$. Then, in subsequent sections, we will show how the constants that appear in the associated autocorrelation functions can be recovered in the same way.

Since the primes (and in particular the p and r primes) are to be treated as being statistically independent, we have from (7) that

$$\overline{r_2}(n) \simeq \overline{B_2}(n) \times \prod_p \left[\sum_{m=0}^{\infty} (m+1) \frac{1}{p^m} \left(1 - \frac{1}{p} \right) \right]. \tag{13}$$

The product and sum converge rapidly and hence each may be taken to infinity. Substituting the expression (11) for $\overline{B_2}(n)$ into this, we have

$$\overline{r_2}(n) \simeq \prod_r \frac{1}{\left(1 + \frac{1}{r}\right)} \prod_p \frac{1}{\left(1 - \frac{1}{p}\right)} \tag{14}$$

$$= L(1) = \frac{\pi}{4}$$

where $L(s)$ is Landau's L-function. This corresponds exactly to the asymptotic result usually proved using lattice-point counting methods.

As another example, we now compute $\overline{r_2^2}(n)$ in the same way. Clearly we have that

$$\overline{r_2^2}(n) \simeq \overline{B_2}(n) \times \prod_p \left[\sum_{m=0}^{\infty} (m+1)^2 \frac{1}{p^m} \left(1 - \frac{1}{p} \right) \right] \tag{15}$$

where now the divergence of the product over p is not cancelled by that in $\overline{B_2}$. For this reason, it is more convenient to consider

$$\overline{r_2^2}(n) \overline{B_2}^2(n) \simeq \overline{B_2}^3(n) \times \prod_p \frac{\left(1 - \frac{1}{p^2}\right)}{\left(1 - \frac{1}{p}\right)^3} \tag{16}$$

$$\simeq \left[\prod_r \frac{1}{1 + \frac{1}{r}} \prod_p \frac{1}{1 - \frac{1}{p}} \right]^3 \prod_p \left(1 - \frac{1}{p^2} \right)$$

since now all prime products converge, and so the leading-order asymptotics is completely determined by the corresponding expression for $\overline{B_2}$. Specifically, recognizing $L(1)$ again and substituting (4),

$$\overline{r_2^2}(n) \simeq \left(\frac{\sqrt{\ln n}}{C_2} \right)^2 \left(\frac{\pi}{4} \right)^3 \prod_p \left(1 - \frac{1}{p^2} \right). \tag{17}$$

Hence, recalling the explicit expression (5) for C_2 and combining this with the product over p-primes above, the result is $\frac{2}{3} \times \zeta(2)^{-1}$, where $\zeta(s)$ is the Riemann zeta-function. Finally, using $\zeta(2) = \pi^2/6$ gives

$$\overline{r_2^2}(n) \simeq \frac{\pi}{4} \ln n. \tag{18}$$
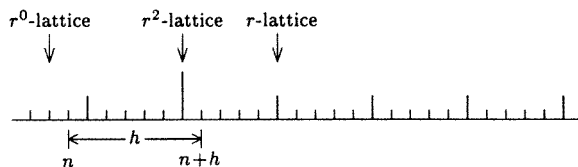
r^0-lattice r^2-lattice r-lattice

Figure 1. The r-lattices

3. Two-point correlations

We are now in a position to use the probabilistic methods developed in the previous section to compute the leading-order asymptotics of the correlation functions $\overline{B_2(n)B_2(n+h)}$ and $\overline{r_2(n)r_2(n+h)}$. Our approach will be to consider the contributions from each prime divisor of h separately. First, we need to ensure that n and $n+h$ have representations as sums of squares. This obviously corresponds to the requirement that every r-prime occurs to an even power in the factorizations of both integers. The probability of this happening is then equal to $\overline{B_2(n)B_2(n+h)}$. Next, we go on to consider the p-primes, since the powers to which these occur, on average, in n and $n+h$ determine $\overline{r_2(n)r_2(n+h)}$.

If every r-prime occurs to an even power in both n and $n+h$, the contribution to both $B_2(n)B_2(n+h)$ and $r_2(n)r_2(n+h)$ is 1, otherwise it is 0. In averaging with respect to n we need to evaluate the probabilities associated with these events. A useful way to think about this is in terms of 'r-lattices'. The r-lattice itself we defined to be the set of integers divisible by r. In the same way, the r^l-lattice consists of all the integers divisible by r^l, and hence is contained within the r^j-lattice for all $j < l$.

