

Gutzwiller's Trace Formula and Spectral Statistics: Beyond the Diagonal Approximation

E. B. Bogomolny

Division de Physique Théorique, Institut de Physique Nucléaire, 91406 Orsay Cedex, France*

J. P. Keating

*School of Mathematics, University of Bristol, Bristol BS8 1TW, United Kingdom
and BRIMS, Hewlett-Packard Labs, Filton Road, Stoke Gifford, Bristol BS12 6QZ, United Kingdom*

(Received 10 April 1996)

We calculate the 2-point spectral correlation function for classically chaotic systems in the semiclassical limit using Gutzwiller's trace formula. The off-diagonal contributions from pairs of nonidentical periodic orbits are evaluated by relating them to the diagonal terms. The behavior we find is similar to that recently discovered to hold for disordered systems using nonperturbative supersymmetric methods. Our analysis generalizes immediately to include parametric statistics and higher-order correlations and to the study of the semiclassical distribution of matrix elements. [S0031-9007(96)00955-6]

PACS numbers: 05.45.+b, 03.65.Sq, 05.40.+j

Gutzwiller's trace formula [1] leads to a semiclassical representation for the density of states $d(E) = \sum_n \delta(E - E_n)$ of a quantum system in terms of the periodic orbits of the underlying classical dynamics:

$$d(E) = \bar{d}(E) + \frac{1}{\pi\hbar} \sum_p \sum_{n=1}^{\infty} T_p A_{p,n} \cos(nS_p/\hbar), \quad (1)$$

where \bar{d} is the mean density [$\bar{d} = O(\hbar^{-f})$ if the system has f degrees of freedom], and the periodic orbits, labeled p , have action S_p (defined here to include the Maslov index), period $T_p = dS_p/dE$, and stability amplitude $A_{p,n} = |\det(M_p^n - I)|^{-1/2}$, M_p being the monodromy matrix that describes the local flow linearized about the p th orbit. This expression provides a natural tool with which to investigate the conjectured link [2,3] between the statistical distribution of the energy levels of classically chaotic systems and the eigenvalue correlations of random matrix theory (RMT). For example, one easily obtains from it a semiclassical approximation for the 2-point spectral correlation function

$$R_2(x) = \langle d(E + x_1)d(E + x_2) \rangle, \quad (2)$$

where $x = x_1 - x_2$ and $\langle \dots \rangle$ denotes an energy average, in the form of a sum over all pairs of periodic orbits. The main questions to be answered are then: Which properties of chaotic orbits are related to the universal features exhibited by complex spectra, such as level repulsion and rigidity, and how can the precise form of the results for the three symmetry-related RMT ensembles—the unitary ensemble (GUE), the orthogonal ensemble (GOE), and the symplectic ensemble (GSE)—be recovered?

The principal development in this direction was Berry's theory [4] for the form factor

$$K(T) = \int_{-\infty}^{\infty} R_2(x) \exp(ixT/\hbar) dx. \quad (3)$$

His analysis was based largely on the diagonal approximation, which involves keeping only the contributions

from pairs of orbits that are identical modulo symmetries. These diagonal terms may be evaluated using the Hannay–Ozorio de Almeida sum rule for ergodic flows [5], and when $T \ll T_H = 2\pi\hbar\bar{d}$ the results coincide with the corresponding RMT expressions. Unfortunately, to calculate $K(T)$ outside this range necessarily involves the evaluation of the off-diagonal terms associated with pairs of nonidentical orbits. The problem is that these are connected with correlations between the actions and stabilities of different orbits, about which we have no *a priori* knowledge [6], except in a few nongeneric examples [7], or in the case of the Riemann zeta function [8].

Our aim here is to outline a method for evaluating the off-diagonal contributions indirectly, but explicitly, by relating them to the diagonal terms. Specifically, we derive expressions for them in terms of a classical zeta function whose analytic structure, associated with the Hannay–Ozorio de Almeida sum rule, determines their form in the semiclassical limit.

