Quantum chaos in systems with ray splitting

Luise Couchman

Acoustics Division, Naval Research Laboratory, Washington, D.C. 20375

Edward Ott and Thomas M. Antonsen, Jr.

*Laboratory for Plasma Research, Department of Electrical Engineering,
and Department of Physics, University of Maryland,
College Park, Maryland 20742*

We consider wave systems in which rays split on reflection from sharp boundaries. Examples include the Schrödinger equation with the potential changing discontinuously across a surface, electromagnetic waves in a region with a discontinuous dielectric constant, elastic media with a clamped or free boundary, etc. By introducing a Monte Carlo treatment of the rays, it is possible to define chaotic rays via the standard Lyapunov number criterion. Numerical implementation of the Monte Carlo ray technique is carried out for the example of elastic media, and is utilized to investigate the extent to which these systems are globally ergodic. It is suggested that results from previous extensive work on quantum chaos without ray splitting can be extended to these ray splitting problems. In particular, we indicate a generalization of the Gutzwiller trace formula to cover ray splitting.

PACS number(s): 05.45.+b

I. INTRODUCTION

Generally, ray equations result from the lowest order in a formal, asymptotic, short-wavelength expansion of a wave equation. In the case where the wave equation is Schrödinger's equation, the ray equations are just the equations of classical mechanics for particle motion. More broadly, the ray equations resulting from a wide class of wave equations (not restricted to quantum-mechanical wave equations) are Hamiltonian. Since ray equations are Hamiltonian, they are subject to the familiar generic phenomena of Hamiltonian systems, including periodic, quasiperiodic, and chaotic behavior. A natural question to ask is what are the consequences of chaotic ray trajectories for the solution of the corresponding wave equation in the semiclassical (i.e., short-wavelength) regime? The field of study addressing this question has been called *quantum chaos* [1,2]. Many interesting quantum-chaos results have been obtained, and, in particular, very general statements and techniques have been formulated concerning both eigenvalue spectra and wave functions.

The purpose of this paper is to consider extending these previous results on quantum chaos to a larger class of wave problems. In particular, we consider problems in which there are sharp boundaries at which an incident ray can split into two (or more) rays. Examples of the type of problem we are considering are shown in Figs. 1(a) and 1(b). The situation in Fig. 1(a) arises if we consider the Schrödinger equation with a constant potential $V(x) = V_1$ in region 1 and another constant potential $V(x) = V_2$ in region 2 with $V(x) = +\infty$ outside regions 1 and 2. Alternatively, one can, for example, also consider electromagnetic waves with regions 1 and 2 occupied by homogeneous media of different dielectric constants and with the outer boundary being a perfectly conducting

surface. As shown in the figure, an incident ray splits, launching both a transmitted ray and a specularly reflected ray. Figure 1(b) shows another possible example, namely waves propagating in an elastic medium which is stress free or clamped at the boundary. In this case the medium supports two wave types, transverse shear waves (denoted S in the figure) and longitudinal compressional waves (denoted P in the figure), and these couple at the boundary. (We assume the S-wave polarization is in the plane of the figure). (The situations shown in Fig. 1 may be viewed as generalizations of the (non-ray-splitting) classical billiard; i.e., the problem of a particle following a planar trajectory of straight line rays
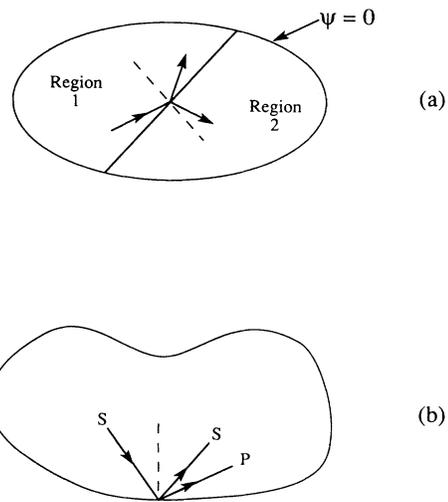
FIG. 1. Two examples of systems which exhibit ray splitting: (a) adjoining bounded regions, each supporting one wave type, or (b) a bounded elastic region.

bouncing specularly off a closed bounding curve (or curves). See, for example, Refs. [3–5]).

The phenomena of ray splitting occurs in these problems due to the fact that there are regions of the medium where its properties change over a distance which is shorter than a wavelength. In particular, the medium changes abruptly at a discontinuity or a boundary. In these regions the short-wavelength expansion of the wave equation breaks down, and the ray may be scattered in a number of directions. Generally speaking, there will be some amount of ray splitting whenever the medium is inhomogeneous. If we consider the problem posed in Fig. 1(a) and imagine that there is a transition layer of thickness Δ separating region 1 from region 2, then the relative amplitudes of the transmitted and scattered rays will depend on the thickness of the transition layer. If the layer is much thicker than a wavelength λ, then almost all of the wave density will be transmitted from region 1 to region 2. This is the usual short-wavelength limit. The amount of reflected wave density in this case is exponentially small in the factor Δ/λ. If the transition layer is comparable to, or smaller than, a wavelength (the case considered in this paper), then the reflected and transmitted wave densities will be comparable. Thus, by studying ray splitting, we examine a phenomenon which is present when the conditions for the usual short-wavelength approximation are not satisfied.

In classical Hamiltonian (without ray splitting) problems there are cases where the rays are chaotic and ergodic over the entire energy surface (e.g., the Sinai billiard and the stadium billiard). There are also cases where there is complete integrability, and there are cases where there is a mixture of chaos and integrability. The latter are very common. Furthermore, analysis of the semiclassical limit of cases with a mixture of integrable and chaotic regions has proven to be much more problematical and tentative than analysis of the two extreme cases of completely integrable or completely chaotic systems. In this regard, intuition might suggest that ray splitting would lead ergodicity to be much more prevalent. [Indeed, the usual concept of Kolmogorov-Arnold-Moser (KAM) tori is absent for orbits that experience ray splitting.] Hence, one might suppose that extensions of results of completely chaotic systems without ray splitting might apply to a much broader range of geometrical configurations with ray splitting. This is one of the prime motivations of our work. Our results generally support the above intuition, but we also find some effects that mitigate ergodicity.

We believe that the much greater tendency for ergodicity and positive Lyapunov exponents for cases with ray splitting implies, for example, that an approximate Wigner distribution of energy-level spacings (e.g., see Refs. [1] and [2]) should apply in a much wider variety of geometrical configurations than would be the case without ray splitting.

In Sec. II we briefly review material on the classical orbits in a singly connected billiard domain. Section III discusses the elastic medium billiard [Fig. 1(b)], which is then used as a numerical example in Sec. III. In particular, in Sec. III we introduce a Monte Carlo treatment of

the ray-splitting problem as a means of defining chaos and examining the tendency for ray ergodicity in such problems. It is found that there are two sources inhibiting ergodicity in the elastic medium billiard problem: (i) critical reflection, and (ii) slow mode conversion for periodic orbits that make only nearly normal reflections. In Sec. IV we discuss situations other than the elastic medium problem (i.e., acoustic and electromagnetic waves) and argue that phenomena analogous to (ii) do not occur in these cases. In Sec. V we generalize the Gutzwiller [6] trace formula to problems with ray splitting. As in the previous work we find that closed orbits are the key. With ray splitting, however, one has to adopt a slightly expanded concept of what we mean by a closed orbit. Section VI presents a concluding discussion.

II. BRIEF OVERVIEW
OF THE BILLIARD PROBLEM

We begin with a brief overview of the classical "billiard" problem. As indicated previously, this is the problem of a ray or point particle bouncing about inside a closed curve which lies in a plane. The bounce, or reflection, is energy conserving, and governed by the law, "angle of incidence equals angle of reflection." (See Fig. 2.).

Since the particle travels in a straight line between reflections, its path, or "orbit," is entirely characterized by specifying the positions and incident angles of the reflections (Fig. 2). An orbit is thus specified by the sequence

$$\begin{bmatrix} \sigma_0 \\ \tau_0 \end{bmatrix}, \begin{bmatrix} \sigma_1 \\ \tau_1 \end{bmatrix}, \begin{bmatrix} \sigma_2 \\ \tau_2 \end{bmatrix}, \dots, \begin{bmatrix} \sigma_n \\ \tau_n \end{bmatrix}. \tag{1}$$

where σ_n is the arclength on the boundary (assumed to be a single connected curve) to the nth bounce point, measured from some fixed reference position, and τ_n is the cosine of the nth bounce angle. (Again see Fig. 2.) We take σ to be normalized such that $\sigma = 1$ corresponds to the full length of the bounding curve. The coordinates of successive bounces are related by the "bounce mapping" M such that

$$\begin{bmatrix} \sigma_{n+1} \\ \tau_{n+1} \end{bmatrix} = M \begin{bmatrix} \sigma_n \\ \tau_n \end{bmatrix}, \tag{2}$$
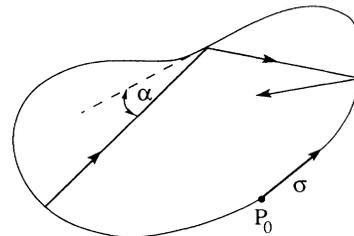
FIG. 2. The classical billiard problem: a ray or point particle bounces about inside a closed curve which lies in a plane. We specify the position of a particular reflection in terms of arclength σ from some point P_0 normalized so that $\sigma = 1$ for a complete circuit, and the angle α of the ray with respect to the tangent to the bounding curve (the "bounce angle").

where M is a nonlinear mapping which may be shown to be area preserving [3–5].

For certain shapes of the bonding curve, the equations of motion of the system are completely *integrable*. As a result there exists a global analytic constant of the motion $F(\sigma, \tau)$ which is conserved along any orbit of the map. Consequently, the successive bounces are restricted to lie on a curve in the (σ, τ) surface of section. For an integrable system, every point in the phase space lies on such an invariant curve. When the system in *nonintegrable*, no such global constant of the motion exists and the successive bounces chaotically fill an area in the (σ, τ) space. For *pseudointegrable* systems, constants of the motion exist for some initial conditions, but not for others. For such systems, the phase space consists of a mixture of invariant curves and chaotic areas.

As an example, the circular billiard (that is, the billiard which has a circle as its boundary curve) is an integrable system. Its (σ, τ) phase space consists entirely of closed orbits and invariant curves characterized by $\tau =$const. (In this case the constant of the motion is angular momentum.) On the other hand, the phase space of the so-called "stadium" billiard, consisting of two equal-radius semicircles connected by straight lines, is entirely chaotic. The entire phase space of the stadium billiard is densely covered by an orbit originating from a typical initial condition.

A class of billiards which varies smoothly from circle to stadium was studied by Benettin and Strelcyn [4]. This oval is illustrated in Fig. 3. It is composed of four circular arcs meeting with continuous slope on the vertices of a square of sidelength two. Opposing arcs have the same radius of curvature. If δ is the offset of the center of curvature of one of the arcs from the square side wall, then the two radii of curvature are related by

$$R_1/R_2 = \delta .$$

For $\delta = 0$, the oval becomes a stadium; for $\delta = 1$, a circle. For this oval with $1 > \delta > 0$, Benettin and Strelcyn observed coexistence of invariant curves and chaotic regions as in other examples of chaotic transition. Figures 4(a) and 4(b) are examples of the phase space of the Benettin-Strelcyn oval, at $\delta = 0.76$, showing a chaotic region [Fig. 4(a)] and some invariant curves resulting from seven (numbered) different integrable orbits [Fig. 4(b)].

