The Spectral Form Factor Is Not Self-Averaging

R. E. Prange*
Laboratoire de Physique Quantique,† Université Paul Sabatier, Toulouse, France
(Received 26 June 1996)

The form factor, $k(t)$, is the spectral statistic which best displays nonuniversal quasiclassical deviations from random matrix theory. Recent estimations of $k(t)$ for a single spectrum found interesting new effects of this type. It was supposed that $k(t)$ is *self-averaging* and thus did not require an ensemble average. We here argue that this supposition sometimes fails and that for many important systems an ensemble average is essential to see detailed properties of $k(t)$. In other systems, notably the nontrivial zeros of Riemann zeta function, it will be possible to see the nonuniversal properties by an analysis of a single spectrum. [S0031-9007(97)02762-2]

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

Recent seminal work by Agam, Altshuler, and Andreev (AAA) [1] connects the energy level statistics of a *single* system, say a particular Sinai billiard, with the statistics of an *ensemble* of systems, namely, with the predictions of *random matrix theory* (RMT) [2]. This long conjectured connection was never before expressed analytically. This work has generated great interest and has led to alternative results by Bogomolny and Keating (BK) [3]. In short, it was proposed that the form factor $k(t)$ of the spectrum of a given system can be calculated quasiclassically to good approximation directly in terms of the classical periodic orbits of the system, and the results show only small deviations from the RMT prediction.

For scaling systems like billiards, it was found that not only does *the spectral form factor approach the universal RMT result* in an appropriate high energy limit, but predictions are made as to *how* this limit is approached. The results of AAA and BK differ in detail from one another, but their gross features, in particular, their scaling at high energy, are the same. (AAA-BK did not point out how the nonuniversal difference between their prediction and RMT scales, but it follows from their results.)

It would obviously be of interest to take the spectrum of some system, obtained either numerically or experimentally, and use it to calculate $k(t)$, so that comparison of the two approximate theories with the numerically "exact" structure factor can be made. The AAA-BK results supply a definite target for such an analysis.

The *main result* of this paper is that *in most cases* such a comparison is *impossible in principle,* at least using presently known methods of data analysis. This is based on the important fact that *the spectral form factor is not self-averaging.* This fact was known, although it evidently deserves more emphasis. What is new here is that we provide the first analysis of a secondary average which shows that it is usually impossible to extract as much from data on a single system as had previously been taken for granted.

Correlation functions and their Fourier transforms—form factors—are ubiquitous in physics. "Two point" correlation functions are the result of averaging "observables" at two "points" of a system or of an ensemble. Often, as for laser speckle pattern, they directly represent real experiments. In the case of speckle pattern, it is widely known that the form factor is *not* self-averaging. Thus laser light scattered from a rough surface shows *very large fluctuations,* fluctuations as large as the average signal, which are *particular* to that surface. These fluctuations disappear under an ensemble average.

Most correlation functions *are* self-averaging. For instance, the electrical conductance is self-averaging unless the electronic states are localized. Breakdown of self-averaging is especially interesting therefore. There is little discussion of self-averaging for the spectral form factor in the literature, and the conclusions reached are contradictory. One work [4] is devoted to proving (incorrectly, as we shall argue) that *all* two point spectral correlation functions including the spectral form factor *are* self-averaging. To be precise, it states that the correlation functions calculated from the energy levels obtained from a typical exemplar of a large random matrix is identical to the RMT ensemble average. Other work [5–7] states clearly, but without much discussion, the correct result that $k(t)$ is *not* self-averaging. The later papers do not cite the earlier one.

The numerical evidence is strong. We reproduce in Fig. 1 the form factor obtained by Eckhardt and Main [8] from some thousand levels of a hydrogen atom in a magnetic field. The large fluctuations about the RMT result attest to the lack of self-averaging.

