Det-Det correlations for quantum maps: Dual pair and saddle-point analyses

S. Nonnenmacher and M. R. Zirnbauer

Citation: Journal of Mathematical Physics **43**, 2214 (2002); doi: 10.1063/1.1462417
View online: http://dx.doi.org/10.1063/1.1462417
View Table of Contents: http://scitation.aip.org/content/aip/journal/jmp/43/5?ver=pdfcov
Published by the AIP Publishing

Det-Det correlations for quantum maps: Dual pair and saddle-point analyses

S. Nonnenmacher
Service de Physique Théorique, CEA-Saclay, 91191 Gif-sur-Yvette Cedex, France

M. R. Zirnbauer
Institut für Theoretische Physik, Universität zu Köln, Zülpicherstr. 77, 50937 Köln, Germany

(Received 25 September 2001; accepted for publication 30 Janunary 2002)

An attempt is made to clarify the ballistic nonlinear sigma model formalism recently proposed for quantum chaotic systems, by looking at the spectral determinant $Z(s)=\mathrm{Det}(1-sU)$ for quantized maps $U\in \mathrm{U}(N)$, and studying the correlator $\omega_U(s)=\int d\theta |Z(e^{i\theta}s)|^2$. By identifying $\mathrm{U}(N)$ as one member of a dual pair acting in the spinor representation of $\mathrm{Spin}(4N)$, the expansion of $\omega_U(s)$ in powers of s^2 is shown to be a decomposition into irreducible characters of $\mathrm{U}(N)$. In close analogy with the ballistic nonlinear sigma model, a coherent-state integral representation of $\omega_U(s)$ is developed. For generic U this integral has $\binom{2N}{N}$ saddle points and the leading-order saddle-point approximation turns out to reproduce $\omega_U(s)$ *exactly*, up to a constant factor. This miracle is explained by interpreting $\omega_U(s)$ as a character of $\mathrm{U}(2N)$, and arguing that the leading-order saddle-point result corresponds to the *Weyl character formula*. Unfortunately, the Weyl decomposition behaves nonsmoothly in the semiclassical limit $N\to\infty$, and to make further progress some additional averaging needs to be introduced. Several schemes are investigated, including averaging over basis states and an ''isotropic'' average. The saddle-point approximation applied in conjunction with these schemes is demonstrated to give incorrect results in general, one notable exception being a semiclassical averaging scheme, for which all loop corrections vanish identically. As a side product of the dual pair decomposition with isotropic averaging, the crossover between the Poisson and CUE limits is obtained. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1462417]

I. INTRODUCTION

One of the striking characteristics of a quantized chaotic Hamiltonian system is found in the correlations inherent in its spectrum at small energy differences. Extensive numerical work has shown that various quantities (such as the nearest-neighbor spacing distribution and the two-level correlation function) of a quantum chaotic system are *universal*: their behavior coincides with that of a Wigner–Dyson random matrix ensemble of the appropriate symmetry class.[1] This property, first noticed in billiards, was found to apply to many chaotic systems, including symplectic *maps*. In contrast, if the dynamics is integrable (in the sense that the $2f$-dimensional phase space foliates into f-dimensional submanifolds invariant under the Hamiltonian flow), the generic behavior of the eigenvalues is expected[2] to be that of independent random variables, so that their correlations are in the Poisson universality class.

The present article will be concerned with quantum maps, i.e., with quantizations of some canonical transformation $\phi: M\to M$ of a compact symplectic manifold M. We assume that the problem of quantization itself has been tackled, so the phase space has been prequantized into a Hilbert space $\mathcal{H}_N$ of dimension $N\sim\hbar^{-1}$, and the quantum map acts on it as a unitary operator.[3,4] With respect to a basis of $\mathcal{H}_N$ this operator is represented by an $N\times N$ unitary matrix $U_{\phi,N}$. The latter has a semiclassical limit, in the sense that traces of its powers can be estimated in terms of

classical periodic points.[4] For a system with one degree of freedom, the Gutzwiller-Tabor trace formula reads

$$\mathrm{Tr}(U_{\phi,N}^n) \overset{N\to\infty}{\sim} \sum_{p\subset \mathrm{Fix}(\phi^n)} N^{\dim(p)/2} A_p \mathrm{e}^{\mathrm{i}N\Phi p}, \tag{1}$$

where p is one component of the set of n-periodic points; for an Anosov system, it is an isolated point $[\dim(p)=0]$, whereas if the dynamics conserves energy, p is one-dimensional. Φ_p and A_p are purely classical quantities related to the dynamics around the set p.

The quantum spectrum consists of the N eigenvalues (pseudo-energies) $\{\mathrm{e}^{\mathrm{i}\theta_j}\}_{j=1,\ldots,N}$ of $U_{\phi,N}$. The first analytical estimates of the two-level correlation function [which is the Fourier transform of the *form factor* $F(n)=|\mathrm{Tr}(U_{\phi,N}^n)|^2$] for such spectra were based on the above trace formula, combined with some known ergodic properties of long periodic orbits.[5] In the present article we focus attention on another statistic, namely the autocorrelation function of the spectral determinant:

$$\Omega_U(\gamma) \overset{\mathrm{def}}{=} \gamma^{-N/2} \int_0^{2\pi} \frac{d\phi}{2\pi} \mathrm{Det}(1-\gamma \mathrm{e}^{\mathrm{i}\phi} U) \mathrm{Det}(1-\mathrm{e}^{-\mathrm{i}\phi} U^\dagger). \tag{2}$$

(The parameter γ will be a complex number close to unity, with the scaling $|\gamma-1|\sim 1/N$.) This correlation function has already been considered[6,7] for chaotic versus integrable quantum maps, and the same universality was observed as for the form factor or the nearest-neighbor distribution. A semiclassical analysis of this correlation function was performed using the Gutzwiller trace formula in Ref. 7.

The computation of correlation functions from the trace formula (1) always requires some sort of averaging. In the semiclassical theory of the form factor $F(n)$ one wants to use the so-called diagonal approximation, neglecting the off-diagonal terms in the double-sum over periodic orbits.[5] To justify this step one must average over energy or some family of systems: one needs slight variations in the classical actions Φ_p to make the phase interferences $\sim \mathrm{e}^{\mathrm{i}N(\Phi_p-\Phi_{p'})}$ average to zero. The need for averaging was emphasized in Refs. 6 and 8, where it was pointed out that the spectrum of an individual quantum system is too noisy to allow universality to be seen in its bare form factor. In addition to the noise problem, there exist some quantum chaotic systems with arithmetic symmetries, which lead to periodic orbit degeneracies and nonuniversal spectral correlations.[9,10] (Such systems are nongeneric, however, in any decent space of smooth maps.)

Thus universal behavior is expected only in the generic case, and to make a correct mathematical statement about universality of the spectral correlations of a general system one ought to define the precise meaning of the word ''generic.'' (In the case of integrable systems, the spectral correlations could sometimes be studied directly, by utilizing the explicit expressions for the eigenvalues; two-point correlations were shown to be Poisson for a rather subtle set of parameters.)[11] We can avoid the issue of genericity by averaging the correlation function over some set of quantum maps. That is, we specify a measure $\mathrm{d}\mathcal{P}_N(U)$ on the unitary group $\mathrm{U}(N)$, and the function to be studied then reads

$$\langle \Omega_U(\gamma) \rangle = \int_{\mathrm{U}(N)} \mathrm{d}\mathcal{P}_N(U) \Omega_U(\gamma).$$

We want this measure to be very concentrated (or ''local'') around the quantum map $U_{\phi,N}$ (see Secs. IV C and VI). In the course of this article, we will also consider cases for which this measure has a broader support (Sec. V), including the case where the measure is the Haar measure on $\mathrm{U}(N)$. We are then dealing with the circular unitary ensemble (CUE), for which the determinant correlation function has been thoroughly investigated.[6,12]

The Gutzwiller trace formula has the attractive feature of relating quantum to classical properties, but its use for estimating the spectral correlation functions still raises questions. The prob-

lem is that the formula is rigorous in general only for times shorter than the Ehrenfest time: $n \leqslant \log N$ (and $N \to \infty$). Yet, long-time traces ($n \sim N$) are needed to obtain spectral correlation functions at the scale of the mean level spacing where universality emerges. The diagonal approximation, which assumes statistical independence of the different periodic orbits, is unsatisfactory at large times where the exponential proliferation of periodic orbits clashes with the finite (N) number of eigenvalues: the classical information is then overcomplete, which implies some sort of hidden correlation between the contributions from classical orbits. A recipe to overcome this difficulty has been devised by Bogomolny and Keating,[13] but so far lacks rigorous justification.

To bypass these problems, a second approach to estimate spectral correlations has recently emerged, inspired by the study of disordered metals. It consists in expressing the correlation function as a quantum field theory (or functional integral) of the type of a nonlinear sigma model (NLσM). One then tries to analyze the functional integral by standard field-theoretic methods such as perturbation expansion, saddle-point analysis and the renormalization group. This approach was first applied successfully to systems with disorder, where the dynamics is governed by a diffusion operator.[14] The formalism was later extended to the ''ballistic case,'' [15,16] and quantum correlation functions were put in relation with the spectrum of the Frobenius–Perron operator (i.e., the evolution operator for classical densities). Although quite elegant, this approach suffered from several drawbacks. Among these are the appearance of unwanted zero modes around the main saddle point, and the problem of ''mode locking.'' [17] Besides, the results do not exactly agree with the correlations calculated numerically for the Riemann zeta function (the prototype of a quantum chaotic spectral determinant);[13] nor do they explain the nongeneric spectral correlations featured by systems with arithmetic symmetries. More recent treatments of the ballistic NLσM have also stressed the need for averaging over a smooth disorder if one wants to avoid the above problems.[18,19]

In an attempt to resolve these uncertainties, we have adapted the latter approach, which had originally been conceived for Hamiltonian systems, to the case of quantum maps $U_{\phi,N}$. Our objective was to prove the universality of the determinant correlation function $\Omega_U(\gamma)$ (Sec. II B) upon averaging wrt a suitable measure on $\mathrm{U}(N)$. This correlation function is easier to treat than the two-level correlation function, as it does not require the use of a supersymmetric representation[15,16] but can be expressed as an ordinary c-number integral over a finite-dimensional manifold $\mathcal{M}_N$ (Sec. II D). We write this integral in the form

$$\Omega_U(\gamma) = \int_{\mathcal{M}_N} dQ \mathrm{e}^{-S(\gamma,U,Q)}, \tag{3}$$

where $S(\gamma,U,\cdot)$ is called the *effective action*.

For the reasons stated, we will consider averages of Ω_U with respect to certain probability measures $\mathrm{d}\mathcal{P}_N(U)$. The averaged correlation function, denoted by $\langle \Omega_U(\gamma) \rangle$, can still be obtained by integrating the Boltzmann weight given by an effective action:

$$\langle \Omega_U(\gamma) \rangle = \int_{\mathcal{M}_N} dQ \mathrm{e}^{-S_{\mathrm{av}}(\gamma,Q)}.$$

To estimate these integrals, we apply the same technique that was used in Ref. 15: we expand $S_{\mathrm{av}}(\gamma,Q)$ up to quadratic order around its saddle-points Q_{crit}, and perform the Gaussian integrals. The result obtained in this way,

$$\langle \Omega_U(\gamma) \rangle_{|\text{s.p. exp.}} = \sum_{Q_{\mathrm{crit}}} \{\mathrm{Det}\delta^2 S_{\mathrm{av}}(\gamma,Q_{\mathrm{crit}})\}^{-1/2} \mathrm{e}^{-S_{\mathrm{av}}(\gamma,Q_{\mathrm{crit}})}, \tag{4}$$

is called the leading-order saddle-point expansion of the integral.

Owing to the absence of a large parameter in front of the action S_{av}, the expansion is *a priori* not justified mathematically. A more careful treatment should in principle include perturbative corrections around each critical point (we actually compute the expansion up to two loops in a particular case, see Sec. IV D).

We have succeeded in computing the leading-order term for a few averaging schemes. For an individual matrix U_N we can actually reproduce the exact value of the correlation function (3) in this way (Sec. III). In Sec. IV C we define a ''semiclassical'' averaging scheme, which we think is a good candidate to obtain universality of correlations;[20] unfortunately, in that case we can only compute the contributions from the two standard saddle points.

In order to test the leading-order saddle-point approximation, we selected a sequence of statistical ensembles [i.e., a sequence of measures $\mathrm{d}\mathcal{P}_N(U)$] for which the averaged correlation function can be computed exactly, and compared the exact result with the saddle-point approximation for the corresponding effective action. All these ensembles are U(N)-rotation invariant, that is, we first average over all bases of $\mathcal{H}_N$ (Sec. V), then possibly over the spectrum of U_N (Secs. V C and VI A). In most cases, the saddle-point expansion of these ensembles yields erroneous results. We still hope that the expansion is better behaved in the case of local averages, like the semiclassical one.

These disappointing results seem to challenge the use of NLσM methods for the study of quantum ballistic systems, unless our understanding and control of these methods significantly improves. In Sec. VI, we introduce a U(N)-isotropic local averaging scheme which we treat by an alternative method; unfortunately, this scheme does not discriminate between the different universal behaviors that are expected for chaotic versus integrable maps. Nevertheless, we use it in Sec. VI A to compute the correlations along a crossover between the Poisson and CUE universality classes.

II. ALGEBRAIC MANIPULATION OF Ω_U

A. Fourier decomposition of Ω_U

We first remind the reader of some known results concerning the correlation function Ω_U.[6,7] The spectral determinant of $U \in \mathrm{U}(N)$ may be expanded as

$$\mathrm{Det}(1-sU)=\sum_{k=0}^{N} s^k a_k(U). \tag{5}$$

The unitarity of U implies a ''self-inversive'' property for the secular coefficients:[21]

$$a_{N-k}(U)=\mathrm{Det}(-U)a_k(\overline{U}).$$

Each coefficient a_k may be obtained from the traces $\{t_l=\mathrm{Tr}U^l\}$ by

$$a_k=-\frac{1}{k}\left(t_k+\sum_{l=1}^{k-1} a_l t_{k-l}\right)=\frac{(-1)^k}{k!}\begin{vmatrix} t_1 & t_2 & t_3 & \dots & t_k \\ 1 & t_1 & t_2 & \dots & t_{k-1} \\ 0 & 2 & t_1 & \ddots & t_{k-2} \\ 0 & 0 & 3 & & \vdots \\ \vdots & \ddots & & & \vdots \\ 0 & \dots & \dots & k-1 & t_1 \end{vmatrix}.$$

Because this dependence is highly nonlinear, the secular coefficients inherit non-Gaussian distributions in the RMT ensembles.[6] However, to compute the ensemble averages of $\Omega_U(\gamma)$ one only needs to know their variances, since

$$\Omega_U(\gamma)=\sum_{k=0}^{N}\ \gamma^{k-N/2}\ \ |a_k|^2=\sum_{k=0}^{N/2}\ (\gamma^{k-N/2}+\gamma^{N/2-k})|a_k|^2. \tag{6}$$

For the Poisson and the CUE ensemble of random matrices, these variances were computed in Ref. 6, and have the following large-N asymptotics:

$$\langle|a_k|^2\rangle_{\text{Poisson}}=\binom{N}{k},\quad \langle|a_k|^2\rangle_{\text{CUE}}=1, \tag{7}$$

$$\langle\Omega_U(e^{ix/N})\rangle_{\text{Poisson}}\sim 2^N,\quad \langle\Omega_U(e^{ix/N})\rangle_{\text{CUE}}\sim N\frac{\sin(x/2)}{x/2}. \tag{8}$$

In Ref. 7, a semiclassical estimation of the $|a_k|^2$ was given for integrable and chaotic quantum maps. The authors used the explicit expression in terms of the traces t_l, and estimated the latter by the Gutzwiller trace formula (1). They made a generalized diagonal approximation treating the traces t_l as statistically independent variables. To obtain the correlation function, one has to estimate the $|a_k|^2$ (and hence the t_k) up to times $k\lesssim N/2$, where the Gutzwiller formula is nonrigorous.