Consider now the two points n and $n+h$ with n free to move through $\mathbb{N}$. It may be helpful to think of a marker of length h moving over $\mathbb{N}$, with $\mathbb{N}$ displayed in terms of the r-lattices, as shown in figure 1. For a non-zero contribution to the correlation functions, we require that both ends of the marker rest on an r^l-lattice with l even.

The first case we shall compute is when $m_r(h) = 2k$. Then for each j in the range $0 \leqslant j \leqslant 2k$, $r^j | n$ implies $r^j | (n+h)$, and so if one end of the h-marker lies on the r^j-lattice then the other does as well. We shall refer to this by saying that the h-marker connects points on the r^j-lattice. This is not true for $j > 2k$.

We now evaluate the mean contribution via an inclusion/exclusion expansion. To first approximation, both ends of the h-marker will lie on the r^0-lattice and the overall contribution is 1. The leading correction to this involves taking into account the possibility of either end of the marker lying on the r-lattice. The probability of one end lying on the r-lattice is $\frac{1}{r}$, and there are two ends but, when $r|h$, (i.e. when $k \geqslant 1$), both ends will hit the r-lattice together. The term we must subtract is therefore $\frac{1}{r}$. If $k = 0$ only one end of the marker can lie on the r-lattice and so then we must subtract twice this value. We now add the probability that an end lies on the r^2-lattice, which will be $\frac{1}{r^2}$ for $k \geqslant 1$, since then both ends hit this lattice together. If $k = 0$ it is again twice this value. One can continue in this way, considering each lattice in turn. For lattices r^j with $j \leqslant 2k$ the contribution at each stage is $\frac{1}{r^j}$; and if $j > 2k$ it is $\frac{2}{r^j}$. The overall contribution for $m_r(h) = 2k$ is therefore

$$1 - \frac{1}{r} + \frac{1}{r^2} - \cdots + \frac{1}{r^{2k}} - \frac{2}{r^{2k+1}} + \frac{2}{r^{2k+2}} - \cdots = \frac{1 - \frac{1}{r^{2k+1}}}{1 + \frac{1}{r}}. \tag{19}$$

Next we consider the case when $m_r(h) = 2k + 1$. The h-marker connects all lattices r^j for $j \leqslant 2k + 1$. For all higher power lattices, if one end of the marker hits an even power

lattice, the other must lie on an odd power lattice; hence r occurs to an odd power in either n or $n + h$ and so there is no contribution. For this case we thus have

$$1 - \frac{1}{r} + \frac{1}{r^2} - \cdots + \frac{1}{r^{2k}} - \frac{1}{r^{2k+1}} = \frac{1 - \frac{1}{r^{2k+2}}}{1 + \frac{1}{r}}. \tag{20}$$

It therefore follows from the last two results that for all values of k, if $m_r(h) = k$ the contribution to both $\overline{B_2(n)B_2(n+h)}$ and $\overline{r_2(n)r_2(n+h)}$ is

$$\frac{1 - \frac{1}{r^{k+1}}}{1 + \frac{1}{r}}. \tag{21}$$

The final factor that must be taken into account in the calculation of $\overline{B_2(n)B_2(n+h)}$ is based on the following observation. For all r-primes, $r^{2\beta} = 1 \bmod 4$ when $\beta \in \mathbb{N}$. And for a sum of two squares, all such primes must occur to even powers in the factorization of n. In addition, for all p-primes $p^\beta = 1 \bmod 4$ for all β. It therefore follows that if n has a representation as a sum of two squares, then

$$n = 2^{m_2(n)} n_1 \qquad \text{with } n_1 = 1 \bmod 4. \tag{22}$$

This means that any representable integer n satisfies $n \bmod 4 \in \{0, 1, 2\}$. We must therefore take into account how the representable integers are distributed $\bmod 2^k$, for each k, and, in particular, account for the distribution of pairs n and $n + h$.

The first point is that it is easily verified that all $k \geqslant m_2(n) + 1$ show the same pair distribution structure, and that these contain all the information about smaller values of k. The general case we therefore need to consider when $m_2(h) = k$ is the pattern for the integers modulo 2^{k+1}. With $n = 2^{m_2(n)} n_1$ and $n_1 = 1 \bmod 4$, as before, the possibilities are shown in table 1, where $\alpha_i = 0, 1, \ldots (2^{k-1-i} - 1)$, and integers with no representation have been omitted. W_i is the probability of $n \bmod 2^{k+1}$ taking the given form, usually $2^i + \alpha_i 2^{i+2}$, shown in the left-hand column. For example if $k = 1$, as already stated $n \bmod 4 \in \{0, 1, 2\}$. The associated probabilities may then be calculated from (8) to be $\{\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\}$.

