Proc. R. Soc. Lond. A **423**, 219–231 (1989)
Printed in Great Britain

Quantum scars of classical closed orbits in phase space

BY M. V. BERRY, F.R.S.
H. H. Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, U.K.

(*Received* 15 *November* 1988)

The way in which quantum eigenstates are influenced by the closed orbits of a chaotic classical system is analysed in phase space $\boldsymbol{x} = (\boldsymbol{q}, \boldsymbol{p})$ through the spectral Wigner function $W(\boldsymbol{x}; E, \epsilon)$. This is a sum over Wigner functions of eigenstates within a range ϵ of energy E. In the classical limit, W is concentrated on the energy surface and smoothly distributed over it. Closed orbits provide oscillatory corrections (scars) for which explicit semiclassical formulae are calculated. Each scar is a fringe pattern decorating the orbit. As $\boldsymbol{x}$ moves off the energy surface the fringes form an Airy pattern with spacing of order $\hbar^{\frac{2}{3}}$. As $\boldsymbol{x}$ moves off the closed orbit the fringes form a complex gaussian with spacing $\hbar^{\frac{1}{2}}$.

1. INTRODUCTION

When the orbits of a classical system are chaotic, it is natural to conjecture that in the semiclassical limit of small Planck's constant $\hbar$ the quantum eigenstates are associated with the whole energy surface explored ergodically by the orbits. This is a special case of the 'semiclassical eigenfunction hypothesis' (Berry 1983); semiclassical states are associated with minimal generic classical invariant sets. In phase space the conjecture implies that the Wigner function of a state is a delta-function concentrated on the surface with the same energy as the state. The conjecture is supported by a theorem of Shnirelman (1974): the quantum expectation value of a smooth operator is the classical microcanonical average for almost all states. Berry (1977*a*) and Voros (1979) formulated the conjecture and showed how it leads to predictions about the wavefunctions in configuration space. Numerical tests of the predictions have been made by McDonald & Kaufman (1988) and Shapiro & Goelman (1986).

Thanks to the computer explorations and theoretical arguments of Heller (1984, 1986) we now know that this picture is too simple. Quantum eigenstates are influenced not only by the energy surface, which is the generic invariant manifold, but by *individual closed orbits*, which are invariant sets of zero measure. The imprints of closed orbits persist up through thousands of states and probably survive into the classical limit. Heller calls these imprints *scars*.

A major step in the theory of scars has been made by Bogomolny (1988) (see also Ozorio de Almeida 1988), who calculated the probability density smoothed over small intervals in energy and in space. The scars that he found were oscillations centred on the closed orbits and superimposed on the dominant contribution from the energy surface.

In this paper my purpose is to show that the scar formulae become very simple when expressed in phase space rather than in configuration space, that is in the

Wigner–Weyl representation (Groenewold 1946; Moyal 1949; Takabayashi 1954; Baker 1958). For any isolated periodic orbit of a system with any number of freedoms, the scar is concentrated near the orbit and surrounded by fringes whose characteristic forms are different off and on the energy surface. This development of Bogomolny's idea extends to chaotic systems my earlier analysis (Berry 1977*b*) of integrable systems, whose Wigner functions are concentrated near invariant tori in phase space, with characteristic fringes away from the tori.

The technique I shall use is a lift into phase space of the analysis by Gutzwiller (1971) of the Green function as a sum over classical paths, with orbits near closed ones being approximated by linear maps on Poincaré sections. This is similar to the method used by Bogomolny (1988), but without the spatial smoothing employed by him (and also by Berry 1977*a*), which is unnecessary. The central result is the scar formula (38).

2. Spectral Wigner function

Let the classical system have N freedoms, and phase-space variables

$$\boldsymbol{x} = (\boldsymbol{q},\boldsymbol{p}) = (q_1,\ldots,p_N), \tag{1}$$

where $\boldsymbol{q}$ are the coordinates and $\boldsymbol{p}$ the momenta. Let the hamiltonian be $H(\boldsymbol{x})$. From the corresponding quantal operators (denoted by carats) we can construct the lorentzian-smoothed *spectral operator*

$$\hat{\Delta}(E;\epsilon) \equiv \delta_\epsilon(E-\hat{H}) \equiv -\pi^{-1}\,\mathrm{Im}\,(E+\mathrm{i}\epsilon-\hat{H})^{-1}. \tag{2}$$

The trace of the spectral operator is the smoothed spectral density

$$\mathrm{Tr}\,\hat{\Delta} \equiv d(E;\epsilon) = \sum_n \delta_\epsilon(E-E_n), \tag{3}$$

where E_n are the eigenvalues. The spatial matrix elements are

$$\langle \boldsymbol{q}_A|\hat{\Delta}|\boldsymbol{q}_B\rangle = \sum_n \delta_\epsilon(E-E_n)\,\psi_n^*(\boldsymbol{q}_B)\,\psi_n(\boldsymbol{q}_A) \tag{4}$$

where $\psi_n(\boldsymbol{q})$ are the eigenfunctions.