If the boundary curve of the billiard has no inflections, it can be parameterized by the angle Ψ of the counterclockwise pointing tangent, measured with respect to a fixed axis (see Fig. 5) and the radius of curvature of the boundary curve at Ψ denoted $R(\Psi)$. Then

$$R(\Psi) = \frac{d\sigma}{d\Psi} . \tag{3}$$

That is,

FIG. 3. The Benettin-Strelcyn oval. P_1, P_2, P_3, and P_4 are points on the vertices of a square of side two. Γ_1, Γ_2, Γ_3, and Γ_4 are circular arcs with common tangents at P_1, P_2, P_3, and P_4. The length δ is measured from the center of Γ_1 to the side $P_1 P_2$. For $\delta = 1$, the oval is a circle; for $\delta = 0$, a stadium.

$$\sigma(\Psi) = \int_{\pi/2}^{\Psi} d\Psi' R(\Psi') . \tag{4}$$

In terms of Ψ and α (the bounce angle defined in Fig. 2) the mapping equations can be written as

$$\left[\int_{\Psi_n}^{\Psi_{n+1}} R(\Psi)\sin\Psi \, d\Psi \right] \left[\int_{\Psi_n}^{\Psi_{n+1}} R(\Psi)\cos\Psi \, d\Psi \right]^{-1}$$
$$= \tan(\Psi_n + \alpha_n) \tag{5}$$

and

$$\Psi_{n+1} - \alpha_{n+1} = \Psi_n + \alpha_n . \tag{6}$$

Thus we obtain an implicit map of (Ψ_n, α_n) to $(\Psi_{n+1}, \alpha_{n+1})$ in terms of the single geometrical quantity $R(\Psi)$ in Eq. (3). Note that Eq. (5), in general, has to be solved numerically for Ψ_{n+1}. Having found Ψ_{n+1} and α_{n+1}, we can determine σ_{n+1} [from Eq. (4)] and $\tau_{n+1} = \cos(\alpha_{n+1})$. Due to the choice of arclength σ and cosine of the bounce angle τ as variables, the mapping is area preserving; i.e., the Jacobian matrix

$$m = \begin{vmatrix} \dfrac{\partial \sigma_{n+1}}{\partial \sigma_n} & \dfrac{\partial \sigma_{n+1}}{\partial \tau_n} \\ \dfrac{\partial \tau_{n+1}}{\partial \sigma_n} & \dfrac{\partial \tau_{n+1}}{\partial \tau_n} \end{vmatrix}$$

has determinant one,

$$\det(m) = 1 .$$

In the notation of Eqs. (2)–(6), the Jacobian matrix m of the mapping is given by [5]
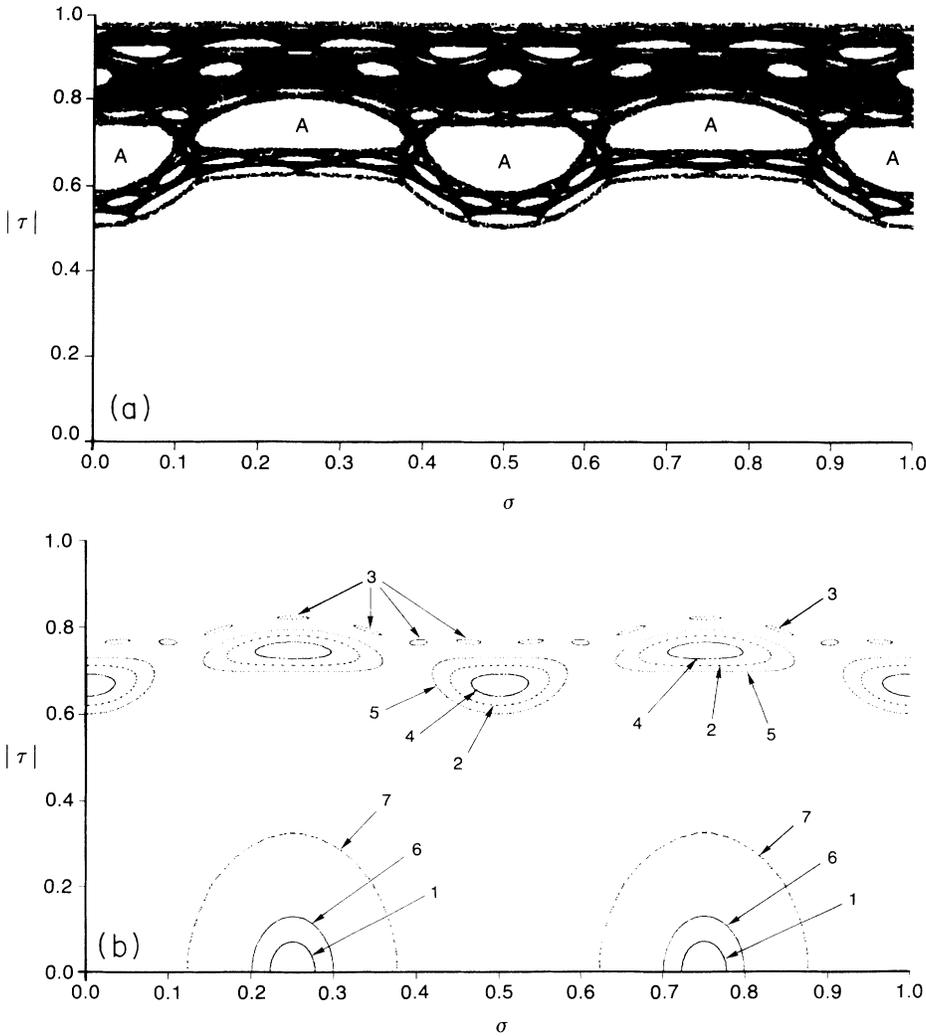
$$\begin{vmatrix} \dfrac{-s_n}{s_{n+1}} + \dfrac{\rho_{n,n+1}}{s_{n+1}R(\Psi_n)} & \dfrac{\rho_{n,n+1}}{s_n s_{n+1}} \\ -\dfrac{\rho_{n,n+1}}{R(\Psi_n)R(\Psi_{n+1})} + \dfrac{s_{n+1}}{R(\Psi_n)} + \dfrac{s_n}{R(\Psi_{n+1})} & \dfrac{s_{n+1}}{s_n} + \dfrac{\rho_{n,n+1}}{s_n R(\Psi_{n+1})} \end{vmatrix} , \tag{7}$$

FIG. 4. (a) An example of the phase space of the Benettin-Strelcyn oval at $\delta = 0.76$: One of the chaotic regions. (b) An example of the phase space of the Benettin-Strelcyn oval at $\delta = 0.76$: Seven of the invariant curves are labeled with the numerals 1–7.

where $\rho_{n,n+1}$ is the direct distance (not arclength) from the nth to the $(n+1)$th bounce, and $s_n = \sin(\alpha_n)$.

III. THE ELASTIC MEDIA PROBLEM AS A "BILLIARD"

A. The governing equations

In order to study the behavior of coupled-wave systems, we consider as an example a bounded two-dimensional elastic region. This problem is analogous to the conventional billiard problem in that it can be treated as a ray tracing problem in the classical (i.e., short-wavelength) limit; however, the medium supports two distinct natural waves, which are coupled at the boundaries. The following is a brief summary of the equations governing the elastic medium problem.

Infinitesimal displacements in the elastic region are governed by the Navier equation,

$$\mu \nabla^2 \mathbf{u} + (\lambda + \mu) \nabla \nabla \cdot \mathbf{u} = \rho \ddot{\mathbf{u}} , \tag{8}$$

where $\mathbf{u}$ is the infinitesimal displacement of the volume element, the dot represents differentiation with respect to

time, λ and μ are the Lamé constants of the elastic material, and ρ is its mass density. If the displacement is expressed as the sum of two parts, due, respectively, to a scalar and a vector potential,

$$\mathbf{u}_p = \nabla \phi , \tag{9a}$$

$$\mathbf{u}_s = \nabla \times \mathbf{\Psi} , \tag{9b}$$
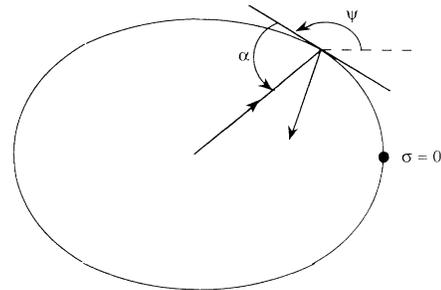
FIG. 5. A convenient coordinate system for computer iteration of the classical billiard problem for a bounding curve: ψ is the angle of the counterclockwise pointing tangent, and α is the bounce angle.

the differential equation [Eq. (8)] separates into two second-order equations, viz.,

$$\nabla^2\Phi = \frac{1}{c_p^2}\ddot{\phi} , \qquad (10a)$$

$$\nabla^2\Psi = \frac{1}{c_s^2}\ddot{\Psi} , \qquad (10b)$$

where $c_p^2 = (\gamma + 2\mu)/\rho$ and $c_s^2 = \mu/\rho$. Let κ be the ratio of the wave speeds

$$\kappa = c_p/c_s = [(\lambda + 2\mu)/\mu]^{1/2} . \qquad (11)$$

The elastic medium thus has two distinct natural waves, sometimes known as P [Eq. (10a)] and S [Eq. (10b)] waves, which are uncoupled except at the boundaries of the medium. The P waves are compressional waves such that planar P waves have displacements parallel to the wave vector [Eq. (9a)]. The S waves are shear waves such that planar S waves have displacements perpendicular to the wave vector [Eq. (9b)]. (Thus S waves possess polarization properties.)

The solution of Eq. (8) is a superposition of S and P waves, which are coupled by the boundaries of the elastic medium. If there are no forces on the boundary, the boundary condition becomes

$$\mathbf{t} = 0 \quad \text{on } \Omega , \qquad (12)$$

where Ω is the bounding surface, and $\mathbf{t}$ is the traction on Ω, given by

$$\mathbf{t} = \hat{\mathbf{n}} \cdot \mathbf{T} . \qquad (13)$$

Here $\hat{\mathbf{n}}$ is a unit vector normal to Ω and $\mathbf{T}$ is the stress tensor

$$\mathbf{T} = \lambda \mathbf{I}(\nabla \cdot \mathbf{u}) + \mu(\nabla \mathbf{u}^T + \nabla \mathbf{u}) , \qquad (14)$$

where $\mathbf{I}$ is the identity tensor.

B. The classical (zero-wavelength) limit

In the small-wavelength limit the problem can be treated as a ray tracing problem. If the curvature of the boundary can be neglected relative to the wavelength, reflection can be treated locally as reflection at a plane boundary. Upon reflection, a ray is split into an S and a P ray.
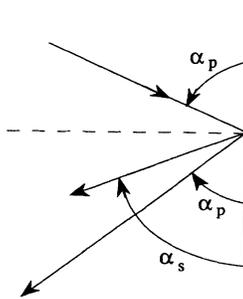
FIG. 6. Reflection of a plane wave propagating in an elastic medium at a boundary. An incident P wave is split into an S wave and a P wave.

For reflection of a plane wave propagating in an elastic medium from a planar stress-free boundary (Fig. 6), one finds that the coupling of S and P waves depends upon the polarization of the S wave. In particular, S waves polarized such that the displacements lie in the plane defined by the wave vector and its projection on the boundary (sometimes designated SV waves) are coupled to P waves. In contrast, S waves of the other polarization (i.e., displacements normal to the plane of Fig. 6) are not coupled to P waves. (These waves are commonly designated SH waves.) For the SV and P waves the following results obtained from Eqs. (13) and (14) apply:

(a) The angle of propagation with respect to a tangent to the boundary for a mode-converted reflected wave (either P to S to S to P) is related to that for the incident wave by

$$\cos(\alpha_P) = \kappa \cos(\alpha_S) , \qquad (15)$$

where α_P and α_S are the ray angles for the P and S waves, respectively. (Note that there is a critical angle given by $\alpha_S = \arccos(\kappa^{-1})$ below which no S to P conversion occurs.)

(b) Defining power flux reflection coefficients as the ratio of the normal component of energy flux in the reflected wave to that in the incident wave, the reflection coefficients for P wave incidence are given by

$$A_{PP} = \left| \frac{\sin 2\alpha_P \sin 2\alpha_S - \kappa^2 \cos^2 2\alpha_S}{\sin 2\alpha_P \sin 2\alpha_S + \kappa^2 \cos^2 2\alpha_S} \right|^2 , \qquad (16)$$

$$A_{PS} = 1 - A_{PP} , \qquad (17)$$

where A_{PP} and A_{PS} represent the power flux reflection coefficients for the case of a reflected P and S wave, respectively, given P wave incidence.

(c) The power flux reflection coefficients for S wave incidence are given by

$$A_{SS} = \left| \frac{\sin 2\alpha_S \sin 2\alpha_P - \kappa^2 \cos^2 2\alpha_S}{\sin 2\alpha_S \sin 2\alpha_P + \kappa^2 \cos^2 2\alpha_S} \right|^2 , \qquad (18)$$

$$A_{SP} = 1 - A_{SS} , \qquad (19)$$

where A_{SP} and A_{SS} represent the power flux reflection coefficients for the case of a reflected P and S wave, respectively, given S wave incidence. Equations (17) and (19) are statements of energy conservation.

IV. NUMERICAL RESULTS IN THE CLASSICAL LIMIT

In this section, we investigate numerically the properties of solutions of the two-wave elastic-medium billiard with displacements in the plane of the billiard [Fig. 1(b)]. To obtain these results we employ the following algorithm:

We follow a ray traveling in a plane in the elastic region through many reflections. Ray splitting is treated by a Monte Carlo approach. That is, at each reflection only one ray is assumed to be emitted, and that ray is taken to be P or S with relative probability given by the appropriate term in the energy equations, Eqs. (16)–(19)

(i.e., either A_{PP}, A_{PS}, A_{SS}, or A_{SP}). If the reflected and incident waves are of the same type, then the angles of incidence and reflection are equal. If not, then the angles are related by Eq. (15).

In principle, one should allow for the generation of two reflected rays at each encounter of the boundary. Further, one should associate with each "daughter" ray an amplitude which decreases on reflection by the appropriate factor given by Eqs. (16)–(19). This procedure, however, would quickly lead to an unmanageably large number of "daughter" rays. We have, instead, followed the above described Monte Carlo method where only one ray is traced. We argue that after a large number of bounces the fraction of time the single Monte Carlo ray spreads in a given region of phase space in a given wave type approaches the value that would be obtained by the aggregate of all "daughter" rays had they been followed. The integrated phase-space density determined from the single ray might then be argued to be directly related to the energy density of a short-wavelength eigenfunction of the time harmonic problem.

