# Partial-differential-equation-based approach to classical phase-space deformations

Emmanuel Tannenbaum

*Harvard University, Cambridge, Massachusetts 02138*

This paper presents a partial-differential-equation-based approach to finding an optimal canonical basis with which to represent a nearly integrable Hamiltonian. The idea behind the method is to continuously deform the initial canonical basis in such a way that the dependence of the Hamiltonian on the canonical position of the final basis is minimized. The final basis incorporates as much of the classical dynamics as possible into an integrable Hamiltonian, leaving a much smaller nonintegrable component than in the initial representation. With this approach it is also possible to construct the semiclassical wave functions corresponding to the final canonical basis. This optimized basis is potentially useful in quantum calculations, both as a way to minimize the required size of basis sets, and as a way to provide physical insight by isolating those effects resulting from integrable dynamics.

## I. INTRODUCTION

Suppose we are given a nearly integrable Hamiltonian $H(\mathbf{q}, \mathbf{p})$, where $(\mathbf{q}, \mathbf{p})$ represents some canonical representation of phase space (not necessarily ordinary position and momentum). We wish to find another canonical representation $(\mathbf{Q}, \mathbf{P})$ in which $H$ is as close as possible to being integrable. Our motivation for this is twofold, and is connected to semiclassical quantum mechanics. First of all, from a purely numerical perspective, a representation in which $H$ is as close as possible to being integrable leads to an optimized semiclassical basis with which to perform quantum calculations. The reason for this is that we can associate with $\mathbf{P}$ a quantum state $|\mathbf{P}\rangle$. If we write $H(\mathbf{Q}, \mathbf{P}) = H^{(0)}(\mathbf{P}) + H^{(1)}(\mathbf{Q}, \mathbf{P})$, then it is clear that $H^{(0)}$ is diagonal in the $\{|\mathbf{P}\rangle\}$ basis, so that only $H^{(1)}$ is available to couple the various basis states. The coupling is given semiclassically by [1],

$$\langle \mathbf{P}' | \hat{H}^{(1)} | \mathbf{P} \rangle = H^{(1)}_{(\mathbf{P}' - \mathbf{P})/2\pi\hbar} \left( \frac{\mathbf{P} + \mathbf{P}'}{2} \right), \qquad (1)$$

where $H^{(1)}_{(\mathbf{P}' - \mathbf{P})/2\pi\hbar}((\mathbf{P} + \mathbf{P}')/2)$ is the $(\mathbf{P}' - \mathbf{P})/2\pi\hbar$ Fourier component at $(\mathbf{P} + \mathbf{P}')/2$ of $H^{(1)}$. A representation in which $H^{(1)}$ is as small as possible will minimize the couplings in the corresponding $\{|\mathbf{P}\rangle\}$ basis, and thus will minimize the size of the basis required to perform a given calculation to some desired accuracy.

Secondly, a canonical representation in which $H$ is as close as possible to being integrable provides physical insight. By incorporating as much of the classical dynamics as possible into an integrable Hamiltonian, this optimal canonical representation can help to isolate those classical and quantum effects resulting from integrable dynamics from those that do not. Thus this representation can isolate nonintegrable quantum and classical effects such as dynamical tunneling and Arnol'd diffusion, respectively. Furthermore, we can also visualize the quantum manifestation of the integrable classical dynamics via the semiclassical prescription for constructing a wave function. Given the generating func-

tion $S(\mathbf{q}, \mathbf{P})$ from the initial basis $(\mathbf{q}, \mathbf{p})$ to the final basis $(\mathbf{Q}, \mathbf{P})$, we have, up to normalization,

$$\langle \mathbf{q} | \mathbf{P} \rangle = \left| \det \frac{\partial^2 S}{\partial \mathbf{q} \partial \mathbf{P}} \right|^{1/2} \exp[iS/\hbar], \qquad (2)$$

where $\det(\partial^2 S / \partial \mathbf{q} \partial \mathbf{P})$ is the well-known Van Vleck determinant.

In the case of action-angle variables, the optimized representation of a nearly integrable Hamiltonian is termed an intrinsic resonance representation (IRR), a term coined by Carioli, Heller, and Moller (CHM) [2,3]. In 1997 they published a paper detailing an algorithm for the construction of such a representation. The idea behind the CHM algorithm is to eliminate all the nonresonant terms of the Hamiltonian via an appropriate canonical transformation. This canonical transformation is obtained via a modified Chapman, Garrett, and Miller method [2–4], which is essentially a Newton-Raphson scheme to find the invariant tori with a desired set of actions for a nearly integrable system. The remaining resonant and near-resonant terms are then reexpressed in the new basis. It is impossible to reduce the angle dependence any further, since this would result in the formation of resonance zones, which prevents a global action-angle description of the Hamiltonian.

In 2001 Tannenbaum and Heller published an alternative algorithm for finding the IRR basis [2]. The method introduced is a partial-differential-equation- (PDE-) based approach which continuously deforms the initial action-angle basis in such a way that the angle dependence of the Hamiltonian is continuously reduced. It amounts to a gradient-descent algorithm in the limit of a first-order perturbation, and was therefore called the GDA method. Formally, the method does not distinguish between resonant, nearly resonant, and nonresonant terms, that is, the evolution is performed on the entire Hamiltonian without any terms neglected. However, the evolution is such that the more nonresonant a term, the more strongly it is affected by the evolution. Thus, the nonresonant terms of the Hamiltonian are essentially removed, the nearly resonant terms are reduced somewhat, while the resonant terms are essentially

unaffected. In Ref. [2], the GDA method was used to semi-classically find the eigenvalues of systems with two, four, and six degrees of freedom, using greatly reduced basis sets.

The GDA method circumvents two main drawbacks of the CHM method. First, the CHM method requires an *a priori* decision as to which terms are resonant and nonresonant. This leads to an ambiguity in the case of near-resonances. It could happen that a given term in the Hamiltonian must be considered resonant in order to get the Newton-Raphson scheme to converge. This leads to a somewhat artificial cut-off criterion, since a nearly resonant Fourier component should in principle still be reduced as much as possible, although not necessarily completely. Thus, unless the Hamiltonian has a few exact or near-resonances, it is not clear that the CHM method will give the optimized torus basis.

Second, the CHM method requires the numerical evaluation of multidimensional integrals, and the numerical inversion of a nonlinear angle map at every iteration step. Furthermore, every iteration step also requires the numerical solution of a system of linear equations. These numerical calculations slow the algorithm down. In contrast, the numerical calculations required by the GDA method are much simpler, so we believe that the GDA method is faster than the CHM approach (although in fairness it should be added that no direct speed comparisons have been made to date).

Despite its advantages, the GDA method, as described in Ref. [2], does not give the overall generating function $S(\mathbf{q},\mathbf{P})$ transforming from the initial $(\mathbf{q},\mathbf{p})$ basis to the final $(\mathbf{Q},\mathbf{P})$ basis. Thus, in Ref. [2], we could compute the energy spectrum arising from an optimized invariant torus basis, but not the corresponding semiclassical wave functions existing on the IRR tori.

This paper is a continuation of the work presented in Ref. [2], and it has several purposes: First, the presentation of the method given in Ref. [2] has been greatly simplified. Secondly, and more importantly, the method has been extended to include the determination of the overall generating function $S(\mathbf{q},\mathbf{P})$. This provides a powerful visualization tool which allows one to actually construct the semiclassical wave functions associated with the optimized canonical basis. Thus, while the full method will be developed in this paper, the numerical examples focus on the overall generating function $S(\mathbf{q},\mathbf{P})$. As mentioned before, numerical examples dealing with the Hamiltonian directly may be found in Ref. [2].

This paper is organized as follows. In Sec. II we obtain the generic evolution equations for the Hamiltonian and the overall generating function, starting from an arbitrary canonical basis $(\mathbf{q},\mathbf{p})$. In Sec. III we consider the case of a nearly integrable Hamiltonian, and obtain the specific form our evolution is to take if we want to optimize the canonical basis used to represent the Hamiltonian. A similar, yet more detailed, derivation may be found in Ref. [2]. In Sec. IV we consider the first-order limit of our PDE approach. We obtain a generalized first-order classical perturbation theory which coincides with standard perturbation theory in an appropriate limit, obtained by taking the evolution parameter to infinity. However, the advantage of our first-order formula is that it remains convergent for all finite values of the evolution pa-

rameter, only possibly diverging as a result of near resonances in the limit to infinity. Section V tests the $S$ evolution equation with a numerical example. We continue in Sec. VI by applying the $S$ evolution equation to actually construct and visualize some semiclassical wave functions associated with optimized canonical bases. Finally, we conclude in Sec. VII with a summary of our results and a discussion of future research plans.