Our interest will focus on the Weyl transform of the spectral operator, which we will call the *spectral Wigner function*:

$$W(\boldsymbol{x};E,\epsilon) \equiv h^N\,\mathrm{Tr}\,\hat{\Delta}\,\delta(\hat{\boldsymbol{x}}-\boldsymbol{x}) \tag{5a}$$

$$= \int \mathrm{d}\boldsymbol{q}'\,\exp\{-\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{q}'/\hbar\}\,\langle \boldsymbol{q}+\tfrac{1}{2}\boldsymbol{q}'|\hat{\Delta}|\boldsymbol{q}-\tfrac{1}{2}\boldsymbol{q}'\rangle \tag{5b}$$

$$= h^N \sum_n \delta_\epsilon(E-E_n)\,W_n(\boldsymbol{x}), \tag{5c}$$

where $\delta(\hat{\boldsymbol{x}}-\boldsymbol{x})$ is defined by its Fourier transform and W_n are the Wigner functions of the eigenstates ψ_n. In all these formulae the effect of the energy smoothing ϵ will depend on its size in relation to inner and outer scales $\Delta E_{\min}$ and $\Delta E_{\max}$. $\Delta E_{\min}$ is the mean level spacing, of order h^N, and $\Delta E_{\max} = h/T_{\min}$ where $T_{\min}$ is the period of the shortest closed orbit of energy E. If $\epsilon \ll \Delta E_{\min}$, W is dominated by the single eigenstate $n(E)$ closest to E, that is

$$W(\boldsymbol{x}; E, \epsilon) \to h^N \delta_\epsilon(E - E_{n(E)})\, W_{n(E)}(\boldsymbol{x}) \quad (\epsilon \ll \Delta E_{\min}). \tag{6}$$

(If E is midway between two eigenstates, W is the average of their Wigner functions.) If $\epsilon \gg \Delta E_{\max}$, W is an average over many states near E, in which as we shall see all oscillatory structure is washed away and we regain the classical limit

$$W(\boldsymbol{x}; E, \epsilon) \to \delta_\epsilon(E - H(\boldsymbol{x})) \quad (\epsilon \gg \Delta E_{\max}). \tag{7}$$

To motivate the analysis that will follow, note first that in terms of W the spectral density (3) is

$$d(E, \epsilon) = h^{-N} \int \mathrm{d}\boldsymbol{x}\, W(\boldsymbol{x}; E, \epsilon). \tag{8}$$

Second, recall that d can be expanded as an average plus corrections in the form of oscillations from the closed orbits (Gutzwiller 1971; Balian & Bloch 1972; see also Ozorio de Almeida 1988). The average is obtained by substituting (7) into (8), that is by integrating over the smoothed energy surface in phase space. It is then natural to expect that the corrections to d will appear as integrals over phase-space functions localized near the closed orbits. These contributions are the scars. They will emerge in the form of oscillatory corrections as (7) is unsmoothed by reducing ϵ.

The analysis will be based on the formula (5*b*), whose lack of symmetry hides the invariance of W under linear canonical transformations of $\boldsymbol{x}$ (Balazs 1980). The fact that in the final scar formulae this invariance will be obvious provides an assurance that the analysis is correct, just as it did in our earlier study of integrable systems (Berry 1977*b*).

We could perform a phase-space gaussian smoothing of our formulae for W, and thereby obtain the Husimi distribution that has been employed by some authors (e.g. Takahashi 1986; Leboeuf *et al.* 1988). This is non-negative (unlike W), and the smoothing may cause classical structures to stand out more clearly as $h \to 0$. But the smoothing brings two disadvantages. First, it destroys the invariance under linear canonical transformations (because of the arbitrariness in the eccentricity and orientation of the gaussian ellipse contours, originating in the arbitrariness in the underlying harmonic oscillator). Second, it destroys fringes that display essential quantal information (about non-locality, for example); after smoothing, this information may still be recoverable (Kano 1965), but this involves an analytic continuation which can be tricky (Takahashi 1986).

3. Semiclassical theory

Writing (2) as

$$\hat{\Delta}(E; \epsilon) = \frac{2}{h} \mathrm{Re} \int_0^\infty \mathrm{d}t \exp\{\mathrm{i}(E - \hat{H})\, t/\hbar\} \exp(-\epsilon t/\hbar), \tag{9}$$

we obtain the spectral Wigner functions (5) as a time integral over the Wigner propagator,

$$K_{\mathrm{W}}(\boldsymbol{x}, t) = \int \mathrm{d}\boldsymbol{q}' \exp(-\mathrm{i}\boldsymbol{p} \cdot \boldsymbol{q}'/\hbar)\, K(\boldsymbol{q} - \tfrac{1}{2}\boldsymbol{q}', \boldsymbol{q} + \tfrac{1}{2}\boldsymbol{q}', t), \tag{10}$$

involving the coordinate propagator,

$$K(q_A, q_B, t) \equiv \langle q_B| \exp(-\mathrm{i}\hat{H}t/\hbar)|q_A\rangle, \tag{11}$$

from q_A to q_B in time t, where the end-points satisfy

$$q = \tfrac{1}{2}(q_A + q_B). \tag{12}$$

The entry point for semiclassical analysis is an approximation for K as a sum over all classical paths (labelled j) starting from q_A and ending at q_B in time t, derived by Van Vleck (1928) (see also Berry & Mount 1972);

$$K(q_A, q_B, t) \approx \sum_j \{h^N \det[R_{AmBn,j}]\}^{\frac{1}{2}} \exp\{\mathrm{i}R_j(q_A, q_B, t)/\hbar + \mathrm{i}\gamma_j\}. \tag{13}$$

In this formula, R_j is the time-dependent action (Hamilton's principal function, see, for example, Synge 1960) along the path j, and

$$R_{AmBn,j} \equiv \partial^2 R_j/\partial q_{Am}\,\partial q_{Bn}. \tag{14}$$

The phase γ_j is a multiple of $\frac{1}{4}\pi$ determined by the focusing of paths close to j; for economy of notation we shall henceforth denote all such focal indices by the generic symbol γ without specifying their precise values, which will change under the transformations of K that we shall soon carry out.

The semiclassical Wigner propagator is obtained by evaluating (10) by the method of stationary phase. Stationary points are defined by

$$\nabla_{q'}[-p\cdot q' + R(q - \tfrac{1}{2}q', q + \tfrac{1}{2}q', t)] = 0, \tag{15}$$

i.e.

$$p = \tfrac{1}{2}(p_A + p_B). \tag{16}$$

Together with (12) this gives the *midpoint rule*: the semiclassical Wigner propagator at x contains contributions from the classical paths j that in time t link phase points x_{Aj}, x_{Bj} centred on x.