B. Representation-theoretic content of Ω_U

We now introduce a more group-theoretic expression for the correlation function. Instead of performing the expansion (5), we will express $\Omega_U(\gamma)$ as a *character* in a certain irreducible representation of U($2N$), which is best described using the physical language of fermions.

Let $\mathcal{F}_N$ be the Fock space for N types of fermions f_i, $f_i^\dagger$. In mathematics $\mathcal{F}_N$ is known as the spinor representation space of the group Spin($2N$). Then, for any $N\times N$ unitary matrix U,

$$\mathrm{Det}(1-U)=\mathrm{Tr}_{\mathcal{F}_N}(-1)^{\sum_i f_i^\dagger f_i}\exp\sum_{i,j=1}^{N} f_i^\dagger(\log U)_{ij}f_j.$$

The exponential on the right-hand side can be shown to be well defined in spite of the multivaluedness of $\log U$. To account for both determinants, we use $2N$ fermions, whose creation operators are denoted by $f_{+j}^\dagger$ and $f_{-j}^\dagger$, $j=1,\ldots,N$. The integration over ϕ in the integral (2) projects on the subspace $\mathcal{F}\underset{\text{def}}{=}\mathrm{Ker}(F_+-F_-)$, where $F_\pm=\Sigma_i f_{\pm i}^\dagger f_{\pm i}$ are the number operators for the two types of fermions. The correlation function reads

$$\Omega_U(\gamma)=\mathrm{Tr}_{\mathcal{F}}\gamma^{(F_++F_--N)/2}\exp\sum_{i,j=1}^{N}(\log U)_{ij}(f_{+i}^\dagger f_{+j}-f_{-j}^\dagger f_{-i}). \tag{9}$$

The operator under the trace belongs to an irreducible representation R of the group U($2N$), realized on the space $\mathcal{F}$, which has dimension $\binom{2N}{N}$. This representation may be defined through its Lie algebra version: any skew-Hermitian $2N\times 2N$ matrix $\mathbf{X}=\left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)$ is represented by the operator

$$R(\mathbf{X})=\sum_{i,j=1}^{N} a_{ij}f_{+i}^\dagger f_{+j}+b_{ij}f_{+i}^\dagger f_{-j}^\dagger+c_{ij}f_{-i}f_{+j}+d_{ij}f_{-i}f_{-j}^\dagger. \tag{10}$$

By exponentiating, $R(\exp\mathbf{X})=\exp R(\mathbf{X})$, we obtain a U($2N$)-representation, which we still denote by R. The correlation function $\Omega_U(\gamma)$ for any $N\times N$ unitary matrix U may be recast as a character in this representation:

$$\Omega_U(\gamma)=\gamma^{-N/2}\mathrm{Det}(U)^{-1}\ \mathrm{Tr}\,R(\Gamma\mathbf{U}) \tag{11}$$

where

$$\mathbf{U} \overset{\text{def}}{=} \begin{pmatrix} U & 0 \\ 0 & U \end{pmatrix}, \quad \Gamma \overset{\text{def}}{=} \begin{pmatrix} \gamma & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{U}(2N).$$

As it stands, the construction assumes $\gamma = e^{i\theta} \in \mathrm{U}(1)$. It can also be used for other values of γ, since R naturally extends to a representation of $\mathrm{GL}(2N, \mathbb{C})$. In the following, matrices in bold print will always be of size $2N \times 2N$.

The assignment $U \mapsto \mathbf{U}$ embeds $\mathrm{U}(N)$ into $\mathrm{U}(2N)$. By this embedding, R restricts to a *reducible* representation of $\mathrm{U}(N)$ on $\mathcal{F}$, which we simply denote by $R(U)$. To express the correlation function, we may also consider the $\mathrm{U}(N)$-representation $R_{-1}(U) \overset{\text{def}}{=} \det(U)^{-1} R(U)$.

In the next section, we decompose $R(U)$ [or equivalently $R_{-1}(U)$] into irreducible representations (irreps) of $\mathrm{U}(N)$, thus expressing the correlation function $\Omega_U(\gamma)$ as a sum of $\mathrm{U}(N)$-characters.

C. Ω_U as a sum of U(N)-characters

The crucial mathematical tool to use is the *dual pair* structure.[22] The subalgebra $\{X_N \otimes \mathbb{I}_2 \,|\, X_N \in \mathfrak{u}(N)\}$ of $\mathfrak{u}(2N)$ commutes with the subalgebra $\{\mathbb{I}_N \otimes x_2 \,|\, x_2 \in \mathfrak{u}(2)\}$, and each is the commutant of the other inside $\mathfrak{u}(2N)$: they are said to form a dual pair. This means that for all $U \in \mathrm{U}(N)$, the operator $R(U)$ commutes with the set

$$J_\uparrow = \sum_i f_{+i}^\dagger f_{-i}^\dagger,$$

$$J_\downarrow = \sum_i f_{-i} f_{+i},$$

$$J_0 = F_+ + F_- - N.$$

The operators J_0, $J^\uparrow$ and $J^\downarrow$ generate an $\mathfrak{su}(2)$ algebra. The equation $J_0 R(U) = R(U) J_0$ implies that $R(U)$ conserves the total number of particles and hence acts inside the subspaces $\mathcal{F}^p = \mathcal{F} \cap \mathrm{Ker}(F_+ + F_- - 2p)$.

The dual pair structure provides us with a prescription[22] to decompose $R(U)$. Inside the reduced Fock space $\mathcal{F}$, we consider the subspace of lowest SU(2) weights, ${}^0\mathcal{F} = \mathcal{F} \cap \mathrm{Ker} J_\downarrow$, and expand it according to its particle content: ${}^0\mathcal{F}^p \overset{\text{def}}{=} {}^0\mathcal{F} \cap \mathcal{F}^p$. Classical results of invariant theory, due mostly to H. Weyl[23] and succinctly summarized by R. Howe,[22] amount to the following statements:

(i) The operator $R(U)$ acts inside each space ${}^0\mathcal{F}^p$, through a certain irrep $\tilde{\rho}_p(U)$ of $\mathrm{U}(N)$. Equivalently, $R_{-1}(U)$ acts on this space through $\rho_p(U) = \mathrm{Det}(U)^{-1} \tilde{\rho}_p(U)$. Furthermore, two irreps ρ_p and $\rho_{p'}$ are inequivalent if $p \neq p'$.

(ii) The image of ${}^0\mathcal{F}^p$ under $(J_\uparrow)^k$ is the space ${}^k\mathcal{F}^{p+k} \subset \mathcal{F}^{p+k}$ which is either trivial (if $k > N - 2p$) or carries the irrep ρ_p (if $k \leqslant N - 2p$). The operators $J_\uparrow$, $J_\downarrow$, J_0 act on this tower of spaces according to the $\mathfrak{su}(2)$-irrep of dimension $N - 2p + 1$.

(iii) The direct sum of these towers exhausts $\mathcal{F}$.

We summarize these statements in the following diagram. All entries in a given row are subspaces containing the same number of fermions; all entries in a given column (or tower) carry the same $\mathrm{U}(N)$-irrep. We only show the case where N is an even integer (the odd-N case being very similar):

$$
\begin{array}{lcccccccc}
\mathcal{F}^N & = & {}^N\mathcal{F}^N & & & & & & \\
 & & \uparrow J_\uparrow & & & & & & \\
\mathcal{F}^{N-1} & = & {}^{N-1}\mathcal{F}^{N-1} & \oplus & {}^{N-2}\mathcal{F}^{N-1} & & & & \\
\vdots & & \uparrow J_\uparrow & & \uparrow J_\uparrow & & \ddots & & \\
\vdots & & \vdots & & \vdots & & & & {}^0\mathcal{F}^{N/2} \\
\vdots & & \uparrow J_\uparrow & & \uparrow J_\uparrow & & \uparrow J_\uparrow & & \\
\mathcal{F}^2 & = & {}^2\mathcal{F}^2 & \oplus & {}^1\mathcal{F}^2 & \oplus & {}^0\mathcal{F}^2 & & \\
 & & \uparrow J_\uparrow & & \uparrow J_\uparrow & & & & \\
\mathcal{F}^1 & = & {}^1\mathcal{F}^1 & \oplus & {}^0\mathcal{F}^1 & & & & \\
 & & \uparrow J_\uparrow & & & & & & \\
\mathcal{F}^0 & = & {}^0\mathcal{F}^0 & & & & & & \\
U(N)-\text{irreps:} & & \rho_0 & & \rho_1 & & \rho_2 & \cdots & \rho_{N/2}
\end{array}
\tag{12}
$$

The leftmost tower on the right-hand side carries the trivial $\mathfrak{u}(N)$-irrep, so all spaces ${}^p\mathcal{F}^p = (J_\uparrow)^p\,{}^0\mathcal{F}^0$ are one-dimensional.

Each irrep ρ_p (or $\tilde{\rho}_p$) may be described by a Young diagram. ρ_p mixes the action of U on p fermions f_+ with the action of $\overline{U}$ on p fermions f_-. Owing to antisymmetrization, it corresponds to the diagram with p rows of length 2 followed by $N-2p$ rows of length one:

$$\tilde{\rho}_p(U_N) = \mathrm{Det}(U_N)\rho_p(U_N) = U_N^{[2^p\,1^{N-2p}]}.$$

In view of the above diagram, the dimensions of the representation spaces ${}^k\mathcal{F}^{k+p}$ follow immediately from those of the spaces $\mathcal{F}^p$:

$$\dim({}^k\mathcal{F}^{k+p}) = \dim(\mathcal{F}^p) - \dim(\mathcal{F}^{p-1}) = \binom{N}{p}^2 - \binom{N}{p-1}^2. \tag{13}$$

By doing the sum over each $\mathfrak{su}(2)$-multiplet we can now express the correlation function (11) in terms of the irreps ρ_p:

$$\Omega_U(\gamma) = \sum_{p=0}^{N/2} \mathrm{Tr}\rho_p(U)\,\frac{\gamma^{p-N/2} - \gamma^{N/2+1-p}}{1-\gamma}, \tag{14}$$

or, making the substitution $\gamma = \mathrm{e}^{\mathrm{i}x/N}$,

$$\Omega_U(\mathrm{e}^{\mathrm{i}x/N}) = \sum_{p=0}^{N/2} \mathrm{Tr}\rho_p(U)\,\frac{\sin\{(x/2)\,[1-(2p-1)/N]\}}{\sin(x/2N)}.$$

For large values of N we may replace the denominator $\sin(x/2N)$ by $x/2N$. A quick comparison shows that this decomposition is actually equivalent to the pedestrian expansion (6) written down in Sec. II A. The squared coefficients $|a_k|^2$ now acquire a representation-theoretic meaning:

$$\forall p \leqslant N/2: \quad |a_p(U)|^2 = \mathrm{Tr}_{\mathcal{F}^p} R_{-1}(U) = \sum_{k=0}^{p} \mathrm{Tr}\rho_k(U), \tag{15}$$

or, equivalently,

$$\mathrm{Tr}\rho_p(U) = |a_p(U)|^2 - |a_{p-1}(U)|^2.$$

As it stands, the decomposition into irreducibles (14) is not very informative if one takes for U the matrix of a quantum map. We have no way *a priori* to estimate the character $\mathrm{Tr}\rho_p(U_N)$ from semiclassical information, except by using the relationship, via the $|a_k|^2$, to the original traces $\mathrm{Tr}(U_N^k)$, as was done in Ref. 7. This decomposition will, however, allow us to obtain rigorous results when adopting a $\mathrm{U}(N)$-isotropic averaging centered around U_N (see Sec. VI).

D. Ω_U as a coherent-state integral

Instead of decomposing the character $\mathrm{Tr}R_{-1}(U)$ into irreducibles, we can rewrite it as an integral over the symmetric space $\mathcal{M}_N=\mathrm{U}(2N)/\mathrm{U}(N)\times\mathrm{U}(N)$. This integral can be interpreted as a variant of the nonlinear sigma model used in Ref. 15 to study the spectral statistics of quantum chaotic Hamiltonians on infinite-dimensional Hilbert spaces. In our case the integral representation is exact, and is well defined mathematically.

To write the character $\Omega_U(\gamma)$ as an integral, one uses the coherent states $R(\mathbf{g})|0\rangle$, where $|0\rangle$ is the vacuum of $\mathcal{F}$ and $\mathbf{g}$ is any matrix in $\mathrm{U}(2N)$.[24] These coherent states provide a resolution of unity on $\mathcal{F}$, i.e., they can be combined to build the orthogonal projector on $\mathcal{F}$, as

$$P_{\mathcal{F}}\overset{\mathrm{def}}{=}\int_{\mathrm{U}(2N)} d\mathbf{g}\quad R(\mathbf{g})|0\rangle\langle 0|R(\mathbf{g})^{-1},$$

where the Haar measure $d\mathbf{g}$ has to be suitably normalized. Let H_N be the block-diagonal subgroup $\mathrm{U}(N)\times\mathrm{U}(N)$ of $\mathrm{U}(2N)$. Then for all $\mathbf{h}\in H_N$, the states $R(\mathbf{g})|0\rangle$ and $R(\mathbf{gh})|0\rangle$ only differ by a phase factor. Therefore, it suffices to integrate over the equivalence classes in $\mathrm{U}(2N)$ modulo H_N:

$$P_{\mathcal{F}}=\int_{\mathrm{U}(2N)/H_N} d[\mathbf{g}]_H R(\mathbf{g})|0\rangle\langle 0|R(\mathbf{g})^{-1}.$$

It is convenient to represent the H_N-equivalence classes (i.e., the points on $\mathcal{M}_N$) by $2N\times 2N$ matrices. To each $\mathbf{g}\in\mathrm{U}(2N)$ one associates $Q_{\mathbf{g}}=\mathbf{g}\Sigma_3\mathbf{g}^{-1}$, where $\Sigma_3=\mathbb{I}_N\otimes\sigma_3$. The set of all these matrices Q is isomorphic to $\mathcal{M}_N$. It is the set of all Hermitian matrices with two eigenvalues, $+1$ and -1, each with multiplicity N. This nonlinear set of matrices is naturally equipped with $\mathrm{U}(2N)$-invariant symplectic structure and metric (and therefore an invariant measure dQ).

