Spectral Oscillations, Periodic Orbits, and Scaling

S. A. Fulling†
Mathematics Dept., Texas A&M University,
College Station, TX, 77843-3368 USA

(Dated: December 13, 2000)

Abstract

The eigenvalue density of a quantum-mechanical system exhibits oscillations, determined by the closed orbits of the corresponding classical system; this relationship is simple and strong for waves in billiards or on manifolds, but becomes slightly muddy for a Schrödinger equation with a potential, where the orbits depend on the energy. We discuss several variants of a way to restore the simplicity by rescaling the coupling constant or the size of the orbit or both. In each case the relation between the oscillation frequency and the period of the orbit is inspected critically; in many cases it is observed that a characteristic length of the orbit is a better indicator. When these matters are properly understood, the periodic-orbit theory for generic quantum systems recovers the clarity and simplicity that it always had for the wave equation in a cavity. Finally, we comment on the alleged "paradox" that semiclassical periodic-orbit theory is more effective in calculating low energy levels than high ones.

†Electronic address: `fulling@math.tamu.edu`; URL: `http://www.math.tamu.edu/~fulling`

I. INTRODUCTION

The basic principle relating classical periodic orbits to quantum spectra has been succinctly stated by Delos and Du [1]:

> The average density of states as a function of energy [is] equal to a smooth monotonic function, related to the volume occupied by the energy-shell in phase-space, plus a sum of sinusoidal oscillations. The wavelength and amplitude of each oscillation are respectively correlated with the period and the stability of a periodic orbit of the system. For given energy resolution ΔE, only those periodic orbits are significant for which the period is less than $2\pi\hbar/\Delta E$.

An example of a precise theorem along these lines can be found in [2]. (Other expository works include [3, 4, 5, 6].)

This principle was simultaneously discovered thirty years ago in three classic research works dealing with three distinct contexts:

1. Balian and Bloch [7] considered a bounded region in space. The differential operator (corresponding to a quantum Hamiltonian H) is the Laplacian with standard (Dirichlet or Neumann) boundary conditions. The relevant classical paths are straight lines with specular reflection at the boundary.

2. Gutzwiller [8, 9] studied a Schrödinger equation in infinite space, with a potential that is responsible for confining particles to compact regions and creating a discrete energy spectrum. The periodic orbits involved are those of the full classical dynamics governed by that potential function. (Balian and Bloch also studied this situation in a later paper [10].)

3. Duistermaat and Guillemin [11], building on earlier papers in the French mathematical literature [12, 13, 14], worked on a compact Riemannian manifold without boundary. The special case of their problem most closely analogous to the other two is that where H is the Laplace–Beltrami operator for that manifold, plus a possible potential function that is treated as a *perturbation*. That is, the manifold itself is responsible for confining the orbits and making the spectrum discrete, and the periodic orbits that appear in the theorem are the closed geodesics of the Riemannian metric, independent of the potential.

Although ultimately the most important thing about these three bodies of work is their similarity, there are some characteristic differences among them.

First, Gutzwiller works with the Schrödinger equation,

$$i\,\frac{\partial\psi}{\partial t} = H\psi \qquad (\hbar = 1), \tag{I.1}$$

the others with the wave equation,

$$-\frac{\partial^2\psi}{\partial t^2} = H\psi \tag{I.2}$$

or its first-order pseudodifferential form

$$i\,\frac{\partial\psi}{\partial t} = \sqrt{H}\,\psi. \tag{I.3}$$

(Indeed, the authors of [11] go so far as to say that "No statement like [our theorem] holds if we replace the first order operator [$\sqrt{H}$] by a higher order operator.")

Second, Gutzwiller and the many physicists who have followed him usually speak of periodicities in the *energy* spectrum, determined (reciprocally) by the *times* (periods) of the classical orbits. But the conclusions of the other works are formulated in terms of periodicity in wave frequency, ω (essentially the *square root* of energy, E), determined by the *lengths* of the orbits. A reader who has not devoted years to the comparative study of these papers might form the impression that the distinction between energy periodicity and frequency periodicity arises from the respective concentration on the Schrödinger equation or the wave equation; in other words, that the square root relating E to ω is merely the square root relating the quantum Hamiltonian to the pseudodifferential operator $\sqrt{H}$ appearing in the wave equation (I.3). That would be wrong. Indeed, in hindsight, it is obviously wrong: The conclusions of the theory are objective statements about the spectrum of the *same* elliptic second-order differential operator H. These facts cannot depend upon which time-dependent equation, (I.1) or (I.2), was used as a technical tool in discovering them.

A third difference is that the orbits studied in [7] and [11] are energy-independent objects, but the closed orbits in a potential vary with the energy. The latter fact means that the parametrization of spectral oscillations by orbits can be meaningful only over rather short energy intervals, in general. (An exception is systems with homogeneous potentials, where the orbits at different energies are related merely by dilation in space. See Sec. VI and [5, text surrounding (1.40)].) Friedrich and Wintgen [15, 3] pointed out that this problem can be at least partially avoided by broadening the point of view to allow variation of a coupling constant with energy. The primary aim of the present paper is to generalize their observation, showing that it has several variants, one that is applicable to any system, and others that may be more appropriate for systems with certain scaling symmetries. In each case the natural variable of the spectral oscillations is a certain power of the energy. It is argued tht the identification of the frequencies of these oscillations with the *periods* of the orbits is rather forced; some other quantity, often a *length*, is more pertinent.

These matters occupy most of the remaining sections of this paper. Sec. IV is a digression to establish a (known) relation among action, energy, and period when the coupling constant is *not* scaled; in the context of scaling, this relation no longer holds, but neither is it necessary. In Sec. II we elucidate the relationship between the wave (I.2)–(I.3) and Schrödinger (I.1) approaches (and their correlates in classical phase space); this point is understood by workers in the field but is seldom spelled out. (To keep the treatments of the two approaches as parallel as possible, we discuss the wave equation in the terminology of relativistic quantum theory; this should not be allowed to obscure the applicability of periodic-orbit analysis in other, more classical, contexts, such as optics and acoustics, where quantization is just a metaphor.) Finally, Sec. IX addresses the frequently remarked-upon counterintuitive fact that the "semiclassical" periodic-orbit method is more effective in reproducing low-lying eigenvalues than eigenvalues in the regime of large quantum numbers.