The results we obtain are very closely related to ones found recently for ensembles of disordered systems using nonperturbative supersymmetric methods [9]. Indeed, making certain simplifying approximations we recover those precisely. This then represents further strong evidence of the deep connections between the quantum properties of individual deterministically chaotic systems and of ensembles of randomly disordered systems. Furthermore, it also emphasizes the close formal relationship between the semiclassical approach, based on the trace formula, and the supersymmetric techniques associated with nonlinear σ models.

We now outline the basic ideas behind our calculation. We begin with the trace formula (1), truncated (smoothly if necessary) so as to include only orbits with periods $T_p < T^*$, where T^* is a parameter whose value will be fixed later in the analysis. Denoting the resulting density by $d_{T^*}(E)$,

one can then define corresponding approximations to the eigenvalues via the quantization condition [10,11]

$$N_{T^*}(E_n(T^*)) = n + 1/2, \quad (4)$$

where

$$N_{T^*}(E) = \int_0^E dT^*(E') dE' = \bar{N}(E) + N_{T^*}^{(\text{osc})}(E)$$

is the associated staircase (counting function),

$$N_{T^*}^{(\text{osc})}(E) = \frac{1}{\pi} \sum_{T_p < T^*} \sum_{n=1}^{\infty} \frac{1}{n} A_{p,n} \sin(nS_p), \quad (5)$$

and for simplicity we have set $\hbar = 1$. (Henceforth, the semiclassical limit will correspond to $\bar{d} \rightarrow \infty$.) Clearly $d_{T^*}(E)$ represents a semiclassical approximation to the exact density of states smoothed on the energy scale $\sim \hbar/T^*$; thus if $T^* \sim T_H$ it can be used to calculate faithful approximations to the true eigenvalues [i.e., $E_n(T^*) \simeq E_n$ for $T^* \geq T_H$] and to study the *average* behavior of the spectral correlations over ranges $x \gg \bar{d}^{-1}$. However, it misses the correlation information related to the singular structure of the exact density of states on scales small compared to the mean level separation. To overcome this difficulty, we define a new *bootstrapped density* $D_{T^*}(E) = \sum_n \delta(E - E_n(T^*))$, which may, in principle, also be calculated using only those orbits with $T_p \leq T^*$, but which does not involve any energy smoothing. We call it a bootstrapped density because its Fourier transform clearly has structure for all times T , not just for $T < T^*$, the role of the *effective orbits* introduced beyond T^* being to generate the correct analytic properties associated with the discreteness of the quantum spectrum. Rewriting D_{T^*} in the form

$$\begin{aligned} D_{T^*}(E) &= d_{T^*}(E) \sum_n \delta(N_{T^*}(E) - n - 1/2) \\ &= d_{T^*}(E) \sum_{k=-\infty}^{\infty} (-1)^k \exp[2\pi i k N_{T^*}(E)], \end{aligned} \quad (6)$$

and substituting into (2) then gives

$$\begin{aligned} R_2(x) &= \left\langle d_{T^*}(E + x_1) d_{T^*}(E + x_2) \sum_{k_1, k_2} (-1)^{k_1 - k_2} \right. \\ &\quad \left. \times \exp\{2\pi i [k_1 N_{T^*}(E + x_1) - k_2 N_{T^*}(E + x_2)]\} \right\rangle, \end{aligned} \quad (7)$$

which is the equation upon which our analysis will be based.

Consider first the $k_1 = k_2 = 0$ term, which we write in the form

$$\langle d_{T^*}(E + x_1) d_{T^*}(E + x_2) \rangle = \bar{d}^2 + R_2^{(\text{diag})}(x), \quad (8)$$

where

$$R_2^{(\text{diag})}(x) = \frac{1}{4\pi^2} \frac{\partial^2}{\partial x_1 \partial x_2} \ln \Delta(x_1, x_2) \quad (9)$$

and

$$\ln \Delta(x_1, x_2) = 4\pi^2 \langle N_{T^*}^{(\text{osc})}(E + x_1) N_{T^*}^{(\text{osc})}(E + x_2) \rangle. \quad (10)$$

