Periodic Orbits, Spectral Statistics and the Riemann Zeros

J.P. Keating*
Basic Research Institute in the Mathematical Sciences
HP Laboratories Bristol
HPL-BRIMS-98-06
February, 1998

semiclassical approximation, quantum chaos, spectral statistics, the Riemann zeros

I review recent developments in the semiclassical theory of spectral statistics based on the trace formula. Applications to the pair-correlation of the Riemann zeros are also discussed.

Internal Accession Date Only
*School of Mathematics, University Walk, Bristol, United Kingdom

PERIODIC ORBITS, SPECTRAL STATISTICS, AND THE RIEMANN ZEROS

J P Keating

School of Mathematics, University Walk,
Bristol BS8 1TW, U.K.,
and Basic Research Institute in the Mathematical Sciences,
Hewlett-Packard Laboratories Bristol, Filton Road,
Stoke Gifford, Bristol BS12 6QZ, U.K.

1. INTRODUCTION

My purpose in this article is to review the background to some recent developments in the semiclassical theory of spectral statistics. Specifically, I will concentrate on approaches based on the trace formula[1,2]; that is, on the link between quantum energy levels and classical periodic orbits. I will also review the closely related theory of the statistics of the zeros of the Riemann zeta function. My hope is to provide an introduction to the introductions of other papers in this volume on the same subjects, and with this in mind will discuss only in outline calculations to be described by them in greater detail.

The statistical properties we seek to understand concern fluctuations in the distribution of the quantum energy levels of a given system in the semiclassical limit. It has been conjectured that these fluctuations are, in this limit, universal, and depend only upon the chaotic nature and symmetries of the system's classical dynamics. For example, Berry and Tabor[3] proposed that the energy levels of classically integrable systems are generically uncorrelated in the semiclassical limit. They also suggested that for classically chaotic systems the levels might be correlated in the same way as the eigenvalues of random matrices. This was confirmed numerically for the distribution of level-spacings in a number of such systems[4,5] and put into the form of an explicit conjecture for all spectral statistics by Bohigas, Giannoni and Schmit[6], who also made detailed studies of several examples. The random matrix conjecture has led to extensive numerical investigations, the results of which may be found in the reviews by Berry[7] and Bohigas[8], to name but two.

One of the goals of the work to be reviewed here is to develop a theory that explains how universality arises in spectral statistics. This is complicated by the fact that some systems exhibit decidedly nonuniversal behaviour. For example, the cat maps are maximally chaotic, but their quantum spectra do not show any signs of being random-matrix correlated[9]. The same is also true for geodesic motion on compact surfaces of constant negative curvature associated with arithmetic groups[10]. Much more is known about this in the case of integrable systems. The harmonic oscillator is an obvious example for which the quantum levels are correlated[3]. Others include rectangular billiards in which the square of the aspect ratio is rational[3,11]. More surprising is the fact[12] that a large class of rectangular billiards for which this number is irrational have also been shown not to have a Poissonian limit, although another large class does. Any theory which hopes to explain universality must also account for these exceptional cases as well.

Another goal is to describe the non-universal deviations from the Poisson and random-matrix forms that occur before the semiclassical limit is reached (and which the conjecture implies disappear in the limit). Put another way, this would be a description of the asymptotic way in which the conjectured limits are approached.

Several theories have been developed to answer these questions. Of these, two are explicitly semiclassical, one being based on the trace formula and the other on field theory[13]. The trace formula was first used by Berry and Tabor[3] to show how the Poisson limit emerges for generic integrable systems. Hannay and Ozorio de Almeida[14] and Berry[15] then extended the approach to recover two-point random matrix correlations for chaotic systems. A key element of their work was the realization that in ergodic systems certain periodic orbit contributions (the *diagonal* terms - see Section 5) can be evaluated using a sum rule (now known as the *Hannay-Ozorio de Almeida sum rule*). Furthermore, Berry[15] also showed how the diagonal terms associated with short orbits can be used to describe some features of the nonuniversal approach to the random-matrix limit as $\hbar \to 0$. These methods were subsequently extended to include, for example, parametric correlations[16,17] and matrix element distributions[18,19]. They also generalize in a trivial way to quantum maps.

Going beyond the diagonal approximation means evaluating the *off-diagonal* terms (Section 6). To do this directly would require more knowledge about correlations between different periodic orbits than we possess at present. However, under certain assumptions, one can compute the off-diagonal contribution indirectly, by relating it to the diagonal terms[20]. This connection is very similar, but not identical to the one which exists between the perturbative and non-perturbative contributions to spectral correlation functions in disordered systems[21,22]. One of the aims of this article to discuss the similarities and differences.

Another aim is to review the links[23,24,25] between the theory of spectral statistics and the statistical distribution of the zeros of the Riemann zeta function (Section 7). The reason for doing this here is that, first, Montgomery[26] has conjectured that the Riemann zeros are correlated like the eigenvalues of matrices in the Gaussian Unitary Ensemble (GUE) of random matrices, and second, there exists a formula relating the zeros to the prime numbers that is the exact analogue of the trace formula. The problem of understanding the statistics of energy levels using periodic orbits is thus identical to that of understanding the statistics of the Riemann zeros in terms of the primes. For the zeros, the analogues of the diagonal periodic orbit terms can again be evaluated explicitly. Moreover, this is a case in which the off-diagonal contributions can be calculated directly as well[24,27,28,29], because we do, by virtue of certain celebrated conjectures due to Hardy and Littlewood[30], have a good understanding of the correlations that exist between the primes. The off-diagonal terms can of course also be calculated indirectly by relating them to the diagonal terms, as in the general semiclassical case, and the fact that these two independent approaches give the same answer represents an important test of the correctness of the indirect method.