II. WAVE AND SCHRÖDINGER DYNAMICS

A Hamiltonian system is defined by a function $H(\mathbf{x},\mathbf{p})$ of $\mathbf{x}\in\Omega$, a d-dimensional region or manifold, and $\mathbf{p}\in\mathbf{R}^d$. Its classical equations of motion are

$$\frac{d\mathbf{x}}{dt} = \nabla_{\mathbf{p}} H, \qquad \frac{d\mathbf{p}}{dt} = -\nabla_{\mathbf{x}} H. \tag{II.1}$$

Energy is conserved: Each trajectory of the system (II.1) remains on a phase-space submanifold

$$H(\mathbf{x},\mathbf{p}) = E \tag{II.2}$$

of dimension $2d-1$.

Suppose that $H(\mathbf{x},\mathbf{p})$ is quadratic in $\mathbf{p}$ and nonnegative, and consider a new Hamiltonian function

$$h(\mathbf{x},\mathbf{p}) \equiv H(\mathbf{x},\mathbf{p})^{1/2}. \tag{II.3}$$

Calling the new time parameter τ, calculate the resulting equations of motion:

$$\frac{d\mathbf{x}}{d\tau} = \nabla_{\mathbf{p}} h = \tfrac{1}{2} H^{-1/2}\frac{d\mathbf{x}}{dt}, \qquad \frac{d\mathbf{p}}{d\tau} = \tfrac{1}{2} H^{-1/2}\frac{d\mathbf{p}}{dt}. \tag{II.4}$$

It follows that the trajectories of the new system in phase space are exactly the same as those of the old one, but parametrized differently. At a fixed energy the two time scales are simply related by

$$\frac{d\tau}{dt} = 2\sqrt{E}. \tag{II.5}$$

The relation between $\mathbf{p}$ and the velocity is changed by the same factor.

Consider the very special case $H = \mathbf{p}^2/2m$. Then $\mathbf{p}$ is a constant vector with magnitude $p=\sqrt{2mE}$, and the velocity in system (II.1) satisfies

$$\mathbf{v} \equiv \frac{d\mathbf{x}}{dt} = \frac{\mathbf{p}}{m}, \qquad v^2 = \frac{2E}{m}, \tag{II.6}$$

while the velocity in system (II.4) satisfies

$$\mathbf{u} \equiv \frac{d\mathbf{x}}{d\tau} = \frac{1}{2\sqrt{E}}\frac{\mathbf{p}}{m}, \qquad u^2 = \frac{1}{2m} = \text{constant}. \tag{II.7}$$

The trajectories of H dynamics in space-time are, of course, those of a nonrelativistic particle: straight lines of arbitrary slope depending on energy. The trajectories of h dynamics in space-time are *independent of energy*, as appropriate for a *relativistic massless particle*. (Clearly, in this context m does not have the physical significance of a mass; rather, it parametrizes the wave speed as $c=(2m)^{-1/2}$.) Passing to a quantum system by the prescription

$$\mathbf{p} \mapsto -i\nabla_{\mathbf{x}}, \qquad E \mapsto i\frac{\partial}{\partial\tau} \tag{II.8}$$

converts the relativistic energy relation $h(\mathbf{x},\mathbf{p}) = E$ to a special case of (I.3), which is equivalent to the wave equation

$$\frac{\partial^2\psi}{\partial t^2} = \frac{1}{2m}\nabla^2\psi \tag{II.9}$$

supplemented by a condition of positive frequency. The nonrelativistic energy relation (II.2) of course becomes the free Schrödinger equation, a case of (I.1).

The energy independence of relativistic trajectories (which extends to the very general systems studied in [11]) leads to a very direct relationship between spectral oscillations and the *singularities* of the fundamental solution of the wave equation (I.3). This is the meaning of the remark of Duistermaat and Guillemin quoted in Sec. I. An up-to-date exposition of that methodology is given in [16].

Now turn to a Hamiltonian including a potential energy function,

$$H(\mathbf{x}, \mathbf{p}) = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}), \tag{II.10}$$

where $V(\mathbf{x})$ is nonnegative. The momentum is no longer constant, and the velocities satisfy

$$v^2 = \frac{2}{m}[E - V(\mathbf{x})] \tag{II.11}$$

(the usual nonrelativistic kinetic-energy relation) and

$$\begin{aligned} u^2 &= \frac{p^2}{2m(p^2 + 2mV(\mathbf{x}))} \\ &= \frac{1}{2m}\left(1 - \frac{V(\mathbf{x})}{E}\right). \end{aligned} \tag{II.12}$$

The trajectories in space-time are now curves, whose slopes at each point are greater than they would have been for a free particle. In particular, in the relativistic case the trajectories stay always inside the local light cones, as appropriate for a relativistic particle with mass (induced here by an $\mathbf{x}$-dependent interaction). The main point is that the traces of these orbits on configuration space are the same as in the nonrelativistic case.

In passing, note that if V were allowed to be negative, the classical relativistic trajectories would be "tachyonic" according to (II.12). Nevertheless, the wave equation (I.2) still obeys hyperbolic causality. This situation was studied by Schroer and others in the context of quantum field theory [17]. We shall not consider this case further, because it interferes with the definition of the square roots in (II.3) and (I.3).

In what follows we restrict attention to Hamiltonians of the form (II.10). Then the quantization prescription (II.8) leads unambiguously to a second-order elliptic partial differential operator

$$H = -\frac{1}{2m}\nabla^2 + V(\mathbf{x}). \tag{II.13}$$

(In more general cases some variant of Weyl quantization must be chosen to resolve operator-ordering ambiguities [18].) The complete specification of the problems (I.1)–(I.3) may also require boundary conditions, which we take to be of the standard Dirichlet, Neumann, or Robin type.

III. SCALING A BILLIARD

Henceforth we adopt the nonrelativistic, or Schrödinger-equation, point of view, and we restore the conventional constant $\hbar$. It is assumed that the spectrum of the operator H is discrete, at least below some threshold. The essence of the Gutzwiller trace theory [1, 3, 4, 5] is that the density of states, $\rho(E)$, contains a term

$$\rho_\gamma(E) = a_\gamma \sin\left[\frac{S_\gamma}{\hbar} + \eta_\gamma\right] \tag{III.1}$$

for each classical periodic orbit γ of energy E; here $S_\gamma(E)$ is the *action* $\oint \mathbf{p}\cdot d\mathbf{x}$ of the orbit γ, and the details of the amplitude $a_\gamma(E)$ and the phase η_γ do not concern us now. (Also, this simplified formulation does not do justice to the complications that arise when the orbits are not isolated. See [5, text surrounding (1.35)] for a summary, and [11] for a technical treatment.)