## II. THE EVOLUTION EQUATIONS

In this section we shall obtain the basic evolution equations for the Hamiltonian $H$ and the overall generating function $S$. The idea is as follows. We start with an initial set of canonical coordinates $(\mathbf{q},\mathbf{p})$, which denote any global representation of phase space, and do not necessarily refer to ordinary position and momentum, respectively. (Nonglobal representations can also work, as long as we remain well within the region of phase space they describe. An example of this is action-angle variables for a system that can dissociate. Action-angle variables should work as a valid representation as long as we remain well below any dissociation threshold.) We continuously deform this system via a series of infinitesimal generating functions. The result is that our canonical representation is evolving with respect to some parameter $t$, and is denoted by $(\mathbf{Q}_t,\mathbf{P}_t)$ at parameter value $t$. For convenience, we will call our parameter $t$ the "time," although it should be understood that it does not represent time in the ordinary sense, but is rather a homotopy parameter. As our canonical pair evolves, the functional dependence of $H$ on the canonical pair changes. In addition, the overall generating function $S(\mathbf{q},\mathbf{P};t)$ connecting the initial $(\mathbf{q},\mathbf{p})$ to the current $(\mathbf{Q}_t,\mathbf{P}_t)$ evolves as well.

Consider an arbitrary set of canonical coordinates. At time $t$, we are at system $(\mathbf{Q}_t,\mathbf{P}_t)$. At time $t+dt$, we are at system $(\mathbf{Q}_{t+dt},\mathbf{P}_{t+dt})$. These are connected by an infinitesimal generating function $F(\mathbf{Q}_t,\mathbf{P}_{t+dt};t)=\mathbf{Q}_t\cdot\mathbf{P}_{t+dt}+dtG(\mathbf{Q}_t,\mathbf{P}_{t+dt};t)$. Therefore,

$$\mathbf{Q}_{t+dt}=\mathbf{Q}_t+dt\boldsymbol{\nabla}_{\mathbf{P}}G(\mathbf{Q}_t,\mathbf{P}_{t+dt};t), \qquad (3)$$

$$\mathbf{P}_t=\mathbf{P}_{t+dt}+dt\boldsymbol{\nabla}_{\mathbf{Q}}G(\mathbf{Q}_t,\mathbf{P}_{t+dt};t). \qquad (4)$$

These equations are just Hamilton's equations of motion for the dynamics governed by $G$, which plays the formal role of a Hamiltonian in this case. The functional dependence of $H$ therefore evolves according to

$$\frac{\partial H}{\partial t}+\{H,G\}=0. \qquad (5)$$

This equation is simply a statement that $H$ is invariant under the canonical transformation induced by $F(dH/dt=0)$, which follows from the fact that $F$ is a time-independent canonical transformation [5].

For what follows in this paper it will prove convenient to represent the dynamics in Fourier space. To this end, assume that $H$ is periodic in each $Q_i$ with period $L_i$. Then we shall choose $G$ to also be periodic in each $Q_i$ with period $L_i$.

Define $V = L_1 \cdots L_D$, and let $\Omega$ denote an arbitrary $D$-dimensional box of side lengths $L_1, \ldots, L_D$. Then,

$$H(\mathbf{Q},\mathbf{P};t) = \frac{1}{V} \sum_{\mathbf{k}} H_{\mathbf{k}}(\mathbf{P};t) e^{2\pi i \mathbf{k} \cdot \mathbf{Q}}, \tag{6}$$

where

$$H_{\mathbf{k}}(\mathbf{P};t) \equiv \int_{\Omega} d\mathbf{Q} H(\mathbf{Q},\mathbf{P};t) e^{-2\pi i \mathbf{k} \cdot \mathbf{Q}}, \tag{7}$$

and similarly for $G(\mathbf{Q},\mathbf{P};t)$. Our componentwise evolution is then

$$\frac{\partial H_{\mathbf{k}}}{\partial t} = 2\pi i \frac{1}{V} \sum_{\mathbf{k}'} \{ (\mathbf{k}' \cdot \nabla_{\mathbf{P}} H_{\mathbf{k}-\mathbf{k}'}) G_{\mathbf{k}'}$$
$$- [(\mathbf{k}-\mathbf{k}') \cdot \nabla_{\mathbf{P}} G_{\mathbf{k}'}] H_{\mathbf{k}-\mathbf{k}'} \}. \tag{8}$$

Note that the evolution preserves the integration limits of our original system. Thus the topology of the original phase space is preserved.

It should be mentioned that degrees of freedom for which $L_i$ is finite can be treated in an action-angle formalism in which $L_i = 1$, while degrees of freedom for which $L_i = \infty$ have their corresponding Fourier sums replaced by integrals.

Finally, the evolution of the generating function $S(\mathbf{q},\mathbf{P};t)$ is given by the Hamilton-Jacobi PDE [5–7]

$$\frac{\partial S}{\partial t} = G\left( \frac{\partial S}{\partial \mathbf{P}}, \mathbf{P}; t \right). \tag{9}$$

### III. CHOOSING $G$

We want an approach that minimizes the dependence of $H$ on $\mathbf{Q}$. The condition that $H$ be independent of $\mathbf{Q}$ is equivalent to the vanishing of the nonzero Fourier components $H_{\mathbf{k}}(\mathbf{P};t)$. We therefore seek to minimize the dependence of $H$ on $\mathbf{Q}$ by choosing $G$ in such a way that the $|H_{\mathbf{k}}(\mathbf{P};t)|$ are continuously decreasing for all $\mathbf{k} \neq \mathbf{0}$.

At some time $t$, we can write $H(\mathbf{Q},\mathbf{P};t) = H^{(0)}(\mathbf{P};t) + H^{(1)}(\mathbf{Q},\mathbf{P};t)$. The idea is that $H^{(0)}$ contains the piece of the Hamiltonian that is dependent only on $\mathbf{P}$, and $H^{(1)}$ contains the remainder. Then,

$$\frac{\partial H}{\partial t} = \nabla_{\mathbf{P}} H^{(0)} \cdot \nabla_{\mathbf{Q}} G - \{H^{(1)}, G\}. \tag{10}$$

In the limit of a first-order perturbation on an integrable Hamiltonian, the relevant equation is

$$\frac{\partial H}{\partial t} = \nabla_{\mathbf{P}} H^{(0)} \cdot \nabla_{\mathbf{Q}} G. \tag{11}$$

In Fourier space, this becomes

$$\frac{\partial H_{\mathbf{k}}}{\partial t} = 2\pi i (\mathbf{k} \cdot \nabla_{\mathbf{P}} H^{(0)}) G_{\mathbf{k}}. \tag{12}$$

Following a similar derivation to the one in Ref. [2], it may be shown that, in the first-order limit, the gradient-descent prescription for minimizing $|H_{\mathbf{k}}|^2 = H_{\mathbf{k}} \bar{H}_{\mathbf{k}}, \mathbf{k} \neq \mathbf{0}$, is to set $G_{\mathbf{k}} = 2\pi i (\mathbf{k} \cdot \nabla_{\mathbf{P}} H^{(0)}) H_{\mathbf{k}}$. Then in the first-order limit we obtain that $\partial H_{\mathbf{k}} \bar{H}_{\mathbf{k}} / \partial t = -8\pi^2 (\mathbf{k} \cdot \nabla_{\mathbf{P}} H^{(0)})^2 H_{\mathbf{k}} \bar{H}_{\mathbf{k}}$, which is clearly negative. For stronger perturbations, this is no longer the gradient-descent prescription. However, for nearly integrable systems (the ones of interest to us in this paper) the perturbation should still be sufficiently weak that the above choice for $G_{\mathbf{k}}$ will shrink the $H_{\mathbf{k}} \bar{H}_{\mathbf{k}}, \mathbf{k} \neq \mathbf{0}$. Therefore, we take $G_{\mathbf{k}} = 2\pi i (\mathbf{k} \cdot \nabla_{\mathbf{P}} H^{(0)}) H_{\mathbf{k}}$, which gives

$$G(\mathbf{Q},\mathbf{P};t) = \nabla_{\mathbf{P}} H^{(0)} \cdot \nabla_{\mathbf{Q}} H = \nabla_{\mathbf{P}} H^{(0)} \cdot \nabla_{\mathbf{Q}} H^{(1)}. \tag{13}$$

Note that the Fourier expansion of $G$ involves terms of the form $\mathbf{k} \cdot \nabla_{\mathbf{P}} H^{(0)}$. A $\mathbf{k}$ for which $\mathbf{k} \cdot \nabla_{\mathbf{P}} H^{(0)} = 0$ is a generalized resonance at $\mathbf{P}$. The integrability of $H^{(0)}$ is destroyed by the resonant terms in $H^{(1)}$. It should also be pointed out here that $G$ has units of $[\text{energy}]^2/[\text{action}] = [\text{power}]$, and so $t$ has units of $[\text{time}]^2$.