AAA-BK ignore the self-averaging issue. Remarkably, although they emphasize that they calculate for a single system, their results do *not* show large fluctuations but rather are characteristic of *ensemble averaged* results. In fact, it was shown [9] how to modify RMT in order to incorporate into the matrix ensemble quasiclassical orbital information about a given system. This method *does* make an ensemble average. The ensemble is perhaps that of all Hamiltonians with the same classical limit. The results [9] are less complete than those of AAA-BK, but agree qualitatively with them. We shall later try to

 0031-9007/97/78(12)/2280(4)$10.00

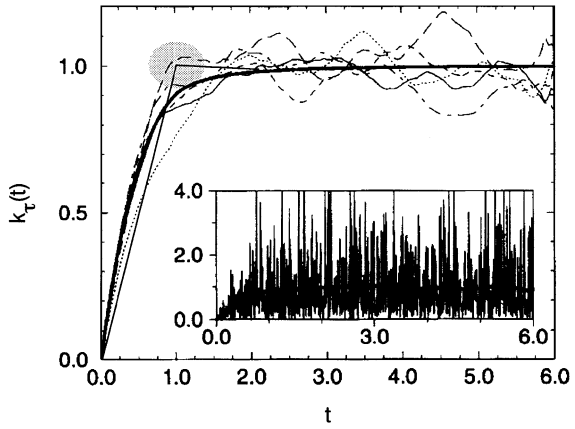

FIG. 1. Spectral form factor $k(t)$ (inset) for the eigenvalues $3997, \ldots, 4996$ of hydrogen in a magnetic field at scaled energy $\epsilon = -0.1$. Main figure: Time smoothed form factor $k_\tau(t)$, $\tau \sim 0.6$, for eigenvalues 1–1000 (dotted), 1001–2000 (dash-dotted), 2001–3000 (long dashed), 3001–4000 (short dashed), and 3997–4996 (continuous). See Ref. [8] for further details. This system is GOE. The GOE form factor is given by the thick line, the GUE by the thin straight lines. The shaded circle indicates schematically where nonuniversal deviations from GUE are predicted.

explain why AAA-BK fail to find the large fluctuations inherent in a form factor calculated for a single system.

It is known to be essential to use an energy average (see below). There is an additional average available for a single system [5,8], namely, a *time* average in which $k(t)$ is averaged over a range of time $t \pm \frac{1}{2}\tau$. Such an average is also illustrated in Fig. 1. We shall argue that for some systems this average, used with care, suffices to achieve agreement with the results of AAA-BK, while other systems require an ensemble average.

Of all the two level spectral correlation functions (which are interrelated by linear transformations), the form factor is of special interest because, according to Berry [10], the dimensionless time t, the argument of $k(t)$, is essentially the relevant classical orbit period in units of the Heisenberg time, $\tau_\hbar = 2\pi\hbar/\Delta$, where Δ is the mean level spacing. In particular, for small t, it has long been established and exploited [11] that the form factor is strongly peaked at the periods of the short periodic orbits, and the weights of these peaks are determined by the stability properties of the orbits. Other correlators smear over the time and are not so directly interpreted. There are other reasons [5] for preferring $k(t)$.

Another nice feature of $k(t)$ is that there are definite *nonuniversal effects* in a *particular range* of t, namely, near $t = 1$. This is the *new result* of the above theories [1,3,9]. These deviations from RMT *disappear* in the high energy limit, and the range over which the deviations occur *shrinks to zero*, as some power, η, of inverse energy. Exactly *which* power determines whether the effects can be observed in numerical calculations.