The phase of each contribution is

$$\begin{aligned} -p\cdot q_j' + R(q - \tfrac{1}{2}q_j', q + \tfrac{1}{2}q_j') &= -p\cdot(q_{Bj} - q_{Aj}) + \int_{q_{Aj}}^{q_{Bj}} p_j(q)\cdot \mathrm{d}q - H_j t \\ &= A_j(x, t) - H_j t. \end{aligned} \tag{17}$$

Here $H_j = H(x_{Aj}(x, t))$ is the energy of the path (which need not be the same as the energy $H(x)$ of the point x). A_j is defined by the following *chord rule* (figure 1): it is the symplectic area of the circuit that starts from x_{Aj}, goes to x_{Bj} along the classical path, and returns straight to x_{Aj} via x. Similar chord and midpoint rules were found in the semiclassical theory of Wigner functions of eigenstates, both integrable (Berry 1977; Balazs 1980; Ozorio de Almeida & Hannay 1982) and non-integrable (Ozorio de Almeida 1988).

The amplitude of each contribution is $2^N D_j^{\frac{1}{2}}$, where (omitting the label j)

$$D = \frac{\det R_{AB}}{\det(R_{AA} - R_{AB} - R_{BA} + R_{BB})}. \tag{18}$$

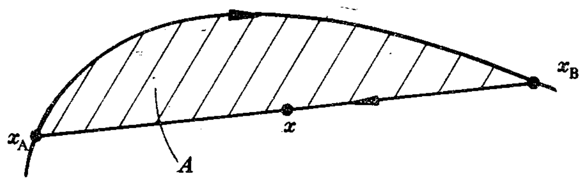

FIGURE 1. Chord construction for phase of contribution to Wigner propagator and spectral Wigner function at x.

(Note that the $N \times N$ matrices R_{AA} and R_{AB} are symmetric and R_{AB} is the transpose of R_{BA}.) This determinant can be expressed in terms of the symplectic map from $\boldsymbol{x}_A$ to $\boldsymbol{x}_B$, that is in terms of the $2N \times 2N$ matrix

$$m \equiv \mathrm{d}\boldsymbol{x}_B/\mathrm{d}\boldsymbol{x}_A \tag{19}$$

for it can be shown (appendix A) that

$$D = (-1)^N/\det(m+I), \tag{20}$$

where I denotes the identity matrix.

Thus the semiclassical Wigner propagator takes a form that is manifestly invariant under linear canonical transformations;

$$K_W(\boldsymbol{x}, t) \approx 2^N \sum_j \frac{\exp\{\mathrm{i}[A_j(\boldsymbol{x}, t) - H(\boldsymbol{x}_{Aj}(\boldsymbol{x}, t))\, t]/\hbar + \mathrm{i}\gamma_j\}}{\{\det(m_j(\boldsymbol{x}, t) + I)\}^{\frac{1}{2}}}. \tag{21}$$

When $t \to 0$ there is only one contributing path, with $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ close to $\boldsymbol{x}$ and m close to the identity. The area $A \to 0$ (cf. figure 1), so

$$K_W(\boldsymbol{x}, t) \to \exp\{-\mathrm{i}H(\boldsymbol{x})\, t/\hbar\} \quad \text{as} \quad t \to 0. \tag{22}$$

This is the obvious semiclassical limit for short times. For larger t, a path contribution will diverge if $\det(m_j + I)$ vanishes, that is if m_j has an eigenvalue -1, because then the midpoint rule at $\boldsymbol{x}, t$ is satisfied not only at $\boldsymbol{x}_A, \boldsymbol{x}_B$ but also for some first-order variations away from $\boldsymbol{x}_A, \boldsymbol{x}_B$. The divergences signal jumps of $\frac{1}{2}\pi$ in the phase γ (which is zero for short times).

Combining (9) and (21) we obtain an integral representation for the spectral Wigner function:

$$W(\boldsymbol{x}; E, \epsilon) \approx \frac{2^{N+1}}{h} \sum_j \mathrm{Re} \int_0^\infty \mathrm{d}t \frac{\exp(-\epsilon t/\hbar)}{\{\det[m_j(\boldsymbol{x}, t) + I]\}^{\frac{1}{2}}} \times \exp\{\mathrm{i}[A_j(\boldsymbol{x}, t) + (E - H_j(\boldsymbol{x}, t))\, t]/\hbar + \mathrm{i}\gamma_j)\}. \tag{23}$$

In the semiclassical limit, there are two sorts of contribution to the integral. First there is the neighbourhood of the endpoint $t \to 0$. For this, the limiting form (22) gives the classical limit (7). Second, there are contributions from those finite t for which the phase is stationary. The condition is

$$\begin{aligned} \mathrm{d}_t[A + (E - H)\, t] &= \mathrm{d}_t[R(\boldsymbol{q}_A, \boldsymbol{q}_B, t) + Et - \boldsymbol{p} \cdot (\boldsymbol{q}_B - \boldsymbol{q}_A)] \\ &= \partial_t R + E - \mathrm{d}_t \boldsymbol{q}_A \cdot \boldsymbol{p}_A + \mathrm{d}_t \boldsymbol{q}_B \cdot \boldsymbol{p}_B - \boldsymbol{p} \cdot \mathrm{d}_t(\boldsymbol{q}_B - \boldsymbol{q}_A) \\ &= E - H(\boldsymbol{x}_A(\boldsymbol{x}, t)) = 0, \end{aligned} \tag{24}$$

where (12) and (16) have been used. Thus we see that these contributions to $W(\boldsymbol{x};E,\epsilon)$ come from the finite classical paths with energy E linking states $\boldsymbol{x}_A, \boldsymbol{x}_B$ with midpoint $\boldsymbol{x}$. The phases are again given by the areas A (figure 1).