The pairs $(n \bmod 2^{k+1}, [n + 2^k] \bmod 2^{k+1})$ available and their probabilities (the product of those given in table 1) are shown in table 2. The number of occurrences of each variety of pair is simply the number of values taken by α_i; the only exception is that in the pair $(2^{k-1}, 2^{k-1} + 2^k)$, the second element has no representation and so makes no contribution. The general weighting factor $W_{m_2(h)}$ is then the sum of the probabilities listed

Table 1. The distribution of representable integers modulo 2^{k+1}.

$n \bmod 2^{k+1}$	Condition	Probability
0	$m_2(n) \geqslant k+1$	$W_\infty = \frac{1}{2^{k+1}}$
$1 + \alpha_0 2^2$	$m_2(n) = 0$	$W_0 = \frac{1}{2^k}$
$2 + \alpha_1 2^3$	$m_2(n) = 1$	$W_1 = \frac{1}{2^k}$
$4 + \alpha_2 2^4$	$m_2(n) = 2$	$W_2 = \frac{1}{2^k}$
$\vdots$	$\vdots$	$\vdots$
$2^{k-2} + \alpha_{k-2} 2^k$	$m_2(n) = k-2$	$W_{k-2} = \frac{1}{2^k}$
2^{k-1}	$m_2(n) = k-1$	$W_{k-1} = \frac{1}{2^k}$
2^k	$m_2(n) = k$	$W_k = \frac{1}{2^{k+1}}$

Table 2. The distribution of pairs of representable integers.

$(n \bmod 2^{k+1}, [n + 2^k] \bmod 2^{k+1})$	Probability	Number of pairs
$(0, 2^k)$	$W_{\infty,k} = \left(\frac{1}{2^{k+1}}\right)^2$	1
$(1 + \alpha_0\, 2^2, 1 + \alpha_0\, 2^2 + 2^k)$	$W_{0,0} = \left(\frac{1}{2^k}\right)^2$	2^{k-1}
$(2 + \alpha_1\, 2^3, 2 + \alpha_1\, 2^3 + 2^k)$	$W_{1,1} = \left(\frac{1}{2^k}\right)^2$	2^{k-2}
$(4 + \alpha_2\, 2^4, 4 + \alpha_2\, 2^4 + 2^k)$	$W_{2,2} = \left(\frac{1}{2^k}\right)^2$	2^{k-3}
$\vdots$	$\vdots$	$\vdots$
$(2^{k-2} + \alpha_{k-2}\, 2^k, 2^{k-2} + \alpha_{k-2}\, 2^k + 2^k)$	$W_{k-2,k-2} = \left(\frac{1}{2^k}\right)^2$	2
$(2^{k-1}, 2^{k-1} + 2^k)$	$W_{k-1,k-1} = 0$	1
$(2^k, 0)$	$W_{k,\infty} = 1$	1

above, multiplied by 2^{k+1}, the total number of pairs; that is,

$$W_{m_2(h)} = \begin{cases} 1 & \text{if } m_2(h) = 0 \\[2mm] \dfrac{2^{m_2(h)+1} - 3}{2^{m_2(h)}} & \text{if } m_2(h) \geqslant 1 . \end{cases} \tag{23}$$

We are now in a position to write down the result for $\overline{B_2(n)\,B_2(n+h)}$. The key point is that to obtain a convergent quantity it is natural to divide by $\overline{B_2}^2(n)$, for which we can use (4). Then

$$\frac{\overline{B_2(n)\,B_2(n+h)}}{\overline{B_2}^2(n)} = W_{m_2(h)} \prod_r \left(1 + \frac{1}{r}\right)^2 \prod_r \frac{1 - \frac{1}{r^{m_r(h)+1}}}{1 + \frac{1}{r}}$$

$$= W_{m_2(h)} \prod_r \left(1 + \frac{1}{r}\right)^2 \prod_{r \nmid h} \frac{1 - \frac{1}{r}}{1 + \frac{1}{r}} \prod_{r \mid h} \frac{1 - \frac{1}{r^{m_r(h)+1}}}{1 + \frac{1}{r}}$$

$$= W_{m_2(h)} \prod_r 1 - \frac{1}{r^2} \prod_{r \mid h} \frac{1 - \frac{1}{r^{m_r(h)+1}}}{1 - \frac{1}{r}} .$$

Here all prime-products converge. Now substituting in the asymptotic formula (4), we have that

$$\overline{B_2(n)\,B_2(n+h)} = \frac{1}{2} W_{m_2(h)} \left(\frac{\overline{B_2}(n)}{C_2}\right)^2 \prod_{r \mid h} \frac{1 - \frac{1}{r^{m_r(h)+1}}}{1 - \frac{1}{r}}$$

$$\simeq \frac{1}{2 \ln n} W_{m_2(h)} \prod_{r \mid h} \frac{1 - \frac{1}{r^{m_r(h)+1}}}{1 - \frac{1}{r}} . \tag{24}$$

As expected, in the semiclassical limit, as $n \to \infty$ this correlation function tends to zero for all fixed h, because the spacing between the levels grows as n increases.