Let us examine what (III.1) asserts for a system of the type studied by Balian and Bloch [7]. The classical paths are straight lines in a region $\Omega \subset \mathbf{R}^d$ with specular reflections at the boundary of Ω. (The distinction among different boundary conditions shows up in the phases η_γ.) In particular, these paths are the *same* for all E. The speed is fixed at $v = \sqrt{2E/m}$ (see (II.6)). If the length of such a closed polygonal path, γ, is L_γ, then its period is

$$T_\gamma(E) = \frac{L_\gamma}{v} = \sqrt{\frac{m}{2E}}\, L_\gamma\,. \tag{III.2}$$

The momentum is parallel to the path with fixed magnitude $p = \sqrt{2mE}$, so the action is

$$S_\gamma(E) \equiv \oint_\gamma \mathbf{p}\cdot d\mathbf{x} = \sqrt{2mE}\, L_\gamma = 2ET_\gamma\,. \tag{III.3}$$

Thus (III.1) becomes

$$\rho_\gamma(E) = a_\gamma \sin\left[\frac{2T_\gamma E}{\hbar} + \eta_\gamma\right], \tag{III.4}$$

which is superficially a sinusoidal function of E with period $\pi\hbar/T_\gamma$. But because T_γ is itself a function of E, whereas L_γ is not, a more accurate picture is obtained by reexpressing (III.4) as

$$\rho_\gamma(E) = a_\gamma \sin\left[\frac{\sqrt{2mE}}{\hbar}\, L_\gamma + \eta_\gamma\right]. \tag{III.5}$$

Thus ρ_γ is exactly (to the extent that a_γ and η_γ can be regarded as constant) a sinusoidal function of $\omega \equiv \sqrt{E}$ with period

$$P_\omega = \sqrt{\frac{2}{m}}\, \frac{\pi\hbar}{L_\gamma}\,, \tag{III.6}$$

or $P_\omega = 2\pi/L_\gamma$ if $\frac{\hbar^2}{2m} = 1$, the most natural normalization when dealing with the wave equation. This is the result obtained by Balian and Bloch [7] working from the relativistic, or wave-equation, point of view: The eigenvalue density oscillates with frequency L_γ as a function of ω.

On the other hand, noting again that L_γ, not T_γ, is independent of E, we have

$$\frac{dS_\gamma}{dE} = \sqrt{\frac{m}{2E}}\, L_\gamma = T_\gamma = \frac{S_\gamma}{2E}\,. \tag{III.7}$$

Approximating $S_\gamma(E)$ locally by a linear function [1], one gets from (III.1) and (III.7)

$$\rho_\gamma(E) \approx a_\gamma \sin\left[\frac{T_\gamma}{\hbar}(E - E_0) + \frac{S_\gamma(E_0)}{\hbar} + \eta_\gamma\right]. \tag{III.8}$$

Thus ρ_γ oscillates *locally* as a function of E with frequency $T_\gamma(E_0)/\hbar$, or period

$$P_E = \frac{2\pi\hbar}{T_\gamma}\,. \tag{III.9}$$

(Note the factor of 2 discrepancy with our initial naive attempt.) This is the sort of statement about spectral oscillations that is most common in the physics literature surrounding the Gutzwiller trace formula (see the next section).

For completeness we record the time-dependent action

$$R_\gamma(T) = S_\gamma(E) - ET_\gamma = ET_\gamma = \frac{m{L_\gamma}^2}{2T}\,, \tag{III.10}$$

where

$$E = \tfrac{1}{2}mv^2 = \frac{mL^2}{T^2} \tag{III.11}$$

has been used. It satisfies

$$\frac{dR_\gamma}{dT} = -E, \tag{III.12}$$

which is dual to (III.7) when (III.10) is interpreted as a Legendre transformation. It can also be obtained as

$$R_\gamma \equiv \oint_\gamma \mathcal{L}\,dt = ET, \tag{III.13}$$

where $\mathcal{L}$ is the Lagrangian, here equal to the kinetic energy.

IV. INTERLUDE: THE ACTION–PERIOD RELATION

In converting (III.1) to (III.8), only the relation

$$\frac{dS_\gamma}{dE} = T_\gamma \tag{IV.1}$$

was used. This identity is frequently cited as "a well known theorem of classical mechanics", but a proof is hard to find in the literature.

Indeed, at second glance it is not even clear what (IV.1) *means* for a generic Hamiltonian (II.10), since S_γ was defined for a single closed orbit γ existing at a particular energy, say E_0. The discussion in Sec. III shows that for a billiard the orbit indeed persists unchanged (in configuration space) as E varies; but when the dynamics involves a potential function the classical path defined by an initial point and direction on γ will generally cease to be a periodic orbit when E deviates from E_0. It is reasonable to expect that generically γ will smoothly evolve with E into a family of nearby periodic orbits $\gamma(E)$ (in general, disjoint from $\gamma(E_0)$). But this picture obviously breaks down at a maximum of the potential in dimension 1 (more generally, at a separatrix), where two families of orbits merge into one. In higher dimensions, moreover, it is not obvious that the embedding of $\gamma(E_0)$ into a family is unique.

The main point of the remainder of this paper is that these existence and uniqueness questions can be avoided by adopting a broader point of view, in which other quantities are varied in addition to the energy. Here, however, we provide a simple derivation of (IV.1) in circumstances where it makes sense.

Assume the existence of a family $\gamma(E)$ of classical closed orbits of (II.1) depending smoothly on the energy, E (as E varies over some interval, possibly small). Assume also that the period T_γ depends smoothly and monotonically on E, so that γ can be alternatively parametrized by T. We shall prove

$$\frac{dR_\gamma}{dT} = -E, \tag{IV.2}$$

which is equivalent to (IV.1) under the Legendre transformation

$$S = R + ET \qquad (S = S(E), \quad R = R(T)). \tag{IV.3}$$

For a classical trajectory connecting two points, $\mathbf{x}$ and $\mathbf{x}'$, and depending on the elapsed time T as a parameter, (IV.2) is the usual Hamilton–Jacobi equation for R as a function of one endpoint (with $E = H(\mathbf{x}, \nabla R)$); but for periodic orbits a different argument is necessary.