The form factor is usually loosely defined as the Fourier transform of the two level correlation function $C(x)$,

$$C(x) = \left\langle \sum_{a,b} \delta\left(\frac{E + x\Delta/2 - E_a}{\Delta}\right) \times \delta\left(\frac{E - x\Delta/2 - E_b}{\Delta}\right)\right\rangle_{\epsilon_0, W}. \quad (1)$$

The spectrum is given by the sequence of levels E_a. The average of E is over an energy range W about a central energy ϵ_0. AAA-BK make no further average. The form factor is $k(t) = \int dx\, e^{2\pi i x t} C(x)$. To be interesting, the results must depend only weakly on the averaging window W, and, indeed, they should not depend much on exactly how the average is made. This is supposed to be the case if (a) there are a large number of energy levels in the average, i.e., $w \equiv W/\Delta \gg 1$ and (b) W is classically small, denoted $W \ll \epsilon_0$. Then Δ and the properties of periodic orbits vary little in the window $\epsilon_0 \pm \frac{1}{2}W$.

Theorists like to clean things up by taking a limit: $\epsilon_0/W \to \infty$, $W/\Delta \to \infty$. This requires an infinite number of levels, and also knowledge of how things scale as $\epsilon_0 \to \infty$. Billiards are prominent scaling systems. For numerical calculations one must study the deviations from the limiting behavior.

To be precise we adopt the definition of Refs. [5–7]. Let the average be Gaussian, i.e., $\langle f(E)\rangle_{\epsilon_0,W} = W^{-1}\int dE\, f(E)\exp[-\pi(E - \epsilon_0)^2/W^2]$. The x integral in the Fourier transform *must also be cut off*, and again a Gaussian is chosen: $\exp[-\pi(x\Delta/2W_x)]^2$ with $W_x = W$. However, AAA tacitly choose an unspecified small W_x or at any rate make approximations requiring $W_x \ll W$. BK, after making approximations using $W_x \le W$, extend the x integral to infinity. This is clearly untenable although it was probably not meant to be taken seriously.

The choice $W_x \ll W$ has virtue. However, the same effect can be achieved by making a time convolution on $k(t)$, i.e., smearing each t by an amount τ, and we choose this route. It is also more general, since then τ can depend on the central time t, if desired.

With our definition, $k(t)$ can be rewritten

$$k(t) = \left| \sum_a F_a e^{2\pi i (E_a/\Delta)t} \right|^2 \quad (2)$$

where $F_a^2 = (\Delta/W)\exp[-\pi(E_a - \epsilon_0)^2/W^2]$. Thus the absolute square of a single sum appears, greatly simplifying the double sum. More general "window" functions, F_a, appropriately normalized, do just as well. (Figure 1 uses a "Hanning" window.) For large t, one expects just the diagonal terms in the double sum to survive, so that in some sense $k \to 1$ in this limit. At very small t, it is expected on general grounds that $k(t) \simeq \delta(t)$. But for small t, the sum can be replaced by an integral, and a Gaussian of width Δ/W replaces the δ function.

Equation (2) makes it clear that for the spectrum of a chaotic system, $k(t)$, for large t, is approximately given

by a *random walk* in the complex plane of about w steps. For the kind of systems envisaged, there will be significant random walk character even for $t \sim 1$. Insofar as the phases of the individual terms in the sum Eq. (2) are indeed independent random variables, one may show that the distribution of values of $y = k(t)$ is $\rho(y) = e^{-y}$. This result is independent of the form of F_a and of W, as long as there are many terms in the sums of Eq. (2), i.e., as long as $w \gg 1$. The mean value (over an ensemble of such random variables) is $\overline{k(t)} = 1$, and the variance equals the mean. These results have been numerically confirmed for the data of Fig. 1 [12].

For $t < 1$, think of the E_a's as ordered. Then it will take some number, $g(t)$, of steps, say, before the knowledge of the original phase is lost. Thus there are $w/g(t)$ effective random walk steps. The effective step length is still of order unity, because the phase changes systematically by an amount $e^{2\pi i t}$ at each step. This gives a formula for the distribution, $\rho[k(t)] = g(t)e^{-k(t)g(t)}$. A natural guess for $g(t)$ gives

$$\rho[k(t)] = \frac{1}{k_E(t)}\, e^{-k(t)/k_E(t)}, \tag{3}$$

where $k_E(t)$ is the ensemble average form factor. Using the quasiclassical approximation, Ref. [6] obtains a result implying Eq. (3) for small t. Equation (3) was in effect used in doing the time average, in Fig. 1.