The spectral Wigner function can be written

$$W(\boldsymbol{x};E,\epsilon) \approx \delta_\epsilon[E-H(\boldsymbol{x})]+\sum_j W_j(\boldsymbol{x};E,\epsilon), \tag{25}$$

where W_j comes from the jth classical path. Scars are associated with the W_j as will now be explained.

4. Scar expansion

It is possible to pursue the stationary-phase analysis just outlined, to obtain the path contributions W_j for arbitrary $\boldsymbol{x}$. More explicit and transparent formulae are however obtained by making a transitional approximation for phase points $\boldsymbol{x}$ close to closed orbits with energy E. Points on these closed orbits are special because they are associated with degenerate stationary points of the integral (23) and so their semiclassical contributions to W are larger.

To see that there is a degeneracy, note first that on a closed orbit with energy E the stationary-phase condition (24) is satisfied when $\boldsymbol{x}_A = \boldsymbol{x}_B = \boldsymbol{x}$ and $t = T$ where T is the period of the orbit. Next note that when t differs to first order from T the midpoint rule can be satisfied by sliding $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ in opposite directions along the orbit, so that the energy $H(\boldsymbol{x}_A(\boldsymbol{x},t))$ remains equal to E. Therefore $\partial_t H(\boldsymbol{x}_A(\boldsymbol{x},t))$ vanishes on a closed orbit. But according to (24) this quantity is the second derivative of the phase of (23), so that we have indeed established that there is a degeneracy.

It is necessary to expand the phase in (23) to third order in

$$\tau \equiv t-T \tag{26}$$

to evaluate the integral asymptotically. Referring to figure 2 we see that the endpoints of the chord centred on $\boldsymbol{x}$ are, in lowest order,

$$\boldsymbol{x}_A = \boldsymbol{x}-\tfrac{1}{2}\tau\dot{\boldsymbol{x}}; \quad \boldsymbol{x}_B = \boldsymbol{x}+\tfrac{1}{2}\tau\dot{\boldsymbol{x}}, \tag{27}$$

where $\dot{\boldsymbol{x}}$ is the phase velocity on the orbit at $\boldsymbol{x}$. The energy of the orbit is

$$\begin{aligned} H(\boldsymbol{x}_A) = H(\boldsymbol{x}_B) &= H(\boldsymbol{x}-\tfrac{1}{2}\dot{\boldsymbol{x}}\tau) \\ &= H(\boldsymbol{x})+\tfrac{1}{8}\tau^2\dot{\boldsymbol{x}}\cdot\nabla\dot{\boldsymbol{x}}\cdot\nabla H \end{aligned} \tag{28}$$

(the term linear in τ vanishes because $\dot{\boldsymbol{x}}$ lies in the energy surface and so is perpendicular to the gradient of H). From Hamilton's equations, the coefficient of τ^2 can be put in the form

$$\begin{aligned} -\dot{\boldsymbol{x}}\cdot\nabla\dot{\boldsymbol{x}}\cdot\nabla H &= -\ddot{\boldsymbol{q}}\cdot\dot{\boldsymbol{p}}+\ddot{\boldsymbol{p}}\cdot\dot{\boldsymbol{q}} \\ &= \begin{pmatrix}\ddot{\boldsymbol{q}}\\ \ddot{\boldsymbol{p}}\end{pmatrix}\begin{pmatrix}0 & -I\\ I & 0\end{pmatrix}\begin{pmatrix}\dot{\boldsymbol{q}}\\ \dot{\boldsymbol{p}}\end{pmatrix} = \ddot{\boldsymbol{x}}J\dot{\boldsymbol{x}}, \\ &= \ddot{\boldsymbol{x}}\wedge\dot{\boldsymbol{x}}, \end{aligned} \tag{29}$$

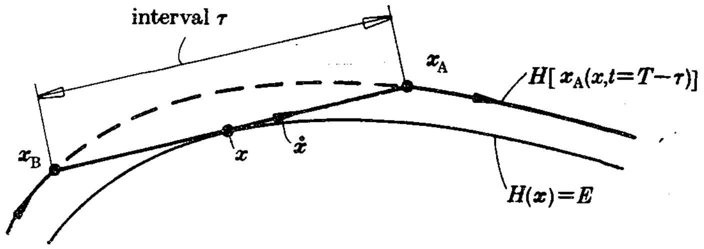

FIGURE 2. Chord construction when x is on a closed orbit and t is close to the orbit period T.

where the last two terms define the $2N \times 2N$ unit symplectic matrix and the symplectic product. Thus (28) becomes

$$H(x_A(x, T+\tau)) = H(x) - \tfrac{1}{8}\tau^2 \ddot{x} \wedge \dot{x}. \tag{30}$$

To get the full lowest-order expansion of the phase in (23), we must now move x off the closed orbit. Define Poincaré surface of section coordinates (figure 3)

$$X = (Q, P) = (Q_1, \ldots, Q_{N-1}, P_1, \ldots, P_{N-1}) \tag{31}$$

transverse to the orbit. On the closed orbit, $X = 0$. Together with the energy H and time t along an orbit these define a new set of canonical variables. The surface contains x and corresponds to $t = 0$. The closed orbit will be assumed to be isolated on each energy surface, as is typical for chaotic systems. It is part of a manifold of closed orbits parameterized by H and filling a two-dimensional region of phase space (the other variable being t).