Since the probabilistic approach used above is essentially heuristic, we now compare the results with those of a numerical computation. In the following we fix h and count the number of pairs $(n, n + h)$ for which each integer in the pair has a representation as a sum of two squares, with n averaged over the range $n \in \{10^7 - 10^5, 10^7 + 10^5\}$. We then take $\overline{B_2(n)\,B_2(n+h)}$ to be the ratio of the number of pairs of representable integers to the total number of pairs tested. Figure 2 shows the ratio of the numerical results obtained to the
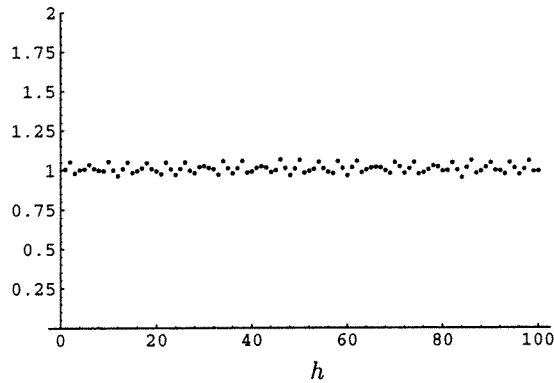
Figure 2. Ratio of (numerics)/(theory) for $\overline{B_2(n)B_2(n+h)}$.

values predicted by directly evaluating the prime product in (24). The value of $\overline{B_2}(n)$ is taken from the Landau result evaluated at $n = 10^7$ (the numerically calculated probability for $\overline{B_2}(n)$ is the same). The good agreement clearly provides support for the approach taken. Interestingly, the ratio appears to have a small oscillatory component of period 12 in h. We suspect that this is related to subdominant contributions to the asymptotics (4) for $\overline{B_2}(n)$, because, as will be seen later, this does not appear to be the case for $\overline{r_2(n)r_2(n+h)}$.

In order to extend the above calculation to $\overline{r_2(n)r_2(n+h)}$, the contributions from the p-primes must now be evaluated. Our approach remains the same: calculating the probability of each possible contribution, which in this case may be obtained from (7); that is, if one end of the h-marker hits the p^k-lattice, the contribution will be $(k+1)$.

Let $m_p(h) = k$. The lattices up to p^k are then connected by the h-marker. We again make an inclusion/exclusion expansion. The first term corresponds to the situation where both ends lie on the p^0-lattice and so the contribution is $(0+1) = 1$. The probability of one end hitting the p-lattice is $\frac{1}{p}$ when $k \geqslant 1$. When this occurs, we subtract the contribution which came from assuming both ends hit p^0 and add to it the new contribution with weighting $(1+1)^2 = 4$, since both ends lie on the p-lattice together. For lattices associated with higher powers, when one end lies on p^j $(j > k)$ the other lies on p^k. In this way we obtain the total contribution as follows:

$$1 + \frac{1}{p}(-1+4) + \frac{1}{p^2}(-4+9) + \frac{1}{p^3}(-9+16) + \cdots + \frac{1}{p^k}(-k^2 + (k+1)^2)$$

$$+ \frac{2}{p^{k+1}}(-(k+1)^2 + (k+1)(k+2))$$

$$+ \frac{2}{p^{k+2}}(-(k+1)(k+2) + (k+1)(k+3)) + \cdots$$

$$= \sum_{j=0}^{k} \frac{2j+1}{p^j} + \frac{2(k+1)}{p^k}\left(\frac{1}{1-\frac{1}{p}}\right) = \frac{\left(1+\frac{1}{p}\right)}{\left(1-\frac{1}{p}\right)^2}\left(1-\frac{1}{p^{k+1}}\right). \qquad (25)$$

Now taking the product over all p-primes, the total contribution may be written in the

form

$$\prod_p \frac{\left(1 + \frac{1}{p}\right)}{\left(1 - \frac{1}{p}\right)^2}\left(1 - \frac{1}{p^{m_p(h)+1}}\right) = \prod_p \frac{1 + \frac{1}{p}}{1 - \frac{1}{p}}\prod_{p|h}\frac{1 - \frac{1}{p^{m_p(h)+1}}}{1 - \frac{1}{p}}.$$ (26)