Thus, for any large W, $k(t)$ suffers large fluctuations, unless t is very small. Increasing W does not change the distribution, but rather makes $k(t)$ vary more rapidly with t. If t is changed by an amount of order Δ/W, then $k(t)$ changes appreciably. For example, one estimates the random walk average

$$\left\langle \left| \frac{d}{dt} \sum_a F_a e^{2\pi i t(E_a - \epsilon_0)/\Delta} \right|^2 \right\rangle = \frac{4\pi^2}{\Delta^2} \sum_a (E_a - \epsilon_0)^2 F_a^2 \sim 2\pi w^2.$$

From this it follows that smearing the time t by an amount τ is like averaging over a number τw independent choices from the distribution $\rho[k(t)]$. Thus the "noise" in the smeared function $k_\tau(t)$ is reduced by a factor $\sqrt{1/w\tau}$. (This holds only for $\tau < 1$.) Of course, $k_\tau(t)$ changes appreciably only when t is varied by an amount of order τ.

In the limit $\tau w \to \infty$ the noise disappears, and $k_\tau(t)$ in that limit becomes self-averaging. For some purposes this is adequate. For example, consider the "proof" [4] that "all" spectral correlation functions of particular members of an ensemble of random matrices are self-averaging (or "ergodic"). Of course, $C(x)$, Eq. (1), is manifestly *not* self-averaging, since it consists of lots of δ functions, while the ensemble average is smooth. But Pandey argued that "observable" quantities involve an integral over x. Presumably he meant to take the limits of such an integral as fixed and independent of ϵ_0. Then in the limit $\epsilon_0 \to \infty$ followed by $W \to \infty$, Pandey indeed finds the energy smeared $C(x) \to$ the RMT prediction. But since the large W limit is taken first, for fixed finite x, it is the same as doing a time average with $\tau W/\Delta \to \infty$.

We need a sharper result. Based on Refs. [1,3,9], we predict for billiards lacking time reversal invariance [Gaussian unitary ensemble (GUE)] that $|\langle k(t)\rangle - k_R(t)| \sim \epsilon_0^{-1/2}$, for $|t - 1| \sim \epsilon_0^{-1/2}$. A modified ensemble average [9] of k is indicated and k_R is the RMT form factor. Here units are chosen so that $\Delta = mass = \hbar = 1$.

We consider the rescaled "signal function" $h(s) \equiv \sqrt{\epsilon_0}[k_\tau(1 + s/\sqrt{\epsilon_0}) - k_R(1 + s/\sqrt{\epsilon_0})]$ which should approach a theoretically predictable limit, of order unity, as $\epsilon_0 \to \infty$ *provided* we can choose a time smearing $\tau(\epsilon_0)$ which leads to a $k_\tau(t)$ sufficiently close to the modified ensemble prediction $\langle k(t)\rangle$.

Put $w \propto \epsilon_0^{\beta}$, and $\tau \propto \epsilon_0^{-\alpha}$. Clearly $\beta \le 1$, so that $W \ll \epsilon_0$ for large ϵ_0. If the time smearing is not to lose the signal, then $\alpha \ge \frac{1}{2}$. The noise in $k_\tau(t)$ then scales as $\epsilon_0^{(\alpha-\beta)/2}$. Multiplying by $\epsilon_0^{1/2}$ to calculate the noise in h gives a noise proportional to $\epsilon_0^{(\alpha-\beta+1)/2}$ where $\frac{1}{2}(\alpha - \beta + 1) \ge \frac{1}{4}$. Thus, we predict that the noise in the signal function *grows* at high energy, and therefore the theoretical signal, of order unity, will be swamped.