2. PERIODIC ORBIT FORMULAE

The foundations of the theory of spectral statistics to be reviewed here rest on Gutzwiller's trace formula[1,2], which provides a semiclassical relationship between the density of states

$$d(E) = \sum_n \delta(E - E_n) \tag{1}$$

of a given quantum system and the periodic orbits of the underlying classical dynamics:

$$d(E) \sim \bar{d}(E) + \frac{1}{\pi\hbar} \mathrm{Re} \sum_p \sum_{n=1}^{\infty} \frac{T_p}{\left|\det\left(M_p^n - I\right)\right|^{\frac{1}{2}}} \exp\left(\frac{i}{\hbar} n S_p\right), \tag{2}$$

where $\bar{d}$ is the mean density; p labels primitive orbits and n their repetitions; T_p is the period of the pth orbit, S_p its action (defined here to include the Maslov index), and M_p the monodromy matrix that describes the flow linearized in its vicinity. Alternatively, this quantum-classical connection can be expressed as a formula for the spectral determinant

$$\Delta(E) = \det\left(E - \hat{H}\right), \tag{3}$$

which, when regularized to ensure convergence, has zeros at the energy levels (eigenvalues of the quantum Hamiltonian $\hat{H}$) E_n; for example, in systems with two degrees of freedom

$$\Delta(E) \sim B(E) \exp\left(-i\pi\bar{N}(E)\right) \prod_p \prod_{m=0}^{\infty} \left(1 - \frac{\exp\left(iS_p/\hbar\right)}{\left|\Lambda_p\right|^{\frac{1}{2}} \Lambda_p^m}\right) \tag{4}$$

where B is a function with no real zeros that is connected with the regularization, $\bar{N}$ is the mean of the eigenvalue counting function

$$N(E) = \int_0^E d(E') dE', \tag{5}$$

and Λ_p is the larger eigenvalue of M_p ($|\Lambda_p|>1$ in strongly chaotic systems - i.e., when all periodic orbits are isolated and unstable). Both equations, (2) and (4), have been written in a form appropriate for flows, but generalize immediately to quantum maps.

Loosely speaking, for most systems (2) and (4) are semiclassical approximations; that is, they represent the leading order asymptotics for d and Δ as $\hbar \to 0$. More precisely, this is the case for Δ, which is a smooth function, in the usual sense of the term. For the density, it is true in the sense that the positions of the singularities on the right-hand side of (2) are semiclassical approximations to the exact energy levels. This follows because these singularities coincide with the zeros of the function on the right-hand side of (4).

There are also systems for which both equations are exact equalities. One class of examples is provided by geodesic motion on compact surfaces of constant negative curvature. For these, (2) is known as the Selberg trace formula and the double product on the right-hand side of (4) is called the Selberg zeta function[31]. The quantum cat maps provide another class of examples[9]. In both of these cases, the classical dynamics is maximally chaotic, in the sense that the systems are Anosov.

When (2) and (4) are not exact, it has been found numerically that the approximation that they represent is good, in the sense that the positions of the singularities/zeros on the classical (right-hand) side are accurate approximations to the quantum energy levels on the scale of the mean level separation[32,33,34,35]. Specifically, it has been found in a wide class of examples, including integrable, chaotic, and mixed maps and flows in both two and three dimensions, that the root means square error in the difference between the positions of the zeros on the left and right-hand sides of (4) is a small fraction (typically less than 10%) of the mean level separation. Of course, one can in principle go further and achieve any desired accuracy by expanding to the appropriate order in $\hbar$ around the periodic orbits in the asymptotic evaluation of the path integral for the Green function that underlies the derivation of the trace formula.

The conclusion is that periodic orbit theory is able to reproduce a discrete spectrum which, on scales of the order of the mean level separation, represents a good approximation to the true quantum energy levels in systems (both maps and flows) whose classical dynamics is integrable, chaotic, or mixed; and that in principle the method can be extended to any desired order of accuracy.

3. SPECTRAL STATISTICS

In most physical problems one is less concerned with computing individual energy levels than with characterising their statistical distribution; that is, one wants to probe correlations in the spectrum of the quantum Hamiltonian in question. Studies of this problem have led to a number of interesting conjectures[3,6], which may be summarized as follows.

a. The statistical correlations in the spectrum of a single, typical, classically integrable system, measured on the scale of a fixed number of mean level spacings, are, in the semiclassical limit $\hbar \to 0$, Poissonian; that is, there are no correlations at all.

b. In typical classically chaotic systems, the corresponding correlations are, in the semiclassical limit, the same as those of the eigenvalues of random matrices. Specifically, the spectral correlations of time-reversal-symmetric systems correspond to those of random real symmetric matrices (the ensemble of which is denoted GOE), and the spectral correlations of non-time-reversal-symmetric systems correspond to those of random complex hermitian matrices (the ensemble of which is denoted GUE).

Measuring energy level correlations in a given system necessarily involves averaging some spectral function over a range of energy (or other suitable classical parameter). These conjectures then imply that for local statistics, the energy average is equivalent to an average over either the Poissonian ensemble, the GOE, or the GUE. Furthermore, since these ensemble averages know nothing about the details of the system in question, the conjectures imply that spectral statistics exhibit *universality*. It is, however, crucial to emphasize the there are exceptions (note the appearance of the word 'typical'). As already mentioned in the introduction, there are families of integrable systems whose spectral statistics are not Poissonian, and strongly chaotic systems whose spectral statistics are not random-matrix. Indeed, our current level of understanding is such that we cannot say what 'typical' (or equivalently 'generic') really means in this context, and so cannot make the conjectures precise in the mathematical sense (although in terms of physics they still have non-trivial content). These systems, whilst singular in some respects, serve to provide an important reminder that any proof of a conjecture like those discussed above must be based on more than just classical integrability or a measure of classical chaos, because these properties alone do not distinguish the exceptional cases.