Due to relation (15), the bounce mapping for the case with ray splitting contains an additional operator T such that the N-bounce mapping has the form

$$(MT)^N \equiv M_1 T_1 M_2 T_2 M_3 T_3 \cdots M_N T_N .$$ (20)

If there is an S to P transition at the nth reflection, T_n has the form

$$T_n(S \rightarrow P) = \begin{vmatrix} 1 & 0 \\ 0 & \kappa \end{vmatrix} .$$ (21a)

For a P to S transition,

$$T_n(P \rightarrow S) = \begin{vmatrix} 1 & 0 \\ 0 & 1/\kappa \end{vmatrix} .$$ (21b)

For no transition,

$$T_n(\text{no transition}) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} .$$ (21c)

The probability that T_n has a particular form depends on the bounce angle α_n via the power flux reflection coefficients A_{PP}, A_{PS}, A_{SS}, and A_{SP}. The resulting numerical model reduces to the conventional billiard problem if ray splitting at the boundaries is turned off.

One might note that, given an odd number of transitions, the N-bounce mapping with ray splitting is not area preserving. As we shall see, however, we are only interested in cases for which an orbit ends in the same (either S or P) wave type as that in which it began. For this case $(MT)^N$ is area preserving.

In addition to tracing the ray, in each case we calculated the Lyapunov numbers of the orbit. These were approximated by calculating the Nth root of the magnitude of the eigenvalues of the Jacobian of the N-bounce mapping for large N. The Jacobian of the N-bounce mapping is calculated by taking the products of the Jacobians of the individual bounce mappings. If the Monte Carlo ray orbit is ergodic over an area and has a Lyapunov number greater than one, then we call the ray orbit chaotic.

A. A two-wave Benettin-Strelcyn oval

As one numerical example, we consider a two-dimensional elastic region analogous to the generalized stadium of Benettin and Strelcyn [4], which was described in Sec. II. Numerical experiments were performed for this geometry with and without ray splitting. The Monte Carlo approach described previously was taken to model the ray splitting. As described in Sec. III, the reflection at a plane stress-free boundary depends on the ratio of wave speeds $\kappa = c_P/c_S$. We choose for our
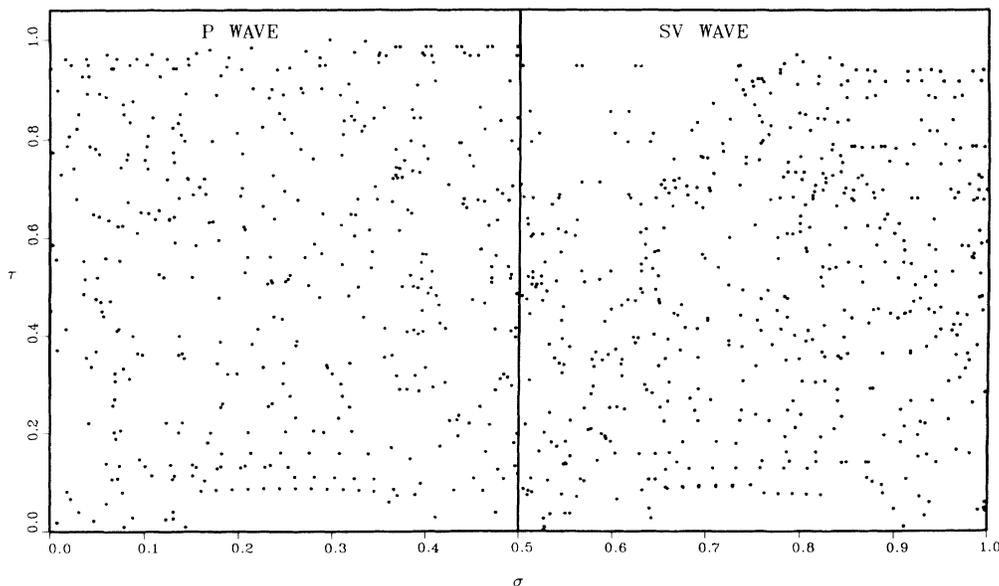
FIG. 7. The phase space for the stadium with ray splitting. The S-wave and the P-wave phase space are shown side by side. Since the results are symmetric about $\sigma = 0.5$ only half of each space is plotted.

TABLE I. Number of iterates in cells of the P-wave phase space of the stadium.

τ \ σ	0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–0.5	0.5–0.6	0.6–0.7	0.7–0.8	0.8–0.9	0.9–1.0
0.8–1.0	4313	4171	4146	4097	3995	4051	4108	4146	4161	4098
0.6–0.8	4437	4135	4164	4182	4145	4216	4216	4005	4150	4179
0.4–0.6	4332	4025	4233	4170	4104	4197	4049	4060	4017	4031
0.2–0.4	4277	4161	4049	3974	4264	4106	4126	4098	4151	3955
0–0.2	4315	4147	4025	4059	4136	4161	4071	4031	4215	4013
(−0.2)–0	4071	4156	4013	4056	4062	4110	4123	4003	4146	4193
(−0.4)–(−0.2)	4231	4196	4180	4225	4062	4125	4102	4042	4091	4188
(−0.6)–(−0.4)	4126	4196	4038	4162	4132	4213	4224	4005	4157	4147
(−0.8)–(−0.6)	4243	4169	4145	4129	4196	4073	4173	4095	4205	4079
(−1.0)–(−0.8)	4108	4184	4183	4086	4145	4036	4081	4148	4204	4160

numerical studies the value

$$\kappa = 1.429 .$$

This value gives a critical angle for incident S waves of

$$\alpha_{Sc} \simeq \pi/4 ,$$

which results in comparable ranges of angle where an incident S wave is totally reflected and where an incident S wave experiences ray splitting. In the plots and tables that follow, the phase space is parameterized in terms of σ, the normalized arclength, and τ, the cosine of the bounce angle, as defined in Fig. 2.

For $\delta=0$, the Benettin-Strelcyn oval is a stadium. Figure 7 shows the phase space for the stadium with ray splitting. 1000 iterates of a single orbit are plotted. (We have plotted only points with values $\sigma \leq 0.5$, since the results are symmetric about $\sigma=0.5$.) Reflections which resulted in a P wave are plotted on the left; those which resulted in an S wave are plotted on the right. As in the case without ray splitting, this oval is filled by a single chaotic orbit. The distribution of the iterates appears to be uniform.

In order to quantify the apparent ergodicity, the phase space was divided by a 10×10 grid into 100 regions. For an orbit with 10^6 bounces, the number of iterates falling into each rectangle of the grid is tabulated in Tables I and

II. The density of iterates is approximately uniform throughout each of the S and P spaces. The standard deviation of the number of iterates per cell was 82.9 and 110.1, for the P and S phase spaces, respectively, with an average number of points per cell of 4133 and 5867, respectively.

These standard deviations are somewhat larger than the values $(4133)^{1/2}=64$ for P space and $(5867)^{1/2}=76$ for S space, which one would expect if the distribution were due to an uncorrelated Gaussian random process. These larger values of the standard deviation can be attributed to long-term correlations in certain portions of the orbit. In particular, it has been shown for a similar geometry which also contains opposing parallel sides [7] that the long-term correlation function of the orbit falls off like $1/T$, rather than exponentially. These long-term correlations are due to the existence of a family of neutrally stable orbits (the two-bounce orbits between opposing straight sides).

The numerical ratio of the number of points in S to P space is $5867 \div 4133 = 1.420$. The theoretically expected average ratio is given by $\kappa = c_P/c_S = 1.429$. This is shown in the Appendix.

As δ increases, the phase space of the Benettin-Strelcyn oval without ray splitting shows increasing numbers of invariant curves. For $\delta < \delta_c \simeq 0.76$ there is a single con-

TABLE II. Number of iterates in cells of the S-wave phase space of the stadium.

τ \ σ	0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–0.5	0.5–0.6	0.6–0.7	0.7–0.8	0.8–0.9	0.9–1.0
0.8–1.0	5766	5832	6079	5878	5919	5855	5838	6011	5907	5749
0.6–0.8	5934	5989	6059	5818	5848	5805	5875	5957	5954	6032
0.4–0.6	5929	5930	5889	5925	5864	5903	5900	6022	6005	5957
0.2–0.4	5735	5717	5780	5831	5814	5874	5897	5977	5847	5946
0–0.2	5846	5679	5625	5749	5919	5936	5617	5646	5637	5972
(−0.2)–0	6021	5643	5608	5841	5980	6034	5766	5562	5737	5794
(−0.4)–(−0.2)	5865	5913	5789	5904	5899	5905	5933	5794	5953	5748
(−0.6)–(−0.4)	5944	5859	5904	5806	5883	5921	5883	5932	5967	5963
(−0.8)–(−0.6)	6017	5837	5904	5923	5857	5912	5935	5955	5975	5901
(−1.0)–(−0.8)	5819	5931	5971	5717	5676	5700	5713	5890	5945	5812

FIG. 8. Two-bounce periodic orbit in the Bennettin-Strelcyn oval.
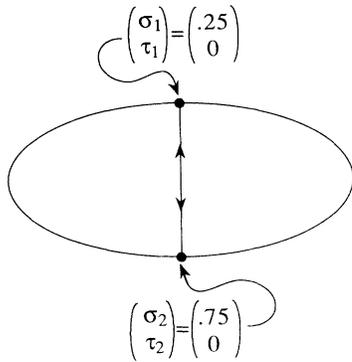
FIG. 9. Four-bounce periodic orbit in the Benettin-Strelcyn oval.

nected ergodic chaotic component which connects the line $\tau=0$ and the line $\tau=1$. As δ is increased to δ_c a KAM surface appears which runs horizontally from $\sigma=0$ to 1, thus dividing the large connected, ergodic, chaotic component mentioned above into two components. Figure 4(a) shows one of these components (the one which includes the line $\tau=1$) for $\delta=0.76$. [Since only 10^5 iterates were used, the orbit plotted in Fig. 4(a) did not reach the largest τ values which are part of the same chaotic region.] Regions encircled by KAM curves are visible as holes in the chaotic regions.

Figure 4(b) shows some of the invariant tori of this billiard with numbers labeling the seven invariant tori, again for the case where ray splitting is absent. The tori in the lower part of the figure, $|\tau|<0.5$, are part of a family of invariant curves which surround the stable two-bounce orbit defined by

$$\begin{bmatrix} \sigma_1 \\ \tau_1 \end{bmatrix} = \begin{bmatrix} 0.25 \\ 0 \end{bmatrix} ; \quad \begin{bmatrix} \sigma_2 \\ \tau_2 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0 \end{bmatrix}$$

(see Fig. 8). These orbits acquire special significance in the case when ray splitting is allowed, as we shall see.

Other families of invariant curves are visible as holes in the chaotic region in Fig. 4(a). For example, the regions marked A correspond to invariant curves surrounding the stable four-bounce orbit shown in Fig. 9.

Figure 10 shows the S and P phase spaces for the same oval ($\delta=0.76$) when ray splitting is allowed at the boundary. The phase spaces of the S and P waves are displayed side by side. The ray splitting has destroyed most of the invariant tori. An initial condition which lies on an invariant curve of the billiard without ray splitting converts energy into other paths and becomes ergodic.