By definition,

$$R = \oint_\gamma \mathbf{p} \cdot d\mathbf{x} - ET = \int_0^T \mathcal{L}\, dt, \tag{IV.4}$$

$$\mathcal{L} = \tfrac{1}{2} m\dot{\mathbf{x}}(t)^2 - V(\mathbf{x}(t)). \tag{IV.5}$$

In the standard way, consider a variation

$$\delta R = \int_0^T \frac{\delta\mathcal{L}}{\delta\mathbf{x}}\, \delta\mathbf{x}\, dt + \mathcal{L}|_T\, \delta t, \tag{IV.6}$$

where

$$\delta\mathcal{L} = m\dot{\mathbf{x}} \cdot \delta\dot{\mathbf{x}} - \nabla V \cdot \delta\mathbf{x} = \frac{d}{dt}(m\dot{\mathbf{x}} \cdot \delta\mathbf{x}) - (m\ddot{\mathbf{x}} + \nabla V) \cdot \delta\mathbf{x}. \tag{IV.7}$$

The last term in (IV.7) vanishes by the equation of motion, and integrating the other term yields

$$m\dot{\mathbf{x}}(T) \cdot \delta\mathbf{x}(T) - m\dot{\mathbf{x}}(0) \cdot \delta\mathbf{x}(0) = -m\dot{\mathbf{x}}(T)^2\, \delta T, \tag{IV.8}$$

because

$$\mathbf{x}(T) = \mathbf{x}(0), \quad \dot{\mathbf{x}}(T) = \dot{\mathbf{x}}(0), \quad (\mathbf{x} + \delta\mathbf{x})(T + \delta T) = (\mathbf{x} + \delta\mathbf{x})(0), \tag{IV.9}$$

so that

$$\begin{aligned} \delta\mathbf{x}(T) &= (\mathbf{x} + \delta\mathbf{x})(T) - \mathbf{x}(T) \\ &= (\mathbf{x} + \delta\mathbf{x})(T + \delta T) - (\dot{\mathbf{x}} + \delta\dot{\mathbf{x}})(T) - \mathbf{x}(0) \\ &= \delta\mathbf{x}(0) - \dot{\mathbf{x}}(T)\, \delta T + O((\delta T)^2). \end{aligned} \tag{IV.10}$$

Putting all the ingredients into (IV.6) yields

$$\begin{aligned} \frac{\delta R}{\delta T} &= -m\dot{\mathbf{x}}(T)^2 + \tfrac{1}{2} m\dot{\mathbf{x}}(T)^2 - V(\mathbf{x}(t)) \\ &= -\tfrac{1}{2} m\dot{\mathbf{x}}(T)^2 - V(\mathbf{x}(t)) \\ &= -E, \end{aligned}$$

as desired.

From (IV.2) follows (IV.1), and then (III.1) leads as before [1] to (III.8) and (III.9), under the additional assumption that T_γ is sufficiently slowly varying for the concept of a local frequency of oscillation of ρ_γ to make sense.

An interesting example is the one-dimensional harmonic oscillator,

$$V(x) = \tfrac{1}{2} m\varpi^2 x^2. \tag{IV.11}$$

The general solution of its equation of motion is

$$x(t) = x_0 \cos \varpi t + \frac{p_0}{m} \sin \varpi t. \tag{IV.12}$$

One calculates

$$T = \frac{2\pi}{\varpi}, \tag{IV.13}$$

$$E = \frac{p_0{}^2}{2m} + \tfrac{1}{2} m\varpi^2 x_0{}^2, \tag{IV.14}$$

$$S = ET = \frac{2\pi E}{\varpi}, \tag{IV.15}$$

$$R = 0. \tag{IV.16}$$

Thus

$$\frac{dS}{dE} = T = \frac{S}{E} \tag{IV.17}$$

in agreement with (IV.1), but in contrast to (III.7). Ironically, our derivation of (IV.1) does not apply to this case, because T is constant instead of parametrizing the orbits. In view of (IV.13) and (IV.14), (IV.2) holds only in the formal sense of $\frac{0}{0}$.

V. SCALING A COUPLING CONSTANT

Consider a Hamiltonian

$$H(\mathbf{x}, \mathbf{p}) = \frac{\mathbf{p}^2}{2m} + \lambda V(\mathbf{x}) \tag{V.1}$$

with a generic potential V and a coupling constant λ. The equations of motion are

$$\frac{d\mathbf{x}}{dt} = \frac{\mathbf{p}}{m}, \qquad \frac{d\mathbf{p}}{dt} = -\lambda \nabla V(\mathbf{x}), \tag{V.2}$$

and any solution curve lies on an energy surface

$$\frac{\mathbf{p}^2}{2m} + \lambda V(\mathbf{x}) = E. \tag{V.3}$$

Let $(\mathbf{x}_0(t), \mathbf{p}_0(t))$ be a solution of (V.2) and (V.3) with $\lambda = \lambda_0$ and $E = E_0$. For any positive real number α, consider

$$\mathbf{x}(t) \equiv \mathbf{x}_0(\alpha t), \qquad \mathbf{p}(t) \equiv \alpha \mathbf{p}_0(\alpha t). \tag{V.4}$$

A short calculation shows that $(\mathbf{x}(t), \mathbf{p}(t))$ satisfies (V.2) and (V.3) with

$$\lambda = \alpha^2 \lambda_0, \qquad E = \alpha^2 E_0. \tag{V.5}$$

In particular, the path $\mathbf{x}(t)$ in configuration space is independent of the parameter $\alpha^2 = E/E_0$. In this sense the same closed orbit exists for all values of the energy. When $V = 0$ this is the familiar billiard orbit reviewed in Sec. III. When the potential is not zero, however, one must pay the price of varying λ — i.e., considering *different* physical systems at different energies — to gain the convenience of fixed orbits.

For the orbit (V.4) it is clear that

$$T = \frac{T_0}{\alpha}\,, \tag{V.6}$$

and one calculates

$$\begin{aligned} S = \oint_\gamma \mathbf{p}\cdot d\mathbf{x} &= \alpha \int_0^T \mathbf{p}_0(\alpha t)\cdot[\alpha\dot{\mathbf{x}}_0(\alpha t)]\,dt \\ &= \alpha \int_0^{T_0} \mathbf{p}_0(\tau)\cdot\dot{\mathbf{x}}_0(\tau)\,d\tau = \alpha S_0\,. \end{aligned} \tag{V.7}$$

Thus (III.1) becomes (by use of (V.5))

$$\begin{aligned} \rho_\gamma(E) &= a_\gamma \sin\left[\frac{S_0}{\hbar}\sqrt{\frac{E}{E_0}} + \eta_\gamma\right] \\ &= a_\gamma \sin\left[\frac{S_0}{\hbar\omega_0}\,\omega + \eta_\gamma\right]. \end{aligned} \tag{V.8}$$

That is, just as for a billiard, we have a globally sinusoidal factor of $\omega \equiv \sqrt{|E|}$ with frequency $S_0/\hbar\omega_0$, or period

$$P_\omega = \frac{2\pi\hbar\omega_0}{S_0}\,. \tag{V.9}$$

Thus the analogue of the L_γ of a billiard is

$$\frac{S_0}{\sqrt{2m|E_0|}}\,. \tag{V.10}$$

For the harmonic oscillator orbit (IV.12), this characteristic length is π times the maximum value of $x(t)$ (not 4 times, which would be the actual length of the orbit). It should be noted that (V.9) or (V.10) depends only on the orbit γ, not on the reference scale arbitrarily chosen to correspond to $\alpha = 1$.