The question then remains as to how to approach a deeper understanding of these issues. One can try to develop a theory based on averaging over a small family of quantum systems (i.e. much smaller than the spaces of random matrices), and such an idea will be described elsewhere in this volume[36]. Otherwise, one can stick to working with a single system (which is the essential element of the conjecture) and use the trace formula to relate the spectral statistics to the classical periodic orbits. More precisely, one may hope that the trace formula will provide a useful link between energy-level correlations and the statistical properties of classical orbits. The basis of this hope is the idea that if, as seems to be the case, the trace formula can be used to generate approximations to the energy levels in a given system that are accurate on the scale of the mean level separation, then it should also be able to describe the correlations between the levels on this scale as well. Put another way, if the approximate spectrum generated by the trace formula exhibits the same correlation statistics as the corresponding exact spectrum, then periodic orbits should be capable of describing these statistics.

This immediately suggests two questions: how can universality emerge from a trace formula that depends on periodic orbits which differ from system to system? And how does the individuality of these orbits influence spectral statistics? The answers lie in the different energy scales involved.

4. ENERGY SCALES

As a function of energy E, each periodic orbit contribution to the trace formula is locally periodic with period h/T_j, where T_j is the orbit period. This is to be compared with the mean level separation $\bar{d}^{-1}$, which Weyl's law implies is of the order of h^f in a system with f classical degrees of freedom. Thus the periodic orbits that contribute to correlations on the scale of $\bar{d}^{-1}$ have period $T \sim T_H = h\bar{d}$, where the Heisenberg time T_H is of the order of h^{1-f}. Clearly as the semiclassical limit is approached the periods of the orbits that determine spectral correlations on universal scales tend to infinity when $f>1$. Conversely, a given periodic orbit contributes to spectral correlations on the energy scale T_H/T_j in units of the mean level spacing.

As an example, consider the two-point correlation function

$$R_2(x) = \frac{1}{\bar{d}^2}\left\langle d(E)d\left(E + x/\bar{d}\right)\right\rangle_E, \tag{6}$$

and its Fourier transform, the form factor

$$K(\tau) = \int_{-\infty}^{\infty}\left(R_2(x) - 1\right)\exp(2\pi i x\tau)dx. \tag{7}$$

The limit (x,τ)-fixed, $\hbar \to 0$ is dominated by semiclassically long orbits, while the limits $x \to \infty$ and $\tau \to 0$ with $\hbar$-fixed are dominated by short orbits.

We can now invoke the following information about periodic orbits. In ergodic systems, long periodic orbits are asymptotically, in the limit $T_j \to \infty$, uniformly dense on the energy shell (the surface of constant energy in phase space) when weighted by their stabilities; that is, the periodic orbits asymptotically approximate the invariant density. Mathematically, this is equivalent to the Hannay-Ozorio de Almeida sum rule[14]:

$$\frac{1}{\Delta T}\sum_{T \le T_j \le T+\Delta T}\frac{f(T_j)}{\left|M_j - I\right|} \sim \frac{f(T)}{T} \tag{8}$$

as $T \to \infty$ and $\Delta T \to 0$.

The basic idea underlying uniformity is that the periodic orbits in chaotic systems are dense and so can be used to describe phase space structures. It follows from Weyl's law that $\bar{d}$ is proportional to the volume of the energy shell, and so measuring spectral correlation lengths in units of the mean level spacing, $\bar{d}^{-1}$, corresponds to normalising this volume. Thus the key point is *that the long periodic orbits of all ergodic systems look the same in these units*. One can view this as saying that universality in quantum level statistics is related to the universality of long periodic orbits in ergodic systems. A further implication is that the way in which universality is approached in spectral statistics as $\hbar \to 0$ is related to the rate of approach to ergodic uniformity in the underlying classical dynamics in the long-time limit, information about which is encoded in the short periodic orbits.

5. THE DIAGONAL APPROXIMATION

Substituting the trace formula (2) into (6) gives a semiclassical approximation to the two-point correlation function:

$$R_2(x) \approx 1 + \left\langle \frac{2}{T_H^2} \sum_p \sum_q \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{T_p T_q}{\left|M_p^m - I\right|^{\frac{1}{2}} \left|M_q^n - I\right|^{\frac{1}{2}}} \times \cos\left(\frac{mS_p - nS_q}{\hbar} + \frac{2\pi m T_p}{T_H} x \right) \right\rangle_E . \tag{9}$$

This can be split up into diagonal contributions $R_2^{(d)}$, for which $mS_p = nS_q$, and off-diagonal contributions $R_2^{(off)}$, for which $mS_p \neq nS_q$, so that

$$R_2(x) \approx 1 + R_2^{(d)}(x) + R_2^{(off)}(x). \tag{10}$$

Such a division is natural for two reasons. First, as functions of E, the diagonal terms are least oscillatory and so survive best the energy average in (9). Thus in appropriate regimes they may be expected to dominate the off-diagonal terms. Second, if the actions of different orbits (i.e. orbits not related by symmetry) are uncorrelated, then the off-diagonal contribution would vanish, being the average of a sum of terms with random phases. This observation will be of importance in the calculation of the off-diagonal contribution to be outlined in Section 6.

The diagonal contribution may be written in the form

$$R_2^{(d)}(x) = \frac{2}{T_H^2} \sum_p \sum_{m=1}^{\infty} g_p \frac{T_p^2}{\left|M_p^m - I\right|} \cos\left(\frac{2\pi m T_p}{T_H} x \right), \tag{11}$$

where g_p is the action-multiplicity of the p^{th} primitive orbit. The key question is how these multiplicities behave. When $x/T_H \to 0$ (as is the case in the limit $\hbar \to 0$ with x-fixed) the sum (11) is dominated by increasingly long orbits. For these, the multiplicity can, in typical systems, be replaced by its mean $\overline{g}$, which takes the values $\overline{g} = 1$ in non-time-reversal-symmetric systems and $\overline{g} = 2$ in the time-reversal-symmetric case, the difference being due to the symmetry between non-self-retracing orbits and their time-reversed twins. Then (11) can be evaluated using the Hannay-Ozorio de Almeida sum rule (8) to give

$$R_2^{(d)}(x) \approx -\frac{\overline{g}}{2\pi^2 x^2}, \tag{12}$$

which coincides precisely with the leading-order $x \to \infty$ asymptotics of the non-oscillatory (in x) contributions to the GUE (when $\overline{g} = 1$) and GOE (when $\overline{g} = 2$) results for the two-point correlation function of the eigenvalues of random matrices. This confirms the argument of the previous section that universality in spectral statistics, in this case represented by (12), is related to ergodic universality in the long periodic orbits of chaotic systems, here expressed through the Hannay-Ozorio de Almeida sum rule (8). Furthermore, it may be seen from (11) that when $x/T_H \to \infty$ the semiclassical approximation is governed by the short-time classical dynamics, and hence can be expressed in terms of the short periodic orbits, or equivalently, the decay of classical correlations.