Our evolution does not formally distinguish between resonances, near-resonances, and nonresonances. The evolution is done on the entire Hamiltonian without any terms neglected. However, the closer a term is to being resonant, the smaller the corresponding Fourier component of $G$, and so the less that term is affected by the evolution.

In the first-order limit, $H^{(0)}(\mathbf{P};t)$ differs from $H^{(0)}(\mathbf{P};0)$ by a correction that is at most first order in $H^{(1)}$. Therefore, if we use $H^{(0)}(\mathbf{P};0)$ instead of $H^{(0)}(\mathbf{P};t)$ in our prescription for choosing $G$ in Eq. (13), we get a discrepancy of at most second order in $H^{(1)}$, so that the two formulations are equivalent to first order. Since our prescription for choosing $G$ was derived from the first-order limit of the evolution of $H$, we see that it is equivalent to use $H^{(0)}(\mathbf{P};0)$ or $H^{(0)}(\mathbf{P};t)$ in Eq. (13). Finally, our $t = 0$ Hamiltonian is usually given as $H(\mathbf{q},\mathbf{p}) = H_0(\mathbf{p}) + V(\mathbf{q},\mathbf{p})$, where $H_0$ is the zeroth-order, integrable Hamiltonian, and $V$ is the perturbation. We can extract the $\mathbf{k} = \mathbf{0}$ Fourier component of $V$, writing $V(\mathbf{q},\mathbf{p}) = V_0(\mathbf{p}) + \tilde{V}(\mathbf{q},\mathbf{p})$. Then $H(\mathbf{Q},\mathbf{P};0) = H_0(\mathbf{P}) + V_0(\mathbf{P}) + \tilde{V}(\mathbf{Q},\mathbf{P})$, so that $H^{(0)}(\mathbf{P};0) = H_0(\mathbf{P}) + V_0(\mathbf{P})$, and $H^{(1)}(\mathbf{Q},\mathbf{P};0) = \tilde{V}(\mathbf{Q},\mathbf{P})$. Therefore, note that in the first-order limit it is equivalent to use $H^{(0)}(\mathbf{P};0)$ or $H_0(\mathbf{P})$ in the prescription for choosing $G$. Once again, this means that it is equivalent to use $H^{(0)}(\mathbf{P};0)$ or $H_0(\mathbf{P})$. In what follows $H^{(0)}(\mathbf{P};t)$, $H^{(0)}(\mathbf{P};0)$, and $H_0(\mathbf{P})$ will all be denoted by $H^{(0)}$, or $H^{(0)}(\mathbf{P})$. When required, we will specify to which $H^{(0)}$ we are referring. Very often, one of the three choices will result in a formulation of the PDEs that is considerably simpler to implement than the others. For example, in Ref. [2] we used the $H^{(0)}(\mathbf{P};t)$ formulation to evolve the Fourier components of $H$, but neglected terms involving second derivatives in $\mathbf{P}$, because they were considered small enough to be unimportant. Thus we solved an approximate system of PDEs instead of the full, exact set. However, had we instead used the $H_0(\mathbf{P})$ formulation, then the resulting evolution would have required only first derivatives in $\mathbf{P}$, so with slight modification the system of approximate PDEs solved in Ref. [2] would have become exact.

We conclude this section by deriving the PDE governing the evolution of $S$, given our prescription for choosing $G$. We have

$$\frac{\partial S}{\partial t}(\mathbf{q},\mathbf{P};t) = G(\mathbf{Q}(\mathbf{q},\mathbf{P};t),\mathbf{P};t)$$

$$= \nabla_{\mathbf{P}}H^{(0)}\cdot\nabla_{\mathbf{Q}}H(\mathbf{Q}(\mathbf{q},\mathbf{P};t),\mathbf{P};t). \quad (14)$$

Now, we know that

$$H(\mathbf{Q},\mathbf{P};t) = H(\mathbf{q},\mathbf{p};0)$$

$$= H(\mathbf{q},(\partial S/\partial\mathbf{q})(\mathbf{q},\mathbf{P};t);0).$$

From $\partial/\partial\mathbf{Q} = (\partial^2 S/\partial\mathbf{P}\partial\mathbf{q})^{-1}\partial/\partial\mathbf{q}$ we obtain

$$\frac{\partial S}{\partial t} = \frac{\partial H^{(0)}}{\partial\mathbf{P}}\cdot\left(\frac{\partial^2 S}{\partial\mathbf{P}\partial\mathbf{q}}\right)^{-1}\frac{\partial H(\mathbf{q},\partial S/\partial\mathbf{q};0)}{\partial\mathbf{q}}\bigg|_{\mathbf{P}}, \quad (15)$$

where

$$\frac{\partial H(\mathbf{q},\partial S/\partial\mathbf{q};0)}{\partial\mathbf{q}}\bigg|_{\mathbf{P}} = \frac{\partial H}{\partial\mathbf{q}}\bigg|_{\mathbf{p}}\left(\mathbf{q},\frac{\partial S}{\partial\mathbf{q}};0\right)$$

$$+ \frac{\partial^2 S}{\partial\mathbf{q}^2}\frac{\partial H}{\partial\mathbf{p}}\bigg|_{\mathbf{q}}\left(\mathbf{q},\frac{\partial S}{\partial\mathbf{q}};0\right). \quad (16)$$

Note that, since $H(\mathbf{q},\mathbf{p};0)$ is simply our initial Hamiltonian, we do not need the evolution of $H$ in order to get the evolution of $S$.

The numerical evolution of $S$ is described in Appendix A. The numerical evolution of $H$ in the case of action-angle variables is discussed at length in Ref. [2]. Since the case for arbitrary canonical pairs is handled similarly, we do not give numerical details for the $H$ evolution in this paper.

## IV. THE FIRST-ORDER LIMIT

From our choice of $G$ in the previous section, it follows that in the limit of a first-order perturbation on an integrable Hamiltonian

$$\frac{\partial H_{\mathbf{k}}}{\partial t} = -4\pi^2(\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)})^2 H_{\mathbf{k}}. \quad (17)$$

Our first-order solution yields $H_{\mathbf{k}}(\mathbf{P};t) = H_{\mathbf{k}}(\mathbf{P};0)\exp[-4\pi^2(\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)})^2 t]$. Note that the more nonresonant a term is the faster the exponential decay. In particular, resonances are not affected at all. It may also be noted that the first-order evolution equation amounts to running the Hamiltonian through a heat equation.

We now turn to the evolution of $S$ in the first-order limit. To this end, write $S(\mathbf{q},\mathbf{P};t) = \mathbf{q}\cdot\mathbf{P} + \tilde{G}(\mathbf{q},\mathbf{P};t)$. In what follows we shall work to first order in $\tilde{G}$ and $H^{(1)}$. Then,

$$\frac{\partial H}{\partial\mathbf{q}}\bigg|_{\mathbf{p}}\left(\mathbf{q},\frac{\partial S}{\partial\mathbf{q}};0\right) = \frac{\partial H}{\partial\mathbf{q}}(\mathbf{q},\mathbf{P};0) + \frac{\partial^2 H}{\partial\mathbf{P}\partial\mathbf{q}}(\mathbf{q},\mathbf{P};0)\frac{\partial\tilde{G}}{\partial\mathbf{q}} \quad (18)$$

and

$$\frac{\partial^2 S}{\partial\mathbf{q}^2}\frac{\partial H}{\partial\mathbf{p}}\bigg|_{\mathbf{q}}\left(\mathbf{q},\frac{\partial S}{\partial\mathbf{q}};0\right) = \frac{\partial^2\tilde{G}}{\partial\mathbf{q}^2}\frac{\partial H}{\partial\mathbf{p}}\bigg|_{\mathbf{q}}\left(\mathbf{q},\mathbf{P}+\frac{\partial\tilde{G}}{\partial\mathbf{q}};0\right)$$

$$= \frac{\partial^2\tilde{G}}{\partial\mathbf{q}^2}\frac{\partial H}{\partial\mathbf{P}}(\mathbf{q},\mathbf{P};0). \quad (19)$$