This factor can then be combined with the result for $\overline{B_2(n)B_2(n+h)}$ to give

$$\overline{r_2(n)r_2(n+h)} = W_{m_2(h)}\prod_r \frac{1 - \frac{1}{r}}{1 + \frac{1}{r}}\prod_p \frac{1 + \frac{1}{p}}{1 - \frac{1}{p}}\prod_{r|h}\frac{1 - \frac{1}{r^{m_r(h)+1}}}{1 - \frac{1}{r}}\prod_{p|h}\frac{1 - \frac{1}{p^{m_p(h)+1}}}{1 - \frac{1}{p}}$$

$$= W_{m_2(h)}L^2(1)\prod_{q>2}\left(1 - \frac{1}{q^2}\right)\prod_{\substack{q|h \\ q>2}}\frac{1 - \frac{1}{q^{m_q(h)+1}}}{1 - \frac{1}{q}}.$$

Finally, using $L(1) = \pi/4$, we have

$$\overline{r_2(n)r_2(n+h)} = \frac{1}{2}W_{m_2(h)}\prod_{\substack{q|h \\ q>2}}\frac{1 - \frac{1}{q^{m_q(h)+1}}}{1 - \frac{1}{q}}$$ (27)

where the q-product includes all odd prime divisors of h. This represents our main result. It holds for all $h \neq 0$. The $h = 0$ result was calculated earlier (18); its divergence as $n \to \infty$ corresponds to the δ-function at the origin of the correlation function.

The consistency of this result can be checked by noting that averaging over a large range of values of h must give

$$\langle \overline{r_2(n)r_2(n+h)}\rangle_h = \bar{r}_2^2(n) = \left(\frac{\pi}{4}\right)^2.$$ (28)

The average of (27) may be calculated directly using (8), which represents the probability that $m_q(h) = m$. The result is

$$\langle \overline{r_2(n)r_2(n+h)}\rangle_h = \frac{1}{2}\left[\frac{1}{2} + \sum_{j=1}^{\infty}\left(\frac{2^{j+1} - 3}{2^j}\right)\frac{1}{2^{j+1}}\right]$$

$$\times \prod_{q>2}\frac{1}{1 - \frac{1}{q}}\left[1 - \sum_{k=0}^{\infty}\frac{1}{q^k}\left(1 - \frac{1}{q}\right)\frac{1}{q^{k+1}}\right]$$

where the cases $m_2(h) = 0$ and $m_2(h) \geqslant 1$ have been considered separately in $W_{m_2(h)}$. Evaluating the sums and rearranging, we then find

$$\langle \overline{r_2(n)r_2(n+h)}\rangle_h = \frac{1}{2}\prod_{q>2}\left(1 - \frac{1}{q^2}\right)^{-1}$$

$$= \left(\frac{\pi}{4}\right)^2$$ (29)

as required.

A further check may be made by numerical computation, as in the previous section for $\overline{B_2(n)B_2(n+h)}$. For each pair of representable integers found in that case, we calculate the number of representations using (7). In this way we can compute $\overline{r_2(n)r_2(n+h)}$ and compare the result with (27). The theoretical data from (27) are shown in figure 3, and the ratio with the numerically calculated values is shown in figure 4 (note the change in scale). The agreement (to within 2%) again supports the use of the probabilistic methods employed. In particular, the ratio shows no discernible, systematic dependence upon h, suggesting that we have captured the correct number-theoretical form of the fluctuations.
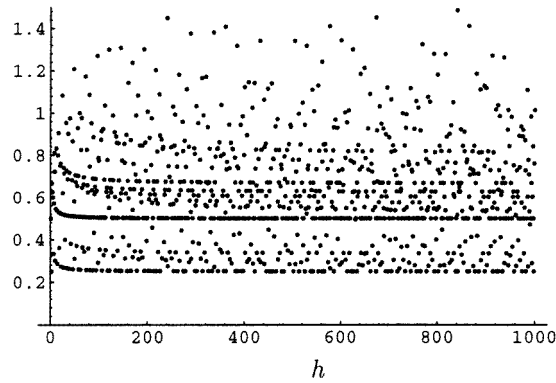
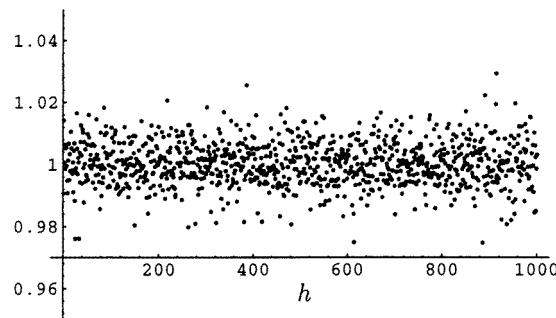**Figure 3.** Plot of $\overline{r_2(n)r_2(n+h)}$ from (27).