It is also clear from the derivation of (12) that spectral universality is related to the behaviour of the multiplicities g_p for long orbits. Indeed, this is a centrally important issue, because in the strongly chaotic systems known to be exceptional in that their level statistics are not random-matrix, the mean multiplicity is not a constant but grows rapidly as a function of the period. Specifically, for geodesic motion on arithmetic surfaces of constant negative curvature and for the cat maps, both of which are fully ergodic, $\bar{g}$ grows with T as the square-root of the total number of periodic orbits of period T. Thus ergodicity alone is not enough to imply (12).

The distribution of orbit multiplities affects the non-universal regime too. As already argued, this regime is governed by the short periodic orbits for which, in time-reversal-symmetric systems, the multiplicties are typically erratic. Replacing g_p by its mean is then a highly questionable approximation.

The semiclassical structure of the two-point correlation function can be expressed succinctly by the following identity:

$$R_2^{(d)}(x) = \frac{2}{T_H^2}\mathrm{Re}\left[\frac{d^2}{ds^2}\ln Z(s) - a(s)\right]_{s=\frac{2\pi i x}{T_H}}, \tag{13}$$

where

$$\frac{1}{Z(s)} = \prod_p \prod_{m=0}^{\infty}\left(1 - \frac{\exp(sT_p)}{|\Lambda_p|\Lambda_p^m}\right)^{(m+1)g_p} \tag{14}$$

and

$$a(s) = \sum_p \sum_{m=0}^{\infty} g_p \frac{(m+1)T_p^2}{\left(|\Lambda_p|\Lambda_p^m \exp(-sT_p) - 1\right)^2}. \tag{15}$$

If $g_p = 1$, Z corresponds to the Ruelle-type zeta-function associated with the spectral determinant of the Fobenius-Perron operator that generates the time evolution of phase-space densities in classical mechanics[37]. Ergodicity, or equivalently the Hannay-Ozorio de Almeida sum rule, implies that $Z(s)$ has a simple pole at s=0, and so (12) can be viewed as a direct consequence of this singularity. The non-universal asymptotic approach to the random-matrix limit is then related to the analytical structure of Z in the rest of the complex plane; for example, in the positions of the nearby singularities of ln$Z$. In dynamical systems terms, this structure is precisely that which determines the decay of classical correlations, and hence the approach to ergodicity.

When $g_p \neq$ constant, Z is not exactly identifiable as a classical zeta function. In cases when $\bar{g} = 1$ it still has a simple pole at s=0, and so the universal limiting result (12) is unchanged. It is, however, not clear to what extent the analytical structure away from s=0 is affected by fluctuations in the mutiplicities. When $\bar{g} = 2$ the structure far away will almost certainly be different from the classical situation. In the exceptional cases, when the mean multiplicity increases with period, Z bears no obvious relation to a classical zeta function, and even the structure around s=0 may be changed, resulting in non-generic quantum spectral statistics.

The periodic orbit sum in (15) converges when s=0 and so

$$\frac{1}{T_H^2} a\left(\frac{2\pi i x}{T_H}\right) \to 0 \tag{16}$$

as $T_H \to \infty$. Hence a contributes to the non-universal approach to the universal limiting regime in R_2, but not to the limit itself. It thus plays the same role in (13) as the analytical structure of Z away from s=0.

The relationship between quantum spectral statistics and the Frobenius-Perron operator was first proposed by analogy with perturbative expressions for the spectral statistics in disordered systems[22], and a programme has been initiated to put it on a firmer footing using a nonlinear sigma-model for chaotic systems[38]. However, there are subtle differences from the results outlined above that would appear to warrant further investigation: the multiplicities g_p do not seem to play the same key role (the trace formula only leads to a classical zeta-function under the assumptions about g_p already stated), and the function a is absent from the formulae corresponding to (13). It is worth repeating that the multiplicities are essential to understanding the exceptional systems that are strongly chaotic (i.e. for which the classical zeta function has a simple pole at s=0, isolated by a gap from other singularities), but for which the spectral correlations are not random-matrix.

The picture for $K(\tau)$ is the same as the one painted above for R_2. The semiclassical formula for K, the fourier transform of (9), can be split into diagonal and off-diagonal contributions, where the diagonal contribution is given by

$$K^{(d)}(\tau) = \frac{1}{T_H^2} \sum_p \sum_{m=1}^{\infty} \frac{g_p T_p^2}{\left|M_p^m - I\right|} \delta\left(\tau - \frac{mT_p}{T_H}\right). \tag{17}$$

Replacing g_p by its mean $\bar{g}$, assuming this to be a constant, and evaluating the periodic orbit sum using the Hannay-Ozorio de Almeida sum rule gives

$$K^{(d)}(\tau) \approx \bar{g}\tau \tag{18}$$

in the limit $\tau T_H \to \infty$. When $\bar{g} = 1$ this coincides with the first term in the Taylor expansion of the GUE form-factor around $\tau = 0$, and when $\bar{g} = 2$ it coincides with the corresponding GOE result. This leads to the important conclusion that the off-diagonal terms do not contribute around $\tau = 0$ and hence that action correlations are negligible for $T << T_H$. In fact, when $\bar{g} = 1$ (18) coincides with the GUE form factor for $\tau \le 1$ and so this conclusion holds for $T \le T_H$.