Now, $H(\mathbf{q},\mathbf{p};0) = H^{(0)}(\mathbf{p};0) + H^{(1)}(\mathbf{q},\mathbf{p};0)$. Then $(\partial H/\partial\mathbf{q})(\mathbf{q},\mathbf{P};0) = (\partial H^{(1)}/\partial\mathbf{q})(\mathbf{q},\mathbf{P};0)$. To first order, $(\partial^2 H/\partial\mathbf{P}\partial\mathbf{q})(\mathbf{q},\mathbf{P};0)\partial\tilde{G}/\partial\mathbf{q} = (\partial^2 H^{(1)}/\partial\mathbf{P}\partial\mathbf{q})(\mathbf{q},\mathbf{P};0)\partial\tilde{G}/\partial\mathbf{q} = 0$. Finally, to first order, $(\partial^2\tilde{G}/\partial\mathbf{q}^2)(\partial H/\partial\mathbf{P})|_{\mathbf{q}}(\mathbf{q},\mathbf{P};0) = (\partial^2\tilde{G}/\partial\mathbf{q}^2)(\partial H^{(0)}/\partial\mathbf{P})(\mathbf{P};0)$. Since each of these terms is first order in either $H^{(1)}$ or $\tilde{G}$, in the first-order limit we take $(\partial^2 S/\partial\mathbf{P}\partial\mathbf{q})^{-1} = \mathbf{1}$. Putting everything together gives us our first-order equation

$$\frac{\partial\tilde{G}}{\partial t} = \frac{\partial H^{(0)}}{\partial\mathbf{P}}\cdot\left(\frac{\partial H^{(1)}}{\partial\mathbf{q}} + \frac{\partial^2\tilde{G}}{\partial\mathbf{q}^2}\frac{\partial H^{(0)}}{\partial\mathbf{P}}\right). \quad (20)$$

In Fourier space, this becomes

$$\frac{\partial\tilde{G}_{\mathbf{k}}}{\partial t} = 2\pi i(\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)})H_{\mathbf{k}}^{(1)}$$

$$-4\pi^2(\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)})^2\tilde{G}_{\mathbf{k}}. \quad (21)$$

Since $\tilde{G}_{\mathbf{k}}(\mathbf{P};0) = 0 \; \forall \; \mathbf{k}$, we obtain

$$\tilde{G}_{\mathbf{k}}(\mathbf{P};t) = \frac{i H_{\mathbf{k}}^{(1)}}{2\pi(\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)})}(1 - e^{-4\pi^2(\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)})^2 t}). \quad (22)$$

Note that $\tilde{G}_{\mathbf{k}}(\mathbf{P};t) = 0$ for all resonant terms. This can be seen by looking at the original ordinary differential equation from which the solution is derived, or equivalently by noting that $\lim_{\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)}\to 0}\{1 - \exp[-4\pi^2(\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)})^2 t]\}/(\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)}) = 0$. We can write

$$\tilde{G}(\mathbf{q},\mathbf{P};t) = \frac{1}{V}\frac{i}{2\pi}\sum_{\mathbf{k}\neq 0}\frac{H_{\mathbf{k}}^{(1)}}{\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)}}$$

$$\times(1 - e^{-4\pi^2(\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)})^2 t})e^{2\pi i\mathbf{k}\cdot\mathbf{q}}. \quad (23)$$

This series is convergent, because the exponential term prevents resonances and near-resonances in the denominator from causing the series to diverge. We can let $t\to\infty$ to get the first-order perturbation theory result

$$\tilde{G}(\mathbf{q},\mathbf{P};\infty) = \frac{1}{V}\frac{i}{2\pi}\sum_{\mathbf{k}\neq 0}\frac{H_{\mathbf{k}}^{(1)}}{\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)}}e^{2\pi i\mathbf{k}\cdot\mathbf{q}}, \quad (24)$$

where the sum is over all nonresonant $\mathbf{k}$. Note, however, that the convergence of the various Fourier components to their $t\to\infty$ limits is not uniform, because the time constant for the exponential term is proportional to $1/(\mathbf{k}\cdot\nabla_{\mathbf{P}}H^{(0)})^2$. This goes to infinity as $\mathbf{k}$ approaches a resonance. The above equation

must be solved for finite $t$, and then the $t\to\infty$ limit must be taken. This prevents any ambiguities in $\tilde{G}$ [6].

We should point out that we have not derived a first-order classical perturbation theory that is necessarily convergent. The standard first-order formula is obtained by taking the $t\to\infty$ limit of our expression, and as can be clearly seen the presence of near-resonances can result in a divergent series. However, the advantage of our formula is that, for all finite $t$, the near-resonances are attenuated by the exponential term in such a way that the series does converge. Thus, this method does not require any *a priori* removal of terms that are assigned as "resonant." Rather, all terms may be included in the first-order expression, and $t$ may be chosen to be as large as possible without any Fourier term exceeding some cutoff criterion. In this way, the nonresonant Fourier terms will essentially have their standard first-order values, while the more nearly resonant terms will still have a non-negligible exponential term present which keeps the overall series convergent. The corresponding nonresonant Fourier terms in the Hamiltonian will then have been essentially removed, while the more nearly resonant terms will have been reduced somewhat, but not completely.

## V. A NUMERICAL EXAMPLE

We chose to test the $S$ evolution numerically on the two-dimensional Pullen-Edmonds Hamiltonian [8], given by

$$H(x,y,p_x,p_y)=\frac{p_x^2+p_y^2}{2}+\frac{1}{2}(\omega_x^2 x^2+\omega_y^2 y^2)+\epsilon x^2 y^2.$$
(25)

This was the two-dimensional system studied numerically in Ref. [2] in testing the GDA method. The results there were compared with those obtained in Ref. [3] using the CHM method. However, while in Ref. [2] the evolution was done on $H$ in the context of obtaining a semiclassical quantum spectrum, in this case it is the PDE for $S$ that is being tested.

We set $\omega_x=\omega_y=1$. The $t=0$ canonical representation is simply the harmonic oscillator action-angle basis, denoted by $(\theta_x,\theta_y,J_x,J_y)$. The arbitrary $t$ canonical representation is denoted by $(\phi_x,\phi_y,I_x,I_y)$, so that $S=S(\theta_x,\theta_y,I_x,I_y;t)$, with $J_x=\partial S/\partial\theta_x$ and $J_y=\partial S/\partial\theta_y$. The transformation to the harmonic representation is obtained by setting $x=\sqrt{J_x/\pi}\cos 2\pi\theta_x$, $p_x=-\sqrt{J_x/\pi}\sin 2\pi\theta_x$, and similarly for $y,p_y$. The result is

$$H(\theta_x,\theta_y,J_x,J_y)=\frac{1}{2\pi}(J_x+J_y)+\frac{\epsilon J_x J_y}{4\pi^2}\{1+\cos 4\pi\theta_x$$
$$+\cos 4\pi\theta_y+\tfrac{1}{2}[\cos 4\pi(\theta_x+\theta_y)$$
$$+\cos 4\pi(\theta_x-\theta_y)]\}.$$
(26)

The only nonvanishing Fourier components are $\pm(2,0)$, $\pm(0,2),\pm(2,2)$, and $\pm(2,-2)$. Note in particular that $\pm(2,-2)$ is a resonance.

Of the three prescriptions for choosing $G$, the simplest one to use is $H^{(0)}=H_0=(1/2\pi)(1,1)\cdot(I_x,I_y)$, so that

$\nabla_{\mathbf{p}}H^{(0)}$ is taken to be $(1/2\pi)(1,1)$. In addition, for this Hamiltonian it is readily verified that

$$\frac{\partial H}{\partial\mathbf{q}}\bigg|_{\mathbf{p}}\left(\mathbf{q},\frac{\partial S}{\partial\mathbf{q}};0\right)=-\frac{\epsilon}{\pi}\frac{\partial S}{\partial\theta_x}\frac{\partial S}{\partial\theta_y}\{\sin 4\pi\theta_x$$
$$+\tfrac{1}{2}[\sin 4\pi(\theta_x+\theta_y)$$
$$+\sin 4\pi(\theta_x-\theta_y)],\sin 4\pi\theta_y$$
$$+\tfrac{1}{2}[\sin 4\pi(\theta_x+\theta_y)-\sin 4\pi(\theta_x-\theta_y)]\}$$
(27)

and

$$\frac{\partial H}{\partial\mathbf{p}}\bigg|_{\mathbf{q}}=\frac{1}{2\pi}(1,1)+\frac{\epsilon}{4\pi^2}\{1+\cos 4\pi\theta_x+\cos 4\pi\theta_y$$
$$+\tfrac{1}{2}[\cos 4\pi(\theta_x+\theta_y)+\cos 4\pi(\theta_x-\theta_y)]\}$$
$$\times\left(\frac{\partial S}{\partial\theta_y},\frac{\partial S}{\partial\theta_x}\right).$$
(28)

We substitute into Eq. (16) and then into Eq. (15) to get the PDE for $S$.