Because only orbits with $T_p < T^* \sim T_H$ (that is, in the range where the diagonal approximation is known to hold) contribute, we assume that the resulting sum over orbit pairs may be replaced by the diagonal terms, for which the pairing involves orbits with exactly the same action and stability [4]. The result can then be expressed in the form $\Delta(x) = |Z_g(ix)|^2$, where

$$Z_g(s) = \exp\left(\sum_{T_p \leq T^*} \sum_{n=1}^{\infty} \frac{g_p}{n^2} |A_{p,n}|^2 e^{nT_p s} \right) \quad (11)$$

and g_p is the number of orbits with period T_p and amplitude A_p . In generic systems the multiplicity g_p is the same for almost all orbits and so, if its value is denoted g , $Z_g = [Z(s)]^g$, where $Z(s)$ is defined by (11) with $g_p = 1$. For systems without time-reversal invariance $g = 1$, and for systems whose dynamics is time-reversal symmetric $g = 2$ [4]. Furthermore, using the Hannay–Ozorio de Almeida sum rule, it is easy to demonstrate that if $sT^* \rightarrow \infty$ [so that the sum in (11) is effectively infinite], then $Z(s) \rightarrow \gamma/s$ as $s \rightarrow 0$, and, hence, one obtains the appropriate RMT expressions: $R_2^{(\text{diag})} \rightarrow -1/(2\pi^2 x^2)$ when $g = 1$ (GUE) and $R_2^{(\text{diag})} \rightarrow -1/(\pi^2 x^2)$ when $g = 2$ (GOE). Since we anticipate taking $T^* \sim T_H$, this result effectively emerges in the limit $x \rightarrow 0$, $x\bar{d} \rightarrow \infty$. We note in passing that there are nongeneric examples of strongly chaotic systems for which $Z(s)$ has a simple pole at $s = 0$, but the multiplicities are not constant [7]. In these cases it is known that the level statistics do not obey the RMT conjecture. Clearly this then implies that it is not only the analytic properties of Z , but the multiplicity structure as well, that determine the behavior of the spectral statistics.

Since in our approach the $k_1 = k_2 = 0$ term corresponds to the usual diagonal approximation, we identify the other terms as representing the off-diagonal contributions, and so denote them by $R_2^{(\text{off})}(x)$. To evaluate these we note first that the Taylor expansion in powers of x of the mean counting function leads to a term $(k_1 - k_2)\bar{N}(E)$ in the phase which is $O(\hbar^{-J})$. Hence the energy average renders negligible any contributions with $k_1 \neq k_2$, and so

$$\begin{aligned} R_2^{(\text{off})}(x) &= \frac{\partial^2}{\partial x_1 \partial x_2} \sum_{k \neq 0} \frac{1}{(2\pi k)^2} \\ &\quad \times \exp[2\pi i \bar{d} k (x_1 - x_2)] \Phi_k(x_1, x_2), \end{aligned} \quad (12)$$

where

$$\begin{aligned} \Phi_k(x_1, x_2) &= \langle \exp\{2\pi i k [N_{T^*}^{(\text{osc})}(E + x_1) \\ &\quad - N_{T^*}^{(\text{osc})}(E + x_2)]\} \rangle. \end{aligned} \quad (13)$$

We now make the key assumption that in generic systems the orbits up to period T^* can be treated as being

statistically independent modulo exact degeneracies, and hence that the energy average of any smooth function of $\exp[iS_p(E)]$, $\langle f \rangle = \langle f(e^{iS_1(E)}, \dots, e^{iS_M(E)}) \rangle$, can be calculated using

$$\langle f \rangle = \int_0^{2\pi} \dots \int_0^{2\pi} f(e^{i\phi_1}, \dots, e^{i\phi_M}) \prod_{j=1}^M \frac{d\phi_j}{2\pi}. \quad (14)$$

This is essentially equivalent to a random phase or strict diagonal approximation. The justification is that the diagonal approximation itself is known to be valid for times $T \ll T_H$ [4].