The matrix elements Q_{ij} are not all independent, and for practical calculations we need to introduce a bona fide coordinate system on $\mathcal{M}_N$. If we denote by $Q_{\mathbf{12}}$, $Q_{\mathbf{22}}$ the two $N\times N$ blocks in the right half of the matrix Q, the entries of the complex matrix $Z=Q_{\mathbf{12}}(Q_{\mathbf{22}}-1)^{-1}$ are good coordinates on the open subset of $\mathcal{M}_N$ where $(Q_{\mathbf{22}}-1)$ is invertible. Geometrically, these $N\times N$ complex coordinates represent a certain stereographic mapping of $\mathcal{M}_N$ onto $\mathbb{C}^{N\times N}$. The matrix Z corresponding to a point $Q_{\mathbf{g}}$ can be extracted from the Gaussian decomposition of $\mathbf{g}$:

$$\mathbf{g}=\begin{pmatrix}1 & Z\\ 0 & 1\end{pmatrix}\begin{pmatrix}A & 0\\ C & D\end{pmatrix}. \tag{16}$$

These complex coordinates also provide a simple definition of the coherent states. Indeed, $R(\mathbf{g})|0\rangle$ is co-linear with

$$|Z\rangle\overset{\mathrm{def}}{=}\exp\left\{\sum_{i,j=1}^{N} f_{+i}^{\dagger}Z_{ij}f_{-j}^{\dagger}\right\}|0\rangle=\exp\left\{R\begin{pmatrix}0 & Z\\ 0 & 0\end{pmatrix}\right\}|0\rangle=R\begin{pmatrix}1 & Z\\ 0 & 1\end{pmatrix}|0\rangle. \tag{17}$$

As it stands, $|Z\rangle$ is not normalized, but has the following properties:

(i) The overlap between two coherent states reads $\langle Z|Z'\rangle=\mathrm{Det}(1+Z^{\dagger}Z')$. In particular, the norm of $|Z\rangle$ is $\mathrm{Det}(1+Z^{\dagger}Z)^{1/2}$.

(ii) The resolution of unity takes the form

$$P_{\mathcal{F}}=\int_{\mathbb{C}^{N\times N}} d\mu_N(Z,Z^\dagger)\frac{|Z\rangle\langle Z|}{\mathrm{Det}(1+Z^\dagger Z)}, \tag{18}$$

where the measure $d\mu_N(Z,Z^\dagger)=C_N\times\mathrm{Det}(1+Z^\dagger Z)^{-2N}\Pi_{i,j=1}^N d^2Z_{ij}/\pi$ is the expression for dQ in the coordinates Z_{ij}. The value of the normalization factor C_N is given at the end of Appendix B.

(iii) The group $\mathrm{U}(2N)$ acts on these coherent states as follows:

$$R\begin{pmatrix} A & B \\ C & D \end{pmatrix}|Z\rangle=\mathrm{Det}(CZ+D)\,|(AZ+B)(CZ+D)^{-1}\rangle. \tag{19}$$

The resolution of unity allows us to write the character (11) as

$$\begin{aligned}\Omega_U(\gamma)&=\gamma^{-N/2}\mathrm{Det}(U)^{-1}\int_{\mathcal{M}_N} d[\mathbf{g}]_H\ \langle 0|R(\mathbf{g})^{-1}R\begin{pmatrix}\gamma U & 0\\ 0 & U\end{pmatrix}R(\mathbf{g})|0\rangle\\ &=\gamma^{-N/2}\int_{\mathbb{C}^{N\times N}} d\mu_N(Z,Z^\dagger)\ \frac{\mathrm{Det}(1+\gamma Z^\dagger UZU^{-1})}{\mathrm{Det}(1+Z^\dagger Z)}\\ &=\int_{\mathcal{M}_N} dQ\ \ \mathrm{e}^{-S(\gamma,U,Q)}.\end{aligned} \tag{20}$$

This expression is the central result of the current section. It is an exact formula, which parallels the ''ballistic'' nonlinear sigma model derived in Ref. 15 for Hamiltonian systems with an infinite-dimensional Hilbert space. In our finite-dimensional framework, the non local field $Q(q',q)$ of 4×4 supermatrices on configuration space is replaced by a ''lattice field'' $Q_{i\alpha,j\beta}$ of 2×2 matrices (with elements indexed by α,β) depending on two discrete positions i,j. The ''effective action'' of the present model,

$$S(\gamma,U,Q)=-\mathrm{Tr}\{\log(1+\gamma Z^\dagger UZU^{-1})-\log(1+ZZ^\dagger)\}+\frac{N}{2}\log\gamma, \tag{21}$$

can be presented[17] in the form

$$S(\mathrm{e}^{\mathrm{i}x/N},U,Q)=-\mathrm{Tr}\log[\cosh(H_{x,U})-\sinh(H_{x,U})Q],$$

with

$$H_{x,U}\overset{\mathrm{def}}{=}\frac{\mathrm{i}x}{4N}\Sigma_3+\frac{1}{2}\log\mathbf{U}.$$

In Ref. 17, this action was further transformed, using the Wigner representation of wave functions, to obtain the same ballistic non-linear sigma model as in Ref. 15. We will not perform these steps, which require some further approximations, but rather try to estimate the integral with the above (purely quantum) effective action.

III. SADDLE-POINT ANALYSIS OF THE ACTION $S(\gamma,U,Q)$

To estimate the field integral of their nonlinear sigma model, the authors in Ref. 15 expand the effective action around two critical points (usually referred to as saddle points in this context). Since there is no large parameter in front of this action, a leading-order saddle-point expansion—see Eq. (4)—is not justified mathematically *a priori*. In the present section we explicitly compute this expansion for the action (21) and compare it to the results of Ref. 15 and the exact correlation function.

The saddle points are determined by requiring the variation of the action to be zero. In the absence of a large parameter, one first needs to understand exactly which action to vary. This point is not entirely obvious: one might be tempted to lift (part of) the denominator $\mathrm{Det}(1+ZZ^\dagger)^{-2N}$ of the measure $d\mu_N(Z,Z^\dagger)$ into the exponent; this modification of the effective action would yield a different saddle-point expansion. However, the requirement of coordinate invariance tells us to keep the U($2N$)-invariant measure dQ as it is, forbidding such manipulations. With this convention the saddle-point expansion of $S(\gamma,U,Q)$ will turn out to yield *the exact* γ*-dependence* for $\Omega_U(\gamma)$. In particular, the problem of "unphysical zero modes" occurring in Refs. 15 and 17 is resolved.

We now describe the saddle-point analysis of $S(\gamma,U,Q)$ in some detail. We first use the fact that the action is invariant under simultaneous rotations of both U and Q:

$$S(\gamma,U,Q)=S(\gamma,VUV^{-1},\mathbf{V}Q\mathbf{V^{-1}}), \tag{22}$$

where we used the shorthand notation $\mathbf{V}=V\otimes\mathbb{I}_2$, for $V\in\mathrm{U}(N)$. Such a V-rotation of Q is an isometry of the Riemannian manifold $\mathcal{M}_N$ and leaves the measure dQ invariant. It therefore suffices to study the simpler situation where U is *diagonal*: $U\equiv D=\mathrm{diag}(\mathrm{e}^{\mathrm{i}\theta_j})$.

One sees from formula (21) that the point $Z=0$ (or equivalently, $Q=\Sigma_3$) is a saddle point, and the quadratic approximation to S for small Z reads

$$S(\gamma,D,Z)\approx\frac{1}{2}N\log\gamma-\mathrm{Tr}(\gamma Z^\dagger DZD^{-1}-Z^\dagger Z)\approx\frac{1}{2}N\log\gamma+\sum_{i,j=1}^{N}|Z_{ij}|^2(1-\gamma\mathrm{e}^{\mathrm{i}(\theta_i-\theta_j)}).$$

This saddle point is the only one on $\mathcal{M}_N$ which is located at a finite Z. It is sometimes called the "perturbative" saddle point in the physics literature. For a generic matrix U, there are N directions Z_{jj} that have a coefficient $(1-\gamma)\sim-\mathrm{i}x/N$; these directions are called "zero modes,"[15] because their coefficient vanishes as $x\to0$. Doing the integral in this quadratic approximation around $Z=0$ yields

$$\Omega_U(\gamma)_{|\Sigma_3}=C_N\frac{\gamma^{-N/2}}{(1-\gamma)^N\Pi_{i\neq j}(1-\gamma\mathrm{e}^{\mathrm{i}(\theta_i-\theta_j)})}. \tag{23}$$

We chose to separate the zero mode contributions from the others.

The existence of a second saddle point was pointed out (in the context of the diffusive nonlinear sigma model) in Ref. 25. It may be exhibited through the change of variable $Z'=1/Z$, which amounts to switching to the stereographic projection of $\mathcal{M}_N$ from the antipodal point. In terms of the new variable Z', the integrand reads

$$\gamma^{N/2}\frac{\mathrm{Det}(1+\gamma^{-1}Z'^\dagger UZ'U^{-1})}{\mathrm{Det}(1+Z'^\dagger Z')},$$

so it has the same structure as the original integrand, but for the replacement $\gamma\to\gamma^{-1}$. Quadratic expansion around $Z'=0$ (or, equivalently, around $Q=-\Sigma_3$) yields

$$\Omega_U(\gamma)_{|-\Sigma_3}=C_N\frac{\gamma^{N/2}}{(1-\gamma^{-1})^N\Pi_{i\neq j}(1-\gamma^{-1}\mathrm{e}^{\mathrm{i}(\theta_i-\theta_j)})}. \tag{24}$$

These two saddle points $Q=\pm\Sigma_3$ (we call them "standard") are the only ones taken into account in the treatment of the ballistic nonlinear sigma model in Refs. 15 and 17. The problem with this approximation is that, in the limit $\gamma\to1$, the sum of the two contributions Eq. (23) and (24) diverges at least as strongly as $1/(1-\gamma)^{N-1}$, whereas the exact correlation function is bounded. This phenomenon was attributed to the $N-1$ "unphysical" zero modes appearing at each saddle

point (as opposed to the single ''ergodic'' zero mode $\Sigma_j Z_{jj}$). More generally, these contributions become singular each time U and γU happen to have common eigenvalues.

We will argue below that this problem with zero modes is actually resolved by taking into account *further* saddle points of the effective action.

A. Weyl character formula

To identify all saddle points, we return to the expression (20) of the integrand. We still study the case where $U=D$ is diagonal, and we write $\Gamma\mathbf{D}\equiv\mathrm{diag}(\gamma D,D)$.

Let ζ be a complex $N\times N$ matrix. The point $Q_{\mathbf{g}}$ of $\mathcal{M}_N$ is a saddle point of the integrand iff the Taylor expansion of $\langle\zeta|R(\mathbf{g}^{-1}\Gamma\mathbf{D}\mathbf{g})|\zeta\rangle$ around $\zeta=0$ contains no term linear in ζ and $\zeta^\dagger$. (Note that this statement is independent of the choice of representative $\mathbf{g}$ for $Q_{\mathbf{g}}$.) Moreover, we do not want the integrand to vanish at $\zeta=0$. If we decompose the unitary matrix as $\mathbf{g}^{-1}\Gamma\mathbf{D}\mathbf{g}=\left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)$, these conditions read

$$b=c=0, \quad \mathrm{Det}(d)\neq 0. \tag{25}$$

This means that the matrix $\mathbf{g}^{-1}\Gamma\mathbf{D}\mathbf{g}$ [for $\gamma=\mathrm{e}^{\mathrm{i}x/N}\in\mathrm{U}(1)$] belongs to the subgroup H_N of $\mathrm{U}(2N)$, which in turn allows $\mathbf{g}$ to be written as the product of a *permutation matrix* $\mathbf{g}_\sigma$ with some element $\mathbf{h}\in H_N$. By $\mathbf{g}_\sigma$ we mean the unitary matrix $(\mathbf{g}_\sigma)_{ij}=\delta_{i,\sigma(j)}$, where σ is a permutation of $\{1,\ldots,2N\}$. To each permutation σ there corresponds a single point $Q_\sigma=\mathbf{g}_\sigma\Sigma_3\mathbf{g}_\sigma{}^{-1}$. Moreover, two permutations σ, σ' lead to the same point if $\sigma=\sigma'\tau$ where τ permutes indices separately inside $\{1,\ldots,N\}$ and $\{N+1,...,2N\}$; this property defines a partition of the symmetric group $\mathfrak{S}_{2N}$ into $\binom{2N}{N}$ equivalence classes, each one corresponding to a saddle point of the integrand.

These classes are in one-to-one correspondence with the sets $S=\sigma(\{1,\ldots,N\})$, so we can write $Q_\sigma=Q_S$. Q_S is then the diagonal matrix with entries $+1$ at the positions $j\in S$, and -1 at the positions $j\in\overline{S}$ (the complement of S in $\{1,\ldots,2N\}$). We partition the set S into $S_1=S\cap\{1,\ldots,N\}$ and $\widetilde{S}_2=S\cap\{N+1,\ldots,2N\}$. In the following we will also use the set $S_2=\{j-N|j\in\widetilde{S}_2\}$, and the sets $\overline{S_1}$ and $\overline{S_2}$ which are the complements in $\{1,\ldots,N\}$ of S_1, resp. S_2. The point Q_S corresponds to the following (coherent) state in $\mathcal{F}$:

$$|S\rangle\overset{\mathrm{def}}{=}R(\mathbf{g}_\sigma)|0\rangle=\pm\prod_{i\in\overline{S_1}} f^\dagger_{+i}\prod_{j\in S_2} f^\dagger_{-j}|0\rangle. \tag{26}$$

The matrix $\mathbf{g}_\sigma$ admits a Gaussian decomposition (16) iff σ is in the trivial class, i.e., $S=\{1,\ldots,N\}$, which explains why only the perturbative saddle point $Q=\Sigma_3$ could be exhibited from the Z-coordinates.

We now compute the leading-order contribution from each saddle point Q_S. In the vicinity of Q_S the integrand in (20) takes the values

$$\langle\zeta|R(\mathbf{g}_\sigma^{-1}\Gamma\mathbf{D}\mathbf{g}_\sigma)|\zeta\rangle/\langle\zeta|\zeta\rangle,$$

where the entries of the matrix ζ are ''small'' (ζ defines a local coordinate system near Q_S). We partition the diagonal matrix $\mathbf{g}_\sigma^{-1}\Gamma\mathbf{D}\mathbf{g}_\sigma$ into two halves: $\mathbf{g}_\sigma^{-1}\Gamma\mathbf{D}\mathbf{g}_\sigma=\mathrm{diag}(\Delta_1,\Delta_2)$. The above integrand then reads

$$\mathrm{Det}(D)\times\frac{\mathrm{Det}(1+\zeta^\dagger\Delta_1\zeta\Delta_2^{-1})}{\mathrm{Det}(1+\zeta^\dagger\zeta)}.$$

Expanding to quadratic order and integrating over $\zeta,\zeta^\dagger$, we obtain from the saddle point Q_S a contribution similar to (23):

$$\Omega_D(\gamma)_{|Q_S}=C_N\gamma^{-N/2}\prod_{\substack{i\in S_1\\ j\in S_2}}\gamma e^{i(\theta_i-\theta_j)}\prod_{\substack{i\in S_1\\ j\in \overline{S_2}}}(1-\gamma e^{i(\theta_i-\theta_j)})^{-1}\prod_{\substack{i\in S_1\\ j\in S_2}}(1-\gamma^{-1}e^{-i(\theta_i-\theta_j)})^{-1}$$

$$\times\prod_{\substack{i\in \overline{S_1}\\ j\in S_1}}(1-e^{i(\theta_i-\theta_j)})^{-1}\prod_{\substack{i\in S_2\\ j\in S_2}}(1-e^{-i(\theta_i-\theta_j)})^{-1}. \tag{27}$$

Note that the product contains a factor, $(1-\gamma)^{-N+2r}$, with $r=\sharp(S_1\cap S_2)$. The most singular case $r=0$ arises for $S_2=\overline{S_1}$, i.e., saddle points of the type $\Pi_{i\in\overline{S_1}}f^\dagger_{+i}f^\dagger_{-i}|0\rangle$.