One can also look at the local frequency with respect to E, following (III.7)–(III.9). From (V.7) and (V.5) one has

$$\frac{dS}{dE} = \frac{S}{2E}\,. \tag{V.11}$$

But this quantity is no longer equal to T, because of the variation of the coupling constant with E. For the harmonic oscillator it equals $\frac{1}{2}T$, by (IV.17). More generally, since the momentum is parallel to the velocity for a system of type (V.1), one can compute

$$S = \oint_\gamma \mathbf{p}\cdot\dot{\mathbf{x}}\,dt = 2\int_0^T [E - \lambda V(\mathbf{x}(t))]\,dt, \tag{V.12}$$

and hence

$$\frac{S}{2E} = T - \frac{\lambda}{E}\int_0^T V(\mathbf{x}(t))\,dt. \tag{V.13}$$

Note also that $S/2E$ scales as $1/\alpha$ and therefore cannot be regarded as a characteristic time of the orbit γ (in contrast to the length (V.10)).

Finally, because nearby nonperiodic orbits also obey the scaling law, the amplitude a_γ in (V.8) (determined by the monodromy of the Poincaré map) is independent of E, whereas the same cannot be said for the a_γ in (III.8). The phase shift η_γ also is constant for Dirichlet and Neumann boundary conditions (not Robin).

VI. SCALING A HOMOGENEOUS POTENTIAL

When the potential in (V.1) is homogeneous of degree ν,

$$V(\beta\mathbf{x}) = \beta^\nu V(\mathbf{x}) \quad (\text{for } \beta > 0), \tag{VI.1}$$

an alternative scaling procedure exists that avoids changing the coupling constant. Assume that the configuration space Ω, if not all of $\mathbf{R}^d$, is invariant under dilations (e.g., a half-space or a cone) with boundary conditions that are either Dirichlet or pure Neumann (not Robin). (If Ω does not satisfy this condition, then it must be dilated along with the orbits, and again the system has been replaced by a one-parameter family.) In place of (V.4) consider

$$\mathbf{x}(t) \equiv \alpha^2\mathbf{x}_0(\alpha^{\nu-2}t), \qquad \mathbf{p}(t) \equiv \alpha^\nu \mathbf{p}_0(\alpha^{\nu-2}t). \tag{VI.2}$$

This is a solution of (V.2) and (V.3) with

$$\lambda = \lambda_0\,, \qquad E = \alpha^{2\nu}E_0\,. \tag{VI.3}$$

The orbits for different values of α are not the same, but they are geometrically similar; one can say that they are fixed if the unit of length is scaled by α^2.

We now have (parallel to (V.6)–(V.8))

$$\alpha = \left(\frac{E}{E_0}\right)^{\frac{1}{2\nu}}, \tag{VI.4}$$

$$T = \alpha^{2-\nu}T_0\,, \qquad S = \alpha^{\nu+2}S_0\,, \tag{VI.5}$$

$$\rho_\gamma(E) = a_\gamma \sin\left[\frac{S_0}{\hbar}\left(\frac{E}{E_0}\right)^{\frac{\nu+2}{2\nu}} + \eta_\gamma\right]. \tag{VI.6}$$

Thus the variable with respect to which the spectral oscillations take place is a peculiar power of E, and the analogue of the characteristic length (V.10) is a rather inconvenient function of the fiducial action and energy. Note that the exponent of $|E|$ in (VI.6) equals 1 for the harmonic oscillator; that it acquires the value $\frac{1}{2}$ (familiar for billiards) only as $\nu \to +\infty$; that it is negative for $-2 < \nu < 0$; and that the cases $\nu = 0$ and $\nu = -2$ are singular.

The analogue of (V.11) is

$$\frac{dS}{dE} = \frac{\nu+2}{2\nu}\,\frac{S}{E}\,, \tag{VI.7}$$

(V.12) remains true, and the analogue of (V.13) is

$$\frac{\nu+2}{2\nu}\,\frac{S}{E} = \frac{\nu+2}{\nu}\left[T - \frac{\lambda}{E}\int_0^T V(\mathbf{x}(t))\,dt\right]. \tag{VI.8}$$

On the other hand, the theorem of Sec. IV applies to the orbit family (VI.2), and so

$$\frac{dS}{dE} = T. \tag{VI.9}$$

Equating (VI.8) and (VI.9) yields the identity

$$\int_0^T \lambda V(\mathbf{x}(t))\, dt = \frac{2ET}{\nu+2} \tag{VI.10}$$

for a periodic orbit of a homogeneous potential of degree ν, which is a rewriting of the virial theorem [19].

The Coulomb potential,

$$V(\mathbf{x}) = -\frac{e^2}{r} \qquad (r^2 \equiv x^2 + y^2 + z^2), \tag{VI.11}$$

has

$$\nu = -1, \qquad \frac{\nu+2}{2\nu} = -\frac{1}{2}\,. \tag{VI.12}$$

Furthermore, in the regime of discrete spectrum and periodic orbits of the atom, E is negative. Thus (VI.7) becomes

$$\frac{dS}{dE} = +\frac{S}{2|E|}\,, \tag{VI.13}$$

which displays a superficial coincidence with the corresponding relation for a billiard, (III.7).