In the regime $\tau T_H \to 0$, the spectral statistics are governed by the short-time classical dynamics and so are non-universal. The form factor can then either be expressed in terms of the individual short periodic orbits, via (17), or by the fourier transform of (13).

Finally, it is worth remarking that the analysis reviewed above relies only upon the existence of a trace formula and the Hannay-Ozorio de Almeida sum rule (8), which itself follows from classical ergodicity. The same approach thus also applies straightforwardly to maps and to integrable systems, where ergodicity on phase-space tori implies a sum rule that corresponds directly to (8). In the latter case, the results coincide exactly with the Poissonian expectation in the universal regime.

6. OFF-DIAGONAL CONTRIBUTIONS

The fact that $R_2(x) \ne R_2^{(d)}(x)$ for any of the random matrix ensembles suggests that if the conjectures reviewed in Section 3 are correct, and if the semiclassical approximation (9) is assumed accurate, then $R_2^{(off)}(x) \neq 0$. But as already discussed, if the periodic orbit contributions to (9) are uncorrelated, this would imply $R_2^{(off)}(x) = 0$. Hence there must be correlations. Unfortunately, we have no a priori knowledge of their origin; that is, there is at present no theory for them based purely on classical dynamics. One can, of course, work backwards by assuming that $R_2(x)$ is given precisely by the appropriate random matrix expressions, setting $R_2^{(off)}(x) = R_2(x) - R_2^{(d)}(x)$, with $R_2^{(d)}(x)$ given by (12), and then fourier-transforming with respect to $\hbar^{-1}$ to obtain a formula for a classical periodic orbit correlation function[39]. Numerical computations support the correctness of the result, but a derivation within classical mechanics is still lacking.

Since a direct evaluation of the sum over off-diagonal orbit pairs in (9) is, for generic systems, beyond our current horizon, we are forced to seek an indirect method of calculation. The basis of such an approach is suggested by the following observations. First, it follows from the general arguments presented in Section 4 that to resolve the quantum spectrum down to the scale of the mean level spacing requires orbits with periods up to the order of the Heisenberg time T_H. Orbits with periods longer than T_H determine spectral structure on scales shorter than the mean spacing, and so should not in principle be part of a theory of long-range statistics. Second, it was argued in Section 5 that, to a first approximation, orbits with periods less than T_H contribute to the semiclassical formula for $R_2(x)$ as if they are uncorrelated (the implication is that action correlations are needed to calculate the contributions from obits with periods larger than T_H). We conclude that a self-consistent semiclassical theory for spectral statistics should be based on orbits with periods up to the order of T_H, and should treat these as if they were uncorrelated.

Such a theory can be constructed as follows[20]. First, we use the fact that the zeros of the function on the right-hand side of (4) are semiclassical approximations to the exact energy levels. The restriction to orbits with periods T_p up to the order of T_H can then be made by truncating the p-product in (4) appropriately. The above arguments imply that the zeros of the truncated product remain good approximations to the energy levels on the scale of the mean level separation. It is a key assumption that this is the case.

The idea is then, essentially, to compute the correlations in the semiclassical spectrum obtained from the truncated product. However, this is complicated by the fact that unlike the exact energy levels, the approximations thus produced are not automatically real. Nevertheless, we can generate real approximations by using the real zeros of the real part of the right-hand side of (4) when truncated. This is semiclassically consistent, because hermiticity implies that the exact quantum spectral determinant is real when E is real and so the semiclassical approximation must be real to leading order in $\hbar$. Taking the real part thus corresponds to rearranging the higher orders. Basically, it can be thought of as imposing the functional equation, somewhat as in the derivation of the Riemann-Siegel lookalike formula[40,41].

The function B in (4) is itself real when E is real, and so the approximations e_n to the energy levels are the real zeros of

$$w(E) = \mathrm{Re}\left[\exp\left(-i\pi\overline{N}(E)\right)f(E)\right], \tag{19}$$

with

$$f(E) = \prod_p^{(T_H)} \prod_{m=0}^{\infty} \left(1 - \frac{\exp\left(iS_p/\hbar\right)}{\left|\Lambda_p\right|^{\frac{1}{2}} \Lambda_p^m}\right), \tag{20}$$

where the product is truncated smoothly to include orbits whose periods are less than or of the order of T_H. This is identical to truncating the p-sum in (2) in the same way, integrating as in (5) to obtain an approximation $n(E)$ to the spectral staircase (counting function), and defining e_n to be the solution of $n(E) = n + 1/2$.

The density associated with the semiclassical approximations e_n (defined as in (1) but with E_n replaced by e_n) can be written

$$\tilde{d}(E) = \frac{d}{dE}\sum_{k=-\infty}^{\infty}\frac{(-1)^k}{2\pi ik}\exp\left(2\pi ik\bar{N}(E)\right)\left(\frac{f^*(E)}{f(E)}\right)^k. \tag{21}$$

Substituting this into (6) then gives an expression for the two-point correlation function of the semiclassical spectrum. The contribution from the k=0 term in (21) can be evaluated by again making the diagonal approximation[20]. The result has the same form as (13), but with the p-product in (14) and the p-sum in (15) truncated to include only orbits with T_p up to the order of T_H. When x>>1 the truncation can be ignored to leading order, and so the k=0 contribution reduces to the diagonal contribution of the previous section. This then gives (12) in the semiclassical limit, which, as already noted, coincides with the leading-order x>>1 asymptotics of the non-oscillatory terms in the corresponding RMT results.

Since the k=0 term in (21) corresponds to the diagonal contribution, the $k \neq 0$ terms must give the off-diagonal contribution. These may be evaluated by substituting (20) for f and then by interchanging the order of the energy average in (6) with the p-product in the ratio of the f-factors. It is this step that corresponds to assuming that the actions of different primitive orbits (of period less than or of the order of T_H) contribute as if they are uncorrelated.

It is crucial to note at this stage that it is primitive orbits whose actions are being treated as uncorrelated. The repetitions of a given primitive orbit are, of course, highly correlated - their actions are integer multiples of the primitive action - and it is a mistake to treat them otherwise.