However, two remnants of the KAM surfaces remain. First, ray splitting does not affect those invariant curves in the S space which fall entirely in the region for which $\tau>\tau_c\equiv\cos(\alpha_c)$, where α_c is the critical angle below which no S to P conversion occurs. This is demonstrated in Fig. 11. Refer to Fig. 4(b), which shows phase-space plots of seven orbits (each with 500 reflections) which are integrable in the absence of ray splitting. The invariant curves are marked with orbit numbers on the plot. Figures 11(a) and 11(b) show the phase-space plots of orbits using the same initial conditions as for orbits 1 and 2 of Fig. 4(b) when ray splitting is allowed. 3000 reflections and 1000 reflections, respectively, are plotted. For these plots the cosine of the critical angle is $\tau_c=0.6998$. Orbits 1 and 2 include values of $\tau<\tau_c$ and become ergodic over the same area when ray splitting is present. On the other hand, when the ray splitting calculation is done for orbit 3, which lies entirely above $\tau=\tau_c$, the non-ray-splitting result of Fig. 4(b) is reproduced. This effect is also demonstrated by the calculated Lyapunov numbers.
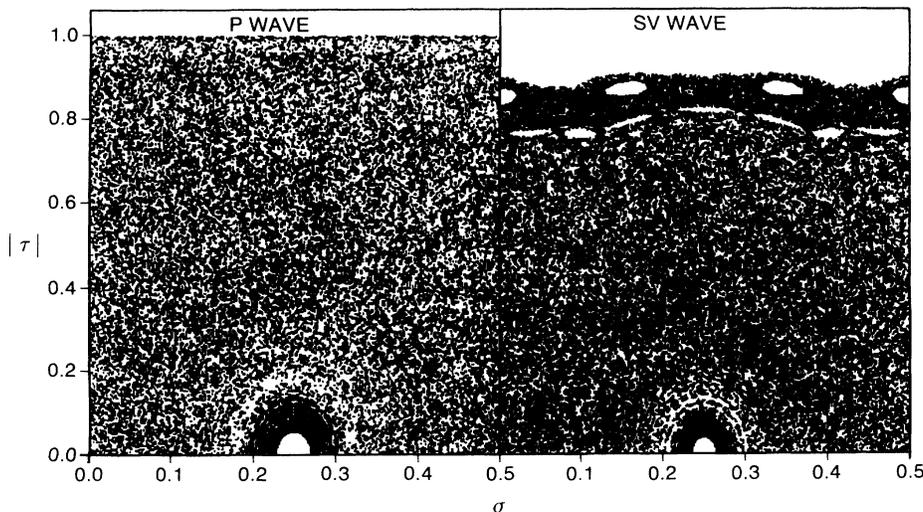
FIG. 10. The S and P phase spaces of the Benettin-Strelcyn oval with ray splitting at $\delta=0.76$.
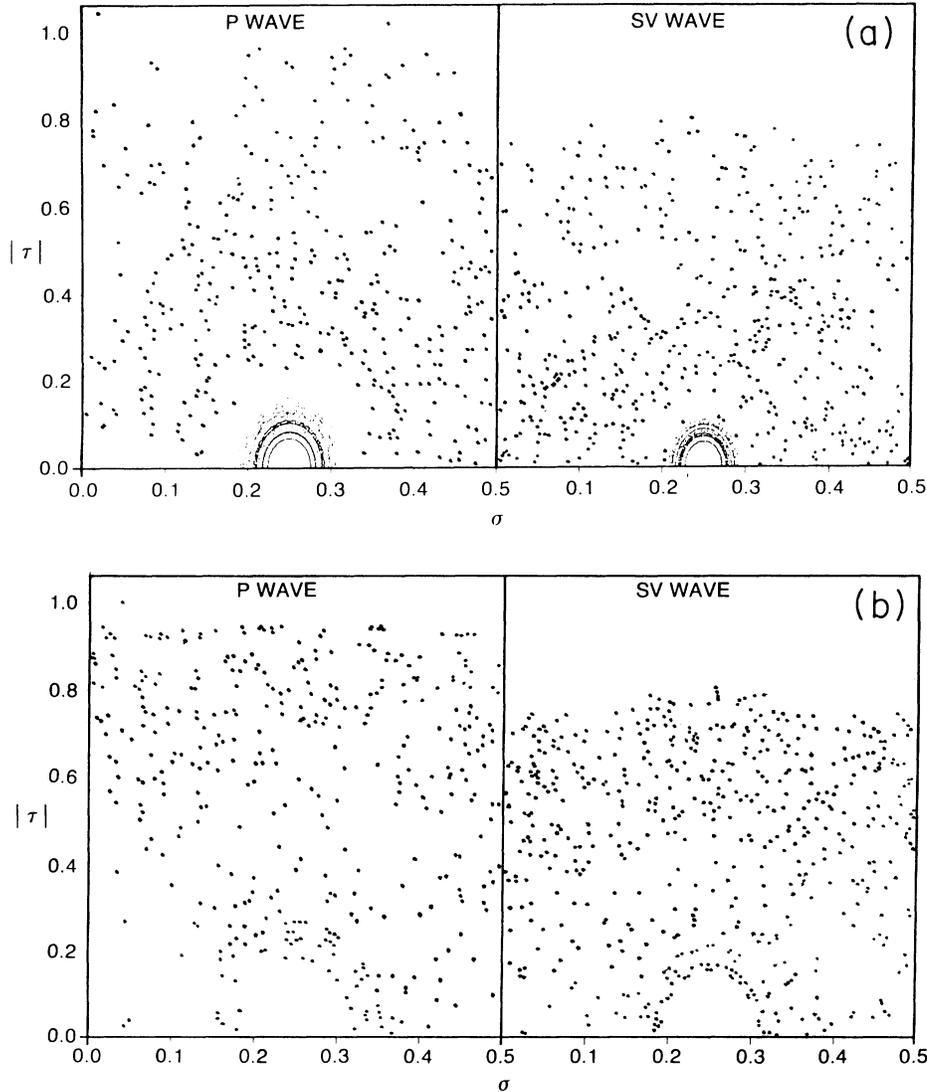
FIG. 11. (a) Orbit resulting from the same initial conditions as orbit 1 of Fig. 4(b) when ray splitting is present. The dot size is smaller in the region near the straight-back-and-forth orbit ($\tau=0$, $\sigma=0.25$). (b) Orbit resulting from the same initial conditions as orbit 2 of Fig. 4(b) when ray splitting is present.

Table III tabulates the Lyapunov numbers of the seven orbits plotted in Fig. 4(b). 3000 reflections were considered in each case. Results are shown both with and without ray splitting for four different values of $\tau_c = \kappa^{-1}$. For the smaller τ_c values more orbits lie entirely above τ_c, and thus remain integrable in the presence of ray splitting. The change from integrability to chaos due to the ray splitting is demonstrated by the increase in the Lyapunov numbers. All the orbits in the table which include values of $\tau < \tau_c$ become chaotic in the presence of ray splitting. All the tabulated orbits which lie entirely in the region $\tau > \tau_c$ remain integrable despite ray splitting.

A second remnant of the KAM surfaces lies in the small-τ regions of the phase space and is clearly seen in Fig. 11(a). That part of phase space, which contains the invariant curves surrounding the stable two-bounce orbits for the case without ray splitting, has become part of

the chaotic region. However, as our analysis below will show, escape from this region is extremely slow (in particular, it is much slower than exponential). This pseudostability may be understood as follows: Snell's law for ray splitting is given by Eq. (15),

$$\cos(\alpha_S) = \frac{c_S}{c_P}\cos(\alpha_P) \ .$$

Thus, near the two-bounce orbits, where α is close to $\pi/2$, the reflection angle change due to ray splitting is small. Furthermore, the probability of transition goes to zero as α approaches $\pi/2$. [See Eqs. (16) and (18).] Thus the orbit tends to remain on an invariant curve of the problem without ray splitting for many bounces before making a transition. Following a transition, the angle α will still be small [cf. Eq. (15)], and hence the orbit again moves on an invariant curve of the one-wave system for many bounces.

TABLE III. Lyapunov numbers with and without ray splitting. The asterisk indicates that this initial condition lies in the same large ergodic component as orbits 2, 4, 5, 6, and 7, and should thus have the same Lyapunov number as those orbits. The larger deviation of this value from that of the other chaotic orbits is due to the fact that this orbit started in the pseudostable region and spent over half of the 1000 iterates used for calculating the Lyapunov numbers in the pseudostable region. When the same orbit is continued for 20 000 bounces, the Lyapunov number is calculated to be 1.20.

Orbit	No ray splitting	Ray splitting			
		$\tau_c = 0.70$	$\tau_c = 0.64$	$\tau_c = 0.61$	$\tau_c = 0.54$
1	1.00	1.08*	1.25	1.21	1.32
2	1.00	1.21	1.31	1.00	1.00
3	1.00	1.00	1.00	1.00	1.00
4	1.00	1.25	1.25	1.00	1.00
5	1.00	1.22	1.29	1.30	1.00
6	1.00	1.21	1.19	1.23	1.19
7	1.00	1.24	1.24	1.29	1.28

B. Analysis of pseudostable orbits

The pseudostable orbits near the stable two-bounce orbits can be understood in terms of the focusing and defocusing effect of opposing curved reflectors. If we consider a ray bouncing back and forth in a plane between opposing mirrors which are spherical segments, allowing wave-type conversion at the mirrors, we can determine the rate at which the ray diverges from near-normal incidence. In what follows, we show that this divergence is much slower than for the case of an exponentially unstable periodic orbit.

Consider a pair of opposing curved mirrors which are spherical segments. (See Fig. 12.) For simplicity, let the radii of curvature of both mirrors have the same value R_c. Let the centers of curvature lie on the same axis. Let the distance between the mirrors on the axis be 1, and assume for convenience in describing the system that the axis lies on the horizontal.

Consider a ray being reflected back and forth between the mirrors at near-normal incidence, in the plane of Fig. 12. Let ζ_n be the (small) angle the ray makes with the horizontal after the nth bounce. Let ξ_n be the vertical
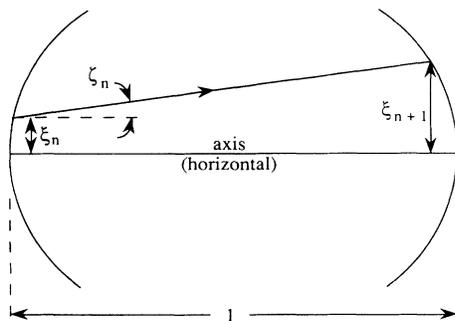
FIG. 12. A pair of opposing curved mirrors which are circular arcs. The separation of the mirrors on the axis is one. The reflection point of a ray is parameterized in terms of the vertical distance from the axis and the angle with respect to the horizontal.

distance of the nth bounce point from the axis.

Let the medium between the mirrors support two wave types S and P with different velocities c_S and c_P, and let $\kappa = c_S / c_P$. Assume that wave-type conversion can occur at either mirror and that the angles of the incident and reflected ways with respect to the tangent are related by

$$\cos\alpha_{inc} = k_n \cos\alpha_{ref} \tag{22}$$

and

$$k_n = \begin{cases} \kappa & \text{if a } P \text{ wave is incident} \\ & \text{and an } S \text{ wave is reflected} \\ 1/\kappa & \text{if an } S \text{ wave is incident} \\ & \text{and a } P \text{ wave is reflected} \\ 1 & \text{if incident and reflected} \\ & \text{wave types are the same .} \end{cases} \tag{23}$$

Then, for small (ζ, ξ), the coordinates of successive reflections are related by

$$\begin{bmatrix} \xi_{n+1} \\ \zeta_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -(1+k_n)\epsilon & k_n - \epsilon(1+k_n) \end{bmatrix} \begin{bmatrix} \xi_n \\ \zeta_n \end{bmatrix}, \tag{24}$$

where $\epsilon = 1/R_c$. We assume $|1-\epsilon| < 1$ so that the problem without ray splitting (i.e., $k_n \equiv 1$) results in the periodic orbit bouncing directly back and forth along the horizontal being stable to small perturbations.

By defining a new variable

$$\beta_n = -\frac{1}{\sin\theta}[\zeta_n + (1-\cos\theta)\xi_n], \tag{25}$$

where $\cos\theta \equiv 1 - \epsilon$, we can rewrite Eq. (24) in the form

$$\begin{bmatrix} \xi_{n+1} \\ \beta_{n+1} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ k_n\sin\theta & k_n\cos\theta \end{bmatrix} \begin{bmatrix} \xi_n \\ \beta_n \end{bmatrix}, \tag{26}$$

Thus, if $k_n = 1$, (ξ_{n+1}, β_{n+1}) is simply a rotation through an angle θ of (ξ_n, β_n). We are led by this to define new

variables r_n and ϕ_n such that

$$\xi_n = r_n \cos\phi_n \; ,$$
$$\beta_n = r_n \sin\phi_n \; . \tag{27}$$

Thus, in the absence of wave-type transitions, r_n does not change and (ξ_n, β_n) lies on the circle of radius r_n.