In the general case $U=VDV^{-1}$, the saddle points are the points $Q_{V,S}=\mathbf{V}Q_S\mathbf{V^{-1}}$, and they lead to the same contributions (cf. the covariance of the action and the measure dQ). The two standard saddle points $Q=\pm\Sigma_3$ are the only ones unaffected by these V-rotations.

It is illuminating to present the result of the approximation (27) in an alternative fashion. For that purpose, we denote the nonzero elements of the diagonal matrix $\Gamma\mathbf{D}$ by $e^{i\phi_\nu}$ ($\nu=1,\ldots,2N$). The sum of contributions (27) can then be rewritten in the form

$$\Omega_D(\gamma)=C_N\gamma^{-N/2}\sum_S\frac{\Pi_{\mu\in\overline{S}_1}\Pi_{\nu\in\tilde{S}_2}e^{i(\phi_\mu-\phi_\nu)}}{\Pi_{\mu\in S}\Pi_{\nu\in\overline{S}}(1-e^{i(\phi_\mu-\phi_\nu)})}. \tag{28}$$

Save for the prefactor C_N, the expression (28) agrees with the result that follows from the *Weyl character formula*[26] for the trace of $R(\Gamma\mathbf{D})$ over $\mathcal{F}$. In general, this formula expresses the character of an element of $U(2N)$ [more generally, $GL(2N,\mathbb{C})$] in some representation R as a sum over all permutations $\sigma\in\mathfrak{S}_{2N}$ [this being the so-called Weyl group of $U(2N)$]. In our case, the terms from the $(2N)!$ elements of $\mathfrak{S}_{2N}$ may be grouped into $\binom{2N}{N}$ classes, according to the equivalence relation described above. Since Weyl's formula is an exact result, the expression (28) remains finite in the limit $\gamma\to1$, which means that the singularities $1/(1-\gamma)^{N-2r}$ of the various terms cancel each other. The complete sum over saddle-point contributions thus solves the problem of "unphysical zero modes," i.e., the divergence problem of the two standard saddle points.

The mathematical reason behind the "almost exactness" of the leading-order saddle-point expansion is as follows. The action of $R(\mathbf{g})$ on coherent states $|Z\rangle$ may be interpreted as the equivariant action of $\mathbf{g}$ on the space of holomorphic sections of a certain complex line bundle $\mathcal{L}_R$ over $\mathcal{M}_N$.[27] This equivariant action can be extended to the (infinite-dimensional) space of square-integrable differential forms of degree $(0,p)$ on the bundle. On the enlarged space, the character becomes a (super)trace, which can still be written as an integral over $\mathcal{M}_N$. Owing to an $N=2$ supersymmetry, the integrand may be continuously deformed without changing the value of the integral. In one limit of the deformation, one gets $\mathrm{Tr}R(\mathbf{g})$; in the other, the integrand *localizes* at the fixed points of $\mathbf{g}$ on $\mathcal{M}_N$, yielding Gaussian integrals around these points.

It turns out that these fixed points coincide with our $Q_{V,S}$, and their (Gaussian) contributions are equal to (27), save for the prefactor C_N. As a result, the leading-order saddle-point approximation (for our nonlocalized integrand) delivers the correct answer (omitting the prefactor). In Sec. IV D, we investigate the higher-order terms of the expansion at $Q=\Sigma_3$ up to two loops: we find that these terms only renormalize the prefactor C_N, without affecting the U- or γ-dependence. We speculate that the (adequately resummed) full series yields the exact answer, including the correct normalization.

To achieve agreement with the Weyl character formula, it was crucial to regard the denominator $\mathrm{Det}(1+Z^\dagger Z)^{-2N}$ as part of the measure (as opposed to lifting it into the action). Indeed, in order for the mechanism of equivariant localization to take effect, the integration measure must be $U(2N)$-invariant—a property not enjoyed by the flat measure $\Pi_{i,j}d^2Z_{ij}$ without the factor $\mathrm{Det}(1+Z^\dagger Z)^{-2N}$.

IV. WHY DO WE NEED AVERAGING?

While the Weyl character formula for $\Omega_U(\gamma)$ constitutes an exact result, it is of no use—at least not as it stands—towards our goal of proving universality of the correlation function. This formula relies on the knowledge of the eigenphases $e^{i\theta_i}$ of U, which are not given *a priori*. It does not exhibit the semiclassical features of the quantum map at all. On the contrary, it is a "purely quantum" decomposition of the correlation function, a complicated reordering of the Fourier decomposition (6).

As was explained in the Introduction, it is not conceivable in general that a universal result for Ω_U can be obtained without doing some kind of averaging over the matrix U. Given the results of the previous section, one might try to perform the averaging term by term in the Weyl decomposition, hoping that most of the terms might average to zero. Such a hope is quickly discouraged by a look at the expression (23): aside from having an Nth-order singularity at $\gamma=1$, whose degree increases each time some $\gamma e^{i(\theta_i-\theta_j)}$ crosses unity, this contribution to $\Omega_U(\gamma)$ is *strictly positive* for real $\gamma<1$. We know that the singularities are artifacts of the Weyl decomposition, as the correlation function $\Omega_U(\gamma)$ itself is uniformly bounded wrt U and γ. Unfortunately, because of the positivity of (23) the singularities can only be removed by reorganizing the entire sum of contributions, not by averaging individual terms.

For this reason, we will adopt a different strategy: we first perform the $d\mathcal{P}_N(U)$ average on the *integrand* of the coherent-state integral, obtaining a new effective action

$$e^{-S_{\rm av}(\gamma,Q)} \overset{\rm def}{=} \langle e^{-S(\gamma,U,Q)} \rangle_{\mathcal{P}_N}. \tag{29}$$

We then estimate the resulting Q-integral by performing a saddle-point approximation on the action $S_{\rm av}(Q)$.[20]

A priori, this approximation is no more justified than the one in the previous section, as $S_{\rm av}$ is preceded by no large parameter either. The absence of a large parameter also implies that averaging and making the saddle-point approximation are noncommuting operations. Therefore, the saddle-point expansion of $S_{\rm av}$ will yield qualitatively different results from the direct expansion for $S(\gamma,U,Q)$. We explained above that averaging the Weyl character formula is hopeless for our aims. The other way around (i.e., performing the expansion after averaging the action) will prove more interesting.

A. Where are the critical points of $S_{\rm av}$?

For any averaging measure $d\mathcal{P}_N$, the two points $Q=\pm\Sigma_3$ remain saddle points of $S_{\rm av}(\gamma,Q)$. In the vicinity of Σ_3, the integrand expands as

$$\left\langle \frac{\mathrm{Det}(1+\gamma Z^\dagger U Z U^{-1})}{\mathrm{Det}(1+Z^\dagger Z)} \right\rangle \approx \exp \mathrm{Tr}(\langle \gamma Z^\dagger \mathrm{Ad}U\cdot Z\rangle - Z^\dagger Z) = e^{-\mathrm{Tr}\, Z^\dagger (\mathbb{I}-\gamma\langle \mathrm{Ad}U\rangle) Z},$$

where $\mathrm{Ad}U\cdot Z \overset{\rm def}{=} UZU^{-1}$ is the adjoint action of U on Z. The approximation is valid for Z small. For larger values of Z, one should add higher cumulants to the right-hand side. However, for the time being we stick to the purely quadratic approximation, and carry out the Gaussian integral to obtain

$$\langle \Omega_U(\gamma)\rangle_{|\Sigma_3} = C_N \gamma^{-N/2} \mathrm{Det}(\mathbb{I}-\gamma\langle \mathrm{Ad}U\rangle)^{-1}. \tag{30}$$

When the averaging is absent (that is, $d\mathcal{P}_N$ is a Dirac δ-measure at U), we recover the contribution (23). The saddle point $Q=-\Sigma_3$ yields the same result, with $\gamma\to\gamma^{-1}$. On setting $\gamma=e^{ix/N}$, the sum of contributions becomes

$$\langle \Omega_U(e^{ix/N})\rangle_{|\Sigma_3\cup-\Sigma_3} = 2C_N \Re\left(\frac{e^{-ix/2}}{\mathrm{Det}(\mathbb{I}-e^{ix/N}\langle \mathrm{Ad}U\rangle)} \right). \tag{31}$$

In the next section, we examine the possible occurrence of further saddle points of $S_{\rm av}$.

1. Searching for other saddle points

In Sec. III A we located the saddle points of the function $Q \mapsto \langle Z(Q)|R(\Gamma\mathbf{U})|Z(Q)\rangle$, using the action of the group $\mathrm{U}(2N)$ on the coherent states $|Z\rangle$. This function may be interpreted as the *Husimi function* (or Q-symbol) of the operator $R(\Gamma\mathbf{U})$ acting on $\mathcal{F}$, and we denote it by $H_{R(\Gamma\mathbf{U})}(Q)$. By the same procedure we can obtain the saddle points of $H_{R(\mathbf{g})}(Q)$ for any nondegenerate matrix $\mathbf{g}\in\mathrm{U}(2N)$; in that case, the saddle points $Q_{\rm crit}$ are given in general by finite matrices $Z_{\rm crit}$ and $Z_{\rm crit}^\dagger$, which are solutions of the saddle-point equations

$$\frac{\partial}{\partial Z_{ij}} H_{R(\mathbf{g})}(Z,Z^\dagger)=0=\frac{\partial}{\partial \overline{Z}_{ij}} H_{R(\mathbf{g})}(Z,Z^\dagger) \qquad (i,j=1,\ldots,N). \tag{32}$$

It is useful to extend $H_{R(\mathbf{g})}$ to a function of two independent complex matrices Z,Z^* (that makes $2N^2$ complex variables). The saddle-point equations pose $2N^2$ constraints on the degrees of freedom Z and Z^*, which yields isolated solutions (Z_i,Z_i^*), provided that the constraints are independent of each other.

The *reality* of these solutions (i.e., $Z_i^*=(Z_i)^\dagger$) is due to a symmetry of the operator $R(\mathbf{g})$, which is not conserved if we replace $R(\mathbf{g})$ by any operator $\mathcal{R}$ on $\mathcal{F}$. For instance, if the representation R is extended to matrices $\mathbf{G}\in\mathrm{GL}(2N,\mathbb{C})$, one can show that the saddle points of $H_{R(\mathbf{G})}(Z,Z^*)$ are real iff $\mathbf{G}$ is a normal matrix (i.e., $\mathbf{G}\mathbf{G}^\dagger=\mathbf{G}^\dagger\mathbf{G}$). We are presently unable to determine the conditions for the saddle points to be real for the most general $\mathcal{R}$. In any case, the saddle points will be real if $\mathcal{R}$ is a Hermitian operator. The Husimi function is then real, and Morse theory applies to it. By Morse's theorem,[28] the number of saddle points (which we assume to be isolated) is at least the sum of all Betti numbers of $\mathcal{M}_N$, which is $\binom{2N}{N}$.[29] This is exactly the number of saddle points we found for $H_{R(\mathbf{X})}(Q)$ when $\mathbf{X}$ is a $2N\times 2N$ Hermitian matrix, so this function is called a *perfect Morse function* for $\mathcal{M}_N$. $\mathbf{X}$ can be joined to $\mathbf{g}\in\mathrm{U}(2N)$ by a continuous path inside the set of nondegenerate normal matrices: this explains why $H_{R(\Gamma\mathbf{U})}$, although a complex function, still has $\binom{2N}{N}$ real saddle points.

Unlike reality, the property that the solutions of (32) are isolated points is robust; (Z_i,Z_i^*) are the common zeros of $2N^2$ polynomials in Z and Z^*, so they are *stable* wrt perturbations of the coefficients, as long as the equations do not become degenerate. In Sec. III A the saddle points of $H_{R(\Gamma\mathbf{U})}(Q)$ were called $Q_{V,S}$. We now switch to such complex coordinates ζ that a saddle point $Q_{V,S}$ is situated at $\zeta=0=\zeta^\dagger$, and perturb $R(\Gamma\mathbf{U})$ in $\mathrm{GL}(\mathcal{F})$ to $\mathcal{R}=R(\Gamma\mathbf{U})+\epsilon\,\delta\mathcal{R}$. Then for ϵ small, $H_{\mathcal{R}}(\zeta,\zeta^\dagger)$ will have an isolated saddle point at $(\zeta_\epsilon,\zeta_\epsilon^*)$, where both ζ_ϵ and ζ_ϵ^* are of order ϵ. Even if it is not real, this saddle point will contribute to the integral over $\mathcal{M}_N$: starting from real coordinates $\Re\zeta_{ij},\Im\zeta_{ij}$, we can locally deform the contour so as to reach the point

$$(\Re\zeta_{ij})^{\rm crit}=(\zeta_{\epsilon,ij}+\zeta_{\epsilon,ji}^*)/2, \qquad (\Im\zeta_{ij})^{\rm crit}=(\zeta_{\epsilon,ij}-\zeta_{\epsilon,ji}^*)/2\mathrm{i},$$

and we can compute the saddle-point expansion of $\int H_{\mathcal{R}}(\Re\zeta,\Im\zeta)$ around it.

The averaged integrands we want to consider are all of the type $H_{\mathcal{R}}(Q)$, where

$$\mathcal{R}=\int_{\mathrm{U}(N)} \mathrm{d}\mathcal{P}_N(V)\ \frac{1}{\mathrm{Det}V} R(\Gamma\mathbf{V}),$$

and $\mathrm{d}\mathcal{P}_N(V)$ is a normalized measure on $\mathrm{U}(N)$. If this measure is very strongly peaked near a matrix U_N, the resulting operator will be a perturbation of $R(\Gamma\mathbf{U}_N)/\mathrm{Det}U_N$, so the above stability arguments apply: the saddle points are then isolated points near the unperturbed ones, and they are "almost real" and hence will lie on the integration contour after a slight contour deformation.

For less concentrated measures $\mathrm{d}\mathcal{P}_N(V)$, the structure of the saddle points can change. In Sec. V we exhibit an averaging scheme for which the saddle points are real but not isolated: they form submanifolds of $\mathcal{M}_N$; this is also the case for $H_{R(\mathbf{g})}$ if $\mathbf{g}$ is degenerate. We do not have a good

estimate of the typical ''width'' of the measure $d\mathcal{P}_N(V)$ above which saddle points can coalesce, spread over higher-dimensional sets, or cease to contribute to the integral (for instance, when they depart too far away from reality).

In general, we are unable to explicitly locate these extra saddle points, even for the relatively narrow averages described in Secs. IV C and IV F; consequently, we cannot do better than stick to the approximation (31) to describe the correlation function. The remaining task then is to investigate the spectrum of the operator $\langle \mathrm{Ad}U\rangle$, which depends on U and on $d\mathcal{P}_N$.