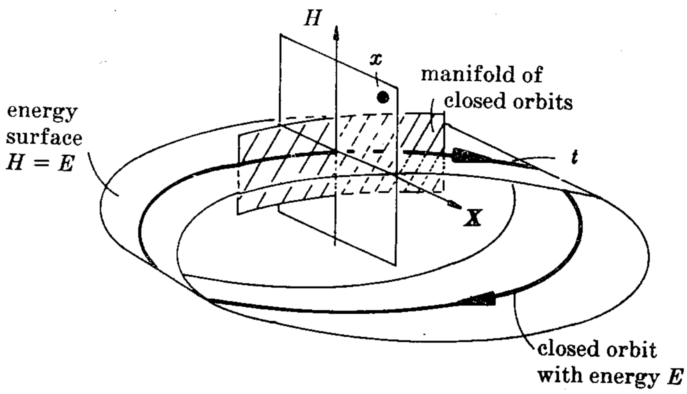

FIGURE 3. Poincaré section coordinates X in phase space near a closed orbit.

When x is not on the orbit, $X \neq 0$ and X_A and X_B lie on an orbit that is not closed (figure 4). The area A contributing to (23) is (cf. figure 1) that of the loop from X_A to X_B along the orbit and straight back via X to X_A in the surface of section. In lowest order A is simply the area

$$S(E) = \oint p \cdot dq \tag{32}$$

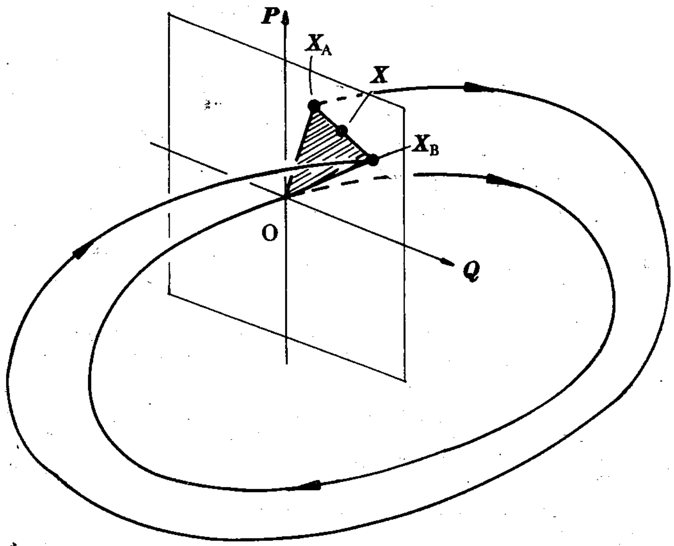

FIGURE 4. Energy surface E showing closed orbit and nearby orbit from X_A to X_B contributing to spectral Wigner function at X.

of the closed orbit. The correction can be calculated as indicated in figure 5, with the result

$$A = \oint \boldsymbol{p}\cdot \mathrm{d}\boldsymbol{q} - \tfrac{1}{2}(\boldsymbol{P}_A\cdot \boldsymbol{Q}_B - \boldsymbol{P}_B\cdot\boldsymbol{Q}_A)$$
$$= S(E) - \tfrac{1}{2}\boldsymbol{X}_A \wedge \boldsymbol{X}_B. \quad (33)$$

A simple explicit expression for $\boldsymbol{X}_A \wedge \boldsymbol{X}_B$ can be obtained in terms of the symplectic Poincaré map defined by

$$\boldsymbol{X}_B = M\boldsymbol{X}_A \quad (34)$$

relating successive intersections with the surface of section of an orbit near the closed one. The expression is (appendix B)

$$\tfrac{1}{2}\boldsymbol{X}_A \wedge \boldsymbol{X}_B = \tilde{\boldsymbol{X}}[J(M-I)/(M+I)]\boldsymbol{X}, \quad (35)$$

where now J is the $2(N-1)\times 2(N-1)$ unit symplectic matrix (cf. (29)).

Combining (33), (35), (30) and (24) we obtain the expansion of the phase in (23) for times $\tau = t - T$ small and $\boldsymbol{x}$ near a closed orbit and near the energy surface E;

$$A(x,t) + (E-H)\,t \approx S - \tilde{\boldsymbol{X}}[J(M-I)/(M+I)]\boldsymbol{X} + [E - H(\boldsymbol{x})]\,\tau + \tfrac{1}{24}\tau^3 \ddot{\boldsymbol{x}} \wedge \dot{\boldsymbol{x}}. \quad (36)$$

In evaluating (23) we can replace the determinant by its value on the closed orbit. This can be expressed in terms of the Poincaré map M because of the trivial dependence of the full map m on the variables H and t;

$$\det(m+I) = 4\det(M+I). \quad (37)$$

In the resulting integral the exponent has two stationary points, corresponding to two non-closed orbits with t slightly smaller and slightly larger than T (figure 2 shows the first of these; the second is got by interchanging $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$). As

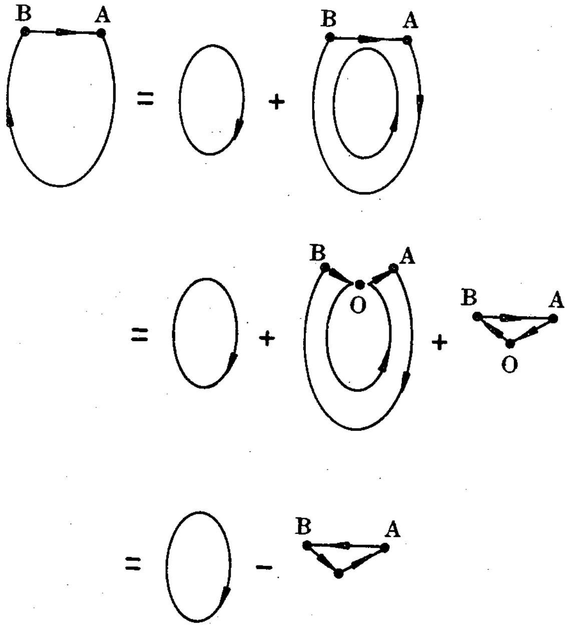

FIGURE 5. Geometry of calculation of symplectic area near a closed orbit. In the second line the middle term vanishes because it is the action of a reducible curve on a manifold of orbits. The last term in the third line is the shaded area in figure 4.