Then the map describing the evolution of r_n and ϕ_n is

$$r_{n+1} = r_n [\cos^2(\theta + \phi_n) + k_n^2 \sin^2(\theta + \phi_n)]^{1/2} \; ,$$
$$\tan\phi_{n+1} = k_n \tan(\theta + \phi_n) \; . \tag{28}$$

We note that $r_{n+1} = r_n$ and $\phi_{n+1} = \phi_n + \theta$ if there is no wave-type transition ($k_n = 1$). In other words, we have

$$\left| \frac{r_{n+1}}{r_n} \right|^2 = \frac{1 + k_n^2 \tan^2(\phi_n + \theta)}{1 + \tan^2(\phi_n + \theta)} \; . \tag{29}$$

Since the angle of incidence is small, the probability of a wave-type transition is also small [see Eq. (19)]. Thus the ray will experience many reflections before undergoing a wave-type transition. A pair of transitions (that is, a change from one wave type to the other, and, after many bounces, a transition back to the original wave type) will result in a change in r given by

$$\left| \frac{r_{\text{before}}}{r_{\text{after}}} \right|^2 = \left| \frac{1 + \kappa^2 \tan^2(\phi_j + \theta)}{1 + \tan^2(\phi_j + \theta)} \right|$$
$$\times \left| \frac{1 + (1/\kappa^2)\tan^2(\phi_k + \theta)}{1 + \tan^2(\phi_k + \theta)} \right| \; , \tag{30}$$

where ϕ_j and ϕ_k are the values of ϕ at the transitions. Since the transitions are infrequent and the angle ϕ_n increases by θ at each reflection, we can consider the values of ϕ_j and ϕ_k to be random, uncorrelated, and uniformly distributed on the interval $[0, 2\pi]$. Thus the expected value of the ratio (30) is

$$q^2 \equiv \int \frac{d\theta_1}{2\pi} \int \frac{d\theta_2}{2\pi} \left| \frac{1 + \kappa^2 \tan^2\theta_1}{1 + \tan^2\theta_1} \right|$$
$$\times \left| \frac{1 + (1/\kappa^2)\tan^2\theta_2}{1 + \tan^2\theta_2} \right|$$
$$= \frac{1}{4} \frac{(1 + \kappa^2)^2}{\kappa^2} \; . \tag{31}$$

We are interested in estimating the rate at which r_n/r_0 grows with n. To this end we consider the following. Suppose the probability of a transition on any particular bounce were a constant, i.e., independent of r_n and ϕ_n. We would then expect r to increase by some factor q in a time equal to the expected time $\bar{n}$ between a pair of transitions. This would give rise to exponential growth of the expected value of r, which we call $\bar{r}$, $\bar{r}^{-1} d\bar{r}/dn = (\ln q)/\bar{n}$. However, the probability of a transition (i.e., $\bar{n}^{-1}$) is *not* constant but depends on r_n and ϕ_n and, in particular, is proportional to r_n^2. Thus a rigorous statistical treatment of the problem would be somewhat complicated. As a heuristic model, we suppose that $\bar{r}$ increases by a factor q in a time the length of which depends inversely on $\bar{r}^2$ ac-

counting for the variation of the transition probabilities with r. The preceding is modeled by the equation

$$\frac{d\bar{r}}{dn} = A\bar{r}^3 \ln q \; ,$$

where $\bar{n}^{-1} = (A\bar{r}^2)$. We thus obtain for $\bar{r}(n)$

$$\bar{r}^2(n) = \bar{r}^2(0)[1 - 2An\bar{r}^2(0)\ln q]^{-1}$$

for $2An\bar{r}^2 \ln q < 1$. Alternatively, we can calculate the number of iterations N necessary for $\bar{r}(n)$ to reach a specific value $\bar{r}_f$,

$$N = \frac{1}{2A \ln q} \left[\frac{1}{\bar{r}(0)^2} - \frac{1}{\bar{r}_f^2} \right] \; .$$

Thus the number of iterations diverges as $\bar{r}(0)^{-2}$ as opposed to $\ln[\bar{r}(0)^{-1}]$ in the case of exponential growth. In this respect the growth of $\bar{r}$ is slower than exponential.

C. Example of a billiard without stable two-bounce periodic orbits

We are led by the above results to look at geometries which have no stable two-bounce orbits in the one-wave case. Such configurations will not have the pseudostable orbits of the ray-splitting problem discussed in the previous section.

The stability conditions for the two-bounce periodic orbits in situations without ray splitting can be derived as follows: As described earlier for a two-bounce closed orbit, the deviation of σ and τ from their initial values after two bounces is given by

$$\begin{bmatrix} \delta\sigma_2 \\ \delta\tau_2 \end{bmatrix} = m^{(2)} \begin{bmatrix} \delta\sigma_0 \\ \delta\tau_0 \end{bmatrix} \; .$$

The deviations $\delta\sigma_2$ and $\delta\tau_2$ remain bounded as the number of cycles goes to infinity if the trace of the two-bounce Jacobian matrix $m^{(2)}$ satisfies

$$|\mathrm{Tr}\, m^{(2)}| < 2 \; .$$

For a normally incident two-bounce orbit between boundaries with radii of curvature R_1 and R_2, $m^{(2)}$ is given by

$$m^{(2)} = \begin{bmatrix} -1 + \dfrac{\rho}{R_2} & -\rho \\[2ex] \dfrac{-\rho}{R_1 R_2} + \dfrac{1}{R_1} + \dfrac{1}{R_2} & -1 + \dfrac{\rho}{R_1} \end{bmatrix}$$
$$\times \begin{bmatrix} -1 + \dfrac{\rho}{R_1} & -\rho \\[2ex] \dfrac{-\rho}{R_1 R_2} + \dfrac{1}{R_1} + \dfrac{1}{R_2} & -1 + \dfrac{\rho}{R_2} \end{bmatrix} \; , \tag{32}$$

where ρ is the travel distance between bounces. Then

$$|\mathrm{Tr}\, m^{(2)}| = 2 \left| 2 \left[1 - \frac{\rho}{R_1} \right] \left[1 - \frac{\rho}{R_2} \right] - 1 \right| < 2 \tag{33}$$

for stability. Thus the periodic two-bounce orbit (without ray splitting) is stable if both of the following two conditions hold:

$$\rho < (R_1 + R_2) \, ,$$ (34a)

and

$$(R_1 - \rho)(R_2 - \rho) > 0 \, .$$ (34b)

Consider now an analogue of the Benettin-Strelcyn oval with threefold symmetry [Fig 13(a)]. This billiard is composed of six circular arcs with alternating radii of curvature, R_1 and R_2. The arcs with smaller radius subtend an angle γ at their center of curvature. We will refer to this billiard as a "tricloid."

(a)

(b)
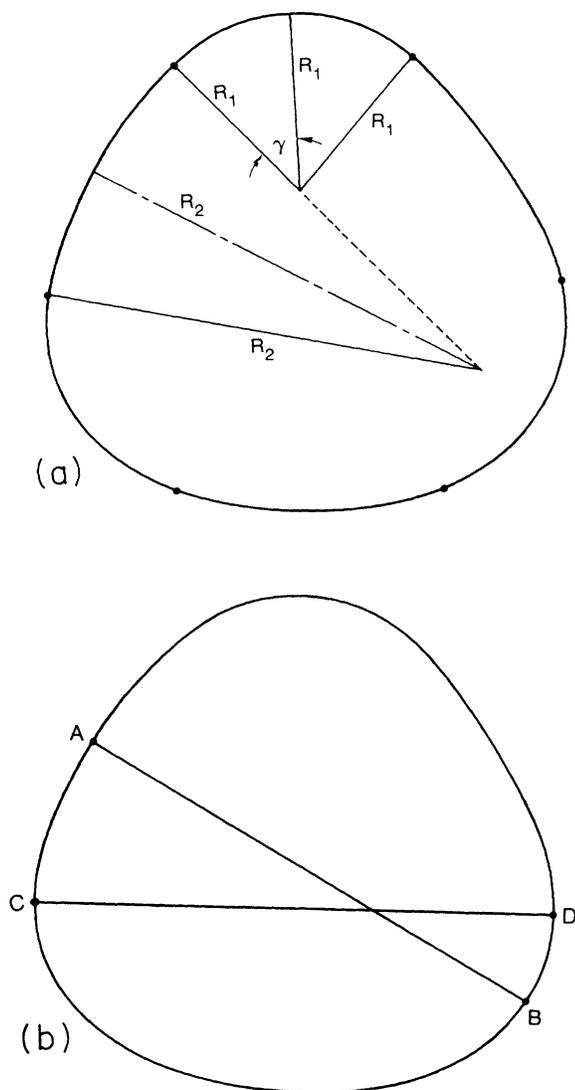
FIG. 13. (a) The "tricloid": a billiard which has no stable two-bounce orbits for some parameter values. The boundary is composed of six circular arcs with alternating radii of curvature R_1 and R_2. The arcs with smaller radius subtend an angle γ at their center of curvature. The slope is continuous at the joints. (b) The two-bounce orbits of the tricloid.

This billiard has two different types of two-bounce orbits, as typified by the paths AB and CD in Fig. 13(b). From the stability criteria, Eqs. (34a) and (34b), we find that in the absence of ray splitting (i) orbit AB is stable if $\gamma > \pi/6$ and $R_2 < \rho$, (ii) orbit CD is stable if $\gamma < \pi/6$, and (iii) both orbits are unstable if $\gamma > \pi/6$ and $R_2 > \rho$. Thus this billiard has parameter ranges both with and without stable two-bounce orbits.

For $\gamma = 0.6$, $R_1 = 1.0$, and $R_2 = 3.0026$, the orbit AB has a path length of $\rho = 3.6969$ and is stable. Figure 14 is a phase-space plot of a chaotic orbit for this billiard without ray splitting. The integrable regions due to the three stable two-bounce AB orbits are visible as gaps in the chaotic region at small τ values. The larger integrable regions are due to three-bounce orbits.

When ray splitting is included (Fig. 15), most invariant tori are destroyed as shown in Fig. 14. As in the case of the Benettin-Strelcyn oval with ray splitting, remnants of the KAM surfaces remain. Those orbits in S space which lie entirely below the critical angle are undisturbed. Furthermore, the pseudostable region near the two-bounce orbits is still visible even when 10^6 iterates are considered. The structure of the pseudostable regions is more clearly visible in Fig. 16, which includes only 10^5 iterates. Point densities for this configuration with 10^6 iterates are tabulated in Tables IV and V. The mean number of iterates falling into a cell was 4131 and 5869 for the P and S phase spaces, respectively. The standard deviations for the two spaces were 257.3 and 381.7, respectively. These high standard deviations are due primarily to the very long correlation times associated with the pseudostable two-bounce orbits. Perusal of the table reveals strong fluctuations in the point densities for small τ values (0 to 0.2 and -0.2 to 0). These are due to slow diffusion into and out of the pseudostable orbits. The numerical ratio of the average S-wave energy density to the average P-wave energy density, $5869/4131 = 1.42$, agrees well with the theoretical value $c_P/c_S = 1.43$ from the Appendix. The Lyapunov number of the 10^6 reflection orbit plotted in Fig. 15 and tabulated in Tables IV and V was calculated to be $1.68 > 1$ (chaos).

Figure 17 is a plot of a chaotic orbit without ray splitting for the tricloid with $\gamma = 0.785$, $R_1 = 1.0$, and $R_2 = 4.3411$. For this billiard the path length of the orbit AB is $\rho = 3.6142$, so there is no stable periodic two-bounce orbit for the system without ray splitting. 10^6 reflections were included.

Figure 18 shows a 10^6 reflection orbit for the same billiard as in Fig. 17, but with ray splitting. The phase space is now uniformly filled by this single chaotic orbit, with the exception of very small integrable regions in the S phase space lying above $\tau = \tau_c$. Point densities are tabulated in Tables VI and VII. The average number of iterates falling into a cell was 4143 and 5857 for the P and S phase spaces, respectively. The standard deviations are of approximately the same magnitude as those found for the orbit tabulated in Tables I and II, which was ergodic. (Again, the numerical ratio of the average S-wave energy density to the average P-wave energy density, $5857/4143 = 1.41$, agrees well with the theoretical value of 1.43.) Thus the elimination of the pseudostable two-
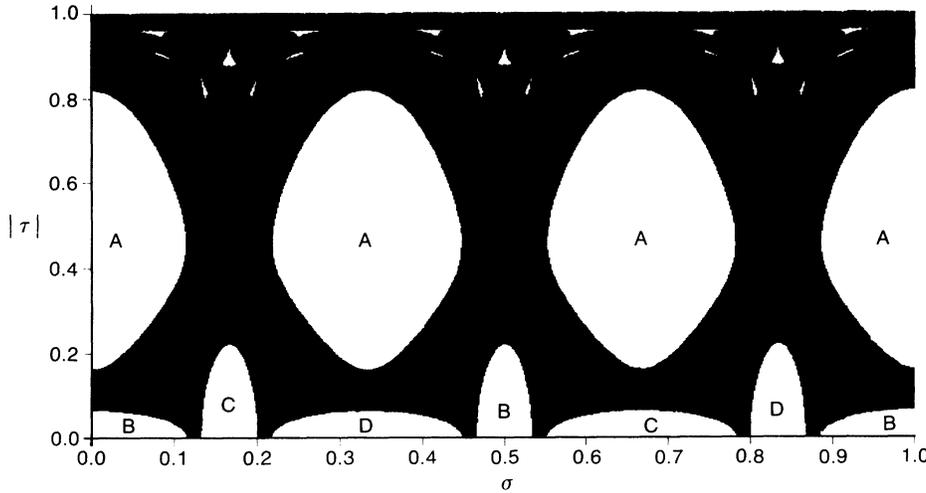
FIG. 14. Phase-space plot of a chaotic orbit of the tricloid without ray splitting for $\gamma = 0.6$, $R_1 = 1$, and $R_2 = 3.0026$. For these parameters the tricloid has a stable two-bounce orbit.

bounce orbits has resulted in more rapid diffusion throughout the phase space (except for orbits lying above τ_c), with the result that a 10^6-reflection orbit originating in the pseudostable region is approximately ergodic. With the exception of the S space above τ_c, this orbit explores the entire phase space approximately uniformly, including the areas occupied by KAM surfaces of the one-wave system, for example, the integrable regions due to the three-bounce orbits, marked A in Fig. 17. The Lyapunov number for the orbit plotted in Fig. 18 and tabulated in Tables VI and VII was calculated to be $3.07 > 1$.