We set $\epsilon=0.05$, and $I_x=I_y=9\pi$. This corresponds to a torus with zeroth-order energy $E=9$ at $t=0$. The system is nearly integrable in this regime [3], yet a study of the semiclassically obtained energy spectrum [2,3] shows clear differences from first-order perturbation theory. We therefore test our PDE beyond the first-order limit with this example.

We set $N=20$, $DX=0.1$, $DT=0.034$, and $GDSZ=9$, which allowed us to propagate the PDE out to a time of 0.306 (the meaning of these parameters is given in Appendix A). The rate of change of $S$ on the grid reaches a minimum at this point [which was determined by tracking $\sqrt{\langle(\partial S/\partial t)^2\rangle}$ on the grid], so the evolution was stopped here. Because our PDE approach is only the gradient-descent prescription in the first-order limit, for finite perturbations there is no reason to expect the evolution to reach steady state. We discuss this issue further in Appendix C.

Figures 1(a)–1(d) show the results of the evolution at times $t=0.0$, 0.1, 0.2, and 0.3. It should be noted that we are not plotting $S$ in these graphs. Rather, we plot $H(\theta_x,\theta_y,\partial S/\partial\theta_x,\partial S/\partial\theta_y)$. There are two reasons for this. First, because we are working in the weakly perturbed, though still beyond first-order, regime, $S$ remains fairly close to the identity transformation. The effect on $H(\theta_x,\theta_y,\partial S/\partial\theta_x,\partial S/\partial\theta_y)$, however, is dramatic, and so it is much more convenient to represent the evolution of $S$ in this indirect fashion. Second, even if the perturbation were sufficiently strong to significantly deform $S$, the only way to determine if the deformation of $S$ is correct is to look at its effect on $H$.

Using the formula derived in Appendix B, it is readily shown that the first-order solution to $H(\theta_x,\theta_y,\partial S/\partial\theta_x,\partial S/\partial\theta_y)$ is
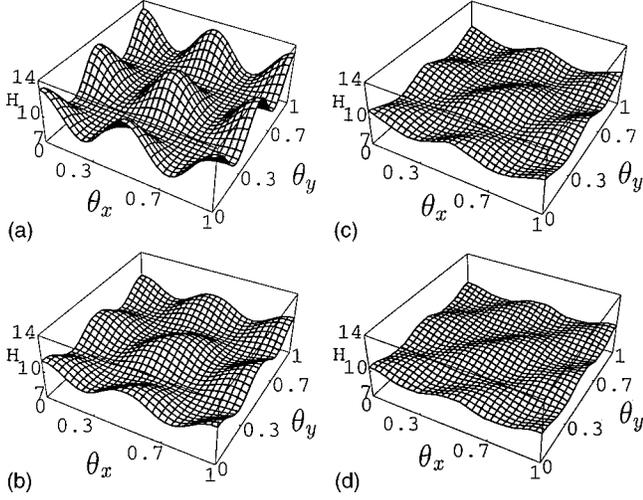
FIG. 1. Plot of $H(\theta_x, \theta_y, \partial S/\partial \theta_x, \partial S/\partial \theta_y)$ at various times for the Pullen-Edmonds Hamiltonian (arbitrary units). (a) $t=0.0$, (b) $t=0.1$, (c) $t=0.2$, and (d) $t=0.3$.

$$H\left(\theta_x, \theta_y, \frac{\partial S}{\partial \theta_x}, \frac{\partial S}{\partial \theta_y}\right) = \frac{1}{2\pi}(I_x + I_y) + \frac{\epsilon I_x I_y}{4\pi^2}\{1 + e^{-4t}$$

$$\times (\cos 4\pi\theta_x + \cos 4\pi\theta_y)$$

$$+ \tfrac{1}{2}[\cos 4\pi(\theta_x - \theta_y)$$

$$+ e^{-16t}\cos 4\pi(\theta_x + \theta_y)]\}. \quad (29)$$

While the perturbation is sufficiently strong that there are clear differences from first-order perturbation theory, the perturbation is still sufficiently weak that the first-order result provides a qualitative and semiquantitative picture of how the evolution should proceed. Letting $t \to \infty$, we see that the long-time limit of the evolution gives

$$H\left(\theta_x, \theta_y, \frac{\partial S}{\partial \theta_x}, \frac{\partial S}{\partial \theta_y}\right)\Bigg|_{t=\infty} = \frac{1}{2\pi}(I_x + I_y)$$

$$+ \frac{\epsilon I_x I_y}{4\pi^2}[1 + \tfrac{1}{2}\cos 4\pi(\theta_x - \theta_y)].$$

$$(30)$$

This is plotted for our specific set of parameters in Fig. 2. Note that the numerical evolution does indeed deform $S$ in
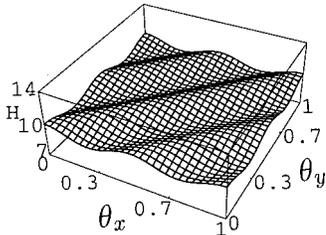


FIG. 2. Plot of $H(\theta_x, \theta_y, \partial S/\partial \theta_x, \partial S/\partial \theta_y)$ for the Pullen-Edmonds Hamiltonian for $t=\infty$ in the first-order limit (arbitrary units).
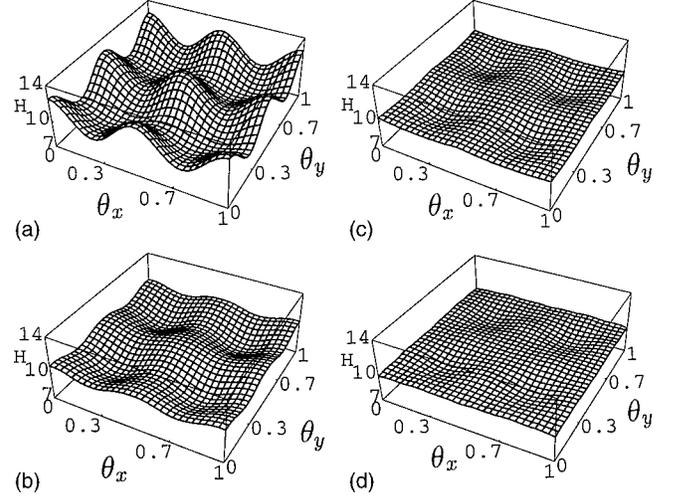


FIG. 3. Plot of $H(\theta_x, \theta_y, \partial S/\partial \theta_x, \partial S/\partial \theta_y)$ at various times for the Pullen-Edmonds Hamiltonian with the $(2,-2)$ resonance removed (arbitrary units). (a) $t=0.0$, (b) $t=0.1$, (c) $t=0.2$, and (d) $t=0.3$.

such a way that the graphs in Figs. 1(a)–1(d) evolve to look like the graph in Fig. 2. Without the $(2,-2)$ resonance, $H$ is integrable at this energy [3], so it is possible to transform to a basis in which $H$ depends only on $(I_x, I_y)$. Instead, the evolution eliminates as much of the nonresonant behavior as possible, but the resonant angle dependence arising from the $(2,-2)$ term remains. In Figs. 3(a)–3(d) we present the results of the evolution on the Hamiltonian obtained by removing the $(2,-2)$ resonance from the Pullen-Edmonds term. We changed $DT$ to 0.033 and $GDSZ$ to 11. All other parameters are otherwise unchanged. In this case, the evolution should flatten $H$, and as Figs. 3(a)–3(d) confirm, this is exactly what happens.

## VI. CONSTRUCTION OF SEMICLASSICAL WAVE FUNCTIONS

We applied this PDE-based approach to construct optimized semiclassical wave functions for Gaussian bump potentials. Thus, we considered a Hamiltonian of the form

$$H(x,y,p_x,p_y) = \frac{p_x^2 + p_y^2}{2} + \lambda \exp[-(ax^2 + 2cxy + by^2)].$$

$$(31)$$

We considered two different potentials. For potential I, we took $\lambda_1 = 1, a = b = 1, c = 0$. For potential II, we took $\lambda_2 = 1, a = b = 3, c = 2$. This corresponds to the potential $\exp[-(x^2 + 5y^2)]$, with the $xy$ axis rotated clockwise by 45°.