$H(\boldsymbol{x}) \rightarrow E$ these coalesce in the degeneracy described at the beginning of this section. The integral is an Airy function (Abramowitz & Stegun 1964). We immediately obtain the central result of this paper: in the spectral Wigner function each closed orbit gives a contribution

$$\begin{aligned} W_{\text{scar}}(\boldsymbol{x};E,\epsilon) &= [2^N/\sqrt{\det(M+I)}]\exp(-\epsilon T/\hbar) \\ &\quad \times \cos\{[S(E)-\tilde{\boldsymbol{X}}[J(M-I)/(M+I)]\boldsymbol{X}]/\hbar+\gamma\} \\ &\quad \times 2/(\hbar^2|\ddot{\boldsymbol{x}}\wedge\dot{\boldsymbol{x}}|)^{\frac{1}{3}}\,\mathrm{Ai}\{2[H(\boldsymbol{x})-E]/(-\hbar^2\ddot{\boldsymbol{x}}\wedge\dot{\boldsymbol{x}})^{\frac{1}{3}}\}. \end{aligned} \tag{38}$$

(In the expansion (25), W_{scar} represents a single term W_j.)

5. DISCUSSION

A number of interesting features of the scar formula (38) will now be enumerated.

(i) According to (9), the weight of a scar, defined as the phase-space integral of its spectral Wigner function, should give the contribution of the closed orbit to the smoothed spectral density. The integral is easily evaluated in $H, t, \boldsymbol{X}$ variables because the trigonometric and Airy factors contribute separately. For the H integration we use

$$\int_{-\infty}^{+\infty} \mathrm{d}u\,\mathrm{Ai}\,(u) = 1. \tag{39}$$

The t integration gives the period T/k, where k is the number of repetitions of the primitive orbit. The X integration involves a $2(N-1)$-dimensional quadratic form. We obtain

$$d_{\text{scar}} = (T/k\pi\hbar)\exp(-\epsilon T/\hbar)\cos(S/\hbar+\gamma)/\sqrt{\det(M-1)}, \quad (40)$$

which is precisely the result already obtained by Gutzwiller (1971).

(ii) The Airy function shows that as x moves off the energy surface in one direction, W_{scar} oscillates with a fringe spacing of order $\hbar^{\frac{2}{3}}$. In the other direction, W_{scar} decays exponentially. The directions of oscillation and decay are determined by the sign of $\ddot{x}\wedge\dot{x}$, a quantity whose geometrical significance has been explained by Balazs (1980). (In one dimension, the closed orbit coincides with the invariant torus, $\ddot{x}\wedge\dot{x}$ is negative where the orbit is convex in phase space and then the fringes decorate the interior of the orbit, see Berry (1977*b*).) These Airy fringes are not confined to the scars. They also occur (lorentzian-smoothed) in a more refined semiclassical theory of the energy-surface term in (25), based on correcting the phase in (22) with the term in t^3. This implies interesting oscillations of wavefunctions near the boundaries of classically allowed regions in configuration space. Details will be published elsewhere.

(iii) The cosine in (38) shows that as x moves off the closed orbit W_{scar} oscillates with a fringe spacing of order $\hbar^{\frac{1}{2}}$. The pattern of these 'transverse' fringes varies with position along the orbit because the Poincaré map M does. But the amplitude does not depend on position along the orbit because $\det(M+I)$ does not.

(iv) For $N=2$, M is a 2×2 matrix with eigenvalues $\exp(\pm\lambda)$ (which are real if the orbit is unstable). The amplitude of the oscillations is then

$$\{\det(M+I)\}^{-\frac{1}{2}} = \{2\cosh(\tfrac{1}{2}\lambda)\}^{-1} \quad (41)$$

and the transverse fringes on the Poincaré section are hyperbolae with (in suitable Poincaré coordinates Q, P) phase

$$-\tilde{X}[J(M-I)/(M+I)]X = -2QP\tanh(\tfrac{1}{2}\lambda). \quad (42)$$

As the orbit bifurcates (e.g. when a parameter is changed) M degenerates, either because $\lambda\to 0$ (e.g. $M\to I$) or because $\lambda\to i\pi$ (e.g. $M\to -I$). As $\lambda\to 0$ the amplitude (41) remains finite, but (42) shows that the fringe spacing diverges indefinitely leading to a divergence in the scar weight (40) (because $\det(M-I) = 4\sinh^2(\frac{1}{2}\lambda)\to 0$). As $\lambda\to i\pi$ the amplitude (41) diverges, the fringe spacing vanishes and the fringe weight remains finite.

(v) As the classical limit $h=0$ is approached, the fringes become too close to be resolved, and (38) can be written

$$W_{\text{scar}}(x;E,\epsilon)\to[2h^{N-1}/k\sqrt{\det(M-I)}]\exp(-\epsilon T/\hbar) \times\cos(S/\hbar+\gamma)\,\delta\{E-H(x)\}\,\delta(X). \quad (43)$$

When combined with (25) this gives the picture of the spectral Wigner function as a distribution concentrated on the energy surface with additional weak concentrations (of order h^{N-1}) on the closed orbits. For scales of resolution between $\hbar^{\frac{2}{3}}$ and $\hbar^{\frac{1}{2}}$ the Airy fringes cannot be resolved but the transverse (off-orbit) fringes can. In this limit,

$$W_{\text{scar}}(\boldsymbol{x};E,\epsilon) \to [2^N/\sqrt{\det(M+I)}]\exp(-\epsilon T/\hbar)\,\delta\{E-H(\boldsymbol{x})\}$$
$$\times\cos\{[S-\tilde{\boldsymbol{X}}[J(M-I)/(M+I)]\,\boldsymbol{X}]/\hbar+\gamma\}, \quad (44)$$

so that the scars decorate the energy surface with oscillations whose amplitude is of order zero and depends only on the stability of the orbit (via M), as argued in a different way by Heller (1984, 1986).