Finally, the average over the m-products in the ratio of the f-factors can be evaluated exactly[20]. The result for the off-diagonal contribution to the two-point correlation function is that when x>>1 and, to take just one case, $g_p = 1$,

$$R_2^{(off)}(x) \approx \frac{2}{T_H^2}\left|\gamma^{-1}Z\left(\frac{2\pi ix}{T_H}\right)^2\right|\mathrm{Re}\left[\exp(2\pi ix)\prod_p \chi_p\left(\frac{2\pi ix}{T_H}\right)\right], \tag{22}$$

with

$$\chi_p(x) = {}_2\phi_1\left(\exp(-T_p x),\exp(-T_p x);\Lambda_p^{-1};\Lambda_p^{-1},|\Lambda_p^{-1}|\exp(T_p x)\right)\frac{|Z_p(0)|^2}{|Z_p(x)|^2}, \tag{23}$$

where γ is the residue of the pole of $Z(s)$ at s=0, ${}_2\phi_1$ is the q-hypergeometric series, and Z_p is the p-th element of the product over primitive orbits in (14). Formally, the p-products in this expression should be truncated so as to include only those orbits with periods T_p up to the order of T_H, but when x>>1 there is no difference to leading order.

When $T_H \to \infty$, $\chi_p \to 1$ and (22) is dominated by the pole of Z at s=0. Hence in this limit (which corresponds to letting $\hbar \to 0$)

$$R_2^{(off)}(x) \to \frac{\cos(2\pi x)}{2\pi^2 x^2}. \tag{24}$$

Combined with the result (12) for the diagonal terms, this then coincides precisely with the exact GUE expression when $\overline{g} = 1$. In the same way, when $\overline{g} = 2$ we recover the leading order $x \gg 1$ asymptotics (rather than the exact form in this case) of the GOE two-point correlation function.

Treating all orbit actions as being uncorrelated, rather than just those of the primitive orbits, gives[20] a formula like (22) but with $\chi_p = 1$. This has the same limit as (22) when $T_H \to \infty$, but a different form for finite T_H. Remarkably, it coincides with the expression for the nonperturbative contribution to the two-point correlation function for disordered systems derived on the basis of the nonlinear σ-model[21]. It would be very interesting to understand the physical origins of the difference.

Given (22) and its analogue for time-reversal-symmetric systems, one can fourier-transform with respect to $\hbar^{-1}$ to obtain a formula for a classical periodic orbit action correlation function, as indicated at the start of this Section. The result takes the form of a sum over pairs of pseudo-orbits (linear combinations of periodic orbits) and generalizes that obtained by Argaman et al[39] in that it includes non-universal effects. It may appear paradoxical that one can obtain orbit correlations from an indirect calculation that ignores them; the point is that they are effectively included by the resummation that underlies the bootstrap formula (21). In the same way, fourier-transforming with respect to x gives an expression for the form-factor, also in terms of a sum of pseudo-orbit-pairs.

It is also worth pointing out that the calculation outlined above is based solely on the trace formula, and thus applies trivially to maps.

7. THE RIEMANN ZEROS

One way of testing the methods described in the previous two sections is to apply them to the zeros of the Riemann zeta function. This may be surprising upon first sight, because there is no proof of any link between the zeta function and a quantum system, but it is not difficult to see on the level of a mathematical 'toy-model'; that is, the analogy is mathematical, rather than physical.

The Riemann zeta-function is defined for Res>1 by a Dirichlet series

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \tag{25}$$

or by an Euler product over the primes p,

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}, \tag{26}$$

and then by analytic continuation to the rest of the complex plane[42,43]. It is a meromorphic function with a single simple pole at s=1, where it has unit residue, and ('trivial') zeros at s=-2, -4, -6, ... The Riemann hypothesis (RH) is that all of the other ('non-trivial') zeros lie on the line Res=1/2. Put another way, the non-trivial zeros lie at points $s_n = 1/2 + iE_n$ where $\mathrm{Im}\, E_n = 0$. Thus, assuming the hypothesis is correct, the set $\{E_n\}$ forms a real and discrete 'spectrum' which can be analysed statistically, in the same way as for energy levels. In fact, one can do this analysis even if the hypothesis is not correct, but for ease of presentation we will assume that it is.

A density of zeros can be defined exactly as in (1), and there is an explicit formula in terms of the primes:

$$d(E) = \bar{d}(E) - \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{\Lambda(n)}{\sqrt{n}}\cos(E\log n), \tag{27}$$

where

$$\bar{d}(E) \approx \frac{1}{2\pi}\log\left(\frac{E}{2\pi}\right) \tag{28}$$

and

$$\Lambda(n) = \begin{cases}\log p & \text{if } n = p^k \\ 0 & \text{otherwise}\end{cases}. \tag{29}$$

The sum in (27) thus runs over primes and prime-powers, and is mathematically the exact analogue of the trace formula (2) (there are problems with its physical interpretation, but this does not affect its application here[25]).

The density can be substituted into (6) leading to a definition of the two-point correlation function of the Riemann zeros. Extensive numerical evidence[44], rigorous results[26,45], and heuristic calculations[20,24,28,29,46] all support the conjecture that this (and other statistics) tends to the corresponding GUE form in the limit $E \to \infty$. Mathematically, we are thus in the same position as when studying the spectral statistics of classically chaotic systems: there is a set of real numbers (assuming RH) whose limiting statistics is GUE and for which there is a trace formula. We can thus follow the analysis outlined in the previous two sections line by line.