Figure 4. Ratio of (numerics)/(theory) for $\overline{r_2(n)r_2(n+h)}$.

It is noticeable that the agreement between the analytical result and the numerics is closer for $\overline{r_2(n)r_2(n+h)}$ than for $\overline{B_2(n)B_2(n+h)}$. This is probably because for B_2 the correlation function depends upon the typical size of n in the averaging range, whereas for r_2 it does not: in this case there is a non-zero limiting distribution.

4. Local smoothing

The result (27) for $\overline{r_2(n)r_2(n+h)}$ incorporates all of the number-theoretical fluctuations associated with the dependence on the prime factorization of h. It was shown at the end of the last section that if the correlation function is averaged over a large range of values of h, the underlying mean corresponds correctly to $\overline{r_2}^2$. Our aim now is to perform this calculation more carefully in order to see how this limit is approached, and hence to characterize the average size of the fluctuations.

To understand the asymptotic dependence on the size of the range, we define a local smoothing by

$$\langle \overline{r_2(n)r_2(n+h)} \rangle_H = \frac{\mathrm{d}}{\mathrm{d}H} \sum_{h=1}^{H} \overline{r_2(n)r_2(n+h)}. \tag{30}$$

By this we mean that the average is the derivative with respect to H of the smooth function that represents the leading-order asymptotics for the sum as $H \to \infty$. It may therefore

be interpreted as the smooth function which when integrated up to H coincides with the asymptotic approximation to the sum.

5. Local Smoothing

The result (27) for $\overline{r_2(n)r_2(n+h)}$ incorporates all of the number-theoretical fluctuations associated with the dependence on the prime factorization of h. It was shown at the end of the last section that if the correlation function is averaged over a large range of values of h, the underlying mean corresponds correctly to $\overline{r_2}^2$. Our aim now is to perform this calculation more carefully in order to see how this limit is approached, and hence to characterize the average size of the fluctuations.

To understand the asymptotic dependence on the size of the range, we define a local smoothing by

$$\langle r_2(n)r_2(n+h)\rangle_H = \frac{\mathrm{d}}{\mathrm{d}H} \sum_{h=1}^{H} \overline{r_2(n)r_2(n+h)}. \tag{31}$$

By this we mean that the average is the derivative with respect to H of the smooth function that represents the leading-order asymptotics for the sum as $H \to \infty$. It may therefore be interpreted as the smooth function which when integrated up to H coincides with the asymptotic approximation to the sum.

We begin by observing that

$$\prod_{\substack{q|h \\ q>2}} \frac{1 - \frac{1}{q^{m_q(h)+1}}}{1 - \frac{1}{q}} = \prod_{\substack{q|h \\ q>2}} \left(1 + \frac{1}{q} + \frac{1}{q^2} + \cdots + \frac{1}{q^{m_q(h)}}\right). \tag{32}$$

This product gives a sum over all the odd divisors of h, that is, all of the divisors of h_o (including 1 and h_o itself), where h_o is the odd part of h, defined by $h = 2^{m_2(h)}h_o$:

$$\prod_{\substack{q|h \\ q>2}} \frac{1 - \frac{1}{q^{m_q(h)+1}}}{1 - \frac{1}{q}} = \sum_{d|h_o} \frac{1}{d}. \tag{33}$$

We now have

$$\langle r_2(n)r_2(n+h)\rangle_H = \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}H} \sum_{h=1}^{H} W_{m_2(h)} \sum_{d|h} \frac{\delta_o(d)}{d} \tag{34}$$

where $\delta_o(d)$ is defined by

$$\delta_o(d) = \begin{cases} 1 & \text{if } d \text{ odd} \\ 0 & \text{if } d \text{ even.} \end{cases} \tag{35}$$

We shall use $\delta_e(d)$ to denote the equivalent function, selecting even values of d.