VII. SCALING OF MIXED TYPE

A mixture of the strategies of Secs. V and VI is appropriate in some situations. Suppose that the potential is a sum of two terms, each of which is homogeneous:

$$V_1(\beta\mathbf{x}) = \beta^{\nu_1} V_1(\mathbf{x}), \qquad V_2(\beta\mathbf{x}) = \beta^{\nu_2} V_2(\mathbf{x}). \tag{VII.1}$$

The equations of motion and the energy equation are

$$\frac{d\mathbf{x}}{dt} = \frac{\mathbf{p}}{m}\,, \qquad \frac{d\mathbf{p}}{dt} = -\lambda_1 \nabla V_1(\mathbf{x}) - \lambda_2 \nabla V_2(\mathbf{x}), \tag{VII.2}$$

$$\frac{\mathbf{p}^2}{2m} + \lambda_1 V_1(\mathbf{x}) + \lambda_2 V_2(\mathbf{x}) = E. \tag{VII.3}$$

(This is *not* a perturbative situation: Neither potential is assumed to be small compared to the other.) Following the usual pattern, suppose that we have a solution with coupling constants λ_{10} and λ_{20} and energy E_0. Perform the scaling (VI.2) appropriate to V_1:

$$\mathbf{x}(t) \equiv \alpha^2 \mathbf{x}_0(\alpha^{\nu_1 - 2} t), \qquad \mathbf{p}(t) \equiv \alpha^{\nu_1} \mathbf{p}_0(\alpha^{\nu_1 - 2} t). \tag{VII.4}$$

Then (VII.4) solves (VII.2) and (VII.3) with

$$\lambda_1 = \lambda_{10}\,, \qquad \lambda_2 = \alpha^{2(\nu_1 - \nu_2)} \lambda_{20}\,, \qquad E = \alpha^{2\nu_1} E_0\,. \tag{VII.5}$$

Equations (VI.4)–(VI.6) continue to apply, with $\nu = \nu_1$.

The point of this transformation is that a fixed orbit can be associated with a one-parameter family of classical situations, in which one coupling constant is held fixed but the other varies with energy. The prototype is the Hamiltonian studied by Friedrich and Wintgen [15, 3],

$$H(\mathbf{x}, \mathbf{p}) = \frac{\mathbf{p}^2}{2m} - \frac{e^2}{r} + \tfrac{1}{2} m\varpi^2 (x^2 + y^2). \tag{VII.6}$$

It describes, in a rotating coordinate system, a hydrogen atom in a constant magnetic field $B = (2mc/e)\varpi$ along the z axis. This is "a real physical system that can be and has been studied in the laboratory". The magnetic field is a continuous variable that is under the experimenter's control; the charge of the proton is not! Therefore, one applies (VII.1), (VII.4), (VII.5) with $\nu_1 = -1$, $\nu_2 = 2$. Thus

$$\lambda_2 = \alpha^{-6} \lambda_{20} \,, \qquad E = \alpha^{-2} E_0 \,, \tag{VII.7}$$

and the bottom-line equation (VI.6) becomes

$$\rho_\gamma(E) = a_\gamma \sin \left[\frac{S_0}{\hbar} \left(\frac{E}{E_0} \right)^{-1/2} + \eta_\gamma \right]. \tag{VII.8}$$

In the notation of Friedrich and Wintgen,

$$\Gamma \equiv \frac{\varpi}{\varpi_0} = \alpha^{-3} \,, \qquad \frac{E}{E_0} = \Gamma^{2/3}, \tag{VII.9}$$

and the contribution ρ_γ to the spectral density is a sinusoidal function of $\Gamma^{-1/3}$. (We write Γ and ϖ for the γ and ω of [3] to avoid notational collisions.) For the beautiful fruits of this approach as applied to (VII.6) we refer to the original papers [15, 3].

From $S = \alpha^{\nu_1 + 2} S_0$ and $E = \alpha^{2\nu_1} E_0$ we still get

$$\frac{dS}{dE} = \frac{\nu_1 + 2}{2\nu_1} \frac{S}{E} \,, \tag{VII.10}$$

but this is not equal to an orbital period in these incompletely homogeneous systems.

Mixed scaling could be applied to a potential with arbitrarily many terms, even an infinite series, but it would be necessary to rescale all but one of the coupling constants. The practical advantage over the procedure of Sec. V is thereby reduced.

VIII. IMPLICATIONS

A central feature of the relation between classical and quantum mechanics is that quantization introduces into each problem a new fundamental scale, set by the quantum of action, $\hbar$. That is, one-parameter families of situations that are equivalent classically become distinct quantum-mechanically. Taking the semiclassical limit refers to motion along one of these families in a certain direction, the opposite direction leading to "deep quantum" behavior. Moving from one family to another, on the other hand, corresponds to various purely classical distinctions, such as integrable versus chaotic, or different values of angular momentum.

When the potential is homogeneous (Sec. VI) or absent (Sec. III), a "situation" simply means a point in phase space. In the more general contexts of Secs. V and VII the space of situations should be enlarged by one dimension, representing the coupling constant λ or λ_2 . In either case the points are grouped into equivalence classes, the classical trajectories. Finally, the trajectories (including, in particular, the closed orbits) fall into families, (V.4) or (VI.2), related by a geometrical similarity and therefore mathematically equivalent. The classical trajectories of a family are exactly the same at all energies, except for a trivial rescaling. But the quantum states at various energies are quite different.

In the study of spectra the most convenient choice of parameter along each family is the energy. What determines whether an energy is large or small? To compare it with $\hbar$ one must construct a quantity with dimensions of action. The dimensions of $\sqrt{2m|E|}$ are action divided by length, so a suitable measure is the product of $\sqrt{2m|E|}$ with some length characteristic of the entire system. The statement "$\hbar$ is small" is meaningless unless translated into such a criterion. In the end one can always choose units in which $\hbar = 1$ and $2m = 1$, and then the only independent physical dimension is length. (Once the quantum of action and the characteristic length of the system are fixed, rescaling m amounts to changing the unit of time. Since m has been totally inert throughout all our considerations, it could have been eliminated at the very beginning, but so that the classical-mechanical equations would look familiar, we did not do so.) The criterion now is whether the length $|E|^{-1/2} \equiv \omega^{-1}$ is large or small relative to the length scale of the system. This restatement makes sense in classical wave theories, such as optics or acoustics, as originally studied by Balian and Bloch [20, 7].