The results are as follows, first the two-point correlation function can be expressed as

$$R_2(x) = 1 + \frac{1}{2\pi^2\bar{d}^2(E)}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{\Lambda(m)\Lambda(n)}{\sqrt{mn}}\left\langle\cos\left(E\log\left(\frac{m}{n}\right) + \frac{x}{\bar{d}(E)}\log m\right)\right\rangle_E, \tag{30}$$

which is the direct analogue of (9). This double sum can then be split up into diagonal (m=n) and off-diagonal ($m \neq n$) contributions. The diagonal part can be evaluated exactly and takes the form[20]

$$R_2^{(d)}(x) = \frac{1}{2\pi^2\bar{d}^2(E)}\mathrm{Re}\left[\frac{d^2}{ds^2}\log\zeta(s) - \sum_p\frac{\log^2 p}{(p^s - 1)^2}\right]_{s=1+ix/\bar{d}(E)} \tag{31}$$

which corresponds to (13). In the limit $E \to \infty$ (and hence from (28) $\bar{d} \to \infty$), (31) is dominated by the pole of the zeta function at s=1, and so

$$R_2^{(d)}(x) \to -\frac{1}{2\pi^2 x^2}. \tag{32}$$

The off-diagonal contribution can be calculated using the methods outlined in Section 6: one can define approximations to the zeros by truncating the product (26) near to the 'Heisenberg prime' when $\log p = 2\pi\bar{d}(E)$, and then calculate the two-point correlation function of these by assuming that the logs of the primes included are statistically uncorrelated. This leads to the result[20]

$$R_2^{(off)}(x) = \frac{1}{2\pi^2\bar{d}^2(E)}\left|\zeta\left(1+\frac{ix}{\bar{d}(E)}\right)\right|^2 \mathrm{Re}\left[\exp(2\pi ix)\prod_p\left(1-\frac{\left(p^{ix/\bar{d}(E)}-1\right)^2}{(p-1)^2}\right)\right] \tag{33}$$

when x>>1, which is the direct analogue of (22). In the limit $E \to \infty$ we again recover the limit (24) due to the pole of the zeta function. Combined with (32) this coincides precisely with the exact GUE expression for R_2.

The importance of the Riemann zeta-function is that it is the one example where we can evaluate the off-diagonal contribution directly from (30) without any reference to the indirect methods reviewed in Section 6. This is because we possess enough information about the correlations between the phases in the off-diagonal terms. The main steps in this calculation are as follows. First, the off-diagonal contribution is clearly given by

$$R_2^{(off)}(x) = \frac{1}{2\pi^2\bar{d}^2(E)}\sum_{m\neq n}\frac{\Lambda(m)\Lambda(n)}{\sqrt{mn}}\left\langle\cos\left(E\log\left(\frac{m}{n}\right)+\frac{x}{\bar{d}(E)}\log m\right)\right\rangle_E . \tag{34}$$

Because of the E-average, only terms with $m \approx n$ contribute. Hence writing $m=n+k$ and expanding the term multiplied by E (the large parameter) to first order in k,

$$R_2^{(off)}(x) = \frac{1}{2\pi^2\bar{d}^2(E)}\sum_n\sum_{k\neq 0}\frac{\Lambda(n)\Lambda(n+k)}{n}\left\langle\cos\left(\frac{Ek}{n}+\frac{x}{\bar{d}(E)}\log m\right)\right\rangle_E . \tag{35}$$

The statistical information corresponding to the periodic orbit action correlations discussed at the beginning of Section 6 is now provided by the Hardy-Littlewood conjecture[30], namely that in the limit $N \to \infty$

$$\frac{1}{N}\sum_{n<N}\Lambda(n)\Lambda(n+k) \to \alpha(k), \tag{36}$$

where

$$\alpha(k) = \sum_{\substack{(r,q)=1\\ r<q}}\exp\left(-2\pi i\frac{r}{q}k\right)\left(\frac{\mu(q)}{\phi(q)}\right)^2, \tag{37}$$

in which the sum (known as a singular series) runs over all rationals, $\mu(q)$ is the Möbius function, and $\phi(q)$ is Euler's totient function. Substituting this into (35), all sums can be performed analytically[20] to give (33) without the assumption that x>>1. This then provides an important check on the correctness of the steps involved in the indirect method outlined in Section 6.

Once again, one can reverse the direction of the calculation and, taking (33) as given by the indirect approach, fourier-transform with respect to E to derive (37); that is, one can recover the Hardy-Littlewood prime correlations from a result derived by ignoring them. This seemingly self-contradictory approach is the direct analogue of the calculation of the periodic orbit action correlation function mentioned at the end of Section 6. Moreover, fourier-transforming (35) with respect to x leads to an expression for the form-factor of the Riemann-zeros in terms of the singular series. This plays the role of the sum over pairs of pseudo-orbits in the corresponding semiclassical result.

8. CONCLUSIONS

It is, perhaps, too soon to try to draw conclusions about the recent developments reviewed here. It is clear that the trace formula provides a method to calculate quantum statistics that is simple and physically transparent. It is also general, in that it applies equally well to a range of related problems. For example, the methods outlined transfer directly[20] to quantum maps, to matrix elements and wavefunctions (for which there is also a trace formula), to parametric statistics, to integrable systems, and to higher-order correlation functions. Moreover, they also apply to the zeros of number-theoretical zeta functions, such as the Riemann zeta-function, where they can be checked against alternative methods of calculation and against numerical computations. Preliminary results support their correctness.

The weaknesses of the semiclassical method described here are that it is hard to see how it can be made rigorous (e.g. how to estimate the errors in the various approximations), and that at this stage it is not known how to go beyond leading-order in x in the calculation of correlation functions. An advantage is that it applies to systems that are non-generic with respect to their spectral statistics. Another advantage is that it applies to mixed systems; for example, it describes the rich critical behaviour associated with orbit bifurcations[47].

At this stage, a pressing problem seems to be to understand the relationship between the results obtained using the trace formula and those derived using supersymmetric methods. Are the differences in the non-universal regime a clue as to the relationship between the two methods? Do they suggest that the supersymmetric approach is best geared to describing the universal regime and the trace formula the non-universal behaviour? Do they point to the effects of the various uncontrolled approximations? Or do they mean that supersymmetric techniques are not yet applicable to individual systems? It is expected that these, and other related questions, will be at the centre of the next round of developments in this rapidly evolving area.