D. Discussion of ergodicity in other cases

We have found that ergodicity is inhibited for those rays which do not experience splitting. In particular, for our elastic medium problem, there were two cases where rays do not split, namely, when the angle of incidence was greater than the critical angle for the S wave, and when the ray was always normally incident (the two-bounce orbit).

The experience we have gained in the elastic medium problem is useful for other cases. As an example, we discuss the case shown in Fig. 1(a). In that case, the media in regions 1 and 2 are assumed to support only one type of propagating wave. For the case, say, of acoustic waves, there is no analogy to the two-bounce orbit of the elastic medium because, aside from the case of critical reflection, incident waves are split by transmission and reflection. Thus, for waves interacting with the boundary, the only source of nonergodicity is critical reflection from the side with the smaller sound speed. One might envision that a situation analogous to the direct two-bounce orbit might arise for the case where the wave being described is an electromagnetic wave with its electric field in the plane of the figure. (We imagine regions 1 and 2 as being occupied by media of different dielectric constant.) In this case, a ray incident at Brewster's angle is not split (it is completely transmitted). If the ray also hits the conducting walls at normal incidence (Fig. 19), then we have a two-bounce orbit without ray splitting. We note, however, that this is a nongeneric situations. This is, only very specially chosen shapes allow such orbits,
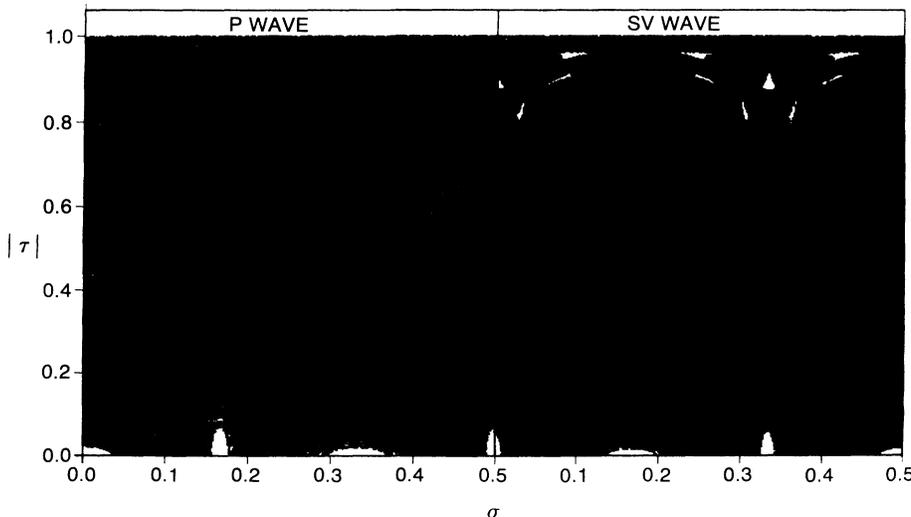
FIG. 15. Orbit resulting from the same initial conditions as those in Fig. 14 when ray splitting is present. 10^6 iterates are plotted.
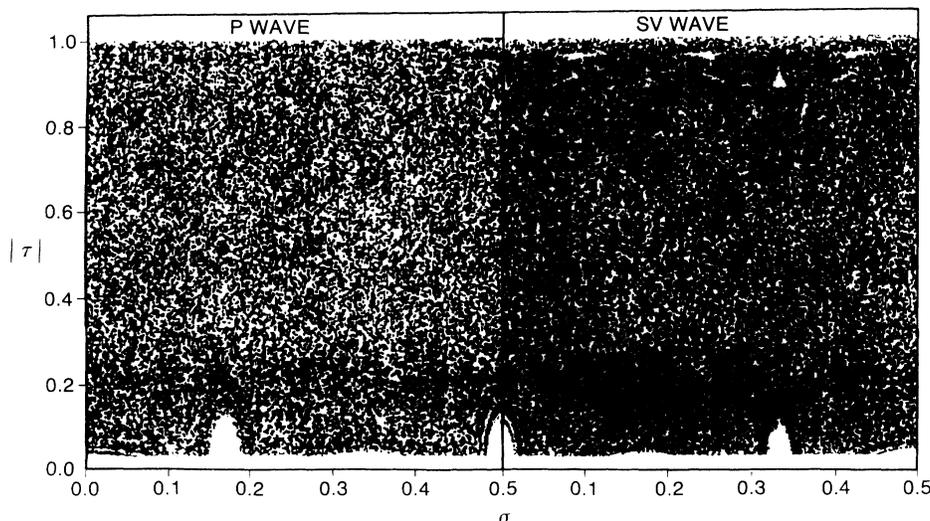
FIG. 16. Same orbit as that of Fig. 15. Only 10^5 iterates are plotted.

and a small change of the wall shape in Fig. 19 eliminates the possibility of this type of orbit. Thus we again find that critical reflection is the only effect likely to spoil ergodicity for rays which interact with the boundary between regions 1 and 2.

V. TRACE FORMULA FOR CHAOTIC BILLIARDS WITH RAY SPLITTING

In the preceding we have been addressing the questions of ergodicity and chaos using our Monte Carlo ray algorithm. In particular, we found Lyapunov numbers greater than one when the Monte Carlo ray fills an area in the τ-σ surface of section, and we called this condition chaos. We believe that this chaos in the Monte Carlo ray picture is analogous to the chaos that occurs in situations without ray splitting (i.e., ordinary billiards) and that results already developed in the field of quantum chaos (without ray splitting) can be straightforwardly generalized to the ray-splitting case. As an example, we indicate the generalization of the Gutzwiller periodic-orbit sum trace formula [6] to include ray splitting. Since the derivation [8] closely parallels Gutzwiller's analysis we shall not give it here; rather, we just report the result and its interpretation.

The "standard" Gutzwiller trace formula [1,6] gives an asymptotic (one hopes) approximation to the oscillatory contribution to the density of states for the Schrödinger wave equation in terms of a summation over classical periodic orbits. If $D(E)$ denotes the exact density of states as a function of the energy E, then the oscillatory contribution $d(E)$ is given by $d(E)=D(E)-\bar{D}(E)$, where $\bar{D}(E)$ is the Thomas-Fermi expression for the smoothed density of states. The Gutzwiller formula for $d(E)$ also has an additional smoothing aspect in that what the asymptotic periodic-orbit sum formula actually yields is $d(E)$ smoothed on some scale [smaller than the scale that produces $\bar{D}(E)$]. The scale of the smoothing for $d(E)$ is set by choosing a cutoff of the periodic-orbit sum; as higher-period periodic orbits are included the smoothing scale becomes finer (but since the sum is asymptotic, one presumably cannot include too many periodic orbits). Recently this approach has been improved upon to obtain actual individual eigenvalues in the semiclassical regime.

Our result involves a summation over closed orbits of the ray splitting problem. We illustrate several such closed orbits in Fig. 20 for the case of the two-region billiard shown in Fig. 1(a). For specificity, consider the case

TABLE IV. Number of iterates in cells of the P-wave phase space for the tricloid with a stable two-bounce orbit.

τ \ σ	0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–0.5	0.5–0.6	0.6–0.7	0.7–0.8	0.8–0.9	0.9–1.0
0.8–1.0	4283	4193	4152	4220	4209	4223	4128	4210	4146	4382
0.6–0.8	4312	4131	4218	4289	4147	4093	4211	4221	4304	4301
0.4–0.6	4524	4315	4267	4229	4248	4242	4189	4304	4154	4206
0.2–0.4	4549	4204	4320	4161	4278	4167	4251	4082	4066	4261
0–0.2	4724	4238	3832	3399	4447	4559	4192	4362	3023	4564
(−0.2)–0	4476	4182	3842	3329	4472	4673	3941	4243	3034	4447
(−0.4)–(−0.2)	4086	4084	4091	3984	4115	4009	4015	4113	4127	4045
(−0.6)–(−0.4)	4076	4127	4101	3934	4000	4080	3928	4002	4055	4037
(−0.8)–(−0.6)	4052	4048	4024	4104	4060	4155	4025	4073	4026	3902
(−1.0)–(−0.8)	4099	4080	4039	4031	3935	3988	4052	4031	3953	4049

TABLE V. Number of interates in cells of the S-wave phase space of the tricloid with a stable two-bounce orbit.

σ τ	0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–0.5	0.5–0.6	0.6–0.7	0.7–0.8	0.8–0.9	0.9–1.0
0.8– 1.0	5563	5233	5401	5691	5318	5272	5771	5481	5199	5463
0.6–0.8	6017	5902	6141	5958	6050	6002	6047	6119	6035	6095
0.4–0.6	6059	6032	5947	6093	6195	5898	6119	6077	6067	6053
0.2–0.4	5922	6137	6116	6130	5996	6093	6032	5970	5803	6053
0–0.2	6365	5957	5660	5124	6146	6451	5886	6057	4803	6400
(−0.2)–0	6107	5911	5628	4944	6084	6311	5749	5930	4750	6293
(−0.4)–(−0.2)	5736	5925	5687	5715	5877	5779	5773	5699	5809	5688
(−0.6)–(−0.4)	5751	5732	5793	5618	5801	5825	5801	5724	5890	5817
(−0.8)–(−0.6)	5743	5737	5733	5789	5849	5739	5788	5763	5817	5788
(−1.0)–(−0.8)	6071	6357	5993	6182	6197	6189	6199	6042	6371	6002

of Schrödinger's equation with two different values of the potential in regions 1 and 2. The orbits shown in Fig. 20 satisfy the following conditions: they close on themselves; they have equal angles of incidence and reflection when they are reflected either from the outer bounding (stadium-shaped) curve or from the boundary separating the two regions of different potential, and, when transmitted through the boundary, their angles of incidence and transmission are determined by conservation of tangential momentum at the boundary [the elastic medium analog of which is Eq. (15)].

With ray splitting the trace formula becomes [11]

$$d(E) \simeq \frac{1}{i\hbar} \sum_k \frac{(\hat{A}_k)^{1/2} T_k}{2\sinh(\lambda_k/2)} \exp\left[i\left(\frac{S_k(E)}{\hbar} + \phi_k \right) \right] ,$$

(35)

where $d(E)$ is the oscillatory contribution to the density of states as a function of energy E, k labels the discrete set of closed orbits, and the other quantities in Eq. (35) have the following meanings.

The quantity $S_k(E)$ is given by $S_k(E) = \oint \mathbf{p} \cdot d\mathbf{q}$, where the integral is taken around closed orbit k and represents the action for this orbit (here $\mathbf{p}$ and $\mathbf{q}$ denote position and momentum). T_k is the primitive period for closed orbit

k. λ_k is the stability exponent for closed orbit k; it gives the exponential rate of increase of linearized perturbations of the orbit at a surface of section across the closed orbit with the perturbed orbits satisfying angle of incidence equal angle of reflection at all reflections and conservation of tangential momentum at all transmissions.

The quantity $\hat{A}_k$ is given by

$$\hat{A}_k = \left[\prod_{i=1}^{r_k} |\hat{\rho}_i|^2 \right] \left[\prod_{j=1}^{t_k} (1 - |\hat{\rho}_j|^2) \right] ,$$

where $\hat{\rho}_i$ is the (amplitude) reflection coefficient at the ith reflection, and r_k and t_k are the number of reflections and transmissions experienced by closed orbit k. Basically, $\hat{A}_k$ is the quantum probability that a particle starting along a leg of the closed orbit completes one circuit of the orbit. For example, for the closed orbit in Fig. 20(a) $r_k = 1$, $t_k = 0$, and $\hat{A}_k$ is $|\hat{\rho}_a|^2$, where $\hat{\rho}_a$ is the reflection coefficient for the angle of incidence for the ray incident on the boundary between the two regions; for the orbit in Fig. 20(b) $r_k = 0$, $t_k = 2$, and $\hat{A}_k$ is $(1 - |\hat{\rho}_b|^2)^2$, where $\hat{\rho}_b$ is the reflection coefficient for the ray incident from region 1 (which is the negative of the reflection coefficient for the ray incident from region 2); for the orbit in Fig. 20(c), $r_k = 1$ and $t_k = 2$.