The $S$ evolution was solved for $P_x = 1.5, P_y = 0.0$, corresponding to a particle momentum of $(1.5, 0.0)$ far away from the potential. Thus, the particle energy is above the bump height for both potentials. We worked on the spatial grid $[-4,4] \times [-4,4]$ with a step size of 0.2. We used a momentum step size of 0.05. Starting at time $t=0$, we propagated out to $t=0.2$ for potential I and $t=0.15$ for potential II using a time step of 0.01. In both cases, the rate of change of $S$ on

(a)

(b)
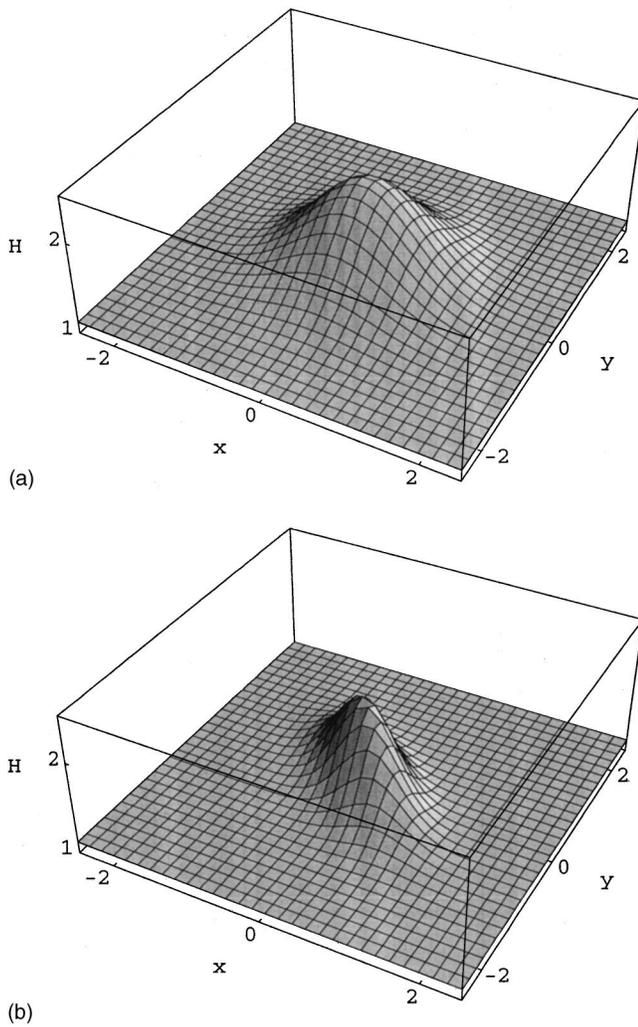
FIG. 4. Plot of $H(x,y,\partial S/\partial x,\partial S/\partial y)$ at $t=0$ for potentials I and II (arbitrary units). (a) Potential I and (b) potential II.

the grid was at least 50% smaller at the end of the evolution than at the beginning. Our choice for $G$ was still able to optimize the canonical basis for this Hamiltonian, even though it was obtained based on first-order considerations, and we were at an energy only slightly above the barrier.

Figures 4(a) and 4(b) plot $H(x,y,\partial S/\partial x,\partial S/\partial y)$ for the two potentials at the start of their respective evolutions. Up to the kinetic energy term $(P_x^2+P_y^2)/2$, these graphs show the shape of the Gaussian bumps for potentials I and II, respectively. Figures 5(a) and 5(b) are probability density plots of the corresponding wave functions. White regions correspond to enhanced probability density, while dark regions correspond to depleted probability density. In both cases, the maximum probability density change from the background value of 1 is from 20% to 30%. Note how the wave functions reveal some of the underlying classical dynamics. For the symmetric Gaussian bump of potential I, classical trajectories are deflected around the bump. Thus, there should be a depletion of probability in the center of the bump, and an enhancement in the region around the bump. As Fig. 5(a) illustrates, this is exactly what happens. A similar situation occurs in Fig. 5(b), corresponding to potential II. This time,



(a)

(b)
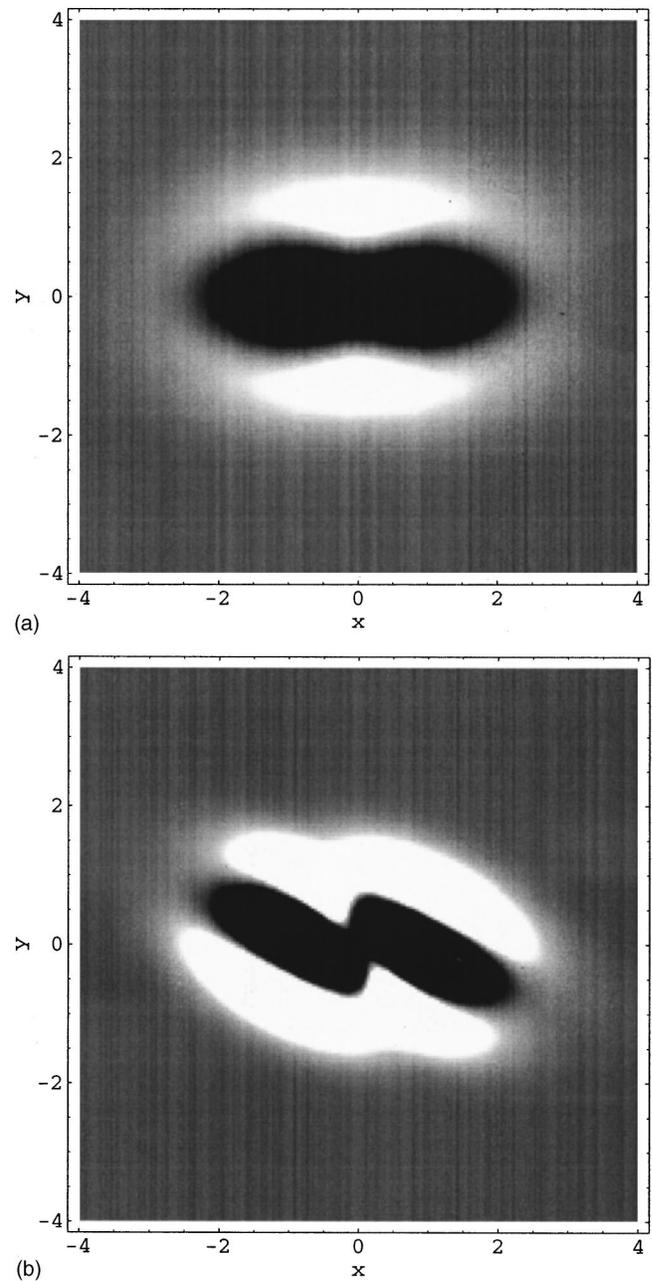
FIG. 5. Probability density plots of the semiclassical wave functions generated from the evolutions on potentials I and II (arbitrary units). (a) The wave function for potential I and (b) the wave function for potential II.

however, the distribution is angled somewhat along the direction of the potential, which should be expected classically.

The one feature that may not be physically intuitive is the symmetry of the probability distribution. It may be shown [6] that the overall generating function for these potentials is not unique, but depends on the choice of boundary conditions at infinity. Our implementation of the PDE for $S$ selects the boundary condition corresponding to a momentum field which angles the trajectories slightly inward at $x=-\infty$, so that the momentum field curves in as trajectories approach the bump, and then curves away as trajectories are deflected

by the bump. While future work would attempt to allow for different boundary conditions numerically, we have found the current implementation of the PDE for $S$ to be the most stable.

## VII. CONCLUSIONS AND FUTURE RESEARCH

This paper presented a PDE-based, phase-space deformation approach to optimize the canonical basis with which to globally represent a nearly integrable Hamiltonian. Because this method reduces to a gradient-descent approach for optimizing the canonical basis in the first-order limit, it was called the GDA method in Ref. [2], where it was initially applied to determine the eigenenergies of systems with two, four, and six degrees of freedom. The GDA method is compatible with any canonical representation of phase space, and allows the construction and visualization of the semiclassical wave functions corresponding to the optimized canonical basis for the given Hamiltonian.