Including the explicit form of $W_{m_2(h)}$, we define $s(H)$ as follows,

$$s(H) = \sum_{h=1}^{H} \left\{ \begin{array}{ll} 1 & \text{if } m_2(h) = 0 \\ 2 - \frac{3}{2^{m_2(h)}} & \text{if } m_2(h) \geqslant 1 \end{array} \right\} \sum_{d|h} \frac{\delta_o(d)}{d} \tag{36}$$

which can be rewritten

$$s(H) = \sum_{h=1}^{H} 2 \sum_{d|h} \frac{\delta_o(d)}{d} - \sum_{h=1}^{H} \delta_o(h) \sum_{d|h} \frac{\delta_o(d)}{d} - \sum_{h=1}^{H} \frac{3\delta_e(h)}{2^{m_2(h)}} \sum_{d|h} \frac{\delta_o(d)}{d}$$

$$= 2 \sum_{h=1}^{H} \sum_{d|h} \frac{\delta_o(d)}{d} - \sum_{h=1}^{H} \delta_o(h) \sum_{d|h} \frac{\delta_o(d)}{d} - 3 \sum_{j=1}^{J} \frac{1}{2^j} \sum_{\substack{h=1 \\ m_2(h)=j}}^{H} \sum_{d|h} \frac{\delta_o(d)}{d}$$

where $J = [\frac{\log H}{\log 2}]$, the largest value of $m_2(h)$ for $h \leqslant H$, and the square brackets denote the integer part of the argument.

As h ranges from 1 to H, each integer $1 \leqslant d \leqslant H$ will occur as a divisor. The number of such occurrences for a given d will be $[\frac{H}{d}]$.

Considering only odd integers $h \leqslant H$, restricts us to counting the number of odd multiples of d less than H. This is simply $[\frac{H}{2d} + \frac{1}{2}]$, the largest integer k such that $(2k-1)d \leqslant H$. Similarly, the number of occurrences of d as a divisor of $1 \leqslant h \leqslant H$ when h is of the form $h = 2^j h_o$ is $[\frac{H}{2^{j+1}d} + \frac{1}{2}]$.

We can now write $s(H)$ as

$$s(H) = \sum_{d=1}^{H} \frac{\delta_o(d)}{d} \left(2 \left[\frac{H}{d} \right] - \left[\frac{H}{2d} + \frac{1}{2} \right] - 3 \sum_{j=1}^{J} \frac{1}{2^j} \left[\frac{H}{2^{j+1}d} + \frac{1}{2} \right] \right). \quad (37)$$

Let $\{x\}$ denote the fractional part of x; that is, $\{x\} = x - [x]$. When H is large, it is reasonable to assume that the mean value of $\{\frac{H}{d}\}$ is $\frac{1}{2}$ (and this can be proved to give the correct leading order approximation to the sum). Approximating the fractional parts of all quantities in this way we have

$$s(H) = \sum_{d=1}^{H} \frac{\delta_o(d)}{d} \left(\frac{2H}{d} - 1 - \frac{H}{2d} - 3 \sum_{j=1}^{J} \frac{1}{2^{2j+1}} \frac{H}{d} \right). \quad (38)$$

Evaluating the j-sum (to infinity, incurring an $O(H^{-2})$ error) leaves

$$s(H) = \sum_{d=1}^{H} \delta_o(d) \left(\frac{H}{d^2} - \frac{1}{d} \right)$$

$$\simeq H \left[\sum_{n=1}^{\infty} \frac{1}{n^2} - \sum_{n=1}^{\infty} \frac{1}{(2n)^2} \right] - \left[\sum_{n=1}^{H} \frac{1}{n} - \sum_{n=1}^{[\frac{H}{2}]} \frac{1}{2n} \right] \quad (39)$$

where the convergent sums are taken to infinity. A direct evaluation thus gives that to leading order

$$s(H) \simeq \frac{\pi^2}{8} H - \frac{1}{2} \ln H. \quad (40)$$

Recalling (34), we have

$$\overline{\langle r_2(n) r_2(n+h) \rangle}_H = \frac{1}{2} \frac{d}{dH} s(H) \quad (41)$$

and so with $s(H)$ calculated as above,

$$\overline{\langle r_2(n) r_2(n+h) \rangle}_H \simeq \frac{d}{dH} \left[\left(\frac{\pi}{4} \right)^2 H - \frac{1}{4} \ln H \right] \quad (42)$$

$$\simeq \left(\frac{\pi}{4} \right)^2 - \frac{1}{4H} \quad (43)$$

in agreement with the result for $\bar{r}_2(n)$ in the limit $H \to \infty$. Finally, normalizing by $\bar{r}_2^2$, we have that to leading order when H is large

$$\frac{\overline{\langle r_2(n) r_2(n+h) \rangle}_H}{\bar{r}_2^2(n)} \simeq 1 - \frac{4}{\pi^2 H}. \quad (44)$$
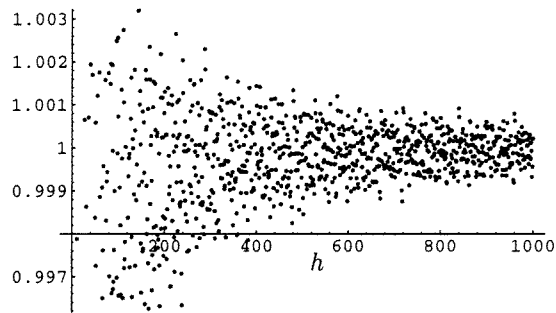
Figure 5. Ratio of numerics/theory for $\langle r_2(n)r_2(n+h)\rangle_H$.