(vi) When projected 'down' $\boldsymbol{p}$ onto $\boldsymbol{q}$, (38) reproduces the configuration-space formulae of Bogomolny (1988).

(vii) $W(\boldsymbol{x};E,\epsilon)$ is the superposition of scars (38) from all the closed orbits. These include all repetitions of each primitive orbit. If T and M refer to the primitive orbit, the scar for the kth repetition is obtained by substituting kT and M^k (in (41) and (42), λ becomes $k\lambda$). Repetitions do not affect the Airy factor in (38), but do change the transverse fringes.

(viii) For chaotic systems the closed orbits are unstable. Long orbits contribute to (38) in a particularly simple way. The amplitude of the transverse fringes decreases with the orbit period T as

$$[\det(M+I)]^{-\frac{1}{2}} \to \exp(-h_{\text{KS}}T) \quad \text{(long orbits)}, \quad (45)$$

where h_{KS} is the Kolmogorov–Sinai entropy (Lichtenberg & Lieberman 1983). In suitable Poincaré coordinates, the phases saturate to

$$-\tilde{\boldsymbol{X}}[J(M-I)/(M+I)]\,\boldsymbol{X} \to -2\boldsymbol{Q}\cdot\boldsymbol{P} \quad \text{(long orbits)}, \quad (46)$$

which is independent of k.

(ix) Smoothing as embodied in ϵ introduces a semiclassical decay factor $\exp(-\epsilon T/\hbar)$ into (38) in addition to the classical decay (45) induced by instability. Large smoothing, that is $\epsilon \gg \Delta E_{\max}$ (defined before equation (7)) obliterates all the scars leaving the classical limit (8). As ϵ is reduced the scars proliferate, and through their superposition W begins to take on the complex oscillatory structure of the individual eigenstates. The superposition varies rapidly with E because of the changing phases $S_j(E)/\hbar$ of the individual contributions.

(x) If $\epsilon \ll \Delta E_{\min}$ (defined before equation (7)) one might hope fully to resolve individual eigenstates (equation (7)) as a sum over scars. But the convergence of this sum is doubtful, in view of the known failure of convergence of its phase-space integral (the unsmoothed closed-orbit sum for the spectral density) in the case of quantized geodesic motion on compact surfaces of constant negative curvature, for which the smoothed series (the Selberg trace formula) is exact for sufficiently large ϵ (Balazs & Voros 1986). The trouble comes from the long closed orbits, which proliferate exponentially with increasing period. These long orbits cover the energy surface densely and uniformly (Hannay & Ozorio de Almeida 1984), which suggests that their combined scars might somehow reproduce the first term (smoothed delta-function on the energy surface) of the semiclassical series (25), in an 'analytic bootstrap'. By analogy with the way in which this idea was implemented for the spectral density (Berry 1985), it might be fruitful to demand that the scar series satisfy the Weyl transform of the operator identity

$$\delta(E-\hat{H}) = \lim_{\epsilon\to 0} 2\pi\epsilon\delta_\epsilon^2(E-\hat{H}), \tag{47}$$

which follows from (2). The transform is

$$W(\boldsymbol{x};E,\epsilon=0) = \lim_{\epsilon\to 0}\left(\frac{2}{h}\right)^{2N}\int \mathrm{d}\boldsymbol{x}_1\int \mathrm{d}\boldsymbol{x}_2 \times W(\boldsymbol{x}_1;E,\epsilon)\, W(\boldsymbol{x}_2;E,\epsilon)\exp\{4\mathrm{i}A(\boldsymbol{x},\boldsymbol{x}_1,\boldsymbol{x}_2)/\hbar\}, \tag{48}$$

where A is the symplectic area of the triangle formed by $\boldsymbol{x}$, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, namely

$$A = \tfrac{1}{2}(\boldsymbol{x}_1-\boldsymbol{x})\wedge(\boldsymbol{x}_2-\boldsymbol{x}). \tag{49}$$

So far I have not succeeded with this approach.

I thank Professor A. M. Ozorio de Almeida and Dr J. H. Hannay for valuable discussions.

Appendix A

This is the derivation of (20). The $2N\times 2N$ matrix m can be written in terms of four $N\times N$ matrices:

$$m = \begin{bmatrix} m_{qq} & m_{pq} \\ m_{pq} & m_{pp}\end{bmatrix}, \quad \text{where } (m_{pq})_{mn} \equiv \partial p_{\mathrm{B}n}/\partial q_{\mathrm{A}m}, \text{ etc.} \tag{A 1}$$

To relate m to the four second derivatives of the action $R(\boldsymbol{q}_\mathrm{A},\boldsymbol{q}_\mathrm{B})$, we differentiate $\boldsymbol{q}_\mathrm{B}(\boldsymbol{q}_\mathrm{A},\boldsymbol{p}_\mathrm{A}(\boldsymbol{q}_\mathrm{A},\boldsymbol{q}_\mathrm{B}))$ and $\boldsymbol{p}_\mathrm{B}(\boldsymbol{q}_\mathrm{A},\boldsymbol{p}_\mathrm{A}(\boldsymbol{q}_\mathrm{A},\boldsymbol{q}_\mathrm{B}))$ with respect to $\boldsymbol{q}_\mathrm{A}$ and $\boldsymbol{q}_\mathrm{B}$. This leads to

$$\det(m+I) = \det\begin{bmatrix} -R_\mathrm{AB}^{-1}T_\mathrm{AA}+I & -R_\mathrm{AB}^{-1} \\ R_\mathrm{BA}-R_\mathrm{BB}R_\mathrm{AB}^{-1}R_\mathrm{AA} & -R_\mathrm{BB}R_\mathrm{AB}^{-1}+I\end{bmatrix}. \tag{A 2}$$

Now add $R_\mathrm{AB}-R_\mathrm{BB}$ times the first 'row' to the second 'row', to get

$$\det(m+I) = \det\begin{bmatrix} -R_\mathrm{AB}^{-1}R_\mathrm{AA}+I & -\mathrm{R}_\mathrm{AB}^{-1} \\ R_\mathrm{BA}+R_\mathrm{AB}-R_\mathrm{AA}-R_\mathrm{BB} & 0\end{bmatrix} = (-1)^N D^{-1}, \tag{A 3}$$

where D is defined by (19).