Finally, the phase quantity ϕ_k includes the sum of the
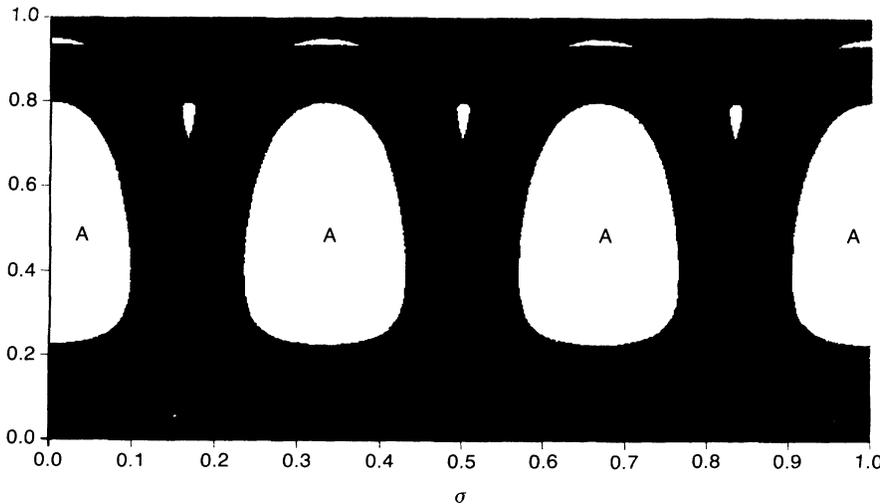
FIG. 17. Phase-space plot of a chaotic orbit of the tricloid without ray splitting for $\gamma = 0.785$, $R_1 = 1$, and $R_2 = 4.3411$. There is no stable two-bounce orbit.
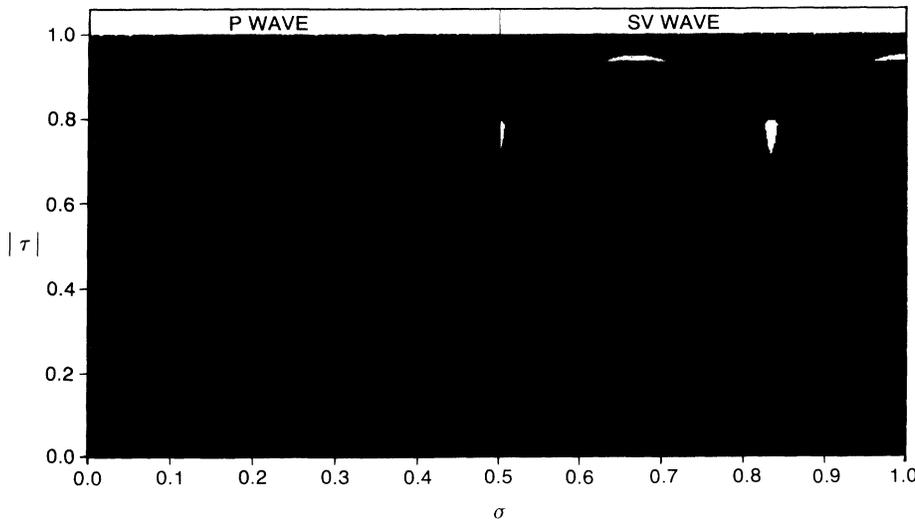
FIG. 18. Orbit resulting from the same initial conditions as those in Fig. 17 when ray splitting is present.

phase of all the reflection coefficients $\hat{\rho}_i$ plus the Maslov contributions. In particular, if $\hat{\rho}_i$ is negative, then the corresponding angle is π. Also, when there is critical reflection, then $\hat{\rho}_i$ has unit magnitude, but is complex, $\hat{\rho}_i = \exp(i\bar{\phi}_i)$, and this $\bar{\phi}_i$ adds to ϕ_k. (Note that there is no phase contribution due to a transmission because the transmission coefficient is always real and positive.)

In the absence of ray splitting $\hat{A}_k \equiv 1$, ϕ_k is the Maslov phase, the closed orbits are the usual periodic orbits, and (35) reduces to the known form without ray splitting. In the case of ray splitting, one can expect many more closed orbits than for a comparable problem without ray splitting. This is because on each encounter with the boundary a ray generates two daughters, each of which could be part of a periodic orbit. Thus in the sum there will be many more terms. Counterbalancing this increase in the number of terms is the fact that the amplitude $\hat{A}_k$ decreases exponentially with the number of ray splittings and hence with the length of the orbit. We emphasize that, although our discussion in this section has been in the context of ray splitting for the Schrödinger problem, a similar result applies in other ray splitting situations, such as the elastic medium problem.

VI. CONCLUSION

All the numerical examples which we studied support the hypothesis that ray splitting greatly increases the incidence of ergodicity. For the case which is ergodic in the absence of ray splitting, ray splitting does not alter the ergodicity. For cases we studied for which the phase space consists of a mixture of chaotic and integrable regions in the absence of ray splitting, the presence of ray splitting dramatically alters the situation. In particular, we found that KAM tori lying partially or entirely in regions of the phase space where ray splitting is possible are destroyed by ray splitting. This includes all orbits which originate as P waves and all S wave orbits which at any point pass below the critical angle for wave conversion.

However, although we found a prevalence of chaos in systems with ray splitting, two remnants of the integrability in the analogous system without ray splitting remain. First, there may be actual KAM surfaces for orbits originating in the wave type with smaller phase velocity. These will be entirely in the region below the critical angle, where no wave-type transition occurs.

Second, for the elastic medium case, the existence of

TABLE VI. Number of iterates in cells of the P-wave phase space of the tricloid without a stable two-bounce orbit.

τ \ σ	0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–0.5	0.5–0.6	0.6–0.7	0.7–0.8	0.8–0.9	0.9–1.0
0.8–1.0	4402	4281	4217	4207	4278	4333	4223	4266	4342	4254
0.6–0.8	4402	4135	4164	4150	4267	4175	4239	4157	4142	4204
0.4–0.6	4378	4077	4181	4206	4283	4257	4170	4106	4037	4114
0.2–0.4	4316	4213	4211	4141	4354	4175	4178	4167	4115	4145
0–0.2	4335	4098	4038	4056	4173	4151	4108	4198	4038	4201
(−0.2)–0	4141	4160	4042	4061	4131	4153	4168	4026	4138	4128
(−0.4)–(−0.2)	3984	4133	4100	4085	4075	4221	4043	4071	4210	4061
(−0.6)–(−0.4)	4026	4156	4008	4119	4056	4099	4002	3953	4190	4004
(−0.8)–(−0.6)	4050	4069	4051	4144	4056	4021	4069	4122	4088	3988
(−1.0)–(−0.8)	4134	4062	4104	4115	4038	4086	4111	4075	4100	4038
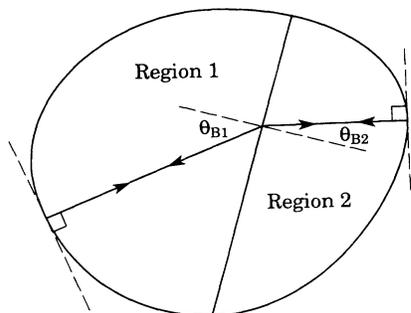
FIG. 19. Two-bounce orbit for electromagnetic waves. θ_{B1} and θ_{B2} are the Brewster angles for the media in regions 1 and 2, respectively.

FIG. 20. Some closed orbits for a ray splitting stadium billiard corresponding to Schrödinger's equation with two homogeneous regions of different potential separated by the diagonal line shown.

stable two-bounce orbits in the no-ray-splitting case results in pseudostable orbits, even in the presence of ray splitting. Although the two-bounce orbits are part of the chaotic region, orbits perturbed from them leave their vicinity by a different process which is very slow.

Thus, with the exception of the case in which the equation governing ray splitting [for example, Eq. (15)] maps an invariant curve of one phase space onto an invariant curve of the other phase space, the following summarizes our findings for two-dimensional elastic domains billiards:

(i) Orbits originating in P space are chaotic for any geometry:

(ii) Orbits originating in S space below $\tau = \tau_c$ are chaotic for any geometry.

(iii) If there are integrable orbits of the case without ray splitting which lie entirely above $\tau = \tau_c$, the S phase space above $\tau = \tau_c$ contains KAM surfaces even in the presence of ray splitting.

(iv) If there are stable two-bounce orbits for the case without ray splitting, there are pseudostable orbits in the case with ray splitting which *are* part of a chaotic region, but for which diffusion out of these orbits is much slower than exponential.

(v) Although we have considered here the particular

example of a bounded elastic domain, the phenomena described above are more general (Sec. IV D).

For a general two-wave system it is easy to see that, if a ray experiences splitting, then any invariant curve associated with it in the case without splitting will be destroyed in all but one exceptional case. This is the case in which the equation which governs the wave-type transition

$$\tau_{\text{after}} = \kappa \tau_{\text{before}} \tag{36}$$

maps an invariant curve for the problem in the absence of ray splitting from one phase space onto an invariant curve in the phase space of the other wave type. This can be seen as follows: Consider a ray launched in one wave type with a value of (σ, τ) that lies on an invariant curve of that phase space. (See Fig. 21.) At each reflection the ray may undergo ray splitting; thus, for each value of τ on the invariant curve indicated by "before" on the figure, the ray also has access to a value $\tau_{\text{after}} = \kappa \tau_{\text{before}}$. Thus the ray also has access to the curve marked "after" on the figure. If the curve marked "after" is not an invariant curve in the absence of ray splitting, later iterates of the wave type 1 space will move off the curve filling an area. Thus the original invariant curve is destroyed by

TABLE VII. Number of iterates in cells of the S-wave phase space of the tricloid without a stable two-bounce orbit.

τ \ σ	0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–0.5	0.5–0.6	0.6–0.7	0.7–0.8	0.8–0.9	0.9–1.0
0.8–1.0	5861	6058	5997	5654	5960	5942	5682	5912	6037	5819
0.6–0.8	5900	5661	5926	5986	5816	5736	5910	5990	5593	5888
0.4–0.6	6015	6003	5952	5915	5814	5895	6084	6042	6066	5993
0.2–0.4	5956	5839	5996	5897	5870	5919	5918	5830	5917	5952
0–0.2	6150	5802	5945	5900	5946	5885	5867	5869	5968	6010
(−0.2)–0	5900	5842	5846	5895	5838	5951	5898	5830	5887	5837
(−0.4)–(−0.2)	5852	5914	5842	5849	5865	5804	5809	5851	5800	5883
(−0.6)–(−0.4)	5713	5825	5839	5711	5865	5861	5846	5743	5901	5790
(−0.8)–(−0.6)	5948	5628	5872	5799	5748	5670	5810	5894	5624	5836
(−1.0)–(−0.8)	5754	5773	5848	5510	5682	5729	5543	5777	5753	5649

Wave Type 1   Wave Type 2
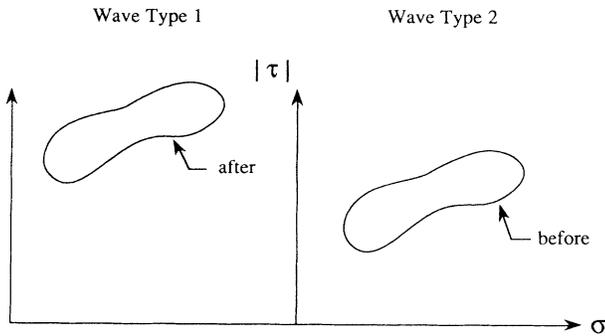
FIG. 21. The invariant curve marked "before" maps onto a curve marked "after" in the other phase space.

ray splitting unless Eq. (36) maps it onto an invariant curve in the other phase space.

Clearly, this is an exceptional situation. One example of such a situation for the case of a two-dimensional elastic region is the case of a circular boundary. In this case the invariant tori are straight lines, $\tau = $ const, in both the S-wave and the P-wave (σ, τ) phase space. Thus wave-type transition according to Eq. (36) simply allows the transition to an invariant curve in the other phase space. The return transition then returns the ray to the original invariant curve. Thus, in this exceptional case, an orbit with ray splitting is restricted to lie on a pair of invariant curves of the case without ray splitting.

In by far the great majority of cases, however, ray splitting destroys most of the invariant curves. Thus the presence of ray splitting greatly increases the prevalence of ergodic behavior. This is strikingly illustrated, for example, by comparing Figs. 17 and 18. The prevalence of ergodicity with ray splitting should have important implications for the wave problem. For example, we conjecture that, since the integrable regions are so small, an approximate Wigner distribution of energy-level spacings should apply well to the situation of Fig. 18, while with the same shape without ray splitting this would not be so. In fact, it appears that much of the work done on quantum chaos in the absence of ray splitting should extend naturally to wave problems with ray splitting. As an example, in Sec. V we have outlined a generalization of the Gutzwiller trace formula to situations with ray splitting.

APPENDIX

We wish to find the ratio of the number of S wave reflections to the number of P wave reflections for an ergodic orbit. For an ergodic Monte Carlo orbit, the number of times it is converted from wave type P to type S must equal the number of times it is converted from wave type S to wave type P, with an additional additive term of $+1$ or -1 if it starts in P and winds up S or vice versa. Thus, for many bounces we have

$$\sum_{j=1}^{N_S} A_{SP}(\alpha_S^{(j)}) \simeq \sum_{k=1}^{N_P} A_{PS}(\alpha_P^{(k)}) \ , \tag{A1}$$

where N_S is the number of reflections which have an incident S wave, N_P is the number of reflections with an incident P wave, $\alpha_S^{(j)}$ is the incident angle at the jth S wave reflection, and $\alpha_P^{(k)}$ is the incident angle at the kth P-wave reflection. Since the orbits are ergodic in the (σ, τ) phase space, and thus are uniformly distributed in τ, the averages in Eq. (A1) over all reflections can be replaced by averages over τ,

$$\frac{2N_S}{\pi} \int_0^{\kappa^{-1}} A_{SP}(\alpha_S) d(\cos\alpha_s)$$

$$= \frac{2N_P}{\pi} \int_0^1 A_{PS}(\alpha_P) d(\cos\alpha_P)$$

$$= \frac{2N_P}{\pi} \int_0^{\kappa^{-1}} A_{SP}(\alpha_S) d(\kappa \cos\alpha_S) \ . \tag{A2}$$

Thus we have

$$N_S = \kappa N_P \ . \tag{A3}$$

[1] M. C. Gutzwiller, *Chaos in Classical and Quantum Mechanics* (Springer, New York, 1990).