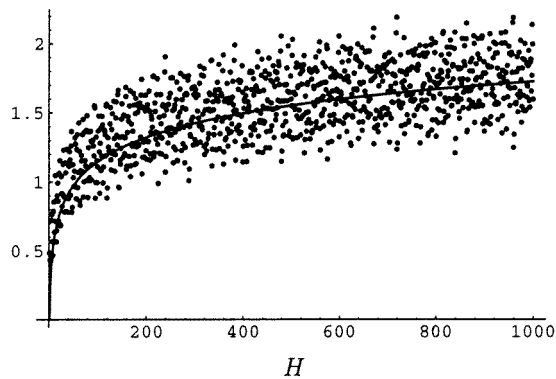**Figure 6.** Plot of $f(H)$ and $\frac{1}{4}\log H$.

Thus, the fluctuation contribution tends to zero as $H \to \infty$. For a truly Poissonian spectrum this term would be identically zero; that is, the normalized correlation function would have a value of 1 for all h. It thus represents a small correction to randomness in the two-point correlations. This is strikingly similar to the behaviour that follows from the Hardy–Littlewood conjecture for the primes [10, 14], which is known to be responsible for the GUE statistics of the Riemann zeros [7, 8].

From the same numerical data used to verify the two point results for B_2 and r_2, we can also verify the predicted behaviour of $\langle r_2(n)r_2(n+h)\rangle_H$. First, the ratio of the numerical results to asymptotic expression (43) is shown in figure 5. To see the second term directly, we also plot

$$f(H) = \frac{\pi^2}{16}H - \sum_{h=1}^{H} \overline{r_2(n)r_2(n+h)} \tag{45}$$

in figure 6, and on the same graph show the curve of $\frac{1}{4}\log H$ which, according to (42), this quantity should be asymptotically described by. The results clearly confirm the presence of the small correction to the Poissonian form.

References

[1] Berry M V 1981 *Ann. Phys.* **131** 163–216
[2] Berry M V 1985 *Proc. R. Soc.* A **400** 229–51

[3] Berry M V 1987 *Proc. R. Soc.* A **413** 183–98

[4] Berry M V and Tabor M 1977 *Proc. R. Soc.* A **356** 375–94

[5] Bleher P M, Dyson F J and Lebowitz J L 1993 *Phys. Rev. Lett.* **71** 3047–50

[6] Bohigas O 1991 Random matrix theories and chaotic dynamics *Les Houches Lecture Series* vol 52, ed M-J Giannoni, A Voros and J Zinn-Justin (Amsterdam: North-Holland) pp 89–199

[7] Bogomolny E B and Keating J P 1995 *Nonlinearity* **8** 1115–31

[8] Bogomolny E B and Keating J P 1996 *Nonlinearity* **9** 911–35

[9] Grosswald E 1985 *Representations of Integers as Sums of Squares* (Berlin: Springer)

[10] Hardy G H and Littlewood J E 1923 *Acta Math.* **44** 1–70

[11] Hardy G H and Wright E M 1984 *An Introduction to the Theory of Numbers* 5th edn (Oxford: Oxford University Press)

[12] Kac M 1959 *Statistical Independence in Probability Theory, Analysis and Number Theory* (Mathematical Association of America)
Kac M 1959 *Probability and Related Topics in Physical Sciences* (New York: Interscience)

[13] Keating J P 1991 *Nonlinearity* **4** 277–307

[14] Keating J P 1993 The Riemann zeta function and quantum chaology *Quantum Chaos* ed G Casati, I Guarneri and U Smilansky (Amsterdam: North-Holland) pp 145–85.

[15] Landau E 1908 *Archiv der Math. und Physik III* **13** 305–15

[16] Major P 1992 *Probab. Theor. Relat. Fields* **92** 413

[17] Sarnak P Personal communication.

[18] Schroeder M R 1984 *Number Theory in Science and Communication* (Berlin: Springer)

[19] Sinai Ya G (ed) 1991 *Advances in Soviet Mathematics* vol 3 (Providence, RI: AMS)