B. Common spectral features of $\langle \mathrm{Ad}U\rangle$

The spectrum of $\langle \mathrm{Ad}U\rangle$ has a few features that are independent of the averaging scheme. Before averaging, the eigenvalue unity occurs in $\mathrm{Ad}U$ with multiplicity N, corresponding to the N-dimensional space spanned by the U-eigenstate projectors $|\psi_j\rangle\langle\psi_j|$ $(j=1,\ldots,N)$, and the remaining N^2-N eigenvalues lie on the unit circle. After averaging, only the uniform mode $\mathbb{I}_N = \Sigma_j |\psi_j\rangle\langle\psi_j|$ is left with eigenvalue at unity, while all other eigenvalues have moved inside the unit disk. As a result, the sum of the contributions (31) stays finite in the limit $\gamma\to 1$. Averaging thus removes the ''unphysical zero mode'' problem associated with the two standard saddle points in Sec. III.

More precisely, the large-N behavior of $\langle \Omega_U(e^{ix/N})\rangle_{|\pm\Sigma_3}$ for finite x mostly depends on the positions of the eigenvalues of $\langle \mathrm{Ad}U\rangle$ *closest to unity*. Within the approximation (31), these eigenvalues are *the relevant dynamical data* of the correlation function.

C. Semiclassical averaging

In Ref. 20, a semiclassical averaging scheme around a quantized map $U_{\phi,N}$ was proposed as a promising candidate to obtain universal spectral statistics, differentiating between integrability versus chaotic behavior of the classical map ϕ. One chooses a finite set of Hamiltonian functions H_j, corresponding to Hamiltonian vector fields Ξ_{H_j} $(j=1,\ldots,r)$, on the classical phase space. These Hamiltonians are quantized on each of the quantum Hilbert spaces $\mathcal{H}_N$, yielding operators $\{\hat{H}_j\}$, which are represented by Hermitian $N\times N$ matrices wrt an orthonormal basis of $\mathcal{H}_N$. An ensemble average is then introduced by

(1) composing $U_{\phi,N}$ with the operator $\exp(-i\Sigma_j t_j \hat{H}_j/\hbar)$, where the ''times'' t_j are real numbers;
(2) averaging over the parameters t_j in a window around the origin of width ϵ using, for instance, the Gaussian weight $(\epsilon^2\pi)^{-r/2} e^{-\Sigma_j t_j^2/\epsilon^2}$.

The width ϵ is taken to be $\hbar$-dependent: $\epsilon\sim\hbar^\alpha\sim N^{-\alpha}$ for some $1>\alpha>0$, so that the probability measure for the classical maps $\exp(\Sigma_j t_j \Xi_{H_j})\circ\phi$ shrinks to a single point, ϕ, in the classical limit $N\to\infty$. The set of Hamiltonians $\{H_j\}$ is chosen once and for all, and is independent of N and the map ϕ. The only constraint on this set is that the second-order differential operator $-\Delta = \Sigma_j \Xi_{H_j}^2$ must be *elliptic*.[20]

As explained in the Introduction, this averaging procedure is introduced in order to suppress the nongeneric spectral statistics of quantum chaotic systems with arithmetic symmetries. In this respect we must mention the results obtained in Ref. 10, where the authors show how nonlinear perturbations of quantum cat maps exhibit generic spectral statistics, as long as one perturbs in both directions of the two-dimensional phase space; in contradistinction, perturbation in a single direction may leave one arithmetic symmetry intact, leading to nongeneric quantum spectral statistics. This need for ''phase-space-isotropy'' of the perturbations is very similar to our ellipticity requirement: Δ is elliptic only if the vector fields Ξ_{H_j} span the whole tangent space at every point of phase space.

Some recent articles[30,31] have dealt with the spectral analysis of the operator $\langle \mathrm{Ad}U\rangle_{\mathrm{semiclas}}$, and obtained interesting results concerning its largest eigenvalues. For a classically chaotic map, these were shown to converge (as $N\to\infty$) to the *Ruelle–Pollicott resonances* of the corresponding

Frobenius-Perron operator.[32] These resonances are inside the unit circle, which means that $\langle \mathrm{Ad}U_N\rangle_{\mathrm{semiclas}}$ has a *finite gap* between unity and the rest of the spectrum, for $N\to\infty$. The huge majority of eigenvalues tend to accumulate on the origin (see the discussion at the end of Sec. V A).

These properties allow us to estimate the contribution from the two standard saddle points for the case of a quantum chaotic map. To lowest order in $1/N$,

$$\langle \Omega_{U_N}(\gamma=\mathrm{e}^{\mathrm{i}x/N})\rangle_{\mathrm{semiclas}|\pm\Sigma_3} \overset{N\to\infty}{\approx} \frac{NC_N}{\mathrm{Det}_\perp(\mathbb{I}-\langle \mathrm{Ad}U\rangle_{\mathrm{semiclas}})}\,\frac{\sin(x/2)}{x/2}, \tag{33}$$

where $\mathrm{Det}_\perp$ means that the determinant is computed after restriction to the traceless matrices, i.e., to the subspace orthogonal to the uniform mode $\mathbb{I}_N$. Apart from the non-universal prefactor, the x-dependence agrees with the CUE result (8) in the limit of large matrices.

In the case of an integrable map, the eigenvalues of $\langle \mathrm{Ad}U_N\rangle_{\mathrm{semiclas}}$ behave differently: some of them populate more and more densely a few curves which connect the origin to some point on the unit circle (including unity). For this reason, one cannot separate unity in $\mathrm{Det}(\mathbb{I}-\mathrm{e}^{\mathrm{i}x/N}\langle \mathrm{Ad}U\rangle_{\mathrm{semiclas}})$ from the rest of the spectrum. All we can say is that the approximation (31) does not yield the CUE formula in that case (in general it does not yield the Poisson answer either).

1. Warning

One might be tempted to present formula (33) as a ''physicist's proof'' of a weak universality conjecture for quantum chaotic maps. The reason why it is not a proof is clear:

(i) As was explained in Sec. IV A 1, there certainly exist other saddle points of (29). The calculation of their contributions is a difficult task, which we have not yet performed. It is far from obvious why these saddle points should be less important than $Q=\pm\Sigma_3$ in the semiclassical averaging scheme.

(ii) As was emphasized before, there is no large parameter in front of the effective action. Without such a parameter, the correction terms of the asymptotic expansion around each saddle point are not small, and their neglect in the formula (33) seems to be unjustified.

The second worry is addressed in the next subsection.

D. Loop expansion

We are now going to investigate those corrections to the formula (33) that result from systematically expanding around the saddle point $Q=\Sigma_3$. The computations will be done up to what is called two-loop order in field-theoretic language.

As a first step, we approximate the integrand by taking the ensemble average inside the determinant:

$$\langle \mathrm{Det}(1+\gamma Z^\dagger \mathrm{Ad}U\cdot Z)\rangle_{\mathrm{semiclas}} \approx \mathrm{Det}(1+\gamma Z^\dagger \langle \mathrm{Ad}U\rangle_{\mathrm{semiclas}}\cdot Z). \tag{34}$$

Although ϵ, the ''width'' of the perturbation, decreases like $\hbar^\alpha$, its effect is strong enough to completely modify the spectrum of $\mathrm{Ad}U$, even in the semiclassical limit. This shows that the above approximation is not necessarily valid if we just suppose that the matrices Z are bounded (uniformly wrt N) in the operator norm on $\mathcal{H}_N$. Using the expansion

$$\mathrm{Det}(1+A)=1+\mathrm{Tr}A+\sum_{j=2}^{N}\mathrm{Tr}(\wedge^j A),$$

Eq. (34) will hold as long as the terms for $j \geqslant 2$ are small compared to $\mathrm{Tr}A$. A sufficient condition for that is $\mathrm{Tr}(|A|) \ll 1$, where $|A| \overset{\text{def}}{=} \sqrt{A^\dagger A}$. Upon the replacement $A = \gamma Z^\dagger \mathrm{Ad}U \cdot Z$, this condition will be met if

$$\mathrm{Tr}(Z^\dagger Z) = \sum_{i,j=1}^{N} |Z_{ij}|^2 \ll 1, \tag{35}$$

uniformly wrt N. It would be desirable to better control the error in (34) for the larger set of matrices Z satisfying (N-uniformly) $\|Z\|_{\mathcal{L}(\mathcal{H}_N)} \leqslant \mathrm{const}$.

Taking (34) for granted, we proceed to the computation of higher loops. To simplify the notation we abbreviate $T \overset{\text{def}}{=} \gamma \langle \mathrm{Ad}U \rangle_{\mathrm{semiclas}}$. Next we formally introduce a parameter M (which will be reset to unity at the end of the calculation) by making in the integrand the replacement

$$\frac{\mathrm{Det}(1+Z^\dagger T Z)}{\mathrm{Det}(1+Z^\dagger Z)} \to \left(\frac{\mathrm{Det}(1+Z^\dagger T Z)}{\mathrm{Det}(1+Z^\dagger Z)} \right)^M .$$

A contribution to the perturbative saddle-point expansion is said to be of n-loop order if it varies as M^{-n} relative to the leading-order term. On rescaling the integration variables to $\zeta = Z\sqrt{M}$ and $\zeta^\dagger = Z^\dagger \sqrt{M}$, the $1/M$ expansion of the integrand looks as follows:

$$d\mu_N(Z,Z^\dagger) \frac{\mathrm{Det}^M(1+Z^\dagger T Z)}{\mathrm{Det}^M(1+Z^\dagger Z)} = C_N \prod_{i,j=1}^{N} \frac{d^2\zeta_{ij}}{\pi M} \mathrm{e}^{-\mathrm{Tr}\zeta^\dagger(1-T)\zeta} (1 + M^{-1} f_1 + M^{-2} f_2 + \cdots)$$

where f_1 and f_2 are the one- and two-loop terms, respectively, and are given by

$$f_1 = \frac{1}{2}\mathrm{Tr}(\zeta^\dagger \zeta)^2 - \frac{1}{2}\mathrm{Tr}(\zeta^\dagger T \zeta)^2 - 2N \mathrm{Tr}\zeta^\dagger \zeta,$$

$$\begin{aligned} f_2 = &-\frac{1}{3}\mathrm{Tr}(\zeta^\dagger \zeta)^3 + \frac{1}{3}\mathrm{Tr}(\zeta^\dagger T \zeta)^3 + \frac{1}{8}[\mathrm{Tr}(\zeta^\dagger \zeta)^2 - \mathrm{Tr}(\zeta^\dagger T \zeta)^2]^2 \\ &+ 2N^2(\mathrm{Tr}\zeta^\dagger \zeta)^2 + N\mathrm{Tr}(\zeta^\dagger \zeta)^2 - N\mathrm{Tr}(\zeta^\dagger \zeta)(\mathrm{Tr}(\zeta^\dagger \zeta)^2 - \mathrm{Tr}(\zeta^\dagger T \zeta)^2). \end{aligned}$$

The Gaussian integral at leading order just yields the result (30). Using standard diagrammatic techniques to do the one-loop integral we find the following expression:

$$\begin{aligned} \frac{1}{2} C_N M^{-N\times N} \mathrm{Det}(1-T)^{-1} \Bigg\{ &-4N \sum_{ij} \left(\frac{1}{1-T} \right)_{ij,ij} + \sum_{ijkl} \left(\frac{1}{1-T} \right)_{ij,kj} \left(\frac{1}{1-T} \right)_{kl,il} \\ &- \sum_{ijkl} \left(\frac{T}{1-T} \right)_{ij,kj} \left(\frac{T}{1-T} \right)_{kl,il} + \sum_{ijkl} \left(\frac{1}{1-T} \right)_{ij,il} \left(\frac{1}{1-T} \right)_{kl,kj} - \sum_{ijkl} \left(\frac{T}{1-T} \right)_{ij,il} \left(\frac{T}{1-T} \right)_{kl,kj} \Bigg\}. \end{aligned}$$

By the relation $(1-T)^{-1} = 1 + T(1-T)^{-1}$ these terms combine to yield the simple answer

$$C_N \int_{\mathbb{C}^{N\times N}} \prod_{i,j=1}^{N} \frac{d^2\zeta_{ij}}{\pi M} \mathrm{e}^{-\mathrm{Tr}\zeta^\dagger(1-T)\zeta} f_1(\zeta,\zeta^\dagger) = C_N M^{-N\times N} \mathrm{Det}(1-T)^{-1}(-N^3).$$

We see that the dependence of the one-loop contribution on T cancels completely, leaving only a constant, $-N^3$. This cancellation is not accidental but continues to higher loop order. By a lengthy but straightforward calculation, the complete perturbative result up to two-loop order can be shown to be

$$\int_{\mathbb{C}^{N\times N}} d\mu_N(Z,Z^\dagger)\,\frac{\mathrm{Det}^M(1+Z^\dagger TZ)}{\mathrm{Det}^M(1+Z^\dagger Z)}$$
$$= C_N M^{-N\times N}\mathrm{Det}(1-T)^{-1}\left(1-M^{-1}N^3+M^{-2}\left(\frac{1}{2}N^6+\frac{7}{12}N^4-\frac{1}{12}N^2\right)+\mathcal{O}(M^{-3})\right).$$

Again, all the T-dependence has disappeared from the loop correction terms. This is true for all M including the case of interest, $M=1$.

The cancellation does not come as a total surprise. The above perturbation expansion, whose low-order terms we have computed, is formally identical to the same expansion *before* averaging. The latter is obtained from the former by simply substituting $\gamma\mathrm{Ad}U$ for $T=\gamma\langle\mathrm{Ad}U\rangle_{\mathrm{semiclas}}$. In the case before averaging we know from Ref. 27 that an index-theoretic mechanism (sometime called *localization*) causes the perturbation expansion to be deformable (by an underlying $N=2$ supersymmetry) to a harmonic oscillator problem (or, equivalently, a Gaussian integral) at $Z=0$. The process of deformation to the Gaussian limit explains why the dependence on $\gamma\mathrm{Ad}U$ is exhausted by the leading-order term. It leads to the Weyl character formula, which implies that the contribution to the character from $Z=0$ (or $Q=\Sigma_3$) is *exactly* given by

$$\int_{\mathbb{C}^{N\times N}} d\mu_N(Z,Z^\dagger)\,\frac{\mathrm{Det}(1+Z^\dagger\gamma\mathrm{Ad}U\cdot Z)}{\mathrm{Det}(1+Z^\dagger Z)}\Big|_{Z=0,\text{all orders}}=\mathrm{Det}(\mathbb{I}-\gamma\mathrm{Ad}U)^{-1},$$

where the normalization constant C_N has now been replaced by unity. The last fact provides the raison d'ĉtre for the N-dependent terms produced by the loop expansion: their role is to cancel, after proper resummation, the prefactor C_N. This property does not depend on the unitarity of $\gamma\mathrm{Ad}U$, so it holds as well after replacing it by its average. Thus, after summing all orders of the perturbation expansion, we expect that the saddle point $Z=0$ contributes to the correlation function as

$$\langle\Omega_U(\gamma)\rangle_{|\Sigma_3,\text{all orders}}=\gamma^{-N/2}\,\mathrm{Det}(\mathbb{I}-\gamma\langle\mathrm{Ad}U\rangle)^{-1}.$$

This perturbative result should be used with some care. Although the function $f(Z,Z^\dagger;T)=\mathrm{Det}(1+Z^\dagger TZ)/\mathrm{Det}(1+Z^\dagger Z)$ is locally well defined, it does not extend to a global smooth function on the manifold $\mathcal{M}_N$ (in particular, this function is NOT the Husimi function of an operator on $\mathcal{F}$). Indeed, setting $Z=zG$ with any invertible matrix G and sending $z\to\infty$ always leads to the same point $Q=-\Sigma_3$ on $\mathcal{M}_N$, regardless of which matrix G we choose, whereas the limit of $f(zG,\overline{z}\overline{G}^\dagger;T)$ as $z\to\infty$ does depend on the choice of G. Thus, the function $f(Z,Z^\dagger;T)$ is not smooth at $Q=-\Sigma_3$.