We do not believe that there can be more sophisticated schemes to extract the signal from the noise. Any such scheme must necessarily involve very high energy data, if ensemble averages are forbidden. Because the noise grows at high energy in the rescaled $h(s)$, it hurts rather than helps to use very high energy data. We conclude that it is impossible in principle to check the theoretical predictions, *absent an actual ensemble average.*

Fortunately, there are systems which have much larger "quasiclassical" corrections to the random matrix results, than do strongly chaotic billiards. Most notable is the spectrum of nontrivial zeros of the Riemann zeta function. Although the hypothetical GUE "Hamiltonian' whose spectrum coincides with the zeros is not known, or even proved to exist, the "quasiclassical" parameters are known [13]. In the language used above, with ϵ_0 having the mathematical meaning of height in the complex plane along the critical line, the mean level spacing shrinks according to $\Delta = 2\pi/\ln(\epsilon_0/2\pi)$, which we denote $\Delta \sim \epsilon_0^{-0}$. This turns out to give the scaling of the nonuniversal effects. We expect, based on the techniques of Ref. [9], that corrections to $k(t)$ of magnitude ϵ_0^{-0} over a width $|t - 1| \sim \epsilon_0^{-0}$ will exist. There should be enough data available [13] to verify this prediction.

GUE billiards with strong, nonisolated orbits will also have effects near $t = 1$ observable in principle, but they are probably too difficult in practice. The weight of a nonisolated orbit in the Gutzwiller formula goes as $\epsilon_0^{-1/4}$ rather than $\epsilon_0^{-1/2}$ as for unstable orbits. We expect corrections near $t = 1$ of magnitude $\epsilon_0^{-1/4}$ over a time window $\epsilon_0^{-1/4}$ thus allowing $\tau = \epsilon_0^{-1/4}$. This

leads to noise of magnitude $\epsilon_0^{-1/8}$ in the rescaled $h(s)$. Thus, with many energy levels in the neighborhood of the 10^8 level, such an effect might be seen in a direct calculation of $k_\tau(t)$ from the spectrum of, say, a stadium billiard containing Bohm-Aharonov flux lines to break time reversal symmetry. In general, a signal of width and strength $\epsilon_0^{-\eta}$ can in principle be extracted from the data on a single system only if $\eta \leq \frac{1}{3}$.

The discontinuity of slope at $t = 1$ in the RMT $k(t)$ for the GUE case plays an essential role in enhancing the nonuniversal effects near $t = 1$. Gaussian orthogonal ensemble (GOE) systems (with time reversal symmetry) unfortunately have very small effects at $t = 1$ and should be even harder to observe. However, Gaussian symplectic ensemble (GSE) systems, (sympletic symmetry) are favorable because the RMT result is singular at $t = 1$ and thus can suffer large corrections from the nonuniversal effects. These two cases have not been worked out in any detail up to now, however.

A similar discussion can be given for $C(x)$ directly. The nonuniversal contributions to $C(x)$ are spread over a wide x scale, however, in contrast to the case of the form factor, making $C(x)$ less convenient than $k(t)$. On the other hand, certain often used linear transforms of $k(t)$ *are* self-averaging, in agreement with Pandey [4]. Among them are $\Sigma^2(L)$ and $\Delta_3(L)$, the number variance and the Dyson-Mehta spectral rigidity. The former is given by $\Sigma^2(L) = \int dt\, k(t)[\sin \pi L t/\pi t]^2$. For $L \sim 1$ or smaller, there is an effective time average over a scale of order unity, thus eliminating the noise for large W. For large L, the contribution of the integral comes from small t, where the nonuniversal signal dominates $k(t)$.

We now ask how AAA-BK can fail to find these large fluctuations. AAA start from a supersymmetric integral representation of $C(x)$, and then make several approximations. One of these relies on $x \ll w$. If consistently imposed, this inequality is equivalent to a time average with $\tau w \gg 1$. In the limit, the noise associated with the lack of self-averaging is suppressed. However, the nonuniversal signal is also suppressed unless $\eta < \frac{1}{3}$.