As was mentioned in the Introduction, the motivation for this work is derived from semiclassical quantum mechanics. In light of the results for the $H$ evolution in Ref. [2], this method may be useful in understanding vibrational dynamics in polyatomic molecules. It may also be useful as a way to construct distorted-wave basis sets for scattering calculations, and hence may find application in mesoscopic physics. It should be noted in this regard that this PDE-based approach contains other methods as subcases. For example, in Ref. [9], Maitra and Heller used one-dimensional WKB wave functions as a distorted-wave basis for computing above-barrier reflection coefficients. It may be shown [6] that for all above-barrier energies our PDE method has a unique steady state which exactly coincides with the one-dimensional WKB wave functions for all above-barrier energies, so that the Maitra-Heller method is contained within the GDA method. In their paper, Maitra and Heller raised the issue of generalizing their technique to higher dimensions, and to action-angle systems. Our PDE-based approach is exactly this generalization [11–31].

At this point, we have not yet applied the GDA method to actual systems. The main difficulty in dealing with vibrational calculations for polyatomics is that it is currently difficult to obtain accurate vibrational potential energy surfaces, even for few-atom polyatomics. However, it would be useful to apply the GDA method to actual systems at some point. In addition, while we wrote in the Introduction that we believe the GDA approach is significantly faster than previous canonical basis optimization algorithms (at least for the $H$ evolution), we also wrote that in fairness no speed comparisons have been made to date. Thus, another potentially useful study would be to compare the GDA method with the various other methods on the market.

To conclude, we should add that Heller has often made the comparison between the separatrix region generated by a local potential bump in one dimension to the resonance zone structure in a Poincaré surface of section of a nearly integrable Hamiltonian [9,10]. Indeed, Heller regards the above-barrier reflection problem as a prototype for the more complicated case of dynamical tunneling between invariant tori

facilitated by resonance zones. Because both types of systems can be treated within the same PDE-based approach, they are, in fact, formally equivalent phenomena.

## APPENDIX A: NUMERICAL PROPAGATION OF $S$

If $H$ is given analytically, then the derivatives of $H$ can also be determined analytically. Therefore, it can be seen from Eqs. (15) and (16) that the numerical propagation of $S$ only requires the numerical evaluation of the partial derivatives of $S$. This is done using centered differences.

At time $t=0$, $S$ is simply the identity transformation, so that $S(\mathbf{q},\mathbf{P};0)=\mathbf{q}\cdot\mathbf{P}$. Thus, $S$ is not periodic in each $q_i$ with period $L_i$. Define $\mathbf{L_n}=(n_1L_1,\ldots,n_DL_D)$, and note, however, that $S(\mathbf{q}+\mathbf{L_n},\mathbf{P};0)=\mathbf{P}\cdot\mathbf{L_n}+S(\mathbf{q},\mathbf{P};0)$. We claim that this property is preserved by the evolution. We shall assume this for what follows, and then prove it at the end of this section. Thus, although $S$ is not periodic in the $q_i$'s, we still need only track $S$ for $\mathbf{q}$ in a $D$-dimensional box of side lengths $L_1,\ldots,L_D$.

The $\mathbf{q}$ grid is given by $\{(n_1L_1/N,\ldots,n_DL_D/N)|n_i=0,\ldots,N-1\}$, giving $N^D$ grid points. We track all $\mathbf{P}$ on a grid of canonical momenta about some central momentum $\mathbf{P}_0$, where our grid consists of all canonical momenta $\mathbf{P_k}=\mathbf{P}_0+DX\mathbf{k}$, with $\mathbf{k}=(k_1,\ldots,k_D)$ satisfying $|k_1|+\ldots+|k_D|\leq GDSZ$. Let us denote this set by $\Omega(\mathbf{P}_0,GDSZ)$. Since our evolution involves a first derivative in $\mathbf{P}$ of $S$, we cannot compute $\partial S/\partial t$ at the boundary of the $\mathbf{P}$ grid. The result is that we can only propagate on $\Omega(\mathbf{P}_0,GDSZ-1)$, so that at each iteration the value of $GDSZ$ shrinks by 1. This collapsing boundary method is described in further detail in Ref. [2], since it also arises naturally in the numerical implementation of the $H$ evolution. The absence of any boundary condition for the $\mathbf{P}$ grid is due to the fact that there are simply no physically natural boundary conditions to impose. This is in contrast to the heat equation, for example, where fixing the temperature at the boundary is physically realized by immersing the system in a constant temperature bath.

Once $\partial S/\partial t$ has been evaluated on all possible grid points, the propagation by some time step $DT$ is done using the explicit Euler method, which means that we set $S(\mathbf{q},\mathbf{P}_k;t+DT)=S(\mathbf{q},\mathbf{P}_k;t)+DT(\partial S/\partial t)(\mathbf{q},\mathbf{P}_k;t)$.

Finally, suppose we are considering a system with unbound degrees of freedom, that is, some of the $L_i=0$. Then we track those $q_i\in\{q_{i0}\pm n\Delta|n=0,\ldots,N_i\}$. At each time step, we can only compute $\partial S/\partial t$ up to $n=N_i-1$, so that after each time step we shrink our set of $q_i$ by decreasing $N_i$

by 1. While in this case it may be possible to impose more natural boundary conditions on the system, in practice it is numerically most stable to have a free boundary, as is done with the $\mathbf{P}$ grid.

We will now prove that $S(\mathbf{q}+\mathbf{L_n},\mathbf{P};t)=\mathbf{P}\cdot\mathbf{L_n}+S(\mathbf{q},\mathbf{P};t)$. To do this, we make the key assumption that the numerical propagation outlined above converges to the exact solution of the PDE in the limit of the time and spatial steps approaching 0. Specifically, given some time $T>0$ to which we wish to propagate the PDE, we divide the time grid into step sizes of length $DT=T/N$, where $N$ is an integer that we let go to $\infty$. At some $\mathbf{q},\mathbf{P}$, we construct $S(\mathbf{q},\mathbf{P};t)$ by constructing $S_n(\mathbf{q},\mathbf{P})\equiv S(\mathbf{q},\mathbf{P};nDT)$, where $n=0,\ldots,N$, and for $t\in[0,T]$ we define $S(\mathbf{q},\mathbf{P};t)$ to be the linear interpolation of $(S_0(\mathbf{q},\mathbf{P}),\ldots,S_N(\mathbf{q},\mathbf{P}))$ on $[0,T]$. Clearly, if we can show that $S_n(\mathbf{q}+\mathbf{L_n},\mathbf{P})=\mathbf{P}\cdot\mathbf{L_n}+S_n(\mathbf{q},\mathbf{P})$, then by interpolation our claim holds $\forall\ t\in[0,T]$. We prove this by induction.

By definition, $S_n(\mathbf{q}+\mathbf{L_n},\mathbf{P})=\mathbf{P}\cdot\mathbf{L_n}+S_n(\mathbf{q},\mathbf{P})$ for $n=0$, so assume the result holds for some $n\geq0$. To prove that it holds for $n+1$, we may note that

$$S_{n+1}(\mathbf{q}+\mathbf{L_n},\mathbf{P})=S_n(\mathbf{q}+\mathbf{L_n},\mathbf{P})+\frac{\partial S}{\partial t}(\mathbf{q}+\mathbf{L_n},\mathbf{P};nDT)DT$$

$$=\mathbf{P}\cdot\mathbf{L_n}+S_n(\mathbf{q},\mathbf{P})$$

$$+\frac{\partial S}{\partial t}(\mathbf{q}+\mathbf{L_n},\mathbf{P};nDT)DT. \qquad (A1)$$

Now, from $S_n(\mathbf{q}+\mathbf{L_n},\mathbf{P})=\mathbf{P}\cdot\mathbf{L_n}+S_n(\mathbf{q},\mathbf{P})$ we obtain that $(\partial S/\partial\mathbf{P})(\mathbf{q}+\mathbf{L_n},\mathbf{P};nDT)=\mathbf{L_n}+(\partial S/\partial\mathbf{P})(\mathbf{q},\mathbf{P};nDT)$, so the periodicity of $G$ then implies that $(\partial S/\partial t)(\mathbf{q}+\mathbf{L_n},\mathbf{P};nDT)=(\partial S/\partial t)(\mathbf{q},\mathbf{P};nDT)$. Then Eq. (A1) becomes

$$S_{n+1}(\mathbf{q}+\mathbf{L_n},\mathbf{P})=\mathbf{P}\cdot\mathbf{L_n}+S_n(\mathbf{q},\mathbf{P})+\frac{\partial S}{\partial t}(\mathbf{q},\mathbf{P};nDT)DT$$

$$=\mathbf{P}\cdot\mathbf{L_n}+S_{n+1}(\mathbf{q},\mathbf{P}), \qquad (A2)$$

thereby completing the induction step, and proving our claim.