This singularity reflects the fact that the cumulants neglected by our basic approximation (34) are small (compared to the terms kept) only for small matrices Z [cf. the discussion following Eq. (34)]. If $Z,Z^\dagger$ (or some matrix elements thereof) are allowed to go to infinity, the approximation clearly loses its validity. To control the error incurred near the saddle point $Q=-\Sigma_3$, one needs to switch to another scheme, by first changing coordinates $Z\to 1/Z$ and $Z^\dagger\to 1/Z^\dagger$ and only afterwards repeating the above steps. The contribution from $Q=-\Sigma_3$ can then be calculated in the same way as the one for $Q=\Sigma_3$. The treatment of further saddle points remains an open problem.

What makes this procedure unsatisfactory is that we are simultaneously working with several approximation schemes, each of which is only locally controlled. To localize the integral at the saddle points in a mathematically rigorous manner, we would need an approximation that is *globally well defined* and well controlled. It is not clear whether such an approximation exists, given the stringent requirement that the integrand should also have the index-theoretic features that allow localization techniques to be used.

V. AVERAGING *U* OVER EIGENBASES

By its definition (2) as a correlation function of spectral determinants, $\Omega_U(\gamma)$ is invariant under any change of basis $U \mapsto VUV^{-1}$, with V an arbitary unitary matrix. In the Q-matrix formulation, this invariance is reflected by the relation $S(\gamma,U,Q)=S(\gamma,VUV^{-1},\mathbf{V}Q\mathbf{V}^{-1})$. Since the transformation $Q \mapsto \mathbf{V}Q\mathbf{V}^{-1}$ has unit Jacobian, we may absorb $\mathbf{V}$ into the integration variable Q and compute $\Omega_U(\gamma)$ by first averaging $\mathrm{e}^{-S(\gamma,U,Q)}$ over all rotations $U \mapsto VUV^{-1}$:

$$\mathrm{e}^{-S_{V\mathrm{av}}(\gamma,U,Q)} \overset{\text{def}}{=} \frac{1}{\mathrm{Vol\ U}(N)} \int_{\mathrm{U}(N)} dV \exp\{-S(\gamma,VUV^{-1},Q)\}, \tag{36}$$

and then integrating $\mathrm{e}^{-S_{V\mathrm{av}}}$ over Q. We saw in Sec. III A that if the matrix $D=V^{-1}UV$ is diagonal, then the saddle points of $S(\gamma,U,Q)$ are situated on the points $Q_{V,S}=\mathbf{V}Q_S\mathbf{V}^{-1}$. Because the locations of these points explicitly *depend* on V, we expect that a smoothing mechanism takes place and the divergences of the individual terms in the Weyl character formula disappear on averaging over V. In fact, as we will see, the expansion obtained by saddle-point analysis of the effective action $S_{V\mathrm{av}}(\gamma,U,Q)$ is qualitatively quite different from Weyl's formula.

A. Analysis around $\pm\Sigma_3$

We first describe $S_{V\mathrm{av}}(\gamma,U,Q)$ near the two saddle points $Q=\pm\Sigma_3$ (cf. Sec. IV A). The V-averaged adjoint operator $\langle \mathrm{Ad}\rangle_V$ has a rather simple spectrum: unity is a simple eigenvalue (associated with $\mathbb{I}_N$), and on the remaining (N^2-1)-dimensional space the operator is proportional to the identity:

$$\langle \mathrm{Ad}U\rangle_V = P_{\mathbb{I}} + (1-P_{\mathbb{I}})\frac{|\mathrm{Tr}U|^2-1}{N^2-1}. \tag{37}$$

($P_{\mathbb{I}}$ is the orthogonal projector on $\mathbb{I}_N$.) We see that $\langle \mathrm{Ad}U\rangle_V$ has a large gap between unity and the second eigenvalue, and this gap has the maximal degeneracy. Assuming that this degenerate eigenvalue is small ($|\mathrm{Tr}U| \ll N$), we get the following leading-order contribution:

$$\langle \Omega_U(\mathrm{e}^{\mathrm{i}x/N})\rangle_{V|\Sigma_3 \cup -\Sigma_3} \sim \frac{2NC_N}{(1-\alpha/N)^{N^2}} \frac{\sin\{x(1/2-\alpha)\}}{x}, \tag{38}$$

$$\text{with} \qquad \alpha \overset{\text{def}}{=} (|\mathrm{Tr}U|^2-1)/N. \tag{39}$$

Within this approximation, the correlation function depends on $U=U_{\phi,N}$ only through the simple quantity $|\mathrm{Tr}U|^2$, which can be estimated semiclassically by the Gutzwiller–Tabor trace formula (1): typically, α is of order $\mathcal{O}(1/N)$ for a chaotic map, and of order $\mathcal{O}(1)$ for an integrable one.

Notice that, due to the high degeneracy of the second eigenvalue, we do not get in general the CUE result (8), although this eigenvalue is far inside the unit circle. This shows that, to obtain the CUE result (33), we not only need a finite gap in the spectrum of $\langle \mathrm{Ad}U\rangle$, but also a fast accumulation of the eigenvalues to the origin. The precise condition on the eigenvalues is $\Sigma_{j=2}^{N^2}[\lambda_j/(1-\lambda_j)] \ll N$. In the present averaging scheme, this means $\alpha \ll 1$.

B. Critical submanifolds

We need to investigate the possible influence of other saddle points of $S_{V\mathrm{av}}(\gamma,U,Q)$; for the present averaging scheme, we will explicitly describe a critical set, which we believe to be exhaustive. The effective action possesses the symmetry $S_{V\mathrm{av}}(Q)=S_{V\mathrm{av}}(\mathbf{W}Q\mathbf{W}^{-1})$ for all $W \in \mathrm{U}(N)$. Therefore, the saddle points are grouped into stationary *submanifolds*, each of them invariant under $\mathrm{U}(N)$.

1. Description of the manifolds

In Appendix A we prove the following statement: for any initial matrix U and any γ, the action (36) is stationary at the points $Q=\mathbf{W}Q_\sigma\mathbf{W}^{-1}$, for all rotations $W\in \mathrm{U}(N)$ and any permutation σ (see Sec. III A for the notations σ, $S=S_1\cup S_2$, etc). Since $\mathrm{U}(N)$ is connected, the points $\mathcal{M}_S \overset{\text{def}}{=} \{\mathbf{W}Q_S\mathbf{W}^{-1} \,|\, W\in \mathrm{U}(N)\}$ form a connected submanifold of $\mathcal{M}_N$.

Let τ be any permutation among N indices. The set $S'=(\tau(S_1),\tau(S_2))$ is in general different from S, and we have $Q_S\neq Q_{S'}$; however, $Q_{S'}\in\mathcal{M}_S$, or equivalently $\mathcal{M}_S=\mathcal{M}_{S'}$. Putting $p=\sharp S_2$ and $r=\sharp(S_1\cap S_2)$, we find that $\mathcal{M}_S$ contains $\binom{N}{p}\binom{p}{r}\binom{N-p}{r}$ different points $Q_{S'}$. The manifolds $\mathcal{M}_S$ are in one-to-one correspondence with the integers (p,r), and their total number is $(N/2+1)^2$ for N even, and $(N+1)(N+3)/4$ for N odd (including the isolated points $\pm\Sigma_3$ in the count).

For generic U and $\gamma\neq 1$ [genericity means here that the matrix $\mathrm{diag}(\gamma U,U)$ is not degenerate], we conjecture that the submanifolds $\mathcal{M}_{(p,r)}$ exhaust all the critical points of the action $S_{V\mathrm{av}}(\gamma,U,Q)$.

2. Contributions of the manifolds

The leading-order contribution of each submanifold $\mathcal{M}_S$ to the Q-integral is calculated by separating the tangent space at Q_S into two parts, one parallel and one transverse to $\mathcal{M}_S$. The integrand in the vicinity of Q_S then reads (to quadratic order)

$$\mathrm{e}^{-S_{V\mathrm{av}}(Q_S)}\mathrm{e}^{-\mathrm{Hess}_T(X_T)+\mathcal{O}(|X_T|^3)},$$

where Hess_T is the Hessian of $S_{V\mathrm{av}}$ around $\mathcal{M}_S$, viewed as a nondegenerate quadratic form on the transverse part of tangent space (coordinatized by X_T). The exact integral over $\mathcal{M}_S$ and the Gaussian integral over the transverse directions yield the contribution

$$\langle\Omega_U(\gamma)\rangle_{|\mathcal{M}_S}=C_N\gamma^{-N/2+p}\frac{\mathrm{Vol}\mathcal{M}_S}{\sqrt{\mathrm{Det}(\mathrm{Hess}_T)}}\mathrm{e}^{-S_{V\mathrm{av}}(\gamma,Q_S)}. \tag{40}$$

In Appendix B, we explicitly compute the volumes of the submanifolds $\mathcal{M}_S=\mathcal{M}_{(p,r)}$:

$$\mathrm{Vol}\mathcal{M}_{(p,r)}=\frac{(\Gamma(1)\cdots\Gamma(r))^2\Gamma(1)\cdots\Gamma(p-r)\Gamma(1)\cdots\Gamma(N-p-r)}{\Gamma(1)\cdots\Gamma(N)}.$$

For all submanifolds $\mathcal{M}_S\neq\{\pm\Sigma_3\}$ (i.e., $0<p<N$), these volumes are N-exponentially small. The quantities Hess_T and $S_{V\mathrm{av}}(Q_S)$ depend on U and γ; we are unable to compute them in general. What we know for sure is that $|\mathrm{e}^{-S_{V\mathrm{av}}}|\leqslant 1$, since $\mathrm{e}^{-S(\gamma,U,Q)}$ has this property.

For a nondegenerate U and $\gamma=\mathrm{e}^{\mathrm{i}x/N}$, the Hessian around $\mathcal{M}_S$ will possess a single eigenvalue that vanishes with x, while all other eigenvalues stay at least of order $\mathcal{O}(1)$. This means that the contribution from $\mathcal{M}_S$ goes like $1/x$ as $x\to 0$. However, the "particle-hole duality" between the submanifolds $\mathcal{M}_{(p,r)}$ and $\mathcal{M}_{(N-p,r)}$ cancels this divergence in the sum of their two contributions (as it does for $\pm\Sigma_3$).

As a result, we conjecture that each contribution $\langle\Omega_U(\gamma)\rangle_{|\mathcal{M}_S\cup\mathcal{M}_S^-}$ is x-uniformly, N-exponentially small compared to that from $\pm\Sigma_3$ for large N, owing to the small volumes of $\mathcal{M}_{(p,r)}$. Since the number of critical submanifolds grows like N^2, we deduce that the leading-order saddle-point expansion for the action $S_{V\mathrm{av}}(U,\gamma,Q)$ can be truncated to (38) for large N.

C. Averaging over random matrix ensembles

We may go further and average e^{-S} not only over the conjugates of a fixed matrix U, but also over the spectrum $\{\mathrm{e}^{\mathrm{i}\theta_j}\}$. For instance, we can average U over all matrices in $\mathrm{U}(N)$, with a weight corresponding to one of the standard random matrix ensembles (Poisson, CUE). The averaged

action will be U(N)-rotation invariant, and its saddle points will still lie on the submanifolds $\mathcal{M}_S$. As a result, the leading-order saddle-point (l.o.s.p.) expansion for such ensemble-averaged actions can again be truncated to the contribution (38), upon replacing the coefficient α by its average $\langle\alpha\rangle_{\text{ensemble}}$ over the ensemble considered.

D. Conclusion: No l.o.s.p. expansion for the *V*-averaged actions

The contribution (38) depends in a very simple manner on the matrix U, namely only on its first trace. This is in contradiction with the fact that *a priori*, all traces up to $\mathrm{Tr}(U^{N/2})$ enter into $\Omega_U(\gamma)$ [cf. Eq. (14)]. By selecting some particular cases, it becomes obvious that the l.o.s.p. expansion (38) deviates strongly from the exact correlation function. The most immediate counterexample is the Poisson ensemble, whose correlation function is given in Eq. (8). For this ensemble, $\langle\alpha\rangle_{\text{Poisson}}=1$, which yields the CUE result (!) when inserted into the formula (38). We are hence forced to abandon the l.o.s.p. expansion for the V-averaged actions.

Nevertheless, we hope that this expansion is still meaningful when the averaging over U is *local* in U(N), which is the case for the semiclassical average in Sec. IV C (but not for the V-average). Hopefully, a local average will still conserve some memory of the ''localization'' property, which entailed the ''almost exactness'' of the l.o.s.p. expansion for $S(\gamma,U,Q)$.

In the next section, we will consider a local averaging scheme different from the semiclassical one. It possesses group-theoretic properties, which will allow us to analyze it from the character decomposition (14) instead of the coherent-state integral.

VI. ISOTROPIC AVERAGING

Starting from a fixed matrix U, one may define an *isotropic* averaging around U, by composing U with the $N\times N$ unitary matrices $\mathrm{e}^{-\mathrm{i}H}$, weighted by $\exp(-\mathrm{Tr}H^2/4\epsilon)dH$ with small ϵ (so that the weight is concentrated at the identity). Isotropy here means that the measure dH is U(N)-invariant. Note that this in sharp contrast with the semiclassical averaging of Sec. IV C, where H was a linear combination of r matrices $\hat{H}_j$, with r independent of N. In the semiclassical case, the perturbation spanned only a r-dimensional submanifold, whereas in the present case the perturbation completely fills the N^2-dimensional ϵ-ball centered at $H=0$.

One can replace the Gaussian weight by any positive normalized U(N)-invariant function of H. For our purposes, it is convenient to use the *heat kernel* on U(N), i.e., the kernel of the regularizing operator $\exp(-\epsilon\Delta)$, where Δ is the (positive) Laplace–Beltrami operator on U(N). The heat kernel centered on U is defined as follows:

$$\forall\,\epsilon>0:\ -\Delta_V K_\epsilon(V,U)=\frac{\partial}{\partial\epsilon}K_\epsilon(V,U)$$

$$\lim_{\epsilon\to+0} K_\epsilon(V,U)=\delta_U(V).$$

Owing to the compactness of U(N), the density $K_\epsilon(\cdot,U)$ for any matrix U converges to the uniform density on U(N) as $\epsilon\to\infty$. Switching ϵ from 0 to ∞ therefore realizes a crossover from the Dirac delta measure $\delta_U(\cdot)$ to the Haar (or CUE) measure. For small values of ϵ, the kernel $K_\epsilon(V,U)=k_\epsilon(VU^{-1})$ is concentrated around $\mathrm{e}^{-\mathrm{i}H}=VU^{-1}\approx 1$ and is approximately given by the Gaussian weight introduced above: $k_\epsilon(\mathrm{e}^{-\mathrm{i}H})\sim\exp(-\mathrm{Tr}H^2/4\epsilon)$.