ACKNOWLEDGEMENTS

All of the recent results reviewed here were obtained in collaboration with Eugene Bogomolny. I have enjoyed too many conversations on this subject to list all of those who have influenced the views expressed. I am grateful to the Isaac Newton Institute for making many of them possible.

REFERENCES

1. Gutzwiller, M.C. *J. Math. Phys.* **12**, 343-358 (1971).
2. Gutzwiller, M.C. *Chaos in classical and quantum mechanics* (Springer, New York, 1990).
3. Berry, M.V. & Tabor, M. *Proc. R. Soc. Lond.* A **356**, 375-394 (1977).
4. MacDonald, S.W. & Kaufman, A.N. *Phys. Rev. Lett.* **42**, 1189-1191 (1979).
5. Berry, M.V. *Ann. Phys.* **131**, 163-216 (1981).
6. Bohigas, O., Giannoni, M.J. & Schmit, C. *Phys. Rev. Lett.* **52**, 1-4 (1984).
7. Berry, M.V. *Proc. R. Soc. Lond.* A**413**, 183-198 (1987).
8. Bohigas, O. *Les Houches Lecture Series* vol. 52 eds. Giannoni, M.J., Voros, A., & Zinn-Justin, J. (Amsterdam: North-Holland), 89-199 (1991).
9. Keating, J.P. *Nonlinearity* **4**, 309-341 (1991).
10. Bogomolny, E.B., Georgeot, B., Giannoni, M.J. & Schmit, C. *Physics Reports* **291**, 220-324 (1997).
11. Connors, R.D. & Keating, J.P. *J. Phys. A* **30**, 1817-1830 (1997).
12. Sarnak, P. *Curr. Dev. Math.* 84-115 (1997).
13. Simons, B.D. in this volume.
14. Hannay, J.H. & Ozorio de Almeida, A.M. *J. Phys. A* **17**, 3429-3440 (1984).
15. Berry, M.V. *Proc. R. Soc. Lond.* A**400**, 229-251 (1985).

16. Goldberg, J., Smilansky, U., Berry, M.V., Schweizer, W., Wunner, G., Zeller, G. *Nonlinearity* **4** 1-14 (1991).
17. Berry, M.V. & Keating, J.P. *J. Phys. A* **27**, 6091-6106 (1994).
18. Wilkinson, M. *J. Phys. A* **21**, 1173-1190 (1988).
19. Eckhardt, B., Fishman, S., Keating, J., Agam, O., Main, J. & Müller, K. *Phys. Rev. E* **52**, 5893-5903 (1995).
20. Bogomolny, E.B. & Keating, J.P. *Phys. Rev. Lett.* **77**, 1472-1475 (1996).
21. Andreev, A.V. & Altshuler, B.L. *Phys. Rev. Lett.* **75**, 902-905 (1995).
22. Agam, O., Altshuler, B.L. & Andreev, A.V. *Phys. Rev. Lett.* **75**, 4389-4392 (1995).
23. Berry, M.V. in *Quantum chaos and statistical nuclear physics* eds. Seligman, T.H. & Nishioka, H., 1-17 (1986).
24. Keating, J.P. in *Quantum Chaos* eds. Casati, G., Guarneri, I. & Smilansky, U., 145-185 (North-Holland, Amsterdam) (1993).
25. Berry, M.V. & Keating, J.P. in this volume,
26. Montgomery, H.L. *Proc. Symp. Pure Math.* **24**, 181-193 (1973).
27. Keating, J.P. in *Quantum Chaos* eds. Cerdeira, H.A., Ramaswamy, R., Gutzwiller, M.C. & Casati, G., 280-290 (World Scientific, Singapore) (1993).
28. Bogomolny, E.B. & Keating, J.P. *Nonlinearity* **8**, 1115-1131 (1995).
29. Bogomolny, E.B. & Keating, J.P. *Nonlinearity* **9**, 911-935 (1996).
30. Hardy, G.H., & Littlewood, J.E. *Acta Math.* **44**, 1-70 (1923).
31. Balazs, N.L. & Voros, A. *Physics Reports* **143**, 109-240 (1986).
32. Sieber, M. & Steiner, F. *Phys. Rev. Lett.* **67**, 1941-1944 (1991).
33. Tanner, G., Scherer, P., Bogomolny, E.B., Eckhardt, B. & Wintgen, D. *Phys. Rev. Lett.* **67**, 2410-2413 (1991).
34. Keating, J.P. & Sieber, M. *Proc. R. Soc. Lond.* A**447**, 413-437 (1994).
35. Schomerus, H. & Haake, F. *Phys. Rev. Lett.* **79**, 1022-1025 (1997).
36. Zirnbauer, M.R. in this volume.
37. Cvitanovic, P. & Eckhardt, B. *J.Phys.A* **24**, L237-L241 (1991).
38. Andreev, A.V., Agam, O., Simons, B.D. & Altshuler, B.L. *Phys. Rev. Lett.* **76**, 3947-3950 (1996).
39. Argaman, N., Dittes, F.M., Doron, E., Keating, J.P., Kitaev, Ya., Sieber, M. & Smilansky, U. *Phys. Rev. Lett.* **71**, 4326-4329 (1993).
40. Keating, J.P. *Proc. R. Soc. Lond.* A**436**, 99-108 (1992).
41. Berry, M.V. & Keating, J.P. *Proc. R. Soc. Lond.* A**437**, 151-173 (1992).
42. Edwards, H.M. *Riemann's Zeta Function* (Academic Press, New York and London, 1974).
43. Titchmarsh, E.C. *The theory of the Riemann zeta-function* (Clarendon Press, Oxford, 1986).
44. Odlyzko, A.M. *Math. of Comp.* **48**, 273-308 (1987).
45. Rudnick, Z. & Sarnak, P. *Duke Math. J.* **81**, 269-322 (1996).
46. Berry, M.V. *Nonlinearity* **1**, 399-407 (1988).
47. Berry, M.V., Keating, J.P. & Prado, S.D. 'Orbit bifurcations and spectral statistics' *J. Phys. A* in the press (1998).