Using the trace formula for N_{T^*} (5) one can thus perform the energy average in (13) by evaluating the integral in (14). We shall give exact results for certain cases later. First, for clarity, we discuss a simple leading-order approximation to $\Phi_k(x)$ based on the relation $\langle \exp[iG(E)] \rangle \approx \exp[-\langle G^2(E) \rangle / 2]$. This is an identity if G is a Gaussian random function with zero mean, and so, since the exponent in (13) behaves like a sum of a large number of independent random terms [cf. (14)], it is expected to be a good approximation for generic systems. From it one obtains

$$\begin{aligned} \Phi_k(x_1, x_2) &= \exp\{-2\pi^2 k^2 [\langle N_{T^*}^{(\text{osc})}(E + x_1) \\ &\quad - N_{T^*}^{(\text{osc})}(E + x_2) \rangle]^2\} \\ &= \left(\frac{\Delta(x)}{L^2} \right)^{k^2}, \end{aligned} \quad (15)$$

where $\Delta(x)$ was defined in (10) and $L = \exp\{2\pi^2 \langle [N_{T^*}^{(\text{osc})}(E)]^2 \rangle\}$. Using the Hannay–Ozorio de Almeida sum rule, it may be seen that $L \sim (T^*)^g$. Since we anticipate taking $T^* \sim T_H$, it then follows that the terms in the k sum decrease rapidly, as $(\bar{d})^{-2gk^2}$. Hence to leading semiclassical order, as $x\bar{d} \rightarrow \infty$, we may retain just the $k = \pm 1$ contributions. If now we choose T^* so that $L = CT_H^g$ (which essentially corresponds to taking $T^* = T_H$), we then have $\Phi_{\pm 1}(x) \approx \Delta(x)/(C\bar{d})^{2g}$, where C is a constant. Thus finally taking $C = 2^{1-g}\gamma^g$, and ignoring terms of higher order in powers of \bar{d}^{-1} ,

$$R_2^{(\text{off})}(x) \approx \frac{\cos(2\pi\bar{d}x)}{2\pi^{2g}\bar{d}^{2g-2}} |\gamma^{-1}Z(ix)|^{2g}, \quad (16)$$

which is directly equivalent to the relationship that was shown by Andreev and Altshuler [9] to hold between the perturbative and nonperturbative parts of R_2 for disordered systems in the limit of large conductance using the supersymmetric approach. The contribution from the pole of $Z(s)$ at $s = 0$ gives immediately the exact GUE expression when $g = 1$, and the leading-order GOE result as $x\bar{d} \rightarrow \infty$ when $g = 2$. Clearly the structure of Z around the pole determines the semiclassical approach to the RMT limit. The value of the constant C is, in principle, fixed by the requirement that $R_2(x) \rightarrow 0$ as $x \rightarrow 0$ for a discrete spectrum. This may be applied directly when $g = 1$, but when $g = 2$ the fact that we have only the leading order $x \rightarrow \infty$ asymptotics means

that the $x \rightarrow 0$ behavior cannot be recovered. In this case the value of C was, for the present, chosen to match the corresponding GOE constant as $\bar{d} \rightarrow \infty$. That we obtain the exact result for the GUE, but not for the GOE, is obviously related to the fact that the diagonal approximation itself is exact for times up to T_H when $g = 1$, but not when $g = 2$. The reasons for this are, however, still to be explained.

It is also worth noting that, as already discussed in [12], $Z(s)$ may be approximately related to a classical zeta function $Z_{\text{cl}}(s)$, since, if the eigenvalues of the monodromy matrix M_p are Λ_p ($|\Lambda_p| > 1$) and Λ_p^{-1} (for simplicity we specialize here to two degree-of-freedom Hamiltonian systems), and ignoring the repetitions [i.e., taking $n = 1$ in (1)], $Z(s) \approx Z_{\text{cl}}(s)$ where

$$\frac{1}{Z_{\text{cl}}(s)} = \prod_p \prod_{k=0}^{\infty} \left(1 - \frac{e^{sT_p}}{|\Lambda_p| \Lambda_p^k} \right)^{k+1}. \quad (17)$$