Schur's lemma ensures that Δ is proportional to the identity on each U(N)-irreducible subspace of $L^2(\mathrm{U}(N))$. As a consequence, its action on each representation matrix $\rho_p(U)$ of Eq. (14) is simply a multiplication by a positive factor, called the quadratic Casimir invariant, which we denote by $\rho_p(\Delta)$. In formulas,

$$\int_{\mathrm{U}(N)} dV\rho_p(V)K_\epsilon(V,U)\overset{\text{def}}{=}\mathrm{e}^{-\epsilon\Delta_U}\rho_p(U)=\mathrm{e}^{-\epsilon\rho_p(\Delta)}\rho_p(U).$$

The factor $\rho_p(\Delta)$ may be computed from the Young diagram of ρ_p; a more direct way is to express Δ in terms of fermionic operators acting on the Fock space $\mathcal{F}$:

$$\Delta|_{\mathcal{F}}=\sum_{i,j=1}^{N}(f_{+i}^{\dagger}f_{+j}-f_{-j}^{\dagger}f_{-i})(f_{+j}^{\dagger}f_{+i}-f_{-i}^{\dagger}f_{-j})=(N+1)(F_{+}+F_{-})-(F_{+}^{2}+F_{-}^{2})-2J_{\uparrow}J_{\downarrow}.$$

Applying this to any element of the subspace ${}^{0}\mathcal{F}^{p}$ (which carries ρ_p) we find

$$\rho_p(\Delta)=2p(N+1-p). \tag{41}$$

On employing the decomposition (14), the heat-kernel averaged correlation function for $\gamma=\mathrm{e}^{\mathrm{i}x/N}$ takes the form

$$\langle\Omega_U(\gamma)\rangle_{\epsilon}\overset{\text{def}}{=}\mathrm{e}^{-\epsilon\Delta_U}\Omega_U(\gamma)=\sum_{p=0}^{N/2}\mathrm{e}^{-2\epsilon p(N+1-p)}\mathrm{Tr}\rho_p(U)\,\frac{\sin[x/2\,(1-(2p-1)/N)]}{\sin(x/2N)}. \tag{42}$$

The effect of the averaging procedure is to damp the large-p traces, which are difficult to estimate from the Gutzwiller trace formula. In the above equation the $\epsilon\to\infty$ behavior is obvious: all traces except the trivial one $\mathrm{Tr}\rho_0(U)=1$ are killed by the exponential, no matter what the matrix U is. It is actually not necessary to set ϵ to ∞ to get the CUE correlation. Since the irreps ρ_p are unitary, their traces are bounded by

$$|\mathrm{Tr}\rho_p(U)|\leqslant\dim\rho_p=\mathrm{Tr}\rho_p(\mathbb{I}).$$

The dimensions of the ρ_p's are given in Eq. (13); for finite p, they are bounded by $\dim\rho_p\leqslant N^{2p}$. In the limit $N,p\to\infty$ with $y=p/N$ fixed, Stirling's formula yields

$$\dim(\rho_{p=Ny})\sim(\pi N)^{-1}\frac{f'(y)}{y(1-y)}\mathrm{e}^{2Nf(y)},$$

where the function $f(y)=-y\log y-(1-y)\log(1-y)$ increases monotonically from $f(0)=0$ to $f(1/2)=\log 2$.

For any sequence $\{U_N\}_{N\in\mathbb{N}}$, if we tune ϵ (possibly varying with N) such that

$$\varepsilon\overset{\text{def}}{=}N\epsilon\gg 1,$$

all the terms making a significant contribution to (42) satisfy $p\ll N$. The x-dependence of all these terms is the same [being given by the CUE correlation $x^{-1}\sin(x/2)$], so the averaged correlation will also have this dependence. Only the prefactor will depend on the matrices U_N explicitly. If ϵ is increased further to $\varepsilon\gg\log N$, the prefactor itself becomes universal.

These statements hold even in the most general case, when the sequence $\{U_N\}$ is completely arbitrary. Therefore, to be able to differentiate between integrable and chaotic quantum maps, one must tune the ''disorder strength'' ε to smaller values, so that contributions from the ''high'' traces $\mathrm{Tr}\rho_p(U_N)$ start to contribute. To recover the Poisson behavior for integrable maps, one actually needs contributions to (42) coming from the whole region $p\lesssim N/2$.

This puts us in a no-win situation. On the one hand, we should tune ε to small enough values so that the high traces $p\sim Ny$ $(y>0)$ survive and Poisson behavior stands a chance to emerge. On the other hand, for a chaotic map we have no control over these high traces (we don't for an integrable map either).

For our purposes, the present averaging scheme is probably ''too algebraic,'' as opposed to the semiclassical average presented in Sec. IV C. To motivate this statement in the spirit of Sec. IV C, let us compare the spectra of the operators $\langle\mathrm{Ad}U\rangle$ for the two schemes:

(i) The spectrum of $\langle \mathrm{Ad}U\rangle_{\mathrm{semiclas}}$ *qualitatively* depends on the nature of the classical dynamics (see Sec. IV C). It has a finite gap for a chaotic map, whereas eigenvalues accumulate near the unit circle for an integrable one.

(ii) In the isotropic scheme, $\langle \mathrm{Ad}U\rangle_{\epsilon}$ is decomposed into the irreps $\rho_0(U)\oplus\rho_1(U)$. Therefore, apart from the single eigenvalue unity, $\langle \mathrm{Ad}U\rangle_{\varepsilon}$ has the eigenvalues $\{\mathrm{e}^{-2\varepsilon}\mathrm{e}^{\mathrm{i}(\theta_i-\theta_j)}\}$, where $\{\mathrm{e}^{\mathrm{i}\theta_j}\}$ are the eigenvalues of U; the eigenvalue $\mathrm{e}^{-2\varepsilon}$ is $(N-1)$-fold degenerate. This spectrum is qualitatively the same for chaotic versus integrable systems.

A. Crossover Poisson-CUE

We now present an application of the above scheme in the area of random matrices. More precisely, we use the isotropic averaging to build a crossover between the Poisson and CUE ensembles, and we derive the transitional determinant correlation function that interpolates between the formulas (8). This crossover, as well as the method used to compute $\langle\Omega_U\rangle$, can be compared to the GOE→GUE crossover studied in Ref. 7.

Our crossover is defined as follows. We start from the Poisson ensemble, then convolute it with the isotropic (heat kernel) measure of width ϵ:

$$\langle\Omega_U(\gamma)\rangle_{\mathrm{Poisson},\epsilon}\overset{\mathrm{def}}{=}\int_{\mathrm{U}(N)}\mathrm{d}\mathcal{P}_{\mathrm{Poisson}}(U)\int_{\mathrm{U}(N)}dV\,K_{\epsilon}(V,U)\Omega_V(\gamma).$$

For $\epsilon=0$, this is the Poisson ensemble. In the large-ϵ limit, the second integral converges U-uniformly to the CUE correlation function, so the output $\langle\Omega_U(\gamma)\rangle_{\mathrm{Poisson},\epsilon}$ does, too.

To calculate the correlation function along the crossover, we will use the decomposition (42) as in the previous section: averaging being a linear operation, we only need to replace the characters $\mathrm{Tr}\rho_p(U)$ by their Poisson averages [see Eqs. (7) and (15)]:

$$\langle\mathrm{Tr}\rho_p(U)\rangle_{\mathrm{Poisson}}=\binom{N}{p}-\binom{N}{p-1}.$$

The asymptotics of these traces in the regime $p,N\to\infty$ with $y=p/N$ fixed, again follows easily from Stirling's formula:

$$\langle\mathrm{Tr}\rho_p(U)\rangle_{\mathrm{Poisson}}\sim(2\pi N)^{-1/2}\frac{f'(y)}{\sqrt{y(1-y)}}\mathrm{e}^{Nf(y)}. \tag{43}$$

The sum over characters therefore approaches the following integral (as $N\to\infty$):

$$\langle\Omega_U(\mathrm{e}^{\mathrm{i}x/N})\rangle_{\mathrm{Poisson},\epsilon}\sim\frac{2N^2}{\sqrt{2\pi N}}\int_0^{1/2}dy\,\frac{f'(y)}{\sqrt{y(1-y)}}\,\frac{\sin(\frac{1}{2}x-yx)}{x}\,\mathrm{e}^{N(f(y)-2\varepsilon y(1-y))}. \tag{44}$$

In the limit $N\to\infty$, this integral is determined by the saddle points (rather, the maximum) of $f_\varepsilon(y)\overset{\mathrm{def}}{=}f(y)-2\varepsilon y(1-y)$ on $[0,\frac{1}{2}]$. Three cases have to be distinguished:

(i)If $\varepsilon<1$, the boundary point $y=\frac{1}{2}$ is a maximum of f_ε and is the only critical point on $[0,\frac{1}{2}]$. Because of the vanishing of the integrand at $y=\frac{1}{2}$, the saddle-point analysis requires some care. For x of order $\mathcal{O}(N^0)$ the result turns out to be independent of x:

$$\langle\Omega_U(\mathrm{e}^{\mathrm{i}x/N})\rangle_{\mathrm{Poisson},\varepsilon}\sim 2^N\mathrm{e}^{-N\varepsilon/2}(1-\varepsilon)^{-3/2},$$

which shows that the Poisson result 2^N is retrieved in the limit $\varepsilon\to0$. The correlation functions starts depending on x on scales of order $x\sim\mathcal{O}(N^{1/2})$.

(ii) If $\varepsilon>1$, the maximum of f_ε is situated at the point $y_\varepsilon \in (0,\frac{1}{2})$ which solves the transcendental equation $f'_\varepsilon(y)=0$. The correlation function depends on $x\sim\mathcal{O}(N^0)$ as

$$\langle \Omega_U(e^{ix/N})\rangle_{\text{Poisson},\varepsilon} \propto \frac{\sin[x(\frac{1}{2}-y_\varepsilon)]}{x}. \tag{45}$$

The flat correlation function has been replaced by an oscillatory function, with the period of oscillation being controlled by the ''frequency shift'' y_ε. When ε becomes large, the shift vanishes as $y_\varepsilon\sim e^{-2\varepsilon}$, so the CUE correlation function is retrieved.

(iii) If $\varepsilon=1$, the correlation function is ''critical'' (in the sense of a phase transition), as the two points $y=\frac{1}{2}$ and y_ε coalesce for $\varepsilon\to 1$ to form a degenerate critical point. In this case the correlation function varies on scales $x\sim\mathcal{O}(N^{1/4})$.

VII. CONCLUSIONS

In this article we have adapted the NLσM approach introduced in Refs. 15 and 16 to the framework of quantized maps on a Hilbert space of dimension $N\sim\hbar^{-1}$. We focused on the spectral determinant correlation function $\Omega_U(\gamma)$ instead of the pair correlation function, thereby obviating the need to introduce supersymmetry; we obtained an *exact* expression for the correlation function as an ordinary integral over a N^2-dimensional complex manifold. Because the manifold is compact and the integrand uniformly bounded, no regularization needs to be introduced (unlike in Ref. 15).

To estimate this integral we expand the integrand around its saddle points, first restricting ourselves to the leading-order perturbative expansion around each point. Owing to the absence of a large parameter in front of the effective action, this approximation is uncontrolled, and the connection between its output and the exact value of the integral seems fortuitous at best.

Yet, for any matrix $U\in \mathrm{U}(N)$, we find that the result from lowest-order saddle-point expansion of the effective action $S(\gamma,U,Q)$ coincides with the exact correlation function, up to a global prefactor:

$$\Omega_U(\gamma)_{|\text{l.o.s.p. exp.}} = C_N \;\; \Omega_U(\gamma)_{\text{exact}}. \tag{46}$$

This remarkable coincidence is linked to a cancellation property of the higher-order terms of the perturbation expansion, which modify only the prefactor, and is explained by the group-theoretic structure of the integrand and the Weyl character formula. Unfortunately, the expansion is of no use for estimating the correlation function of quantized maps in the semiclassical limit.

We argue that a decent semiclassical estimate of the correlation function $\Omega_U(\gamma)$ can only be reached if one takes an average over a set of unitary matrices in the vicinity of U. To estimate this averaged correlation, we first average the integrand $e^{-S(\gamma,U,Q)}$ over U, and then perform the saddle-point expansion of the output. Because averaging and saddle-point expansion are operations that do not commute, this procedure yields an expansion different from that of the ''individual'' action. At the same time, averaging *a priori* breaks the group-theoretic structure, and with it the exactness (modulo prefactor) of the leading-order saddle-point expansion. Moreover, the explicit computation of saddle points and their contributions is, in general, a nontrivial task for a general averaging scheme.

We have been able to locate the complete set of critical points only for a certain type of average, namely averaging over all bases of Hilbert space. This produces a $\mathrm{U}(N)$-invariant effective action, the critical points of which are grouped into submanifolds, and are independent of the matrix U we started from (as long as its spectrum is nondegenerate). Two of these submanifolds are isolated points; we conjectured that the contributions from these two ''standard'' saddle points, which can be computed explicitly, always dominate the leading-order saddle-point expansion.

The contributions from these two points are unfortunately ''too simple'' to constitute a good approximation of the correlation function, except in some exceptional cases, which we do not truly

understand. If we average over $U \in \mathrm{U}(N)$ with the Poisson measure, the saddle-point result strongly differs from the exact one. We are thus led to conclude that the leading-order saddle-point expansion of rotation-averaged effective actions does not yield a good estimate of the full integral.

What happens in the case of a *local* average, i.e., when the weight of the probability measure is concentrated near the quantized map U_N, is unclear. For one thing, we are only able to exhibit the two standard saddle points of the averaged action, but there surely exist many more.

In the case of the ''semiclassical'' averaging scheme, expansion around these saddle points yields results similar to those obtained in Ref. 15, except that the ''resonances'' we identify are eigenvalues of a quantum operator. Yet, these resonances for large N seem related to the (classical) Ruelle–Pollicott resonances,[30,31] in particular they indicate whether the classical dynamics is chaotic or integrable.

To connect these resonances with the determinant correlation function on a rigorous footing, we need two nontrivial assumptions to be fulfilled. First, we must assume that the leading-order saddle-point expansion of the (local average) $S_{\mathrm{semiclas}}(\gamma, U_N, Q)$ makes sense, i.e., gives a good approximation of the exact result; the two-loop calculation around $\pm\Sigma_3$ in Sec. IV D seems to support this assumption. Second, hindered by our inability to compute the contributions from further saddle points, we are forced to assume that the full expansion can be truncated to the two standard saddle points, or at least that this truncation provides a reasonable approximation. We presently see no way to prove these assumptions.