## APPENDIX B: AN ADDITIONAL FIRST-ORDER RESULT

In this section we derive the first-order result for $H(\mathbf{q},\partial S/\partial\mathbf{q};0)$. Following the procedure in Sec. IV, we write $S=\mathbf{q}\cdot\mathbf{P}+\widetilde{G}(\mathbf{q},\mathbf{P};t)$. We also write $H(\mathbf{q},\mathbf{p};0)=H^{(0)}(\mathbf{p};0)+H^{(1)}(\mathbf{q},\mathbf{p};0)$. Using $\mathbf{p}=\partial S/\partial\mathbf{q}=\mathbf{P}+\nabla_\mathbf{q}\widetilde{G}(\mathbf{q},\mathbf{P};t)$, we get to first order that

$$H(\mathbf{q},\mathbf{p};0)=H^{(0)}(\mathbf{P};0)+\nabla_\mathbf{P}H^{(0)}(\mathbf{P};0)\cdot\nabla_\mathbf{q}\widetilde{G}(\mathbf{q},\mathbf{P};t)$$

$$+H^{(1)}(\mathbf{q},\mathbf{P};0). \qquad (B1)$$

For simplicity, we assume that $G$ was chosen using $H^{(0)}(\mathbf{P})=H^{(0)}(\mathbf{P};0)$. As mentioned before, in the first-order limit all three prescriptions for choosing $G$ are equivalent.

Using the other two prescriptions will lead to at most second-order corrections in our final result. From Eq. (23) we get that

$$\nabla_\mathbf{P}H^{(0)}\cdot\nabla_\mathbf{q}\widetilde{G}(\mathbf{q},\mathbf{P};t)=-\frac{1}{V}\sum_{\mathbf{k}\neq\mathbf{0}}H_\mathbf{k}^{(1)}$$

$$\times(1-e^{-4\pi^2(\mathbf{k}\cdot\nabla_\mathbf{P}H^{(0)})^2t})e^{2\pi i\mathbf{k}\cdot\mathbf{q}}$$

$$=-H^{(1)}+\frac{1}{V}\sum_{\mathbf{k}\neq\mathbf{0}}H_\mathbf{k}^{(1)}$$

$$\times e^{-4\pi^2(\mathbf{k}\cdot\nabla_\mathbf{P}H^{(0)})^2t}e^{2\pi i\mathbf{k}\cdot\mathbf{q}}, \qquad (B2)$$

and so we obtain

$$H\left(\mathbf{q},\frac{\partial S}{\partial\mathbf{q}}(\mathbf{q},\mathbf{P};t);0\right)=H^{(0)}(\mathbf{P};0)+\frac{1}{V}\sum_{\mathbf{k}\neq\mathbf{0}}H_\mathbf{k}^{(1)}(\mathbf{P};0)$$

$$\times e^{-4\pi^2(\mathbf{k}\cdot\nabla_\mathbf{P}H^{(0)})^2t}e^{2\pi i\mathbf{k}\cdot\mathbf{q}}. \qquad (B3)$$

## APPENDIX C: PROPAGATION TIME

Recall from Eq. (12) that the first-order expression for the evolution of $H$ in Fourier space is

$$\frac{\partial H_\mathbf{k}}{\partial t}=2\pi i(\mathbf{k}\cdot\nabla_\mathbf{P}H^{(0)})G_\mathbf{k}. \qquad (C1)$$

This equation was then used to obtain the gradient-descent prescription for choosing $G$ in the first-order limit. For weak perturbations, this prescription no longer coincides with the gradient-descent approach, but should still shrink the nonzero Fourier components of $H$. This will occur as long as the right side of Eq. (C1) [or Eq. (12)] is sufficiently dominant compared to the remaining terms in the full PDE for $H$. Note then that for resonances and near-resonances this condition does not hold. However, for sufficiently nonresonant terms this condition does hold. Thus, in general, for a weak perturbation, our PDE-based approach starts out by decreasing the more nonresonant terms of $H$. The $\mathbf{Q}$ dependence of $H$ starts decreasing, and so the rate of change of $S$ decreases as well as the evolution proceeds. Eventually, the sufficiently nonresonant terms of $H$ are reduced to a point where higher-order terms become important, so that our first-order gradient-descent prescription for choosing $G$ will no longer work to reduce the $\mathbf{Q}$ dependence of $H$. The rate of change of $S$ then begins to increase after this point, and eventually the PDE becomes numerically unstable. By tracking $\sqrt{\langle(\partial S/\partial t)^2\rangle}$ on the grid, it is possible to stop the evolution where the rate of change of $S$ reaches its minimum, and consequently where the canonical basis has been optimized.

Of course, the weaker the perturbation, the closer a given $\mathbf{k}$ must be to a resonance for our choice of $G$ to no longer work to reduce the corresponding $H_\mathbf{k}$. Furthermore, the weaker the perturbation, the longer it is possible to propagate the PDE before $\sqrt{\langle(\partial S/\partial t)^2\rangle}$ reaches its minimum, and the more closely this minimum will correspond to a steady state. It would be interesting to develop a simple criterion to estimate at what time this minimum occurs, and how far away the system is from steady state at the minimum.

[1] C. Jaffé and P. Brumer, J. Chem. Phys. **82**, 2330 (1985); C. Jaffé, *ibid.* **88**, 7603 (1988).

[2] E. Tannenbaum and E. J. Heller, J. Phys. Chem. A **105**, 2803 (2001).

[3] M. Carioli, E. J. Heller, and K. B. Moller, J. Chem. Phys. **106**, 8564 (1997).

[4] S. Chapman, B. C. Garrett, and W. H. Miller, J. Chem. Phys. **64**, 502 (1976).

[5] Herbert Goldstein, *Classical Mechanics*, 2nd ed. (Addison-Wesley, Reading, MA, 1980).

[6] E. Tannenbaum, Ph.D. thesis, Harvard University, 2002.

[7] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd ed. (Springer-Verlag, New York, 1989).

[8] R. A. Pullen and A. R. Edmonds, J. Phys. A **14**, L477 (1981).

[9] N. T. Maitra and E. J. Heller, Phys. Rev. A **54**, 4763 (1996).

[10] E. J. Heller, J. Phys. Chem. **99**, 2625 (1995).

[11] O. Brodier, P. Schlagheck, and D. Ullmo, Phys. Rev. Lett. **87**, 064101 (2001).

[12] R. L. Warnock and R. D. Ruth, Physica D **26**, 1 (1987).

[13] A. A. Stuchebrukhov and R. A. Marcus, J. Chem. Phys. **98**, 6044 (1993).

[14] M. Kaasalainen and J. Binney, Phys. Rev. Lett. **73**, 2377 (1994).

[15] D. M. Leitner and P. G. Wolynes, Phys. Rev. Lett. **79**, 55 (1997).

[16] B. Ramachandran and K. G. Kay, J. Chem. Phys. **99**, 3659 (1993).

[17] E. L. Sibert III, J. Chem. Phys. **88**, 4378 (1988).

[18] E. L. Sibert III and A. B. McCoy, J. Chem. Phys. **105**, 469 (1996).

[19] R. E. Wyatt, Adv. Chem. Phys. **73**, 231 (1989).

[20] R. E. Wyatt, C. Iung, and C. Leforestier, J. Chem. Phys. **97**, 3458 (1992); **97**, 3477 (1992).

[21] S. M. Lederman and R. A. Marcus, J. Chem. Phys. **88**, 6312 (1988).

[22] S. M. Lederman, S. J. Kippenstein, and R. A. Marcus, Chem. Phys. Lett. **146**, 7 (1988).

[23] R. T. Lawton and M. S. Child, Mol. Phys. **37**, 1799 (1979); **44**, 709 (1981).

[24] M. J. Davis and E. J. Heller, J. Chem. Phys. **75**, 246 (1981).

[25] D. Farelly and T. Uzer, J. Chem. Phys. **85**, 308 (1986).

[26] A. M. Ozorio de Almeida, J. Phys. Chem. **88**, 6139 (1984).

[27] W. H. Miller, Adv. Chem. Phys. **25**, 69 (1974).

[28] Edward Ott, *Chaos in Dynamical Systems* (Cambridge University Press, New York, 1993).

[29] James E. Bayfield, *Quantum Evolution: An Introduction to Time-Dependent Quantum Mechanics* (Wiley, New York, 1999).

[30] Martin C. Gutzwiller, *Chaos in Classical and Quantum Mechanics* (Springer-Verlag, New York, 1990).

[31] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. (Cambridge University Press, New York, 1992).