We are now in a position to consider the statistics of the third RME: the GSE. The way in which our method is constructed allows us to make use of a remarkable theorem of Mehta and Dyson [13], which states that if every alternate level is removed from a spectrum in which the correlations are GOE, then the correlations in the resulting sequence are the same as those of the GSE. This can be implemented in our semiclassical scheme by first taking $D_{T^*}(E) = d_{T^*}(E) \sum_n \delta(N_{T^*}(E) - 2n - 1/2)$, second by replacing \bar{d} in the above formula by $2\bar{d}$ (because the original spectrum has twice the density of the new one), and finally by assuming, as in the GOE calculation, that the orbits come in time symmetric pairs. The result when $x\bar{d} \rightarrow \infty$ is that

$$R_2^{(\text{diag})}(x) = \frac{1}{4\pi^2} \frac{\partial^2}{\partial x^2} \ln |Z(ix)|, \quad (18)$$

$$R_2^{(\text{off})}(x) = \bar{d} \frac{\cos(2\pi\bar{d}x)}{4} |\gamma^{-1}Z(ix)|, \quad (19)$$

where $Z(s)$ is defined by (11) with $g_p = 1$ or (approximately) by (17). Again, the simple pole at $s = 0$ gives the leading order GSE asymptotics.

The method and results described above generalize in a number of different directions (for a more complete discussion, see [14]). First they extend directly to parametric correlations. Consider, for example, the correlation function $R_2(x, \phi) = \langle d(E + x/2, \bar{\phi} + \phi/2) d(E - x/2, \bar{\phi} - \phi/2) \rangle$, where ϕ is a parameter in the Hamiltonian and the average extends over both E and $\bar{\phi}$. In this case the diagonal terms may be evaluated by assuming that the action derivatives $\partial S_p / \partial \phi$ are Gaussian distributed with variance αT for orbits with $T_p \approx T$ [15]. Applying the result exactly as above it follows that for all three ensembles the semiclassical approximation to $R_2(x, \phi)$ is obtained from the result for $R_2(x)$ by replacing x by $y = x + i\alpha\phi^2/2$ in Z , but not in the $\cos(2\pi\bar{d}x)$ factor. When $\bar{d} \rightarrow \infty$ this agrees with the large ϕ limit of

the exact expressions given in [16]. The importance of the off-diagonal contribution in this case illustrated by the fact that when $x = 0$ it coincides exactly with the diagonal term.

$$\tilde{R}_3(x_1, x_2, x_3) = -\frac{1}{2\pi^3} \sin[2\pi\bar{d}(x_2 - x_3)] |\gamma^{-1} Z(i(x_2 - x_3))|^2 \text{Im} \left[\frac{Z'(i(x_2 - x_1))}{Z(i(x_2 - x_1))} + \frac{Z'(i(x_1 - x_3))}{Z(i(x_1 - x_3))} \right] + (1 \rightarrow 2, 3),$$

where $(1 \rightarrow 2, 3)$ denotes a sum of two terms, each the same as the first, but with x_1 interchanged with x_2 and x_3 , respectively. As $x_j \rightarrow 0$ this expression is dominated by the pole of $Z(s)$ at $s = 0$, and we recover the exact GUE formula for \tilde{R}_3 , as above. It is worth noting that in this case the diagonal contributions are all zero.

Our method also applies to correlations involving densities weighted by, for example, diagonal matrix elements. The diagonal contributions from the corresponding trace formula can again be evaluated and applied, using a direct extension of the method outlined above, to calculate the off-diagonal terms [14]. In this case the factor \bar{d}^{s-2} in (16) is replaced by $\bar{d}^s / \langle d(\hat{A}, E) \rangle^2$, and so if the microcanonical average of \hat{A} is zero, and thus $\langle \langle d(\hat{A}, E) \rangle \rangle = 0$, these off-diagonal terms vanish at the leading order, thus justifying the previous use of the diagonal approximation [17].

Up to now we have based our discussion on the approximate formula (15). Perhaps the most important extension of the method outlined above is that the calculation can also be performed exactly, with the correlations between orbit repetitions n in (1) included correctly. This may be achieved by noting that $N_{T^*}^{(\text{osc})}(E) = -\text{Im} \ln Z_{T^*}(E) / \pi$, where $Z_{T^*}(E)$ is the semiclassical Selberg-type zeta function with its Euler product truncated at T^* . Substituting this into (13), the integral in (14) can be performed exactly. For example, the result for non-time-reversal-symmetric systems may be written in the form

$$R_2^{(\text{off})} = \frac{1}{2\pi^2} |\gamma^{-1} Z_{\text{cl}}(ix)|^2 \text{Re} \left[\exp(2\pi i \bar{d}x) \prod_p \phi_p(x) \right], \tag{20}$$