The length scale can be set by the geometry of the boundary, if there is one. Otherwise it must be a characteristic of the potential. At first sight this is a simple matter: Quantities such as $\nabla^2 V/V$ (evaluated at a minimum of the potential, say) are independent of the coupling constant and characterize the spatial scale of the potential, as distinct from its strength. However, if V is homogeneous, (VI.1) shows that the distinction between the spatial scale and the coupling constant is a mirage; and Sec. VII suggests that this ambiguity can be imported into more general potentials as well. For a homogeneous potential there is an alternative way to set the scale of the system, exploiting (VI.1) to "transmute" the coupling constant into a length, $x_0 \equiv \beta^{-1}$. For example, if $d = 1$ and the potential energy is λx^ν, then λ has dimensions [energy][length]$^{-\nu}$ and can be written

$$\lambda \equiv \frac{\hbar^2}{2m} x_0{}^{-(\nu+2)}, \tag{VIII.1}$$

so that

$$\lambda x^\nu = \frac{\hbar^2}{2m x_0^2} \left(\frac{x}{x_0}\right)^\nu . \tag{VIII.2}$$

It is instructive to look at everyone's favorite homogeneous potentials, the harmonic oscillator (IV.11) and the hydrogen atom (VI.11), in the light of the foregoing remarks. If we treat these systems according to Sec. V, we are taking a fixed "classical situation" to be a fixed ratio E/λ, with $\lambda \propto \varpi^2$ or e^2 respectively. We expect to encounter semiclassical behavior as $|E| \to \infty$ along one of the diagonal lines in Figs. 1 and 2, and indeed this leads into the region of large quantum numbers in each case. There is no unique way to associate a natural length with either of these potentials, but lengths independent of λ can be built out of derivatives of V at an arbitrarily chosen point. If we treat the systems according to

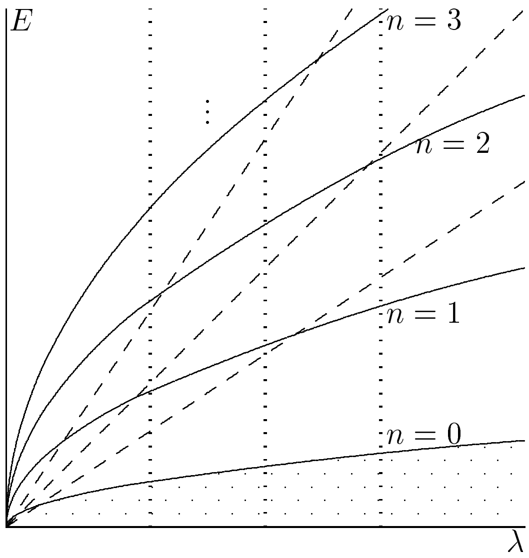

FIG. 1: Solid curves: Energy levels of the harmonic oscillator, $E_n = (n + \frac{1}{2})\hbar\varpi$, as functions of coupling constant, $\lambda \propto \varpi^2$. No quantum states exist in the shaded area. Dashed lines: Loci of typical classical orbits fixed in space (Sec. V). Dotted lines: Loci of typical homogeneously scaled orbits (Sec. VI). The direction of increasing E along either set of lines leads into the semiclassical region of large quantum numbers.

Sec. VI, we keep λ constant and therefore move along one of the vertical lines in the figures. From

$$x_0 = \lambda^{-1/(\nu+2)} \qquad (\hbar^2 = 2m), \tag{VIII.3}$$

we have

$$x_0^2 = \frac{1}{\varpi}, \qquad E_n x_0^2 = n + \frac{1}{2} \tag{VIII.4}$$

for the energy levels of the oscillator, and

$$x_0^2 = e^{-4}, \qquad E_n x_0^2 \propto -\frac{1}{n^2} \tag{VIII.5}$$

for those of the atom. The semiclassical regime of large quantum number n is reached by going vertically upward in either of the figures. For the oscillator this is again the limit of large E, albeit in a different direction in the E–λ plane. But for the Coulomb potential this limit corresponds to $E \to 0$, not $|E| \to \infty$. (In view of (VI.13) this is still the direction of increasing S.) This result perhaps should not be surprising in view of the negative exponent in (VI.6), (VI.13). But it indicates that the identification of "semiclassical" with "large energy" can be ambiguous, even within the theory of the same physical system.

Let us summarize the main points:

1. The concept of "spectral oscillations" (associated with the basic formula (III.1), which we took as given) has meaning only if the orbit structure is at least locally independent of energy.

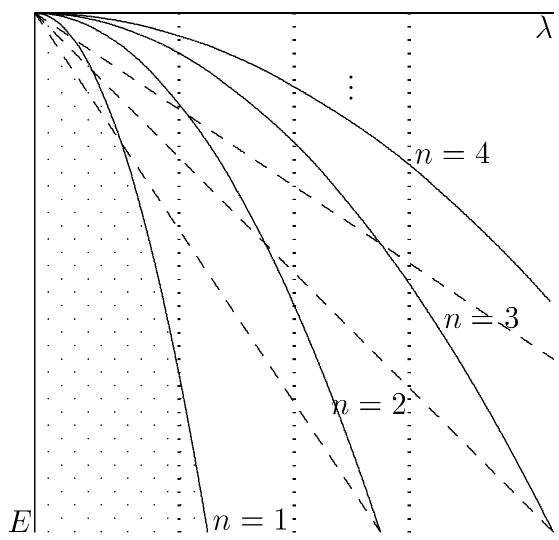

FIG. 2: Solid curves: Energy levels of the hydrogen atom, $E_n \propto -e^4/n^2$, as functions of coupling constant, $\lambda \propto e^2$. No quantum states exist in the shaded area. Dashed lines: Loci of typical classical orbits fixed in space (Sec. V). Dotted lines: Loci of typical homogeneously scaled orbits (Sec. VI). The semiclassical region of large quantum numbers is reached by moving along the dashed lines in the direction of increasing $|E|$ ($E \to -\infty$), or along the dotted lines in the direction of decreasing $|E|$ ($E \to 0$).

2. *The orbits can be "nailed down" by scaling the coupling constant along with E.*

3. For a homogeneous or partially homogeneous potential there are other options for fixing the orbit structure, involving dilations.

4. The wave frequency $\omega = E^{1/2}$ is a better variable than E for describing the oscillations globally. (In the homogeneous case the exponent $\frac{1}{2}$ is replaced by $(\nu + 1)/2\nu$.) Correspondingly, the best descriptor of the oscillations caused by a particular orbit is a characteristic length (V.10) of the orbit, not the orbital period.

5. Even the local frequency of spectral oscillation with respect to energy, dS/dE, is not equal to the orbital period when the coupling constant is varied. (In passing we supplied a proof of the equality when λ is constant.)

IX. A FINAL COMMENT

When periodic-orbit theory was first applied numerically to the spectra of concrete systems, early authors expressed surprise that it gave accurate results "even" for the lowest eigenvalues. As a semiclassical method, the technique had been expected to be applicable primarily in the regime of “large quantum numbers” and hence high energy. Later it became clear that the method was generally practical *only* in the low-energy regime, because the number of periodic orbits increases roughly exponentially with period, and as energy increases it becomes necessary to consider increasingly long orbits in order to resolve individual eigenvalues [4, 5]. Long after this fact has been accepted, it is still often regarded as “paradoxical”.