BK, in an indirect and imaginative way, extend the diagonal approximation originally used [10] for $t \ll 1$ to all values of t. It is assumed that the energy average by itself validates the operation used to give the diagonal approximation which is equivalent to the idea that the actions of all orbits (unrelated by symmetry) may be separately averaged. In addition, statistics of classical orbits are invoked in the form of the Hannay-Ozorio de Almeida sum rule [14]. BK's result for $k(t)$ is formally a sum of δ functions, $\delta(t - t_p)$ at the orbit periods, but (we argue), these δ functions should have a width Δ/W. Except for t very small, the spacing between the t_p's will be much smaller than this, since the spacing decreases exponentially with increasing t_p. (This is the exponential proliferation of orbits in chaotic systems.) Invoking the sum rule [10,14] yields a smooth, non-noisy result. This is in accord with previous work [7] where it was stated that the neglected off-diagonal contributions are responsible for the lack of self-averaging.

Thus, the results announced by AAA-BK cannot be obtained from single systems unless the exponent $\eta \leq \frac{1}{3}$. This probably precludes direct confrontation of their results with experiments on single systems, although numerical experiments on the Riemann zeta can be studied. However, their important argument that the statistics of a single system approaches that of RMT is not affected by the results of this paper. And their results remain of interest whenever an ensemble of wave systems with the same classical parameters can be found.

Valuable discussions with O. Agam, B. Altshuler, T. Antonsen, Jr., E. Bogomolny, O. Bohigas, S. Fishman, B. Georgeot, J. Keating, E. Ott, D. Poilblanc, and C. Sire are gratefully acknowledged. We thank E. Bogomolny and B. Eckhardt for calling the work of M. Lombardi to our attention, who then provided us with further details. We thank B. Eckhardt and J. Main for permission to use their figure. We thank J. Bellissard for hospitality at the Université Paul Sabatier. The work was supported in part by NSF Grant No. DMR 9625549.

*Permanent address: Department of Physics, University of Maryland, College Park, MD 20742.
†Unité Mixte de Recherche 5626.

[1] O. Agam, B. L. Altshuler, and A. V. Andreev, Phys. Rev. Lett. **75**, 4389 (1995); A. V. Andreev, O. Agam, B. D. Simons, and B. L. Altshuler, Phys. Rev. Lett. **76**, 3947 (1996). Also see A. L. Andreev and B. L. Altshuler, Phys. Rev. Lett. **75**, 902 (1995); O. Agam and S. Fishman, Phys. Rev. Lett. **76**, 726 (1996).
[2] M. L. Mehta, *Random Matrices* (Academic Press, New York, 1990).
[3] E. B. Bogomolny and J. P. Keating, Phys. Rev. Lett. **77**, 1472 (1996).
[4] A. Pandey, Ann. Phys. **119**, 170–191 (1979).
[5] L. Leviander *et al.*, Phys. Rev. Lett. **56**, 2449 (1986); M. Lombardi *et al.*, Comments At. Mol. Phys. **25**, 345 (1991).
[6] N. Argaman, Y. Imry, and U. Smilansky, Phys. Rev. B **47**, 4440 (1993). Their footnote [28].
[7] N. Argaman *et al.*, Phys. Rev. Lett. **71**, 4326 (1993).
[8] B. Eckhardt and J. Main, Phys. Rev. Lett. **75**, 2300 (1995).
[9] R. E. Prange, Phys. Rev. Lett. **77**, 2447 (1996).
[10] M. V. Berry, Proc. R. Soc. London A **400**, 229 (1985).
[11] U. Eichmann, K. Richter, D. Wintgen, and W. Sandner, Phys. Rev. Lett. **61**, 2438 (1988).
[12] B. Eckhardt (private communication).
[13] A. M. Odlysko, Math. Comput. **48**, 273–308 (1987); (unpublished).
[14] J. H. Hannay and A. M. Ozorio de Almeida, J. Math. Phys. A **17**, 3429 (1984).