with

$$\phi_p(x) = {}_2\phi_1(a_p, b_p; c_p; q_p, z_p) \frac{|Z_p(0)|^2}{|Z_p(ix)|^2}, \tag{21}$$

where $a_p = b_p = \exp(-iT_p x)$, $c_p = q_p = \Lambda_p^{-1}$, $z_p = |\Lambda_p|^{-1} \exp(iT_p x)$, ${}_2\phi_1(a, b; c; q, z)$ is the q -hypergeometric series, and $Z_{\text{cl}}(s) = \prod_p Z_p(s)$ is the classical zeta function defined in (17) (for details see [14]). The product in (20) converges when x is real and takes the value 1 when $x = 0$. This result appears to represent a refinement of the expressions presented in [9] and [12] and rederived above by ignoring correlations between orbit repetitions in the trace formula. Evidence that such corrections are meaningful is that for the 2-point correlations of the zeros

The second generalization is to higher orders of correlation. For example, the irreducible off-diagonal contributions to the 3-point correlation function when $g = 1$ are given in terms of the same function Z :

of the Riemann zeta function $\zeta(s)$ [8] the corresponding result

$$R_2^{(\text{off})}(x) = \frac{1}{2\pi^2} |\zeta(1 + ix)|^2 \times \text{Re} \left[e^{2\pi i \bar{d}x} \prod_p \left(1 - \frac{(1 - p^{ix})^2}{(p - 1)^2} \right) \right]$$

can be shown [14] to be exactly equivalent to the Hardy-Littlewood conjecture for the pairwise distribution of the prime numbers [8].

The authors wish to thank to S. Fishman for attracting their attention to the results in [9]. J.P.K. is also grateful to the Institut Henri Poincaré, Paris, for generous support.

*Unité de Recherche Universités Paris 11 et Paris 6, Associée au CNRS.

[1] M. C. Gutzwiller, *J. Math. Phys.* **12**, 343 (1971).
 [2] O. Bohigas, M.-J. Giannoni, and C. Schmit, *Phys. Rev. Lett.* **52**, 1 (1984).
 [3] M. V. Berry, *Proc. R. Soc. London A* **413**, 183 (1987).
 [4] M. V. Berry, *Proc. R. Soc. London A* **400**, 229 (1985).
 [5] J.H. Hannay and A.M. Ozorio de Almeida, *J. Math. Phys. A* **17**, 3429 (1984).
 [6] N. Argaman *et al.*, *Phys. Rev. Lett.* **71**, 4326 (1993).
 [7] J.P. Keating, *Nonlinearity* **4**, 309 (1991); E. Bogomolny, F. Leyvraz, and C. Schmit, *Commun. Math. Phys.* **176**, 577 (1996); W. Luo and P. Sarnak, *Commun. Math. Phys.* **161**, 419 (1994).
 [8] J.P. Keating, in *Quantum Chaos*, edited by G. Casati, I. Guarneri, and U. Smilansky (North-Holland, Amsterdam, 1995), p. 145; E.B. Bogomolny and J.P. Keating, *Nonlinearity* **8**, 1115 (1995); *ibid.* (to be published).
 [9] A.V. Andreev and B.L. Altshuler, *Phys. Rev. Lett.* **75**, 902 (1995).
 [10] E. Bogomolny and C. Schmit, *Nonlinearity* **6**, 523 (1993).
 [11] R. Aurich *et al.*, *Phys. Rev. Lett.* **68**, 1629 (1992).
 [12] O. Agam, B.L. Altshuler, and A.V. Andreev, *Phys. Rev. Lett.* **75**, 4389 (1995).
 [13] M.L. Mehta and F.J. Dyson, *J. Math. Phys.* **4**, 713 (1963).
 [14] E.B. Bogomolny and J.P. Keating (to be published).
 [15] M. V. Berry and J.P. Keating, *J. Phys. A* **27**, 6167 (1994).
 [16] B.D. Simons and B.L. Altshuler, *Phys. Rev. B* **48**, 5422 (1993).
 [17] B. Eckhardt *et al.*, *Phys. Rev. E* **52**, 5893 (1995).