We suggest that, in hindsight, this phenomenon is merely an instance of the familiar principle that when a calculation is very stable and involves some kind of smoothing or averaging, the inverse calculation is likely to be unstable (highly dependent on the details of the input) and hence difficult. There is a loose analogy with the solution of an initial-value problem for the heat equation. For large positive time the problem is very easy to solve and very insensitive to the details of the initial data; consequently, reconstructing the data from the final solution is hopeless. For very small time an approximate solution to the backwards problem, adequate for some practical purpose, may be feasible; of course, the forward problem is less trivial in that case.

In spectral asymptotics we are interested in deducing the spectrum of a differential operator from its geometry, or vice versa. Here “geometry” is meant in an extended sense, including not only the literal geometry of the region Ω where the wave functions are defined, but also the potential function (or other coefficient functions in the operator H). The classical periodic orbits of the Hamiltonian $H(\mathbf{x}, \mathbf{p})$ are also aspects of the geometry of the system.

In the “old” spectral asymptotics, associated with the names Weyl, Thomas–Fermi, Minakshisundaram, Schwinger–DeWitt, etc., one associated the (high-E) asymptotic behavior of the density of states with the global geometry of the operator: the volume of Ω, the integrated curvature of its boundary, the integral of the Ricci curvature of Ω if it is a manifold, the integral of the potential V over Ω, and so on. Via the asymptotics of the heat kernel, the passage from the spectrum to the geometry is rigorously asymptotic. In the inverse direction, however, the geometry does not determine a genuine asymptotic expansion of the eigenvalue density in powers of E^{-1}, precisely because of the presence of the oscillatory terms that are the subject of the periodic-orbit theory. The formal expansion of the eigenvalue density becomes literally asymptotic only when some kind of averaging is performed, such as Lorentzian smoothing [20] or Riesz means [21].

A complementary situation has always existed at the low end of the spectrum. If we know Ω and H exactly, then it is relatively easy to construct the lowest-lying eigenvalues and eigenfunctions, by variational methods, numerical methods, etc. Larger eigenvalues are harder. Furthermore, merely from a knowledge of the lowest energies one would not expect to be able to reconstruct Ω or H. (Complete knowledge of a single eigen*function* is a different matter, however.)

The “new” spectral asymptotics of periodic orbits extends this picture. Decreasing the width over which the eigenvalue distribution is averaged, one supplements the power-law asymptotics with the longest-wavelength oscillatory components (III.1), characterized by the lengths or periods of the shortest classical orbits. This is still an asymptotic (high-energy) matter. It is best thought of as a prediction of classical behavior (well-defined orbits for wave packets) in the high-energy, large-action regime. It still involves an averaging of the spectrum, albeit on a smaller scale. The stable direction of prediction is *from* the spectrum *to* the classical, geometrical description.

The counterpart of this at low energy is that from the classical orbits (and their associated amplitudes and phases, cf. (III.1)) one can predict the formation of discrete eigenvalues (resonant frequencies), with greatest precision near the bottom of the spectrum. In some sense this involves an averaging over geometrical information. More precise geometrical information in principle allows precise construction of longer-period orbits and hence more complete reconstruction of the spectrum (both improved resolution and extension to higher energies). The stable direction of prediction is *from* the geometry *to* the spectrum.

It is misleading to think of the regime of low energy (or small quantum numbers) as the "deep quantum regime", as if the duality between the spectrum of quantum eigenvalues and the spectrum of classical periodic orbits is irrelevant there. A better phrase is "resonant regime": There, resonant behavior emerges on the spectral side, out of the geometry and classical mechanics, much as, in the opposite "classical regime", classical-mechanical behavior emerges on the geometrical side, out of the quantum substrate.

ACKNOWLEDGMENTS

The calculation in Sec. IV developed in conversations with David Garfinkle at the Parkerfest (Oakland University, 2000).

REFERENCES

[1] J. B. Delos and M.-L. Du, *IEEE J. Quantum Electron.* **24**, 1445 (1988).
[2] M. Combescure, J. Ralston, and D. Robert, *Commun. Math. Phys.* **202**, 463 (1999).
[3] H. Friedrich and D. Wintgen, *Phys. Reports* **183**, 37 (1989).
[4] M. Berry, in *Chaos et Physique Quantique / Chaos and Quantum Physics*, Les Houches Session LII, 1989, ed. by M.-J. Giannoni, A. Voros, and J. Zinn-Justin, Elsevier, 1991, pp. 251–303.
[5] J. Bolte, *Semiclassical Trace Formulae and Eigenvalue Statistics in Quantum Chaos*, Lectures at the Third International Summer School/Conference at the University of Maribor, Slovenia, 1996 (chao-dyn/9702003).
[6] D. Robert, *Helv. Phys. Acta* **71**, 44 (1998).
[7] R. Balian and C. Bloch, *Ann. Phys. (NY)* **69**, 76 (1972).
[8] M. C. Gutzwiller, *J. Math. Phys.* **11**, 1791 (1970).
[9] M. C. Gutzwiller, *J. Math. Phys.* **12**, 343 (1971).
[10] R. Balian and C. Bloch, *Ann. Phys. (NY)* **85**, 514 (1974).
[11] J. J. Duistermaat and V. W. Guillemin, *Invent. Math.* **29**, 39 (1975).
[12] J. Chazarain, *Invent. Math.* **24**, 65 (1974).
[13] Y. Colin de Verdiére, *Comp. Math.* **27**, 159 (1973).
[14] M. Cotsaftis, *Compt. Rend. Acad. Sci. Paris* **274**, A911 (1973).
[15] D. Wintgen and H. Friedrich, *Phys. Rev. A* **36**, 131 (1987).
[16] Yu. Safarov and D. Vassiliev, *The Asymptotic Distribution of Eigenvalues of Partial Differential Operators*, Transl. Math. Monog. 155, American Mathematical Society, 1997.
[17] B. Schroer, *Phys. Rev. D* **3**, 1764 (1971).
[18] S. A. Fulling, in *Relativity, Particle Physics and Cosmology*, Proceedings of the Richard Arnowitt Fest, ed. by R. E. Allen, World Scientific, 1999, pp. 329–342.
[19] H. Goldstein, *Classical Mechanics*, Addison–Wesley, 1950, Sec. 3.4.

[20] R. Balian and C. Bloch, *Ann. Phys. (NY)* **60**, 401 (1970).
[21] L. Hörmander, *Acta Math.* **121**, 193 (1968).