[2] F. Haake, *Quantum Signatures of Chaos* (Springer, New York, 1991).

[3] G. D. Birkhoff, Acta Math. **50**, 359 (1927).

[4] G. Benettin and J. M. Strelcyn, Phys. Rev. A **17**, 773 (1978).

[5] M. V. Berry, Eur. J. Phys. **2**, 91 (1981).

[6] M. C. Gutzwiller, J. Math. Phys. **8**, 1979 (1967); **10**, 1004 (1969); **11**, 1791 (1970); **12**, 343 (1971).

[7] A. Zacherl, T. Geisel, J. Nierwetberg, and G. Radons, Phys. Lett. A **114**, 317 (1986).

[8] L. Couchman, Ph.D. thesis, University of Maryland, 1991. This reference contains greater detail on the derivation of the trace formula (Sec. V of the present paper) and also uses closed orbits to address scarring and the Δ_3 statistic. These latter two proceed in a way analogous to the non-ray-splitting analysis of Ref. [9] (for the case of scars) and Ref. [10] (for the case of the Δ_3 statistic).

[9] E. B. Bogomolny, Physica D **31**, 169 (1988).

[10] M. V. Berry, Proc. R. Soc. London Ser. A **400**, 229 (1985)

[11] Equation (35) is for the case of isolated closed orbits that are hyperbolic without reflection; i.e., the eigenvalue governing perturbations about the orbit is positive and greater than one in magnitude. When the eigenvalue is less than -1 (hyperbolic with reflection), the sinh function in (35) is replaced by a cosh function.
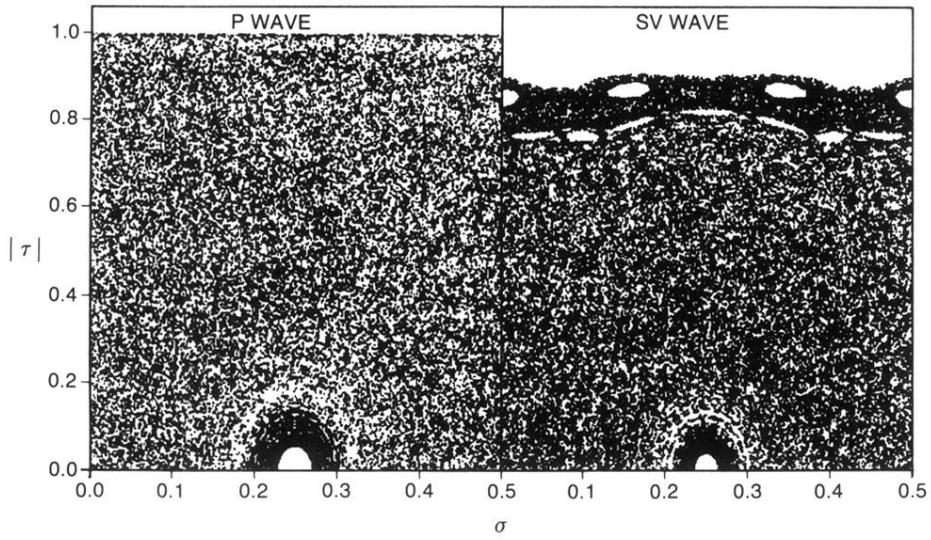
FIG. 10. The S and P phase spaces of the Benettin-Strelcyn oval with ray splitting at δ=0.76.
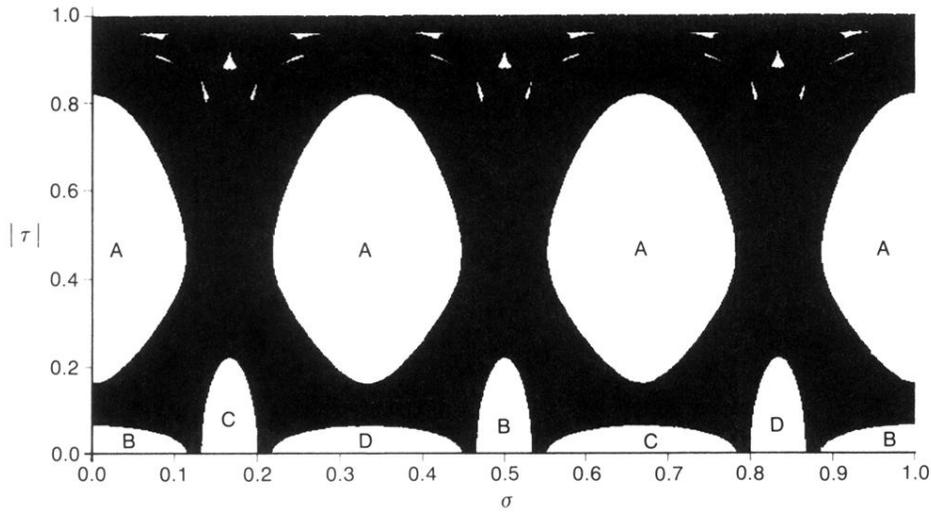
FIG. 14. Phase-space plot of a chaotic orbit of the tricloid without ray splitting for $\gamma = 0.6$, $R_1 = 1$, and $R_2 = 3.0026$. For these parameters the tricloid has a stable two-bounce orbit.
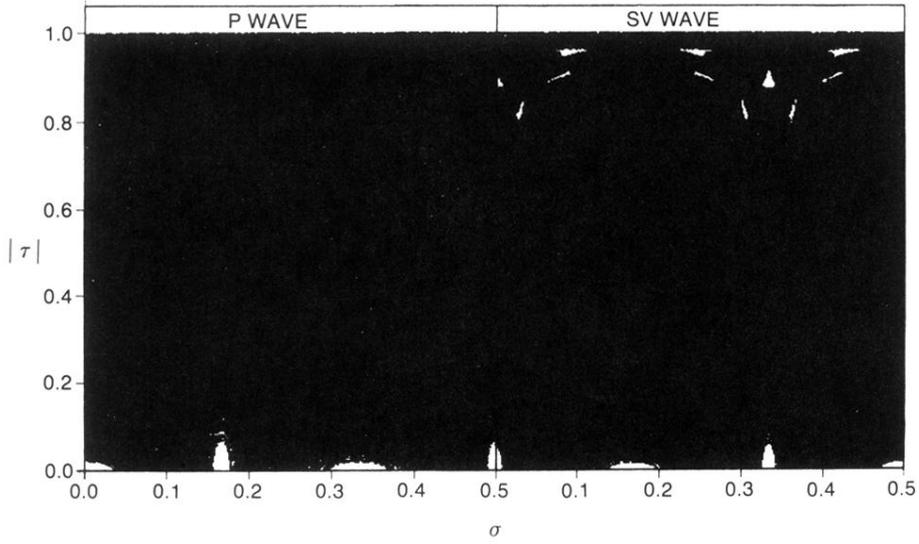
FIG. 15. Orbit resulting from the same initial conditions as those in Fig. 14 when ray splitting is present. 10^6 iterates are plotted.
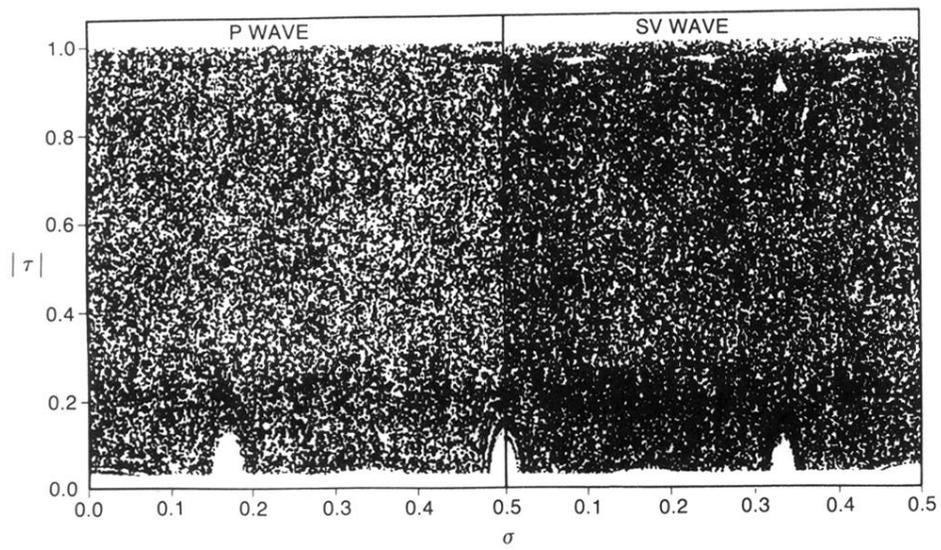
FIG. 16. Same orbit as that of Fig. 15. Only 10^5 iterates are plotted.
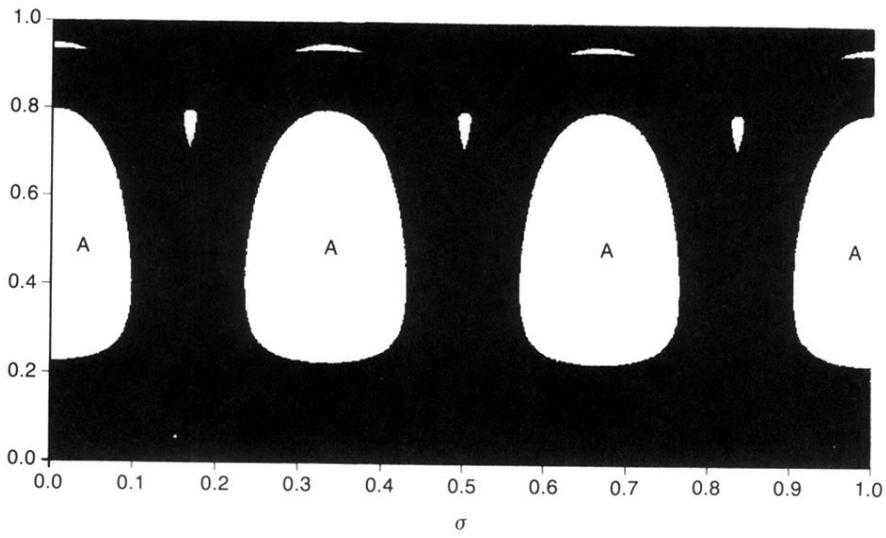
FIG. 17. Phase-space plot of a chaotic orbit of the tricloid without ray splitting for $\gamma = 0.785$, $R_1 = 1$, and $R_2 = 4.3411$. There is no stable two-bounce orbit.
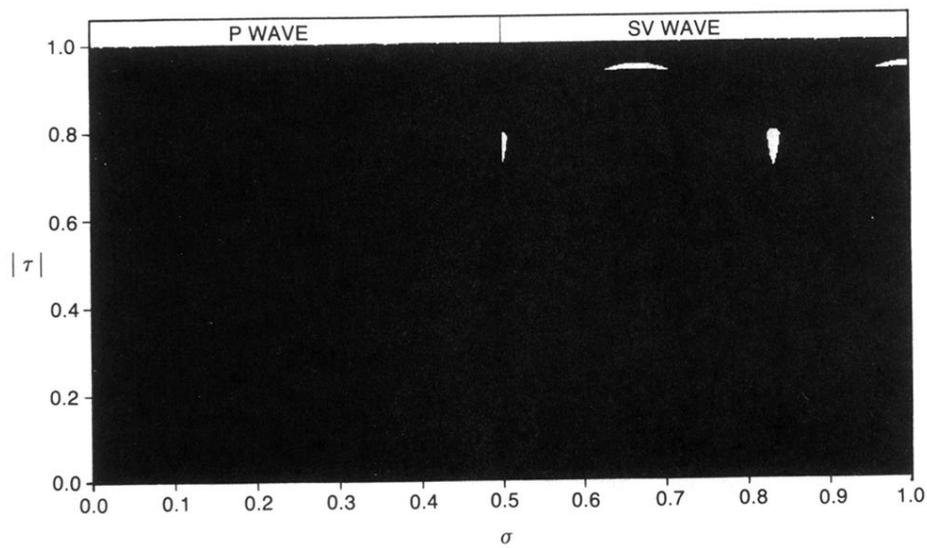
FIG. 18. Orbit resulting from the same initial conditions as those in Fig. 17 when ray splitting is present.

FIG. 4. (a) An example of the phase space of the Benettin-Strelcyn oval at $\delta = 0.76$: One of the chaotic regions. (b) An example of the phase space of the Benettin-Strelcyn oval at $\delta = 0.76$: Seven of the invariant curves are labeled with the numerals 1–7.