Appendix B

This is the derivation of (35). From (34) and the midpoint rule

$$X = \tfrac{1}{2}(X_\mathrm{A}+X_\mathrm{B}) \tag{B 1}$$

we find

$$X_\mathrm{A} = 2(I+M)^{-1}X; \quad X_\mathrm{B} = 2M(I+M)^{-1}X, \tag{B 2}$$

so that

$$X_\mathrm{A}\wedge X_\mathrm{B} \equiv \tilde{X}_\mathrm{A}JX_\mathrm{B} = \tilde{X}NX, \tag{B 3}$$

where

$$N = 4(I+\tilde{M})^{-1}JM(I+M)^{-1}. \tag{B 4}$$

We can replace N by its symmetrization;

$$N \to N_s = 2(I+\tilde{M})^{-1}JM(I+M)^{-1} - 2\tilde{M}(I+\tilde{M})^{-1}J(I+M)^{-1} \tag{B 5}$$

(by using $\tilde{J} = -J$). Now M is symplectic, that is (Lichtenberg & Lieberman 1983)

$$\tilde{M}JM = J, \tag{B 6}$$

so

$$\tilde{M} = -KM^{-1}J$$

and

$$(I+\tilde{M})^{-1} = -JM(I+M)^{-1}J. \tag{B 7}$$

Substitution into (B 5) gives

$$N_s = 2J\{M^2(I+M)^{-2} - (I+M)^{-2}\}$$

$$= 2J(M-I)/(M+I), \tag{B 8}$$

so that (B 3) becomes

$$X_A \wedge X_B = 2X[J(M-I)/(M+I)]X. \tag{B 9}$$

References

Abramowitz, M. & Stegun, I. A. 1964 *Handbook of mathematical functions*. Washington, D.C.: National Bureau of Standards.

Baker, G. A. Jr 1958 *Phys. Rev.* **109**, 2198–2206.

Balazs, N. L. 1980 *Physica* A **102**, 236–254.

Balazs, N. L. & Voros, A. 1986 *Physics Rep.* **143**, 109–240.

Balian, R. & Bloch, C. 1972 *Annls Phys.* **69**, 76–160.

Berry, M. V. 1977*a* *J. Phys.* A **10**, 2083–2091.

Berry, M. V. 1977*b* *Phil. Trans. R. Soc. Lond.* A **287**, 237–271.

Berry, M. V. 1983 Semiclassical mechanics of regular and irregular motion. In *Chaotic Behavior of Deterministic Systems, Les Houches Lectures* XXXVI (ed. G. Iooss, R. H. G. Helleman & R. Stora), pp. 171–271. Amsterdam: North-Holland.

Berry, M. V. 1985 *Proc. R. Soc. Lond.* A **400**, 229–251.

Berry, M. V. & Mount, K. E. 1972 *Rep. Prog. Phys.* **35**, 315–397.

Bogomolny, E. B. 1988 *Physica* D **31**, 169–189.

Groenewold, H. J. 1946 *Physica* **12**, 405–460.

Gutzwiller, M. C. 1971 *J. math. Phys.* **12**, 343–358.

Hannay, J. H. & Ozorio de Almeida, A. M. 1984 *J. Phys.* A **17**, 3429–3440.

Heller, E. J. 1984 *Phys. Rev. Lett.* **53**, 1515–1518.

Heller, E. J. 1986 In *Quantum chaos and statistical nuclear physics* (ed. T.-H. Seligman & H. Nishioka), pp. 162–181. Lecture notes in Physics 263. Berlin: Springer.

Kano, Y. 1965 *J. math. Phys.* **6**, 1913–1915.

Leboeuf, P., Saraceno, M., Baranger, M., Mahoney, J. & Provost, D. 1988 (In the press.)

Lichtenberg, A. J. & Lieberman, M. A. 1983 *Regular and stochastic motion*. New York: Springer.

McDonald, S. W. & Kaufman, A. N. 1988 *Phys. Rev.* A **37**, 3067–3086.

Moyal, J. E. 1949 *Proc. Camb. phil. Soc.* **45**, 99–124.

Ozorio de Almeida, A. M. 1988 *Hamiltonian systems: chaos and quantization*. Cambridge University Press.

Ozorio de Almeida, A. M. & Hannay, J. H. 1982 *Annls Phys.* **138**, 115–154.

Shapiro, M. & Goelman, G. 1984 *Phys. Rev. Lett.* **53**, 1714–1717.

Shnirelman, A. I. 1974 *Usp. mat. Nauk.* **29** (6), 181–182.

Synge, J. L. 1960 *Encyclopedia of physics* (ed. H. Flugge), vol. 3/1, pp. 1–225. Berlin: Springer-Verlag.

Takabayashi, T. 1954 *Proc. theor. Phys. Japan* **11**, 341–373.

Takahashi, K. 1986 *J. phys. Soc. Japan* **55**, 762–779.

Van Vleck, J. H. 1928 *Proc. natn. Acad. Sci. U.S.A.* **14**, 178–188.

Voros, A. 1979 In *Stochastic behaviour in classical and quantum hamiltonian systems* (ed. G. Casati & J. Ford), pp. 326–333. Lecture notes in Physics 93. Berlin: Springer.