APPENDIX A: PROOF OF CRITICALITY OF THE SUBMANIFOLDS $\mathcal{M}_S$

To prove that the V-averaged integrand $\mathrm{e}^{-S_{V\mathrm{av}}(\gamma,U,Q)}$ is stationary on the submanifolds $\mathcal{M}_S \subset \mathcal{M}_N$, we employ the coherent-state formulation of the Q-integral. The point Q_S corresponds to the state $|S\rangle = R(\mathbf{g}_\sigma)|0\rangle$, and the points in a neighborhood of Q_S may be parametrized as $R(\mathbf{g}_\sigma)|\zeta\rangle$, where ζ runs through the $N\times N$ matrices (with small coefficients) and $|\zeta\rangle$ is the corresponding coherent state. The permutation $\sigma \in \mathfrak{S}_{2N}$ is chosen in such a way as to interchange the sets $\overline{S_1}$ and $\widetilde{S}_2 = S_2 + N$, and to keep S_1 and $\widetilde{\overline{S_2}} = \overline{S_2} + N$ fixed.

We write the $2N\times 2N$ matrix $\mathbf{g}_\sigma^{-1}\mathbf{\Gamma}\mathbf{U}\mathbf{g}_\sigma$ in the block form $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$, and first compute the value of the integrand in the vicinity of Q_S *before averaging*:

$$\frac{\langle\zeta|R(\mathbf{g}_\sigma^{-1}\mathbf{\Gamma}\mathbf{U}\mathbf{g}_\sigma)|\zeta\rangle}{\langle\zeta|\zeta\rangle} = \mathrm{Det}(D)\left(1+\mathrm{Tr}(D^{-1}C\zeta + BD^{-1}\zeta^\dagger) + \mathcal{O}(|\zeta|^2)\right).$$

Then we perform the V-average on U (recall that $V \in \mathrm{U}(N)$ acts on $U \in \mathrm{U}(N)$ by conjugation: $U \mapsto VUV^{-1}$), and study its output on the right-hand side of the above equation. To first order in ζ and $\zeta^\dagger$, we need the averages $\langle \mathrm{Det}(D)\rangle_V$, $\langle D^{-1}C\ \mathrm{Det}(D)\rangle_V$ and $\langle BD^{-1}\mathrm{Det}(D)\rangle_V$. By decomposing the sets $\{1,\ldots,N\} = S_1 \cup \overline{S_1}$ and $\{N+1,\ldots,2N\} = \widetilde{S}_2 \cup \widetilde{\overline{S_2}}$, the $N\times N$ matrices B, C, D may be written in block form:

$$B = \begin{pmatrix} \gamma U_{S_1\overline{S_1}} & 0 \\ 0 & U_{S_2\overline{S_2}} \end{pmatrix}, \quad C = \begin{pmatrix} \gamma U_{\overline{S_1}S_1} & 0 \\ 0 & U_{\overline{S_2}S_2} \end{pmatrix}, \quad D = \begin{pmatrix} \gamma U_{\overline{S_1}\overline{S_1}} & 0 \\ 0 & U_{\overline{S_2}\overline{S_2}} \end{pmatrix},$$

where each entry $U_{ss'}$ is a matrix of size $\sharp s \times \sharp s'$, whose indices take values in the sets s, s'. Thus the V-averaged coefficients of the term linear in ζ are the following matrix elements:

$$\langle \mathrm{Det}(U_{\overline{S_1}\overline{S_1}})\mathrm{Det}(U_{\overline{S_2}\overline{S_2}})(U^{-1}_{\overline{S_1}\overline{S_1}}U_{\overline{S_1}S_1})_{ik}\rangle_V, \quad \langle \mathrm{Det}(U_{\overline{S_1}\overline{S_1}})\mathrm{Det}(U_{\overline{S_2}\overline{S_2}})(U^{-1}_{\overline{S_2}\overline{S_2}}U_{\overline{S_2}S_2})_{lm}\rangle_V,$$

where we have displayed only the dependence on U (and omitted the γ-dependence). We now use the invariance of the Haar measure dV under (left) multiplication by any unitary matrix and any diagonal unitary matrix $\delta = \mathrm{diag}(\delta_1,\ldots,\delta_N)$ in particular. Under such a left translation, the above matrix elements acquire extra factors δ_i/δ_k (resp. δ_l/δ_m). Hence

$$\langle \mathrm{Det}(D)(D^{-1}C)_{ik}\rangle_V = \langle \mathrm{Det}(D)(D^{-1}C)_{ik}\rangle_V \delta_i/\delta_k \quad \text{for any} \quad \delta_i, \delta_k .$$

Since $i \in \overline{S_1}$ and $k \in S_1$ (resp. $l \in \overline{S_2}$ and $m \in S_2$) are never equal and the ratio δ_i/δ_k may take any value in U(1), we conclude

$$\langle \mathrm{Det}(D)(D^{-1}C)_{ik}\rangle_V = 0.$$

By the same reasoning, the terms linear in $\zeta^\dagger$ vanish after V-averaging.

We have thus shown that the point Q_S on $\mathcal{M}_N$ is a critical point of the V-averaged action $S_{V\mathrm{av}}$, Eq. (36). By the U(N)-invariance of $S_{V\mathrm{av}}$, it follows that the whole submanifold $\mathcal{M}_S$ is critical for the V-averaged action, no matter what U is.

APPENDIX B: VOLUMES OF THE CRITICAL SUBMANIFOLDS

We treat the general case with $\sharp S_1 = \sharp \overline{S_2} = N-p$, $\sharp \overline{S_1} = \sharp S_2 = p$, $\sharp(S_1 \cap S_2) = \sharp(\overline{S_1} \cap \overline{S_2}) = r$, and to build Q_S we use the same permutation σ as in the previous appendix.

The manifold $\mathcal{M}_S$ is given by the set of states $\{R(\mathbf{V}\mathbf{g}_\sigma)|0\rangle | V \in \mathrm{U}(N)\}$. These states may be written (up to normalization) in the form $R(\mathbf{g}_\sigma)|\zeta_V\rangle$ where the coherent state $|\zeta_V\rangle$ is determined by the matrix

$$\zeta_V = \begin{pmatrix} V_{S_1\overline{S_1}} V_{\overline{S_1}\overline{S_1}}^{-1} & 0 \\ 0 & V_{S_2\overline{S_2}} V_{\overline{S_2}\overline{S_2}}^{-1} \end{pmatrix} \overset{\mathrm{def}}{=} \begin{pmatrix} \zeta^{(1)} & 0 \\ 0 & \zeta^{(2)} \end{pmatrix} \tag{B1}$$

according to Eqs. (16) and (17). The block structure of this matrix derives from the sets (S_1, S_2) vertically, and $(\overline{S_1}, \overline{S_2})$ horizontally.

When V runs through U(N), the upper-left matrix $\zeta^{(1)}$ takes all possible values in $\mathbb{C}^{(N-p)\times p}$. The matrix $\zeta^{(2)}$ is not independent of $\zeta^{(1)}$. For a fixed $\zeta^{(1)}$, we need to identify the remaining degrees of freedom in $\zeta^{(2)}$, which is quite easy to do if $\zeta^{(1)}=0$, i.e., if V has the structure $V = \mathrm{diag}(V_{S_1S_1}, V_{\overline{S_1}\overline{S_1}})$. The matrices $V_{S_2\overline{S_2}}$ and $V_{\overline{S_2}\overline{S_2}}$ in this case block decompose as

$$V_{S_2\overline{S_2}} = \begin{pmatrix} V_{12,\overline{12}} & 0 \\ 0 & V_{\overline{12},12} \end{pmatrix}, \quad V_{\overline{S_2}\overline{S_2}} = \begin{pmatrix} V_{\overline{12},\overline{12}} & 0 \\ 0 & V_{\overline{12},\overline{12}} \end{pmatrix}$$

where the index 12 refers to the set $S_1 \cap S_2$, etc. The degrees of freedom of the lower-right part of ζ_V are thus two matrices, $\zeta^{(11)} \overset{\mathrm{def}}{=} V_{12,\overline{12}} V_{\overline{12},\overline{12}}^{-1} \in \mathbb{C}^{r\times(N-p-r)}$, and $\zeta^{(\overline{11})} \overset{\mathrm{def}}{=} V_{\overline{12},\overline{12}} V_{\overline{12},\overline{12}}^{-1} \in \mathbb{C}^{(p-r)\times r}$. They are independent of each other, and take all possible values in their respective vector spaces. Since the subgroup U($N-p$)$\times$U(p) of U(N) acts transitively on the submanifold $\zeta^{(1)}=0$ of $\mathcal{M}_S$, there exists a natural choice of invariant measure on that submanifold. It has the factorized form

$$\mathrm{Det}(1+\zeta^{(11)\dagger}\zeta^{(11)})^{N-p} \prod_{i,j} d^2\zeta_{ij}^{(11)}/\pi \times \mathrm{Det}(1+\zeta^{(\overline{11})\dagger}\zeta^{(\overline{11})})^{p} \prod_{i,j} d^2\zeta_{ij}^{(\overline{11})}/\pi .$$

The matrix $\zeta^{(1)}$ parametrizes a coset space U(N)/U($N-p$)$\times$U(p), with the corresponding invariant measure being $\mathrm{Det}(1+\zeta^{(1)\dagger}\zeta^{(1)})^N \Pi_{i,j} d^2\zeta_{ij}^{(1)}/\pi$. By group invariance arguments, the volume element of $\mathcal{M}_S$ (normalized so that it agrees with the Riemannian measure inherited from the Riemannian manifold $\mathcal{M}_N$) is the product of the measures for $\zeta^{(1)}$, $\zeta^{(11)}$, and $\zeta^{(\overline{11})}$ above. Using this fact and the result[33]

$$I(m,n) \overset{\mathrm{def}}{=} \int_{\mathbb{C}^{m\times n}} \prod_{i=1}^{m} \prod_{j=1}^{n} \frac{d^2Z_{ij}}{\pi} \mathrm{Det}(1+Z^\dagger Z)^{-n-m} = \frac{\Gamma(1)\cdots\Gamma(n)\ \ \Gamma(1)\cdots\Gamma(m)}{\Gamma(1)\cdots\Gamma(n+m)},$$

we obtain the volume of $\mathcal{M}_S$:

$$\mathrm{Vol}\mathcal{M}_S = I(p,N-p)I(r,p-r)I(r,N-p-r).$$

A similar integral yields the normalization factor C_N of the measure $d\mu_N(Z,Z^\dagger)$ on the full manifold $\mathcal{M}_N$:

$$\frac{1}{C_N} = \int_{\mathbb{C}^{N\times N}} \prod_{i,j=1}^{N} \frac{d^2 Z_{ij}}{\pi} \mathrm{Det}(1+Z^\dagger Z)^{-2N-1} = \frac{\Gamma(2)\cdots\Gamma(N+1)}{\Gamma(N+2)\cdots\Gamma(2N+1)}.$$

[1] O. Bohigas, "Random matrix theory and chaotic dynamics," in *Chaos et physique quantique* (École d'été des Houches, Session LII, 1989), edited by M. J. Giannoni, A. Voros, and J. Zinn-Justin (North Holland, Amsterdam, 1991).
[2] M. V. Berry and M. Tabor, Proc. R. Soc. London, Ser. A **356**, 375 (1977).
[3] S. Zelditch, Ann. Inst. Fourier **47**, 305 (1997).
[4] M. Tabor, Physica D **6**, 195 (1983).
[5] M. V. Berry, Proc. R. Soc. London, Ser. A **400**, 229 (1985).
[6] F. Haake, M. Kuś, H.-J. Sommers, H. Schomerus, and K. Życzkowski, J. Phys. A **29**, 3641 (1996).
[7] S. Kettemann, D. Klaklow, and U. Smilansky, J. Phys. A **30**, 3643 (1997).
[8] R. Prange, Phys. Rev. Lett. **78**, 2280 (1997).
[9] E. Bogomolny, F. Leyvraz, and C. Schmit, Commun. Math. Phys. **176**, 577 (1996).
[10] J. P. Keating and F. Mezzadri, Nonlinearity **13**, 747 (2000).
[11] J. Marklof, "The Berry-Tabor conjecture," in *Proceedings of the 3rd European Congress of Mathematics*, Barcelona 2000 (Birkhäuser, Basel, 2001).
[12] M. R. Zirnbauer, J. Phys. A **29**, 7113 (1996).
[13] E. B. Bogomolny and J. P. Keating, Phys. Rev. Lett. **77**, 1472 (1996).
[14] K. B. Efetov, Adv. Phys. **32**, 53 (1983).
[15] A. V. Andreev, B. D. Simons, O. Agam, and B. L. Altshuler, Nucl. Phys. B **482**, 536 (1996).
[16] B. A. Muzykantskii and D. E. Khmel'nitskii, JETP Lett. **62**, 76 (1995).
[17] A. Altland, C. R. Offer, and B. D. Simons, "Quantum Chaos: Lessons from Disordered Metals," in *Supersymmetry and Trace Formulae, Chaos and Disorder*, edited by I. V. Lerner, J. P. Keating, and D. E. Khmel'nitskii (Kluwer Academic/Plenum, Dordrecht, 1999).
[18] D. Taras-Semchuk and K. B. Efetov, Phys. Rev. B **64**, 115301 (2001).
[19] I. V. Gornyi and A. D. Mirlin, "Wave function correlations on the ballistic scale: Exploring quantum chaos by quantum disorder," cond-mat/0105103.
[20] M. R. Zirnbauer, "Pair correlations of Quantum Chaotic maps from Supersymmetry," in *Supersymmetry and Trace Formulae, Chaos and Disorder*, edited by I. V. Lerner, J. P. Keating, and D. E. Khmel'nitskii (Kluwer Academic/Plenum, Dordrecht, 1999).
[21] E. Bogomolny, O. Bohigas, and P. Lebœuf, J. Stat. Phys. **85**, 639 (1996).
[22] R. Howe, Trans. Am. Math. Soc. **313**, 539 (1989).
[23] H. Weyl, *The Classical Groups: Their Invariants and Representations* (Princeton University Press, Princeton, NJ, 1946).
[24] A. M. Perelomov, *Generalized Coherent States and their Applications* (Springer, New York, 1986).
[25] A. V. Andreev and B. L. Altshuler, Phys. Rev. Lett. **75**, 902 (1995).
[26] A. Knapp, *Representation Theory of Semisimple Groups* (Princeton University Press, Princeton, NJ, 1986).
[27] M. Stone, Nucl. Phys. B **314**, 557 (1989).
[28] J. Milnor, *Morse Theory*, (Princeton University Press, Princeton, NJ, 1963).
[29] J. Leray, C. R. Acad. Sci. Paris **228**, 1902 (1949).
[30] C. Manderfeld, J. Weber, and F. Haake, "Classical versus Quantum Time Evolution of Densities at Limited Phase-Space Resolution," nlin.CD/0107020.
[31] S. Fishman, "Wave Functions, Wigner Functions and Green Functions of Chaotic Systems," in *Supersymmetry and Trace Formulae, Chaos and Disorder*, edited by I. V. Lerner, J. P. Keating, and D. E. Khmel'nitskii (Kluwer Academic/Plenum, Dordrecht, 1999).
[32] D. Ruelle, Phys. Rev. Lett. **56**, 405 (1986).
[33] L. K. Hua, *Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains* (American Mathematical Society, Providence, RI, 1963).